

Міністерство освіти і науки України
Національний аерокосмічний університет
«Харківський авіаційний інститут»
Міністерство освіти і науки України
Національний аерокосмічний університет
«Харківський авіаційний інститут»

Кваліфікаційна наукова
праця на правах рукопису

КРИВЕНКО СЕРГІЙ СТАНІСЛАВОВИЧ

УДК 004.932:528.8:616-071(043)

ДИСЕРТАЦІЯ

МЕТОДИ ПРОГНОЗУВАННЯ ТА ІНТЕЛЕКТУАЛЬНОЇ ОБРОБКИ
ЗОБРАЖЕНЬ У СИСТЕМАХ ДИСТАНЦІЙНОГО ЗОНДУВАННЯ ТА
МЕДИЧНОЇ ДІАГНОСТИКИ

05.12.17 – Радіотехнічні та телевізійні системи

17 – Електроніка та телекомунікації

Подається на здобуття наукового ступеня доктора технічних наук

Дисертація містить результати власних досліджень. Використання ідей, результатів і текстів інших авторів мають посилання на відповідне джерело. Всі частини тексту дисертації, під час написання яких використовувалися технології штучного інтелекту, перевірені та відредаговані мною особисто



Сергій КРИВЕНКО

Науковий консультант Лукін Володимир Васильович, доктор технічних наук,
професор

Харків – 2025

АНОТАЦІЯ

Кривенко С.С. Методи прогнозування та інтелектуальної обробки зображень у системах дистанційного зондування та медичної діагностики. – Кваліфікаційна наукова праця на правах рукопису.

Дисертація на здобуття наукового ступеня доктора технічних наук за спеціальністю 05.12.17 – радіотехнічні та телевізійні системи. – Національний аерокосмічний університет «Харківський авіаційний інститут» Міністерства освіти і науки України, Харків, 2025.

Дисертаційну роботу присвячено ефективним інтелектуальній фільтрації та стисненню із втратами зображень дистанційного зондування Землі та медичної діагностики в умовах присутності складних завад і обмежених обчислювальних ресурсів з урахуванням візуальної якості.

Сучасний етап розвитку науки та технологій супроводжується стрімким зростанням обсягів даних, що генеруються у різних галузях людської діяльності, зокрема у дистанційному зондуванні Землі (ДЗЗ) та медичній діагностиці. Ці дані відіграють критично важливу роль у прийнятті рішень, що впливають на суспільне здоров'я, безпеку та збереження природного середовища. Водночас ефективне використання цих даних ускладнюється через обмеження у швидкості передачі інформації, ємності сховищ, а також необхідність забезпечення високої якості обробки та інтерпретації отриманих сигналів та зображень.

Дистанційне зондування Землі є основним джерелом отримання інформації про стан довкілля, прогнозування кліматичних змін, а також моніторингу надзвичайних ситуацій. Для досягнення цієї мети необхідні методи, що забезпечують: автоматичну обробку мультиспектральних та гіперспектральних зображень; адаптацію алгоритмів обробки до шумів і спотворень; швидку передачу великих обсягів даних з супутників.

Сфера медичної діагностики зазнає значних змін під впливом розвитку телемедичних технологій. У цьому контексті виникає необхідність у забезпеченні збереження діагностичної якості медичних зображень: при їх компресії; розробці алгоритмів для автоматичного виявлення патологій на основі обробки сигналів;

стандартизації процесів фільтрації та відновлення біомедичних сигналів для зменшення впливу шумів.

Реалізація підходів обробки потребує вирішення низки технічних викликів оптимізації алгоритмів для: роботи на пристроях із обмеженими ресурсами; забезпечення узгодженості між якістю оброблених даних і їх об'ємом; створення універсальних моделей, які можуть адаптуватися до різних типів сигналів і зображень.

У дисертаційній роботі вирішено важливу науково-практичну проблему ефективного інтелектуального оброблення (фільтрації та стиснення із втратами) зображень дистанційного зондування Землі та медичної діагностики в умовах присутності складних завад і обмежених обчислювальних ресурсів з урахуванням візуальної якості.

У вступі обґрунтовується актуальність теми дисертаційної роботи. Визначено мету роботи, основні задачі та методи досліджень. Сформульовано наукову новизну і практичне значення отриманих результатів.

В першому розділі проведено аналіз сучасного стану методів фільтрації та стиснення із втратами. Визначено основні задачі фільтрації та стиснення зображень, критерії якості обробки зображень. Окремо розглянуто візуальну якість зображень, як основу оцінки ефективності майбутніх досліджень. Також розглянуто шуми на зображеннях, їхню природу на зображеннях дистанційного зондування та медичної діагностики, а також існуючі засоби їхнього придушення.

В другому розділі розглянуто питання, що стосуються фільтрації зображень. В першу чергу розглянуто питання детектування текстур на спотворених завадами зображеннях за допомогою комбінування декількох параметрів локальної активності. Крім цього, розглянуто адаптивну фільтрацію зображень, спотворених сигнально-залежними та просторово-корельованими завадами, за допомогою алгоритмів машинного навчання. В останньому підрозділі розглянуто питання оцінки ефективності фільтрації для різних типів зображень.

В третьому розділі розглянуті питання, що стосуються двоетапних процедур забезпечення бажаної якості зображень. Крім того, розгадаються питання

стиснення зображень із забезпеченням візуально непомітних спотворень. Окрему увагу приділено адаптивному стисненню багатоканальних зображень дистанційного зондування Землі. Також розглянуто питання дивних зображень дистанційного зондування і їхніх властивостей.

У четвертому розділі розглядаються підходи до стиснення зображень, що спотворені шумом (зашумлених зображень) для кодерів на основі ДКП. Ці алгоритми ґрунтуються на прогнозуванні середньоквадратичних похибок для заданого кроку квантування з використанням простого аналізу складності зображення. Проаналізовано статистичні та просторово-спектральні характеристики спотворень, внесених кодерами на основі ДКП, і показано, що вони залежать від кроку квантування та локального вмісту. Показано, що MSE спотворень можна досить легко спрогнозувати. Ці прогнози є достатньо точними і можуть бути використані для належного встановлення кроку квантування. Також розглядається підхід до прогнозування візуальної якості. Продемонстровано, що можна оцінити MSE-HVSM в обмеженій кількості блоків 8x8 пікселів, а потім передбачити значення цієї метрики для всього зображення для розглянутого кодеру. Досліджено вплив вмісту зображення та кількості аналізованих блоків.

В п'ятому розділі розглянуто питання стиснення медичних зображень, в тому числі спотворених шумами різної природи. Зокрема, проаналізовані особливості стиснення медичних зображень, що сформовані сучасними медичними системами реєстрації зображень. Також виявлені особливості шумів, що присутні на більшості медичних зображень. Розроблено підходи і надано рекомендації щодо стиснення таких зображень із урахування візуальної якості.

В шостому розділі розглянуто питання програмно-алгоритмічних реалізацій та прискорення при стисненні та фільтрації зображень. Зокрема, розглянуто питання автоматичного стиснення в задачі розпізнавання цілей, а також запропоновано алгоритмічну реалізацію деяких із розроблених методів фільтрації зображень, спотворених просторово корельованим шумом із невідомим спектром і стандартним відхиленням, яке змінюється для різних областей зображення, за допомогою комбінації згорткових нейронних мереж.

Вперше запропоновано метод виділення текстурних ділянок на зображеннях, спотворених шумом, з використанням методів машинного навчання.

Вперше показано можливість прогнозування ефективності фільтрації у разі сигнально-залежних завад на зображеннях при використанні традиційних та візуальних метрик якості.

Удосконалено двоетапний метод забезпечення необхідної візуальної якості при стисканні зображень із втратами.

Отримали подальший розвиток методи прогнозування ефективності стиснення із втратами зображень, спотворених шумом.

Удосконалено методи обробки зашумлених медичних зображень.

Удосконалено методи прискорення обробки медичних зображень та зображень дистанційного зондування.

Практична значимість отриманих результатів дисертації підтверджується їх використанням. Розроблені методи фільтрації радіолокаційних зображень, що спотворені сигнально залежними та/або просторово-корельованими завадами, а також метод прогнозування ефективності фільтрації радіолокаційних зображень для кількох метрик якості дозволили підвищити якість радіолокаційних зображень у відповідності до різних критеріїв, а також забезпечити прогнозування характеристик фільтрації та вибір найкращих типу та параметрів фільтру. Це підтверджено відповідним актом впровадження, наданим Інститутом радіофізики та електроніки імені О. Я. Усикова НАН України.

Метод прогнозування показників якості первинних рентгеновських зображень забезпечив стиснення зображень без втрати діагностично важливої інформації. Також було запроваджено метод стиснення медичних зображень із урахування спотворень, що вносяться апаратурою формування зображень, та забезпеченням найкращих з точки зору візуальної якості стиснення. Це підтверджено відповідним актом впровадження, наданим КНП ХОР «Обласна дитяча клінічна лікарня».

Розроблений двоетапний підхід до придушення шуму та метод адаптації до просторово корельованого шуму було використано при розробці програмного

забезпечення медичних приладів для підвищення ефективності вторинної обробки і аналізу зашумлених даних, отриманих з медичних датчиків. Це підтверджено відповідним актом впровадження, наданим ТОВ « Біопромінь ».

Методи оцінки якості медичних сигналів за допомогою методів стиснення оптичних зображень та двоетапні процедури стиснення зображень з втратами для забезпечення бажаної якості стиснення зображення з втратами відповідно до метрик візуальної якості дозволили суттєво спростити процедуру оцінки ефективності фільтрації та стиснення зображень, що спотворені шумами складної природи, а також прискорити процедуру стиснення медичних та оптичних зображень із збереженням інформаційних ознак, що дозволяє використовувати наявні автоматичні методи класифікації зображень. Це підтверджено відповідним актом впровадження, наданим ТОВ НК КОЛІБРІ.

Ключові слова: ОРТ, ШІ, шум, зображення, медична діагностика, дистанційне зондування, стиснення, фільтрація, якість, ДКП, класифікація, обробка сигналів.

Список публікацій здобувача за темою дисертації

Праці, в яких опубліковано основні наукові результати дисертації:

1. С. К. Абрамов, Р. А. Кожемякин, С. С. Кривенко, Н. Н. Пономаренко, and В. В. Лукин, “Эффективность фильтрации однозглядовых РСА-изображений при пространственно-коррелированных помехах,” *Радіoeлектронні і комп’ютерні системи*, no. 3, pp. 18–25, 2012.

2. В. В. Лукин, А. В. Науменко, and С. С. Кривенко, “Адаптивная фильтрация РСА-изображений с использованием нейросетевых детекторов границ,” *Радіoeлектронні і комп’ютерні системи*, no. 4, pp. 121–129, 2012.

3. С. С. Кривенко, Е. О. Колганова, and В. В. Лукин, “Фильтрация изображений с использованием преобразования Хаара в блоках,” *Системи обробки інформації*, no. 8, pp. 63–69, 2013.

4. V. V. Lukin, S. S. Krivenko, E. O. Kolganova, and M. K. Chobanu, "Acceleration and efficiency analysis of dct-based filtering," *Telecom Rad Eng*, vol. 72, no. 7, pp. 613–626, 2013, doi: 10.1615/TelecomRadEng.v72.i7.50.

5. S. Abramov, S. Krivenko, A. Roenko, V. Lukin, I. Djurovic, and M. Chobanu, "Prediction of filtering efficiency for DCT-based image denoising," in *2013 2nd Mediterranean Conference on Embedded Computing (MECO)*, Budva, Montenegro: IEEE, Jun. 2013, pp. 97–100. doi: 10.1109/MECO.2013.6601327.

6. D. V. Fevraleev, S. S. Krivenko, V. V. Lukin, R. Marques, and F. D. Medeiros, "Combining level sets and orthogonal transform for despeckling sar images," *Авиационно-космическая техника и технология*, no. 2, pp. 103–112, 2013.

7. R. A. Kozhemiakin, S. S. Krivenko, V. V. Lukin, R. C. P. Marques, F. N. S. Medeiros, and B. Vozel, "Efficiency analysis of combined despeckling of single-look sar images," *AKTT*, no. 5, pp. 102–111, 2013.

8. V. V. Lukin *et al.*, "Approaches to automatic data processing in hyperspectral remote sensing," *Telecom Rad Eng*, vol. 73, no. 13, pp. 1125–1139, 2014, doi: 10.1615/TelecomRadEng.v73.i13.10.

9. S. Krivenko, V. Lukin, B. Vozel, and K. Chehdi, "Filtering efficiency prediction for images corrupted by signal-dependent noise," in *12th International Conference Modern Problems Of Radio Engineering, Telecommunications, and Computer Science*, 2014, pp. 615–617. Accessed: Feb. 08, 2024. [Online]. Available: <https://hal.science/hal-01111060/>

10. S. Krivenko, V. Lukin, B. Vozel, and K. Chehdi, "Prediction of DCT-based denoising efficiency for images corrupted by signal-dependent noise," in *2014 IEEE 34th International Scientific Conference on Electronics and Nanotechnology (ELNANO)*, Kyiv, Ukraine: IEEE, Apr. 2014, pp. 254–258. doi: 10.1109/ELNANO.2014.6873920.

11. V. V. Lukin *et al.*, "An approach to prediction of signal-dependent noise removal efficiency by dct-based filter," *Telecom Rad Eng*, vol. 73, no. 18, pp. 1645–1659, 2014, doi: 10.1615/TelecomRadEng.v73.i18.40.

12. O. S. Rubel, R. O. Kozhemiakin, S. S. Krivenko, V. V. Lukin, B. Vozel, and K. Chehdi, “A method for predicting denoising efficiency for color images,” in *2015 IEEE 35th International Conference on Electronics and Nanotechnology (ELNANO)*, Kyiv, Ukraine: IEEE, Apr. 2015, pp. 304–309. doi: 10.1109/ELNANO.2015.7146897.

13. A. Naumenko, S. Krivenko, V. Lukin, and K. Egiazarian, “Texture region detection by trained neural network,” in *2016 9th International Kharkiv Symposium on Physics and Engineering of Microwaves, Millimeter and Submillimeter Waves (MSMW)*, Kharkiv, Ukraine: IEEE, Jun. 2016, pp. 1–4. doi: 10.1109/MSMW.2016.7538174.

14. A. Naumenko, S. Krivenko, N. Ponomarenko, A. Zelensky, and V. Lukin, “Texture detection in noisy images by combining several local parameters,” in *2015 Second International Scientific-Practical Conference Problems of Infocommunications Science and Technology (PIC S&T)*, Kharkiv, Ukraine: IEEE, Oct. 2015, pp. 230–233. doi: 10.1109/INFOCOMMST.2015.7357321.

15. **С. С. Кривенко, А. В. Наumenко, and В. В. Лукин, “Обнаружение текстурных участков SVM-классификатором на изображениях при наличии помех,” *Радіoeлектронні і комп’ютерні системи*, no. 2, pp. 50–57, 2015.**

16. A. V. Naumenko, V. V. Lukin, M. S. Zriakhov, and S. S. Krivenko, “Feature aggregation for noisy image to improve ‘texture/non-texture’ classification,” in *2017 IEEE 37th International Conference on Electronics and Nanotechnology (ELNANO)*, Kyiv, Ukraine: IEEE, Apr. 2017, pp. 461–464. doi: 10.1109/ELNANO.2017.7939808.

17. **S. K. Abramov, A. Naumenko, V. V. Lukin, S. S. Krivenko, and I. Kaluzhynov, “Neural network based detection of heterogeneities in noisy images,” *Telecom Rad Eng*, vol. 79, no. 19, pp. 1691–1705, 2020, doi: 10.1615/TelecomRadEng.v79.i19.20.**

18. B. Vozel *et al.*, “Processing of multichannel Remote-sensing images with prediction of performance parameters,” in *Environmental Applications of Remote Sensing*, IntechOpen, 2016. Accessed: Feb. 08, 2024. [Online]. Available: <https://www.google.com/books?hl=uk&lr=&id=Wm-QDwAAQBAJ&oi=fnd&pg=PA373&dq=Processing+of+Multichannel+Remote->

Sensing+Images+with+Prediction+of+Performance+Parameters&ots=ClwhdRf6aV&sig=1jtYRgPSY3irBSaTTqyoJQ-L3n4

19. V. V. Lukin et al., “Application of filtering efficiency prediction to hyperspectral data pre-processing,” *Ukrainian journal of remote sensing*, no. 7, pp. 4–11, 2015.

20. S. Krivenko, O. Kolganova, and V. Lukin, “Lossy compression of images corrupted by spatially correlated noise,” in *2016 13th International Conference on Modern Problems of Radio Engineering, Telecommunications and Computer Science (TCSET)*, Lviv, Ukraine: IEEE, Feb. 2016, pp. 698–702. doi: 10.1109/TCSET.2016.7452157.

21. M. Zriakhov, S. Kryvenko, and B. Лукін, “Peculiarities of multispectral image lossy compression,” *Вісник Хмельницького національного університету. Серія: Технічні науки*, vol. 331, no. 1, pp. 370–376, Feb. 2024, doi: 10.31891/2307-5732-2024-331-56.

22. V. Lukin, B. Kovalenko, S. Kryvenko, V. Naumenko, and B. Vozel, “Prediction of Optimal Operation Point Existence and Its Parameters in BPG-Based Automatic Lossy Compression of Noisy Images,” in *Current Overview on Science and Technology Research Vol. 9*, Dr. G. Bucci, Ed., B P International (a part of SCIENCEDOMAIN International), 2022, pp. 1–36. doi: 10.9734/bpi/costr/v9/4316A.

23. S. S. Krivenko, V. V. Lukin, and E. Kolganova, “Peculiarities of lossy compression of optical and medical images corrupted by noise,” in *2016 IEEE 36th International Conference on Electronics and Nanotechnology (ELNANO)*, Kyiv, Ukraine: IEEE, Apr. 2016, pp. 251–256. doi: 10.1109/ELNANO.2016.7493060.

24. S. Krivenko, V. Lukin, O. Krylova, and V. Shutko, “Visually Lossless Compression of Retina Images,” in *2018 IEEE 38th International Conference on Electronics and Nanotechnology (ELNANO)*, Kiev: IEEE, Apr. 2018, pp. 255–260. doi: 10.1109/ELNANO.2018.8477459.

25. S. Krivenko, M. Zriakhov, V. Lukin, and B. Vozel, “MSE prediction in DCT-based lossy compression of noise-free and noisy remote sensing,” in *2018 14th International Conference on Advanced Trends in Radioelectronics, Telecommunications*

and Computer Engineering (TCSET), Lviv-Slavske, Ukraine: IEEE, Feb. 2018, pp. 883–888. doi: 10.1109/TCSET.2018.8336338.

26. V. Naumenko, V. Lukin, and S. Kryvenko, “Comparison of Noisy Medical Image Lossy Compression by BPG and AGU,” in *Integrated Computer Technologies in Mechanical Engineering - 2022*, vol. 657, M. Nechyporuk, V. Pavlikov, and D. Kritskiy, Eds., in *Lecture Notes in Networks and Systems*, vol. 657. , Cham: Springer Nature Switzerland, 2023, pp. 636–648. doi: 10.1007/978-3-031-36201-9_53.

27. V. Lukin, B. Kovalenko, S. Kryvenko, N. Ponomarenko, K. Egiazarian, and J. Astola, “Compression of Noisy Images Taking into Account Visual Quality: A Comprehensive Study,” *Research Developments in Science and Technology Vol. 8*, pp. 89–109, 2022.

28. B. Kovalenko, V. Lukin, S. Kryvenko, V. Naumenko, and B. Vozel, “Prediction of Parameters in Optimal Operation Point for BPG-based Lossy Compression of Noisy Images,” *UJRS*, vol. 9, no. 2, pp. 4–12, Jun. 2022, doi: 10.36023/ujrs.2022.9.2.212.

29. B. Kovalenko, V. Lukin, S. Kryvenko, V. Naumenko, and B. Vozel, “BPG-Based Automatic Lossy Compression of Noisy Images with the Prediction of an Optimal Operation Existence and Its Parameters,” *Applied Sciences*, vol. 12, no. 15, p. 7555, Jul. 2022, doi: 10.3390/app12157555.

30. B. Kovalenko, V. Lukin, V. Naumenko, and S. Krivenko, “Analysis of noisy image lossy compression by BPG using visual quality metrics,” in *2021 IEEE 3rd International Conference on Advanced Trends in Information Theory (ATIT)*, Kyiv, Ukraine: IEEE, Dec. 2021, pp. 20–25. doi: 10.1109/ATIT54053.2021.9678575.

31. С. К. Абрамов, Н. В. Бурцев, С. С. Кривенко, А. Н. Земляченко, and В. В. Лукин, “Автоматическое сжатие в окрестности оптимальной рабочей точки изображений с шумом кодерами типа SPIHT и JPEG2000,” *Авиационно-космическая техника и технология*, no. 2, pp. 99–109, 2016.

32. S. Krivenko, V. Lukin, and B. Vozel, “Prediction of Introduced Distortions Parameters in Lossy Image Compression,” in *2018 International Scientific-Practical Conference Problems of Infocommunications. Science and Technology (PIC S&T)*,

Kharkiv, Ukraine: IEEE, Oct. 2018, pp. 447–451. doi: 10.1109/INFOCOMMST.2018.8632130.

33. S. Krivenko, M. Zriakhov, V. Lukin, and B. Vozel, “MSE and PSNR prediction for ADCT coder applied to lossy image compression,” in *2018 IEEE 9th International Conference on Dependable Systems, Services and Technologies (DESSERT)*, Kyiv, Ukraine: IEEE, May 2018, pp. 613–618. doi: 10.1109/DESSERT.2018.8409205.

34. S. S. Krivenko, O. Krylova, E. Bataeva, and V. V. Lukin, “Smart lossy compression of images based on distortion prediction,” *Telecom Rad Eng*, vol. 77, no. 17, pp. 1535–1554, 2018, doi: 10.1615/TelecomRadEng.v77.i17.40.

35. V. Lukin, A. Zemliachenko, S. Krivenko, B. Vozel, and K. Chehdi, “Lossy compression of remote sensing images with controllable distortions,” *Satellite Information Classification and Interpretation*, 2018, Accessed: Feb. 11, 2024. [Online]. Available:

https://www.google.com/books?hl=uk&lr=&id=lzj8DwAAQBAJ&oi=fnd&pg=PA27&dq=Lossy+compression+of+remote+sensing+images+with+controllable+distortions&ots=irka3mNKud&sig=ljJ_EiGSZkYlW4tNWjpzfsvU3dg

36. S. Krivenko, V. Lukin, and O. Krylova, “Visually Lossless Compression of Dental Images,” in *2019 IEEE 39th International Conference on Electronics and Nanotechnology (ELNANO)*, Kyiv, Ukraine: IEEE, Apr. 2019, pp. 394–399. doi: 10.1109/ELNANO.2019.8783218.

37. S. Krivenko, M. Zriakhov, N. Kussul, and V. Lukin, “Prediction of Visual Quality for Lossy Compressed Images,” in *2019 IEEE 15th International Conference on the Experience of Designing and Application of CAD Systems (CADSM)*, Polyana, Ukraine: IEEE, Feb. 2019, pp. 1–5. doi: 10.1109/CADSM.2019.8779266.

38. O. Krylova, L. Kryvenko, S. Krivenko, and V. Lukin, “A Fast Noniterative Visually Lossless Compression of Dental Images Using AGU-M Coder,” in *2021 IEEE 16th International Conference on the Experience of Designing and Application of CAD Systems (CADSM)*, Lviv, Ukraine: IEEE, Feb. 2021, pp. 6–10. doi: 10.1109/CADSM52681.2021.9385257.

39. S. Krivenko, V. Lukin, O. Krylova, L. Kryvenko, and K. Egiazarian, “A Fast Method of Visually Lossless Compression of Dental Images,” *Applied Sciences*, vol. 11, no. 1, Art. no. 1, Jan. 2021, doi: 10.3390/app11010135.

40. S. S. Krivenko, S. K. Abramov, V. V. Lukin, B. Vozel, and K. Chehdi, “Lossy DCT-based compression of remote sensing images with providing a desired visual quality,” in *Image and Signal Processing for Remote Sensing XXV*, L. Bruzzone, F. Bovolo, and J. A. Benediktsson, Eds., Strasbourg, France: SPIE, Oct. 2019, p. 36. doi: 10.1117/12.2532726.

41. F. Li, S. Krivenko, and V. Lukin, “A Modified Algorithm of Visual Quality Providing in Lossy Compression of Remote Sensing Images,” in *2021 IEEE International Conference on Information and Telecommunication Technologies and Radio Electronics (UkrMiCo)*, Odesa, Ukraine: IEEE, Nov. 2021, pp. 16–21. doi: 10.1109/UkrMiCo52950.2021.9716684.

42. S. Krivenko, D. Demchenko, I. Dyogtev, and V. Lukin, “A Two-Step Approach to Providing a Desired Quality of Lossy Compressed Images,” in *Integrated Computer Technologies in Mechanical Engineering*, vol. 1113, M. Nechyporuk, V. Pavlikov, and D. Kritskiy, Eds., in *Advances in Intelligent Systems and Computing*, vol. 1113, Cham: Springer International Publishing, 2020, pp. 482–491. doi: 10.1007/978-3-030-37618-5_41.

43. S. Krivenko, F. Li, V. Lukin, B. Vozel, and O. Krylova, “Prediction of Visual Quality Metrics in Lossy Image Compression,” in *2020 IEEE 40th International Conference on Electronics and Nanotechnology (ELNANO)*, Kyiv, Ukraine: IEEE, Apr. 2020, pp. 478–483. doi: 10.1109/ELNANO50318.2020.9088819.

44. F. Li, S. Krivenko, and V. Lukin, “A Two-step Procedure for Image Lossy Compression by ADCTC With a Desired Quality,” in *2020 IEEE 11th International Conference on Dependable Systems, Services and Technologies (DESSERT)*, Kyiv, Ukraine: IEEE, May 2020, pp. 307–312. doi: 10.1109/DESSERT50317.2020.9125000.

45. F. Li, S. S. Krivenko, and V. V. Lukin, “Analysis of two-step approach for compressing texture images with desired quality,” *AKTT*, no. 1, pp. 50–58, Jan. 2020, doi: 10.32620/aktt.2020.1.08.

46. V. Lukin et al., “Lossy Compression of Multichannel Remote Sensing Images with Quality Control,” *Remote Sensing*, vol. 12, no. 22, p. 3840, Nov. 2020, doi: 10.3390/rs12223840.

Праці, що підтверджують апробацію матеріалів дисертації:

47. F. Li, S. Krivenko, and V. Lukin, “Adaptive two-step procedure of providing desired visual quality of compressed image,” in *Proceedings of the 2020 4th International Conference on Electronic Information Technology and Computer Engineering*, Xiamen China: ACM, Nov. 2020, pp. 407–414. doi: 10.1145/3443467.3443791.

48. F. Li, S. Krivenko, and V. Lukin, “An Approach to Better Portable Graphics (BPG) Compression with Providing a Desired Quality,” in *2020 IEEE 2nd International Conference on Advanced Trends in Information Theory (ATIT)*, Kyiv, Ukraine: IEEE, Nov. 2020, pp. 13–17. doi: 10.1109/ATIT50783.2020.9349289.

49. F. Li, S. Krivenko, and V. Lukin, “A Fast Method for Visual Quality Prediction and Providing in Image Lossy Compression by SPIHT,” in *Integrated Computer Technologies in Mechanical Engineering - 2020*, vol. 188, M. Nechyporuk, V. Pavlikov, and D. Kritskiy, Eds., in *Lecture Notes in Networks and Systems*, vol. 188. , Cham: Springer International Publishing, 2021, pp. 17–29. doi: 10.1007/978-3-030-66717-7_2.

50. F. Li, S. Krivenko, and V. Lukin, “Two-step providing of desired quality in lossy image compression by SPIHT,” *Радіоелектронні і комп’ютерні системи*, no. 2, pp. 22–32, Apr. 2020, doi: 10.32620/reks.2020.2.02.

51. F. Li, S. Krivenko, and V. Lukin, “An Automatic Optimization Method for BPG Compression Based on Visual Perception,” in *Mathematical Modeling and Simulation of Systems*, vol. 344, S. Shkarlet, A. Morozov, A. Palagin, D. Vinnikov, N. Stoianov, M. Zhelezniak, and V. Kazymyr, Eds., in *Lecture Notes in Networks and Systems*, vol. 344. , Cham: Springer International Publishing, 2022, pp. 213–225. doi: 10.1007/978-3-030-89902-8_17.

52. A. Kharkov, V. Oliinyk, V. V. Lukin, and S. S. Krivenko, “Blind estimation of noise variance for 1d signal denoising,” *Telecom Rad Eng*, vol. 79, no. 7, pp. 567–581, 2020, doi: 10.1615/TelecomRadEng.v79.i7.30.

53. V. Abramova, S. Krivenko, V. Lukin, and O. Krylova, “Analysis of Noise Properties in Dental Images,” in *2020 IEEE 40th International Conference on Electronics and Nanotechnology (ELNANO)*, Kyiv, Ukraine: IEEE, Apr. 2020, pp. 511–515. doi: 10.1109/ELNANO50318.2020.9088768.

54. M. Ponomarenko, O. Miroshnichenko, V. Lukin, S. Krivenko, and K. Egiazarian, “Blind denoising of dental X-ray images,” *ei*, vol. 35, no. 9, pp. 299-1-299–6, Jan. 2023, doi: 10.2352/EI.2023.35.9.IPAS-299.

55. V. Abramova, V. Lukin, S. Abramov, S. Kryvenko, P. Lech, and K. Okarma, “A Fast and Accurate Prediction of Distortions in DCT-Based Lossy Image Compression,” *Electronics*, vol. 12, no. 11, p. 2347, May 2023, doi: 10.3390/electronics12112347.

56. V. Lukin, S. Krivenko, I. Kaluzhinov, O. Krylova, and L. Kryvenko, “Lossy Compression of Remote Sensing and Dental Images Corrupted by Spatially Correlated Noise,” in *Integrated Computer Technologies in Mechanical Engineering - 2021*, vol. 367, M. Nechyporuk, V. Pavlikov, and D. Kritskiy, Eds., in *Lecture Notes in Networks and Systems*, vol. 367. , Cham: Springer International Publishing, 2022, pp. 1003–1014. doi: 10.1007/978-3-030-94259-5_77.

57. G. Proskura, O. Rubel, S. Kryvenko, and V. Lukin, “On classifier performance for remote sensing images compressed by different coders,” *aktt*, no. 1, pp. 67–77, Feb. 2023, doi: 10.32620/aktt.2023.1.07.

58. V. Lukin, S. Abramov, S. Krivenko, A. Kurekin, and O. Pogrebnyak, “Analysis of classification accuracy for pre-filtered multichannel remote sensing data,” *Expert Systems with Applications*, vol. 40, no. 16, pp. 6400–6411, Nov. 2013, doi: 10.1016/j.eswa.2013.05.061.

59. V. Lukin, F. Li, G. Proskura, S. Kryvenko, and B. Vozel, “Peculiarities of Classification of Lossy Compressed Multichannel Remote Sensing Images Using Trained Neural Networks,” in *Digital Ecosystems: Interconnecting Advanced Networks with AI Applications*, vol. 1198, A. Luntovskyy, M. Klymash, I. Melnyk, M. Beshley, and A. Schill, Eds., in *Lecture Notes in Electrical Engineering*, vol. 1198. , Cham: Springer Nature Switzerland, 2024, pp. 127–143. doi: 10.1007/978-3-031-61221-3_7.

60. V. Lukin, F. Li, J. Zhu, and S. Kryvenko, “Peculiarities of SVM-Based Classification of BPG Compressed Three-Channel Images,” in *Proceedings of the 8th International Symposium of Space Optical Instruments and Applications*, vol. 191, H. P. Urbach, D. Li, and D. Yu, Eds., in Springer Proceedings in Physics, vol. 191., Singapore: Springer Nature Singapore, 2024, pp. 352–365. doi: 10.1007/978-981-97-6718-2_31.

Праці, які додатково відображають наукові результати дисертації:

61. L. Kryvenko, O. Krylova, V. Lukin, and S. Kryvenko, “Intelligent visually lossless compression of dental images,” *Adv. Opt. Technol.*, vol. 13, p. 1306142, Feb. 2024, doi: 10.3389/aot.2024.1306142.

62. L. Kryvenko, O. Krylova, S. Krivenko, and V. Lukin, “A Fast Noniterative Visually Lossless Compression of Dental Images Using BPG Coder,” in *2022 11th Mediterranean Conference on Embedded Computing (MECO)*, Budva, Montenegro: IEEE, Jun. 2022, pp. 1–6. doi: 10.1109/MECO55406.2022.9797114.

63. B. Bondžulić, V. Lukin, D. Bujaković, F. Li, S. Kryvenko, and B. Pavlović, “On Visually Lossless JPEG Image Compression,” in *2023 Zooming Innovation in Consumer Technologies Conference (ZINC)*, Novi Sad, Serbia: IEEE, May 2023, pp. 113–118. doi: 10.1109/ZINC58345.2023.10174090.

64. V. Abramova, S. Kryvenko, S. Abramov, V. Makarichev, and V. Lukin, “A Fast Procedure for Image Lossy Compression by ADCTC Using Prediction of Distortions’ MSE,” in *Integrated Computer Technologies in Mechanical Engineering - 2022*, vol. 657, M. Nechyporuk, V. Pavlikov, and D. Kritskiy, Eds., in Lecture Notes in Networks and Systems, vol. 657., Cham: Springer Nature Switzerland, 2023, pp. 625–635. doi: 10.1007/978-3-031-36201-9_52.

65. F. Li, S. Kryvenko, and V. Lukin, “Lossy Compression of Three-Channel Remote Sensing Images Based on Controlled Visual Quality,” in *Integrated Computer Technologies in Mechanical Engineering - 2022*, vol. 657, M. Nechyporuk, V. Pavlikov, and D. Kritskiy, Eds., in Lecture Notes in Networks and Systems, vol. 657., Cham: Springer Nature Switzerland, 2023, pp. 649–661. doi: 10.1007/978-3-031-36201-9_54.

66. V. Lukin, S. Krivenko, F. Li, S. Abramov, and V. Makarichev, “On Image Complexity in Viewpoint of Image Processing Performance,” in *2022 3rd International Workshop on Intelligent Information Technologies and Systems of Information Security*, 2022, pp. 458–473. Accessed: Feb. 12, 2024. [Online]. Available: <https://ceur-ws.org/Vol-3156/paper35.pdf>

67. B. Bondžulić, N. Stojanović, V. Lukin, S. A. Stankevich, D. Bujaković, and S. Kryvenko, “Target acquisition performance in the presence of JPEG image compression,” *Defence Technology*, p. S2214914723003306, Dec. 2023, doi: 10.1016/j.dt.2023.12.006.

68. F. Li, S. Kryvenko, and V. Lukin, “Remote Sensing Image Lossy Compression Based on JPEG with Controlled Visual Quality,” in *Proceedings of the 7th International Symposium of Space Optical Instruments and Applications*, vol. 295, H. P. Urbach and H. Jiang, Eds., in Springer Proceedings in Physics, vol. 295. , Singapore: Springer Nature Singapore, 2023, pp. 8–19. doi: 10.1007/978-981-99-4098-1_2.

69. S. Kryvenko and V. Lukin, “Peculiarities of noisy image lossy compression,” *Вісник Хмельницького національного університету. Серія: Технічні науки*, vol. 333, no. 2, pp. 278–283, Apr. 2024, doi: 10.31891/2307-5732-2024-333-2-44.

70. B. Bondžulić, N. Stojanović, V. Lukin, and S. Kryvenko, “JPEG and BPG visually lossless image compression via KonJND-1k database,” *Vojnotehnički glasnik*, vol. 72, no. 3, pp. 1214–1241, 2024, doi: 10.5937/vojtehg72-50300.

71. F. Li, V. Lukin, S. Kryvenko, B. Bondzulich, D. Bujakovic, and B. Pavlovic, “Strange Images in Remote Sensing and Their Properties,” *UJRS*, vol. 10, no. 2, pp. 12–18, Jun. 2023, doi: 10.36023/ujrs.2023.10.2.240.

ANNOTATION

Kryvenko S.S. Methods of forecasting and intellectual processing of images in remote sensing and medical diagnostics systems. - Proficiency scientific treatise on the rights of the manuscript.

A thesis submitted in fulfilment of the Doctor of Technical Sciences in the specialty 05.12.17 – radio engineering and television systems. – National Aerospace University “Kharkiv Aviation Institute” of the Ministry of Education and Science of Ukraine, Kharkiv, 2025.

The dissertation is devoted to increasing the efficiency of intelligent filtering and lossy compression of images of remote sensing of the Earth and medical diagnostics in the presence of complex interference and limited computing resources, taking into account visual quality.

The current stage of development of science and technology is accompanied by a rapid growth in the volume of data generated in various fields of human activity, in particular in remote sensing of the Earth (RES) and medical diagnostics. These data players play a critically important role in decisions that affect public health, safety and preservation of the natural environment. At the same time, the effective use of these data is complicated by limitations in the speed of information transmission, storage capacity, and it is also necessary to ensure high quality processing and interpretation of the received signals and images.

Remote sensing of the Earth is the main source of information about the state of the environment, forecasting climate change, and monitoring emergency situations. To achieve this goal, methods are provided that provide automatic processing of multispectral and hyperspectral images; adaptation of processing algorithms to noise and distortion; rapid transmission of large amounts of data from satellites.

The field of medical diagnostics is undergoing significant changes under the influence of the development of telemedical technologies. In this context, it is necessary to ensure the preservation of the diagnostic quality of medical images during their compression; separate development of algorithms for automatic restoration of pathologies

based on signal processing; standardization of filtering and restoration processes of biomedical signals to reduce the impact of noise.

The implementation of processing approaches requires solving a number of technical problems: optimizing algorithms for operation on devices with limited resources; ensuring consistency between the quality of processed data and their volume; creating universal models that can adapt to different types of signals and images.

The dissertation solves an important scientific and technical problem of increasing the efficiency of intelligent filtering and lossy compression of remote sensing and medical diagnostics images in the presence of complex interference and limited computing resources, taking into account visual quality.

The introduction justifies the relevance of the topic of the dissertation. The purpose of the work, main tasks and research methods are determined. The scientific novelty and practical significance of the results obtained are formulated.

The first section analyzes the current state of lossy filtering and compression methods. The main tasks of image filtering and compression, image processing quality criteria are determined. The visual quality of images is separately considered as the basis for assessing the effectiveness of future research. Noise in images, their nature in remote sensing and medical diagnostics images, as well as existing means of their suppression are also considered.

The second section considers issues related to image filtering. First of all, the issue of detecting textures in images distorted by noise by combining several parameters of local activity is considered. In addition, adaptive filtering of images distorted by signal-dependent and spatially correlated noise is considered using machine learning algorithms. The last subsection considers the issue of evaluating the effectiveness of filtering for different types of images.

The third section considers issues related to two-stage procedures for ensuring the desired image quality. In addition, the issue of image compression with visually imperceptible distortions is solved. Special attention is paid to adaptive compression of multi-channel images of remote sensing of the Earth. The issue of strange remote sensing images and their properties is also considered.

The fourth section considers approaches to compressing images distorted by noise (noisy images) for encoders based on DCT. These algorithms are based on predicting the mean square errors for a given quantization step using a simple analysis of image complexity. The statistical and spatial-spectral characteristics of distortions introduced by DCP-based encoders are analyzed, and it is shown that they depend on the quantization step and local content. It is shown that the MSE of distortions can be predicted quite easily. These predictions are quite accurate and can be used to properly set the quantization step. An approach to predicting visual quality is also considered. It is demonstrated that it is possible to estimate the MSE-HVSM in a limited number of 8x8 pixel blocks, and then predict the value of this metric for the entire image for the considered encoder. The influence of the image content and the number of analyzed blocks is studied.

The fifth section considers the issue of compression of medical images, including those distorted by noise of various nature. In particular, the features of compression of medical images formed by modern medical image registration systems are analyzed. The features of noise present in most medical images are also revealed. Approaches are developed and recommendations are given for compressing such images taking into account visual quality.

The sixth section considers the issue of software and algorithmic implementations and acceleration in image compression and filtering. In particular, the issue of automatic compression in the target recognition task is considered, and an algorithmic implementation of some of the developed methods for filtering images distorted by spatially correlated noise with an unknown spectrum and standard deviation that varies for different image regions is proposed using a combination of convolutional neural networks.

For the first time, a method for extracting texture areas in images distorted by noise using machine learning methods has been proposed.

For the first time, the possibility of predicting the efficiency of filtering in the case of signal-dependent interference in images using traditional and visual quality metrics has been shown.

A two-stage method for ensuring the required visual quality when compressing images with loss has been improved.

Methods for predicting the efficiency of lossy compression of images distorted by noise have been further developed.

Methods for processing noisy medical images have been improved.

Methods for accelerating the processing of medical images and remote sensing images have been improved.

The method for predicting the quality indicators of primary X-rays ensured image compression without loss of diagnostically important information. A method for compressing medical images was also introduced, taking into account the distortions introduced by the imaging equipment and ensuring the best visual quality of compression. This is confirmed by the relevant implementation act provided by the Regional Children's Clinical Hospital.

The developed two-stage approach to noise suppression and the method of adaptation to spatially correlated noise were used in the development of software for medical devices to improve the efficiency of secondary processing and analysis of noisy data obtained from medical sensors. This is confirmed by the relevant implementation act provided by LLC "Biopromin".

Methods for assessing the quality of medical signals using optical image compression methods and two-stage lossy image compression procedures to ensure the desired quality of lossy image compression in accordance with visual quality metrics have significantly simplified the procedure for assessing the efficiency of filtering and compression of images distorted by noise of a complex nature, as well as accelerating the procedure for compressing medical and optical images while preserving information features, which allows using existing automatic image classification methods. This is confirmed by the relevant implementation act provided by LLC NC KOLIBRI.

Keywords: OOP, AI, noise, image, medical diagnostics, remote sensing, compression, filtering, quality, DCT, classification, signal processing.

ЗМІСТ

ПЕРЕЛІК УМОВНИХ ПОЗНАЧЕНЬ ТА СКОРОЧЕНЬ.....	24
ВСТУП.....	26
РОЗДІЛ 1 ФІЛЬТРАЦІЯ ТА СТИСНЕННЯ ЗОБРАЖЕНЬ ДИСТАНЦІЙНОГО ЗОНДУВАННЯ ЗЕМЛІ ТА МЕДИЧНИХ ЗОБРАЖЕНЬ. МЕТРИКИ ЯКОСТІ ОБРОБКИ ЗОБРАЖЕНЬ	37
1.1 Задачі фільтрації та стиснення зображень.....	37
1.2 Критерії якості обробки зображень.....	42
1.3 Візуальна якість зображень.....	47
1.4 Шуми на зображеннях	51
1.5 Методи фільтрації зображень	56
1.6 Методи стиснення зображень із втратами.....	64
1.7 Цілі дослідження	75
РОЗДІЛ 2 ІНТЕЛЕКТУАЛЬНА ФІЛЬТРАЦІЯ ЗОБРАЖЕНЬ	78
2.1 Методи детектування та збереження текстур	78
2.1.1 Виявлення текстурних ділянок SVM-класифікатором на зображеннях, спотворених завадами.....	78
2.1.2 Виявлення текстурних ділянок нейронною мережею на зображеннях, спотворених завадами.....	90
2.2 Адаптивна фільтрація зображень за допомогою алгоритмів машинного навчання в умовах сигнально-залежних та просторово-корельованих завад.....	96
2.2.1 Аналіз ефективності комбінованого придушення спеклу в однопоглядових зображеннях РСА	96
2.2.2 Використання нейронних мереж для адаптивної фільтрації РСА-зображень та аналіз її ефективності.....	111
2.3 Прогнозування ефективності фільтрації зображень.....	120
2.3.1 Прогнозування ефективності фільтрації сигнально-залежного шуму на оптичних зображеннях та одноканальних зображеннях ДЗЗ	120

2.3.2 Прогнозування ефективності фільтрації для попередньої обробки гіперспектральних зображень	135
Висновки до розділу 2	141
РОЗДІЛ 3 РОЗРОБКА ТА ВДОСКОНАЛЕННЯ АЛГОРИТМІВ СТИСНЕННЯ НЕЗАШУМЛЕНИХ ЗОБРАЖЕНЬ.....	144
3.1 Двоетапні процедури забезпечення необхідної якості.....	144
3.2 Стиснення зображень із забезпеченням візуально непомітних спотворень ...	174
3.3 Адаптивне стиснення багатоканальних зображень	186
3.4 Аномальні (дивні) зображення в дистанційному зондуванні, їх властивості.	199
Висновки до розділу 3	209
РОЗДІЛ 4 АЛГОРИТМИ СТИСНЕННЯ ЗАШУМЛЕНИХ ЗОБРАЖЕНЬ ТА ПРОГНОЗУВАННЯ ХАРАКТЕРИСТИК СТИСНЕННЯ.....	212
4.1 Аналіз характеристик внесених спотворень при стисненні зображень	212
4.2 Інтелектуальне стиснення зображень із втратами з забезпеченням візуально непомітних спотворень	236
Висновки до розділу 4	248
РОЗДІЛ 5 МЕТОДИ ІНТЕЛЕКТУАЛЬНОГО СТИСНЕННЯ МЕДИЧНИХ ЗОБРАЖЕНЬ	250
5.1 Використання сучасних кодерів для роботи з медичними зображеннями	250
5.1.1 Стиснення незашумлених медичних зображень.....	250
5.1.2 Стиснення медичних зображень з урахуванням шумів	258
5.2 Удосконалення методів стиснення із втратами в умовах шумів різної природи та рівнів	274
Висновки до розділу 5	279
РОЗДІЛ 6 ПРОГРАМНО-АЛГОРИТМІЧНА РЕАЛІЗАЦІЯ ТА ПІДВИЩЕННЯ ШВИДКОДІЇ ПРИ СТИСНЕННІ ТА ФІЛЬТРАЦІЇ ЗОБРАЖЕНЬ.....	282
6.1 Реалізація автоматичного стиснення із втратами при детектуванні цілей	282

6.2 Покращення швидкодії фільтрації зображень	299
Висновки до розділу 6	318
ВИСНОВКИ.....	320
СПИСОК ВИКОРИСТАНИХ ДЖЕРЕЛ.....	323
ДОДАТОК А ДОСЛІДЖЕННЯ ДВОЕТАПНИХ ПРОЦЕДУР ЗАБЕЗПЕЧЕННЯ НЕОБХІДНОЇ ЯКОСТІ ДЛЯ КОДЕРА ADST.....	363
ДОДАТОК Б РЕЗУЛЬТАТИ КЛАСИФІКАЦІЇ ЗОБРАЖЕНЬ ДЛЯ НМ-КЛАСИФІКАТОРА.....	373
ДОДАТОК В АНАЛІЗ ПРОСТОРОВОЇ КОРЕЛЯЦІЇ ШУМУ В МЕДИЧНИХ СИСТЕМАХ РЕЄСТРАЦІЇ ЗОБРАЖЕНЬ.....	379
ДОДАТОК Д СПИСОК ПУБЛІКАЦІЙ ЗДОБУВАЧА ЗА ТЕМОЮ ДИСЕРТАЦІЇ	388
ДОДАТОК Е ДОДАТКОВІ МАТЕРІАЛИ ДО РОЗДІЛУ 1	399
ДОДАТОК Ж ДОДАТКОВІ МАТЕРІАЛИ ДО РОЗДІЛУ 2	412
ДОДАТОК К ДОДАТКОВІ МАТЕРІАЛИ ДО РОЗДІЛУ 3.....	418
ДОДАТОК Л ДОДАТКОВІ МАТЕРІАЛИ ДО РОЗДІЛУ 4.....	430
ДОДАТОК М ДОДАТКОВІ МАТЕРІАЛИ ДО РОЗДІЛУ 5	444
ДОДАТОК Н АКТ ПРО ВПРОВАДЖЕННЯ РЕЗУЛЬТАТІВ ДОКТОРСЬКОЇ ДИСЕРТАЦІЇ В ІНСТИТУТІ РАДІОФІЗИКИ ТА ЕЛЕКТРОНІКИ ІМ. О.Я. УСИКОВА НАН УКРАЇНИ.....	458
ДОДАТОК П АКТ ПРО ВПРОВАДЖЕННЯ РЕЗУЛЬТАТІВ ДОКТОРСЬКОЇ ДИСЕРТАЦІЇ В КНП ХОР «ОБЛАСНА ДИТЯЧА КЛІНІЧНА ЛІКАРНЯ»	460
ДОДАТОК Р АКТ ПРО ВПРОВАДЖЕННЯ РЕЗУЛЬТАТІВ ДОКТОРСЬКОЇ ДИСЕРТАЦІЇ В ТОВ НК «КОЛІБРІ»	462
ДОДАТОК С АКТ ПРО ВПРОВАДЖЕННЯ РЕЗУЛЬТАТІВ ДОКТОРСЬКОЇ ДИСЕРТАЦІЇ В ТОВ «БІОПРОМІНЬ».....	464

ПЕРЕЛІК УМОВНИХ ПОЗНАЧЕНЬ ТА СКОРОЧЕНЬ

КС	–	Коефіцієнт стиснення
ДКП	–	Дискретно-косинусне перетворення
ДСП	–	Дискретно-синусне перетворення
ФЩЙ	–	Функція щільності ймовірності
SVM, МОВ	–	Support vector machines, Машина опорних векторів
ВЛД	–	Відносна локальна дисперсія
КР	–	Квазірозмах
ROC	–	Receiver operating characteristic curve, крива операційної характеристики приймача
TPR	–	True positive rate, чутливість
SPC	–	Specificity, специфічність
MCC	–	Matthew's correlation coefficient, коефіцієнт кореляції Метью
ACC	–	Assurasy, точність
НМ	–	Штучна нейронна мережа
SAR, PCA	–	Synthetic aperture radar, Радар із синтезованою апертурою
АБГШ	–	Адитивний білий гаусів шум
ДЗЗ, ДЗ	–	Дистанційне зондування Землі
ЕКП	–	Еквівалентна кількість поглядів
СКП, MSE	–	Середньоквадратична помилка, mean squared error
КВ	–	Коефіцієнт варіації
КШС	–	Крива швидкість-спотворення
СКВ	–	Середньоквадратичне відхилення
ВСП	–	Варіаційно-стабілізуюче перетворення
PSNR	–	Peak signal-to-noise ratio, пікове співвідношення сигнал/шум

IPSNR	–	Improvement of PSNR, покращення пікового співвідношення сигнал/шум
ПКС	–	Параметр, контролюючий стиснення
КК	–	Крок квантування
КМ	–	Коефіцієнт масштабування
БНП	–	Кількість біт на один піксель
ОРТ	–	Оптимальна робоча точка
НОР	–	Незалежний і однаково розподілений
СКШ	–	Середньо корельований шум
ВКШ	–	Сильно (велико) корельований шум
МЯЕ	–	Метрика якості з еталоном
ФЯ	–	Фактор якості
ЛПР	–	Ледь помітна різниця
ТРСС	–	Загальна ймовірність правильної класифікації
КТ	–	Комп'ютерна томографія
МРТ	–	Магнітно-резонансна томографія
УЗД	–	Ультразвукова діагностика
ШІ	–	Штучний інтелект

ВСТУП

Обґрунтування вибору теми дослідження. Сучасний світ стикається з величезним обсягом цифрових даних, які потребують ефективної обробки та аналізу. У цьому контексті методи прогнозування та інтелектуальної обробки сигналів і зображень відіграють ключову роль, особливо в галузях дистанційного зондування Землі та медичної діагностики. Ці технології дозволяють оперативно отримувати цінну інформацію для прийняття рішень у складних і динамічних середовищах, таких як охорона здоров'я, екологічний моніторинг та управління ресурсами.

Україна має розвинену інфраструктуру для ДЗЗ, що підтверджено низкою урядових ініціатив і програм. Постанова Президії НАН України № 77 від 1997 року визначає пріоритетним завданням використання ДЗЗ для вивчення природних ресурсів та вирішення екологічних проблем. Ці технології дозволяють швидко аналізувати стан ґрунтів, водних ресурсів та лісових масивів, а також контролювати наслідки надзвичайних ситуацій. Завдяки автоматизації аналізу зображень та прогнозуванню на основі супутникових даних суттєво скорочуються витрати на наземні дослідження. Використання алгоритмів машинного навчання забезпечує точне виявлення змін у довкіллі, що є важливим для сталого розвитку.

Методи інтелектуальної обробки медичних зображень, таких як аналіз даних комп'ютерної томографії (КТ), магнітно-резонансної томографії (МРТ), ультразвукової діагностики (УЗД) є критично важливими для діагностики важких захворювань, наприклад, раку. Такі технології дозволяють підвищити точність і швидкість постановки діагнозу. Вони знижують навантаження на медичний персонал, автоматизуючи рутинні задачі та забезпечуючи стандартизацію діагностики. Прогнозування на основі даних медичних сигналів і зображень сприяє впровадженню персоналізованих підходів до лікування, що є важливим аспектом сучасної охорони здоров'я. Стратегія розвитку охорони здоров'я України до 2030 року акцентує увагу на цифровізації медицини як ключового пріоритету.

Слід зазначити окремо існуючі законодавчі ініціативи. До них можна віднести Закон України "Про наукову і науково-технічну діяльність", що стимулює розвиток наукових досліджень і впровадження інноваційних технологій у різних галузях. Постанова Кабінету Міністрів України №1188 від 2019 року визначає напрями розвитку цифрових технологій, включно з впровадженням методів обробки великих даних у медицині та ДЗЗ, а програма Євросоюзу "Горизонт Європа" підтримує інтеграцію українських дослідників у міжнародні проекти в галузях високих технологій.

Розробка методів прогнозування та обробки сигналів актуальна через:

- потребу в оперативності. Швидке реагування на екологічні та медичні виклики потребує автоматизованих систем обробки даних;
- глобальну інтеграцію. Українські технології у сфері ДЗЗ та медицини мають великий потенціал для інтеграції у міжнародні проекти;
- підвищення ефективності. Використання таких методів дозволяє знижувати людський фактор, оптимізуючи процеси прийняття рішень.

Обробка зображень відіграє ключову роль у забезпеченні ефективності комп'ютерного зору в військових безпілотних системах, таких як дрони та автономні наземні транспортні засоби. Фільтрація зображень є важливим етапом, оскільки вона дозволяє зменшити вплив шуму, спричиненого датчиками, атмосферними умовами або електромагнітними перешкодами, що значно покращує якість вхідних даних для системи аналізу. Сучасні методи нелінійної фільтрації допомагають зберегти ключові текстурні та просторові особливості, необхідні для ідентифікації об'єктів і виконання складних задач, наприклад, виявлення цілей або перешкод. Стиснення зображень із мінімальними втратами якості є критичним для безпілотних систем, оскільки обмежена пропускна здатність каналів передачі даних вимагає оптимізації обсягу переданих зображень. Використання новітніх методів стиснення зображень дозволяє зберігати візуальну якість при значному зменшенні обсягу даних. Збереження візуальної якості є важливим для завдань автоматичного розпізнавання об'єктів, де система повинна точно визначати положення, розміри та типи об'єктів, навіть у складних умовах

освітлення або часткового перекриття. Наприклад, високоякісні зображення необхідні для ідентифікації військової техніки або рухомих об'єктів у реальному часі. У військових додатках обробка зображень також сприяє підвищенню точності навігації, забезпечуючи розпізнавання місцевості, картографування та обхід перешкод в умовах, де GPS може бути недоступним або подавленим.

Таким чином, поєднання фільтрації, стиснення та збереження або покращення візуальної якості зображень є ключовими факторами для підвищення надійності, точності та швидкості роботи комп'ютерного зору у військових безпілотних системах. Це забезпечує перевагу у складних бойових умовах, де ефективність аналізу візуальної інформації має вирішальне значення.

Обробка зображень є критично важливою складовою як для дистанційного зондування, так і для медичної діагностики, оскільки вона забезпечує виділення ключових даних, необхідних для прийняття рішень. Щодо міжнародних програм, то у рамках програм Horizon 2020 і Horizon Europe активно фінансуються дослідження в обох цих галузях, що свідчить про їхню значущість для глобальних наукових та соціальних викликів. Спільність проблем обробки зображень у цих сферах обумовлена кількома ключовими факторами:

- шуми та артефакти. У ДЗЗ зображення можуть бути спотворені атмосферними умовами, калібрувальними помилками, внаслідок самого принципу формування тощо. У медичних зображеннях, наприклад, КТ чи МРТ, спотворення спричиняються артефактами, рухами пацієнта або технологічними обмеженнями. У обох випадках важливою задачею є розробка методів видалення шумів без втрати ключової інформації;

- обмеженість ресурсів. Обидві галузі стикаються з проблемою ефективного використання даних за умов обмеженої пропускної здатності каналів зв'язку або обчислювальних ресурсів. Наприклад, у космічних програмах Copernicus використовуються оптимізовані алгоритми для стискання зображень перед передачею на Землю. Аналогічно, в медицині алгоритми стискання використовуються для зберігання великих обсягів даних пацієнтів або в телемедичних системах;

– аналіз великих даних. У ДЗЗ обробка супутникових даних спрямована на ідентифікацію змін у ландшафті чи моніторинг катастроф. У медичній діагностиці схожі алгоритми аналізу зображень застосовуються для виявлення патологій, таких як пухлини або порушення кровообігу. У обох випадках використовуються, в тому числі, методи штучного інтелекту (ШІ) для обробки великих масивів даних;

– глобальні виклики. Horizon Europe ставить за мету вирішення глобальних проблем, включаючи зміни клімату та забезпечення здоров'я населення. У ДЗЗ це включає моніторинг природних ресурсів та прогнозування катастроф, а у медичній сфері — діагностику та попередження поширення інфекцій, зокрема завдяки цифровим технологіям та інструментам ШІ;

– міждисциплінарність. Horizon 2020 демонструє важливість міждисциплінарного підходу до досліджень. Так, програми інноваційного охорони здоров'я (Innovative Health Initiative) поєднують біомедицину, обробку зображень і цифрові технології, тоді як космічні програми, такі як SatCen, інтегрують дані ДЗЗ із наземними спостереженнями для розв'язання питань безпеки та гуманітарної допомоги.

Виходячи з викладеного вище, ефективна інтелектуальна фільтрація та стиснення із втратами зображень дистанційного зондування Землі та медичної діагностики в умовах присутності складних завад і обмежених обчислювальних ресурсів з урахуванням візуальної якості є важливою та актуальною науково-практичною проблемою.

Мета і завдання дослідження. Метою дослідження є підвищення ефективності методів обробки зображень у дистанційному зондуванні Землі та медичній діагностиці, зокрема фільтрації та стиснення, із урахуванням візуальної якості. Для досягнення вказаної мети в роботі вирішуються наступні завдання:

1) розробка методів та засобів детектування текстур на спотворених завадами зображеннях за допомогою комбінування декількох параметрів локальної активності.

2) дослідження адаптивної фільтрації зображень, спотворених сигнально-залежними та просторово-корельованими завадами, за допомогою алгоритмів машинного навчання.

3) розробка методів оцінки та прогнозування ефективності фільтрації для різних типів зображень і шумів.

4) розробка та застосування двоетапних процедур забезпечення бажаної якості зображень.

5) розробка методів та засобів стиснення зображень із забезпеченням візуально непомітних спотворень.

6) дослідження адаптивного стиснення багатоканальних зображень дистанційного зондування Землі

7) дослідження дивних зображень дистанційного зондування і їхніх властивостей

8) аналіз статистичних та просторово-спектральних характеристик спотворень, внесених кодерами на основі дискретно-косинусного перетворення (ДКП), і доведення, що вони залежать від кроку квантування та локального вмісту.

9) доведення, що середньо-квадратична похибка (MSE) спотворень можна легко спрогнозувати, а також що ці прогнози є достатньо точними і можуть бути використані для належного встановлення кроку квантування.

10) дослідження прогнозування візуальної якості в обмеженій кількості блоків 8x8 пікселів для сучасних кодерів

11) дослідження стиснення медичних зображень, спотворених шумами різної природи.

12) розробка підходів і рекомендацій щодо стиснення медичних зображень із урахування візуальної якості.

13) аналіз автоматичного стиснення в задачі розпізнавання цілей

14) розробка методів та алгоритмів фільтрації зображень, спотворених з просторово-корельованим шумом із невідомим спектром.

Об'єкт дослідження – одно- та багатоканальні зображення, що формуються аерокосмічними та медичними системами.

Предмет дослідження – методи обробки, фільтрації та стиснення зображень із втратами із забезпеченням високої візуальної якості в радіотехнічних системах обробки даних ДЗЗ та системах медичної діагностики.

Методи дослідження. Для вирішення поставлених завдань було використано методи теорії ймовірностей, математичної статистики, нелінійної фільтрації сигналів, числового моделювання та машинного навчання.

Наукова новизна отриманих результатів полягає у тому, що:

1) вперше запропоновано метод виділення текстурних ділянок на зображеннях, спотворених шумом, з використанням методів машинного навчання, що використовує локальні характеристики, визначені в межах ковзного вікна, завдяки чому вдалося покращити надійність детектування текстурних фрагментів;

2) вперше показано можливість прогнозування ефективності фільтрації у випадку сигнально-залежних завад на зображеннях при використанні традиційних та візуальних метрик якості, який забезпечує високу точність: коефіцієнт детермінації сягає 0,98 для традиційних метрик та 0,91 для метрик, що враховують візуальну якість;

3) удосконалено двоетапний метод забезпечення необхідної візуальної якості при стисканні зображень із втратами, який забезпечує досягнення необхідного рівня якості відповідно до класичних метрик та метрик, що враховують візуальну якість, і дозволяє знизити дисперсію залишкових помилок майже на порядок порівняно з одноетапними підходами;

4) отримали подальший розвиток методи прогнозування ефективності стиснення із втратами зображень, спотворених шумом, що передбачає поканальний аналіз мультиспектральних даних з подальшим вибором найкращих параметрів стиснення для кожного каналу, що дозволяє покращити результати класифікації шляхом придушення шумових і високочастотних компонент;

5) удосконалено методи обробки зашумлених медичних зображень шляхом аналізу просторово-спектральних характеристик шумів, що дозволило забезпечити

сліпе придушення шуму з досягнення приросту PSNR до 2,7 дБ відносно існуючих підходів;

б) удосконалено методи контролю якості обробки медичних зображень та зображень дистанційного зондування, що дозволило гарантувати високий ступінь стиснення (у 10–40 разів), зокрема, для медичних зображень при збереженні усієї діагностично цінної інформації.

Особистий внесок здобувача. Усі основні результати дисертаційної роботи отримано автором самостійно. У роботах [1], [2], [3] автором вивчені питання характеристик ДКП-фільтрів при роботі в умовах сигнально-залежних та просторово корельованих шумів. В роботах [4], [5] автором досліджені порогові операції при фільтрації зображень із застосування часткового перекриття блоків зображення. Підхід Level Set при фільтрації зображень радарів із синтезованою апертурою (РСА), спотворених спекл-шумом досліджено в роботах [6], [7]. В роботах [8], [9], [10], [11], [12] автором розглянуто питання прогнозування ефективності фільтрації одноканальних та багатоканальних зображень для різних методів придушення шумів. В роботах [13], [14], [15], [16], [17] автор розробив підхід до детектування текстур на зображенні за допомогою методів машинного навчання. Декілька підходів до стиснення гіперспектральних зображень розроблено автором у роботах [18], [19], [20], [21]. Прогнозування оптимальної робочої точки та вивчення характеристик стиснення зображення в її околах розглянуто автором в роботах [22], [23], [24], [25], [26], [27], [28], [29], [30], [31]. У роботах [24], [32], [33], [34], [35] автор запропонував декілька підходів прогнозування якості стиснених зображень, а також дослідив ефективність цих методів у випадку медичних зображень [36], [37], [38], [39] та зображень дистанційного зондування [37], [40], [41]. У роботах [42], [43], [44], [45], [46], [47], [48], [49], [50], [51] автором запропоновані двоетапні процедури стиснення зображень та оцінено їхню ефективність. Аналіз характеристик різних типів шумів, що спотворюють зображення дистанційного зондування та медичні зображення, здійснено автором у роботах [52], [53], [54], [55], [56]. Обчислення характеристик класифікації незашумлених та зашумлених стиснених зображень

здійснено автором у роботах [57], [58], [59], [60]. Межі коефіцієнта стиснення зображень зубів із використання методу ледь помітних спотворень отримано автором у роботах [61], [62]. Також питання ефективності кодерів, пришвидшення обчислень та використання наданих ресурсів було вивчено автором у роботах [63], [64], [65], [66]. У роботі [67] автором здійснено порівняльний аналіз метрик візуальної якості в задачі розпізнавання цілі в умовах складного фону. Пошук оптимальних характеристик сучасних кодерів для стиснення без візуально помітних втрат здійснено автором у роботах [68], [69], [70]. Криві швидкості спотворення для дивних зображень були отримані автором у роботі [71].

Публікації.

Основні результати дисертаційної роботи викладено в 71 науковій роботі, у тому числі в 4 монографіях, 15 статтях у наукових виданнях, що входять до переліку наукових фахових видань категорій А та Б, 14 статей в наукових періодичних виданнях інших держав (США, Швейцарія), які індексується в Scopus (7 статей належать до журналів першого та другого квартилів), а також 38 тезах доповідей у збірниках праць наукових конференцій, 27 з яких проіндексовано у базах даних Web of Science Core Collection та Scopus.

Апробація матеріалів дисертації. Основні положення дисертаційної роботи доповідалися і обговорювалися на Міжнародній конференції «The Experience of Designing and Application of CAD systems in Microelectronics» (2013, 2021, 2023, Поляна-Свялява, Україна), Всеукраїнській науково-технічній конференції «Інтегровані комп'ютерні технології в машинобудуванні» (2016, 2019-2022, Харків, Україна), Міжнародній конференції «Problems of infocommunications. Science and Technology» (2015, 2017, 2018, Харків, Україна), Міжнародній конференції «Mediterranean Conference on Embedded Computing» (2013, 2017-2023, Будва, Чорногорія), Міжнародній конференції «International Scientific Conference on Electronics and Nanotechnology» (2014-2020, Київ, Україна), Міжнародній конференції «International Scientific Conference on Electronics and Nanotechnology» (2016, 2018, 2020, 2022, Славське, Україна), Міжнародній конференції «International Conference on Dependable Systems, Services and Technologies» (2018,

2020, 2024, Афіни, Греція), Міжнародній конференції «Computing in Cardiology» (2021, Брно, Чехія), Міжнародній конференції «Ukraine Conference on Electrical and Computer Engineering» (2017, Київ, Україна).

Структура дисертації. Дисертація складається зі вступу, шести розділів, висновків і додатків. Повний обсяг дисертації становить 465 сторінок, в тому числі 84 рисунків по тексту, з них 8 – на 8 окремих сторінках, 54 таблиці по тексту, з них 4 – на 4 окремих сторінках, список використаної літератури складається з 317 найменувань на 40 сторінках, 13 додатків на 102 сторінках, в тому числі – основний текст на 296 сторінках.

Зв'язок роботи з науковими програмами, планами, темами. Дослідження, результати яких представлені у цій дисертаційній роботі, проводились у рамках наукових напрямів кафедр інформаційно-комунікаційних технологій імені О.О. Зеленського та конструкції авіаційних двигунів Національного аерокосмічного університету «ХАІ», пов'язаних із цифровою обробкою сигналів та зображень, та були відображені у звітах з НДР: українсько-польська НДР «Методи інтелектуальної обробки зображень та відео на основі метрик візуальної якості для перспективних застосувань» (2021, № держреєстрації 0121U113841), держбюджетна НДР «Формування технологій створення адаптивних систем керування ГТД пасажирських і транспортних літаків» (2019-2020, № держреєстрації 0119U100942), держбюджетна НДР «Методи інтелектуальної комп'ютерної обробки великих даних в системах дистанційного зондування, мультимедіа та телекомунікацій» (2018-2020, № держреєстрації 0118U003020), українсько-французька НДР «"Розумна" обробка великих об'ємів даних у застосуванні до багатоканальних зображень дистанційного зондування» (2017-2018, № держреєстрації 0117U005010), держбюджетна НДР «Методи підвищення якості сервісу засобами цифрової обробки багатоканальних даних в системах мультимедіа, дистанційного зондування та телемедицини» (2015-2016, № держреєстрації 0115U000827), держбюджетна НДР «Методи обробки багатоканальних зображень у мобільних системах на основі прогнозування та машинного навчання» (2021-2023, № держреєстрації 0121U112176),

держбюджетна НДР «Інтелектуальні методи обробки багатоканальних зображень на основі карт візуальної уваги та баз зображень із спотвореннями» (2024-2026, № держреєстрації 0124U001094).

Практичне значення отриманих результатів дисертаційного дослідження складають:

1) розроблено метод виявлення текстур на базі методів машинного навчання, що використовує локальні параметри, розраховані у ковзному вікні, в якості вхідних даних, що дозволило підвищити надійність детектування текстурних ділянок;

2) запропоновано використовувати метод Level Set для локалізації неоднорідностей зображень, який дозволяє адаптивно застосовувати частотно-залежну фільтрацію до однорідних ділянок, забезпечуючи збереження важливих деталей;

3) розроблено модифікацію вейвлет-фільтра Хаара та інтегрально-диференціального методу, які показали найкращі результати як за кількісними метриками (MSE, MSSIM), так і за якістю збереження меж;

4) розроблено метод прогнозування якості фільтрації зображень, спотворених шумами різної природи. Запропонований підхід забезпечує високе значення коефіцієнту детермінації (до 0,98 для класичних метрик та 0,91 для метрик візуальної якості) при прогнозуванні значень приросту якості зображень;

5) запропоновано двоетапну процедуру стиснення зображень, яка дозволяє забезпечити бажаний рівень якості зображень згідно з метриками PSNR-HVS-M і PSNR, що дозволяє зменшити дисперсію залишкових помилок приблизно на порядок в порівнянні із одноетапними методами;

6) проаналізовано просторово-спектральний склад шумів у медичних зображеннях; запропоновані алгоритми для сліпого придушення шуму забезпечують покращення PSNR до 2,7 дБ у порівнянні з існуючими методами;

7) визначено підхід для стиснення зображень зубів із використанням кодера ADCT, що дозволяє досягти високих значень коефіцієнта стиснення (10-40 разів) без втрат діагностично важливої інформації;

8) запропоновано методику поканального аналізу мультиспектральних даних, що дозволяє визначити оптимальні параметри стиснення для кожного спектрального каналу, забезпечивши покращення результатів класифікації завдяки придушенню шумів і високочастотних варіацій.

Здобувач висловлює щиру подяку провідному вченому в галузі цифрової обробки сигналів та зображень, Лауреату Державної премії України в галузі науки і техніки, доктору технічних наук, професору Лукіну Володимирі Васильовичу.

РОЗДІЛ 1 ФІЛЬТРАЦІЯ ТА СТИСНЕННЯ ЗОБРАЖЕНЬ ДИСТАНЦІЙНОГО ЗОНДУВАННЯ ЗЕМЛІ ТА МЕДИЧНИХ ЗОБРАЖЕНЬ. МЕТРИКИ ЯКОСТІ ОБРОБКИ ЗОБРАЖЕНЬ

1.1 Задачі фільтрації та стиснення зображень

У епоху цифрових технологій здатність ефективно обробляти та передавати великі обсяги візуальних даних є надзвичайно важливою. Придушення шумів і стиснення зображень є двома основними завданнями в цій галузі, особливо в таких сферах, як дистанційне зондування та медична діагностика, де якість зображення є життєво важливою для виконання вимірювань та прийняття рішень, а розмір даних може бути непомірно великим. Придушення шумів, що присутні на зображенні, передбачає не тільки боротьбу з шумом, але й мінімізацію внесених при цьому спотворень, тоді як стиснення зображення має на меті зменшити розмір файлу зображення без суттєвої шкоди для його якості. Ці два завдання нерозривно пов'язані (рис. 1.1), оскільки артефакти стиснення часто створюють шум, а придушення шумів може допомогти у створенні високоякісних стиснених зображень.



Рисунок 1.1 – Актуальні задачі обробки зображень у телекомунікаційних системах

Шум на зображеннях стосується небажаних випадкових коливань інформації про яскравість або колір, які спотворюють зображення [72]. У дистанційному зондуванні цей шум може виникати через такі фактори навколишнього середовища, як атмосферні завади, обмеження сенсора або електромагнітні перешкоди тощо [73]. Подібним чином у медичній візуалізації шум може бути внесений через апаратні обмеження (наприклад, для апаратів МРТ), рух пацієнта або методи візуалізації з низькими дозами, які використовуються для мінімізації радіаційного опромінення під час рентгенівських променів і комп'ютерної томографії [74].

Поширені типи шуму включають [75]:

- адитивний білий гаусів шум (АБГШ): часто вноситься електронними схемами, він має нормальний розподіл;
- пуассонівський шум: виникає в умовах слабкого освітлення або під час отримання зображень із низькою кількістю фотонів;
- шум солі та перцю: призводить до спотворення чорно-білих пікселів, як правило, через помилки передачі або пилу на датчиках зображення;
- спекл-шум: поширений у когерентних системах формування зображень, таких як радіолокаційні та ультразвукові, цей шум є мультиплікативним і часто впливає на медичні зображення та зображення дистанційного зондування.

Класична фільтрація зображень включає [76]:

- лінійні фільтри, включаючи неадаптивні та адаптивні фільтри [77];
- нелінійні фільтри в ковзному вікні, включаючи неадаптивні та локально адаптивні [77];
- нелокальні методи, включаючи нелокальний усереднюючий фільтр, нелокальний сігма-фільтр, фільтр BM3D [78].

До більш просунутих методів фільтрації можна віднести методи на базі ортогональних перетворень та фільтрацію на базі нейронних мереж (НМ). Методи ортогональних перетворень використовують як вейвлет-перетворення, так і дискретне косинусне перетворення (ДКП) для придушення шуму [79]. З розвитком

глибокого навчання фільтрація зображень досягла нових рівнів ефективності та точності. Сучасні методи часто використовують згорткові нейронні мережі (ЗНМ), автокодери та генеративні змагальні мережі (ГЗМ) для вивчення складних моделей шуму та вмісту зображень [80]. Але НМ потребують значних обчислювальних ресурсів, що ускладнює їх використання на пристроях з обмеженими ресурсами. Крім того, велика кількість параметрів робить моделі чутливими до перенавчання, особливо за наявності обмеженого обсягу навчальних даних. НМ часто працюють як "чорна скринька", що ускладнює інтерпретацію результатів та оцінку їхньої ефективності в контексті фільтрації.

Дистанційне зондування передбачає отримання зображень Землі із супутників або бортових платформ. Ці зображення використовуються в різних сферах, включаючи моніторинг клімату, міське планування та сільське господарство [81].

Стиснення зображень має на меті зменшити розмір файлів зображень, зберігаючи якомога більше початкової інформації. Стиснення може бути з втратами (де деяка інформація втрачається) [82] або без втрат (де зображення може бути ідеально реконструйовано). У медичній візуалізації, де точність діагностики є критичною, перевага часто віддається стисненню без втрат, але розмір таких зображень залишається неприйнятно великим. У дистанційному зондуванні, де величезні обсяги даних повинні передаватися із супутників на наземні станції, часто використовується стиснення з втратами, щоб збалансувати ефективність передачі даних і збереження якості [83].

Наведемо декілька прикладів. Датасет Dental Panoramic Caries Segmentation [84] містить сегментацію карієсу зубів у панорамних рентгенівських знімках зубів. Розмір одного зображення складає 2943x1435 пікселів. Приклади медичних зображень наведено в Додатку Е, рис. Е.1. Треба також додати, що як медичні зображення, так і зображення дистанційного зондування можуть бути багатоканальними (так, для КТ-зображень типовим є 64 або 128 каналів, в той час, як для гіперспектральних зображень дистанційного зондування кількість каналів сягає сотень).

Іншим прикладом є датасет Chikusei, що містить гіперспектральні зображення поверхні, сегментовані на 19 класів. Розмір одного зображення складає 2517x2335 пікселів, а загальна кількість каналів складає 128 у спектральному діапазоні від 363 нм до 1018 нм [85]. Приклад зображення наведено в Додатку Е, рис. Е.2. При цьому розмір одного зображення сягає 735 Мбайт, що накладає певні обмеження в питаннях передачі та зберігання таких обсягів.

Традиційні методи стиснення із втратами включають:

- JPEG та його модифікації [86]: найпоширеніший метод стиснення з втратами, JPEG використовує ДКП для зменшення просторової надмірності. Втім, JPEG може створювати блокові артефакти та розмиття, особливо при високому ступені стиснення, що може перешкоджати виконанню таких важливих завдань, як, наприклад, виявлення уражень у медичних зображеннях або класифікація типів місцевості у зображеннях ДЗЗ;

- JPEG2000 та його модифікації [87]: покращений у порівнянні з JPEG, JPEG2000 використовує стиснення на основі малохвильового перетворення. Він підтримує стиснення як із втратами, так і без втрат і працює краще за вищих коефіцієнтів стиснення, що робить його придатним для таких додатків, як стиснення супутникових зображень, де збереження деталей у стиснених зображеннях та забезпечення заданого коефіцієнту стиснення (КС) має вирішальне значення;

- кодери AGU, AGU-M, ADCT, ADCT-M, що використовують деякі модифікації ДКП-перетворення та враховують особливості зорової системи людини[88];

- інші сучасні методи, зокрема, кодери на основі стандарту H.265, а саме BPG, HEIF, AVIF [88].

Шум у даних дистанційного зондування може надходити з кількох джерел, як атмосферні перешкоди, шуми апаратури, помилки калібрування датчика та передачі. Значною проблемою є спекл-шум, який часто зустрічається на радіолокаційних зображеннях, і смугастий шум, виявлений у супутникових даних.

Методи усунення шумів для зображень дистанційного зондування часто використовують багат шарові підходи (наприклад, вейвлет-перетворення), які фіксують як локальні, так і глобальні структури зображення.

Щодо стиснення в дистанційному зондуванні, то супутникові зображення високої роздільної здатності вимагають ефективного стиснення для зберігання та передачі. Методи стиснення на основі вейвлетів, такі як JPEG2000, широко застосовуються для супутникових зображень [89], оскільки вони забезпечують бажані коефіцієнти стиснення з мінімальною втратою деталей. Методи, засновані на глибокому навчанні, які поєднують стиснення та фільтрації, все частіше використовуються, оскільки вони забезпечують досить високу ефективність у обробці шумних даних із високою роздільною здатністю.

Медична візуалізація покладається на такі методи, як рентгенівські промені, МРТ, комп'ютерна томографія та ультразвукове дослідження внутрішніх структур тіла. Якість цих зображень має першочергове значення для точної діагностики, і будь-який шум може приховати важливі деталі. Шум у медичних зображеннях може виникати від самого обладнання для візуалізації (наприклад, обмеження сенсора при низькій дозі рентгенівського випромінювання) або від руху пацієнта. Під час сканування МРТ шум може впливати на видимість невеликих та/або слабконтрастних структур, що ускладнює діагностику таких станів, як пухлини. Методи придушення шумів повинні досягати балансу між зменшенням шуму та збереженням діагностичних функцій.

Окремо наведемо приклади зображень дистанційного зондування (ДЗ) та медичних зображень, що спотворені шумами різної природи. В Додатку Е, на рис. Е.3 та рис. Е.4, наведено тестові зображення рентгенівських знімків зубів, що використовувалися в дослідженнях, та їхні збільшені фрагменти. В Додатку Е, на рис. Е.5 наведено приклади однопоглядного зображення радару із синтезованою апертурою (РСА). Візуальний аналіз цих зображень демонструє проблему наявності шумів попри те, що ці зображення сформовані сучасними системи отримання зображень. Причому, навіть не озброєним оком видно складний характер шуму, що спотворює ці зображення. Крім того, різна якість різних каналів

багатоканальних зображень так само обумовлює необхідність застосування попередньої обробки. Приклад наведено в Додатку Е, на рис. Е.6, де показано залежність якості зображення від номеру каналу у системі AVIRIS.

Таким чином, можна вважати, що проблеми, притаманні зображенням ДЗ також притаманні і медичним зображенням, а саме:

- існує тенденція зростання об'ємів даних, які потребують негайної або відкладеної обробки (передача, зберігання, інша обробка, причому часто оперативна та автоматична, отримування корисної інформації та надання її потенційним користувачам у зручному для них вигляді);
- незважаючи на всі зусилля, якість первинних зображень часто недостатньо висока, насамперед через наявність шуму;
- візуальна якість виходить на перший план, відповідно треба застосовувати не традиційні метрики типу PSNR (MSE), а принаймні декілька різнопланових метрик та показників, що характеризують якість зображень та ефективність обробки з різних сторін;
- шуми мають складну природу та характер взаємодії із зображеннями, що обмежує застосовність більшості розроблених раніше методів обробки та вимагає створення нових методів та засобів.

1.2 Критерії якості обробки зображень

Середня квадратична похибка (MSE) і пікове відношення сигнал/шум (PSNR) є широко використовуваними кількісними показниками в області обробки зображень для оцінки якості реконструйованих, стиснутих або оброблених зображень [90]. Обидва показники є фундаментальними для порівняння двох зображень та є важливими інструментами для оцінки ефективності різних алгоритмів обробки зображень, включаючи методи стиснення, усунення шумів і відновлення. Менший MSE вказує на меншу похибку, тобто оброблене зображення ближче до оригіналу. PSNR є похідним від MSE, воно визначається в децибелах

(дБ), що полегшує порівняння якості зображення в різних методах обробки. Вищі значення PSNR вказують на кращу якість. Логарифмічний характер PSNR означає, що невелика зміна MSE може призвести до суттєвої зміни PSNR. Наприклад, зниження MSE з 10 до 5 збільшує PSNR більш суттєво, ніж зменшення зі 100 до 95. Типові значення PSNR для високоякісних зображень коливаються від 35 до 50 дБ.

Припустимо, що маємо безшумне зображення I_{ij}^t , $i=1, \dots, I_{im}$, $j=1, \dots, J_{im}$, розміром $I_{im}J_{im}$, де ij — індекси пікселів. Зображення після стиснення з втратами даним кодером позначається як I_{ij}^c , $i=1, \dots, I_{im}$, $j=1, \dots, J_{im}$. Зображення після фільтрації позначається як I_{ij}^f , $i=1, \dots, I_{im}$, $j=1, \dots, J_{im}$. Тоді метрики можуть бути розраховані для декількох варіантів. У формулах нижче припускаємо, що зображення 8-бітне.

Варіант порівняння істинного зображення I^t да стисненого зображення I^c . Тоді MSE та PSNR можна представити як:

$$MSE_{tc} = \sum_{i=1}^{I_{im}} \sum_{j=1}^{J_{im}} (I_{ij}^c - I_{ij}^t)^2 / (I_{im}J_{im}) \quad (1.1)$$

$$PSNR_{tc} = 10 \log \left(\frac{255^2}{MSE_{tc}} \right) \quad (1.2)$$

Варіант порівняння істинного зображення I^t да стисненого зашумленого зображення I^{nc} . Тоді MSE та PSNR можна представити як:

$$MSE_{tn} = \sum_{i=1}^{I_{im}} \sum_{j=1}^{J_{im}} (I_{ij}^{nc} - I_{ij}^t)^2 / (I_{im}J_{im}) \quad (1.3)$$

$$PSNR_{tn} = 10 \log \left(\frac{255^2}{MSE_{tn}} \right) \quad (1.4)$$

Варіант порівняння зашумленого зображення I^n да стисненого зашумленого зображення I^{nc} . Тоді MSE та PSNR можна представити як:

$$MSE_{cn} = \sum_{i=1}^{I_{im}} \sum_{j=1}^{J_{im}} (I_{ij}^{nc} - I_{ij}^n)^2 / (I_{im}J_{im}) \quad (1.5)$$

$$PSNR_{cn} = 10 \log \left(\frac{255^2}{MSE_{cn}} \right) \quad (1.6)$$

Варіант порівняння істинного зображення I^t да фільтрованого зашумленого зображення I^f . Тоді MSE та PSNR можна представити як:

$$MSE_{tf} = \sum_{i=1}^{I_{Im}} \sum_{j=1}^{J_{Im}} (I_{ij}^t - I_{ij}^f)^2 / (I_{Im} J_{Im}) \quad (1.7)$$

$$PSNR_{tf} = 10 \log \left(\frac{255^2}{MSE_{tf}} \right) \quad (1.8)$$

Решта метрик в нашому дослідженні буде використовувати аналогічну систему індексів, якщо не буде визначене інше.

Досить часто отримані та оброблені зображення підлягають певному аналізу з метою видобування з них корисної інформації. Це може бути класифікація, сегментація, виявлення певних типів об'єктів або ділянок тощо. Тоді PSNR не здатне характеризувати ці операції (хоча їх ефективність неявним чином залежить від PSNR).

Показники ефективності класифікації зображень важливі для оцінки того, наскільки точно модель може класифікувати зображення за попередньо визначеними класами. Метрики (кількісні критерії), які зазвичай використовуються в класифікації зображень, включають точність, влучність, повноту, оцінку F1 і площу під кривою робочих характеристик приймача (AUC-ROC) [91].

У метриках класифікації терміни TP (True Positive), FP (False Positive), TN (True Negative) та FN (False Negative) використовуються для оцінки якості роботи класифікаційної моделі:

- TP (True Positive, істинно позитивне). Кількість випадків, коли модель правильно передбачила позитивний клас. Наприклад, якщо модель спрогнозувала, що пацієнт хворий, і це дійсно так.

- FP (False Positive, хибно позитивне). Кількість випадків, коли модель передбачила позитивний клас, але насправді це не так. Наприклад, якщо модель

спрогнозувала, що пацієнт хворий, але він здоровий. FP також називають помилковою тривоною.

– TN (True Negative, істинно негативне). Кількість випадків, коли модель правильно спрогнозувала негативний клас. Наприклад, якщо модель передбачила, що пацієнт здоровий, і це дійсно так.

– FN (False Negative, хибно негативне). Кількість випадків, коли модель передбачила негативний клас, але насправді це не так. Наприклад, якщо модель передбачила, що пацієнт здоровий, але він насправді хворий. FN також називають пропущеним виявленням.

Загальну чисельність позитивного класу позначають P, а загальну кількість негативного класу позначають N.

Точність (Accuracy) – це відношення правильно класифікованих екземплярів до загальної кількості екземплярів. Він дає загальне уявлення про ефективність моделі в усіх класах. Точність розраховується наступним чином:

$$Accuracy = \frac{TP+TN}{P+N} \quad (1.9)$$

Точність корисна для збалансованих наборів даних, де кожен клас однаково представлений, але вона може ввести в оману для незбалансованих наборів даних. Наприклад, якщо 90% зображень належать до одного класу, модель може досягти 90% точності, просто передбачивши превалюючий клас для всіх зображень, незалежно від справжніх міток інших класів.

Влучність та повнота особливо цінні для вимірювання ефективності в бінарній класифікації та багатокласових сценаріях із дисбалансом класів. Вони допомагають оцінити здатність моделі правильно ідентифікувати екземпляри певного класу і часто використовуються в медичних візуалізаціях, виявленні аномалій та інших програмах, де хибно-позитивні або хибно-негативні результати є важливими.

Влучність (Precision) – це відношення правильно передбачених позитивних випадків до загальної кількості прогнозованих позитивних випадків:

$$Precision = \frac{TP}{TP+FP} \quad (1.10)$$

Велике значення влучності вказує на низький рівень хибнопозитивних результатів, тобто модель не часто неправильно класифікує інші класи як цільовий клас. Цей показник особливо актуальний у завданнях, де помилкові спрацьовування небажані, наприклад, виявлення спаму.

Повнота (Recall, або чутливість) — це співвідношення правильно передбачених позитивних випадків до всіх фактичних позитивних випадків:

$$Recall = \frac{TP}{TP+FN} \quad (1.11)$$

Велике значення повноти вказує на те, що модель ефективно фіксує справжні позитивні випадки.

Оцінка F1 — це середнє гармонійне значення влучності та повноти, що забезпечує єдиний показник, який врівноважує обидва. Це особливо корисно, коли є нерівномірний розподіл класів, оскільки він штрафує екстремальні значення в влучності або повноти. Оцінка F1 розраховується як:

$$F1 = 2 * \frac{Precision * Recall}{Precision + Recall} \quad (1.12)$$

Велике значення F1 вказує на те, що модель досягає гарного балансу між влучністю та повнотою, що робить її надійною метрикою у випадках, коли як хибнопозитивні, так і хибнонегативні результати є важливими.

AUC-ROC (площа під кривою робочих характеристик приймача) — це показник, який оцінює здатність моделі розрізняти класи за різними пороговими значеннями класифікації. Крива ROC показує повноту проти частоти помилкових

позитивних результатів (FPR) за різних порогових значень. AUC підсумовує ефективність моделі, причому оцінка 1,0 вказує на ідеальну класифікацію, а 0,5 представляє випадкове вгадування. AUC-ROC особливо корисний для оцінки класифікаторів у сценаріях із незбалансованими даними, оскільки він дає уявлення про ефективність моделі в діапазоні порогових значень, а не в одному пороговому значенні.

Тут треба коротко обговорити вписування кривої в діаграми розсіювання, а також особливості формування діаграм розсіювання. Показники ефективності підгонки кривої є важливими для кількісного визначення того, наскільки добре математична модель або функція апроксимує набір точок даних. Щоб оцінити якість відповідності моделі, використовуються різні показники, кожен з яких фіксує трохи інший аспект точності або надійності моделі. Загальні показники включають середню квадратичну помилку (MSE), середньоквадратичну помилку (RMSE), R^2 (коефіцієнт детермінації), скоригований R-квадрат. Існує добре розроблена теорія регресії та критерії якості підгонки [92]. Найбільш відомими з них є R^2 і RMSE. R^2 для компактних і гарно підігнаних даних прагне до одиниці, тоді як RMSE має бути якомога меншим.

1.3 Візуальна якість зображень

Основним обмеженням PSNR є те, що вона спирається виключно на стандартне кількісне порівняння і фактично не враховує особливості зорової системи людини (HVS). Це може бути проблематично, оскільки два зображення з однаковим PSNR можуть мати суттєво різну візуальну якість. Людські спостерігачі більш чутливі до спотворень у низькочастотних областях (гладкі області та межі) і менш чутливі до високочастотного шуму або текстур, які менш помітні людському оку. Отже, PSNR часто не відповідає середнім оцінкам думки (MOS) у суб'єктивних оцінках якості зображення.

Щоб дати оцінку, яка більше відповідає результатам оцінки якості зорової системи людини на зображенні, деякі дослідники поєднали характеристики HVS із

чисто математичними алгоритмами. Наступні показники є типовими HVS-метриками.

PSNR-HVS (Peak Signal-Noise Ratio-Human Visual System) [93] — це модифікована версія традиційної метрики PSNR, адаптованої для врахування характеристик HVS. PSNR-HVS покращує це шляхом зважування різних частотних компонентів, щоб відобразити чутливість HVS до певних типів спотворень, особливо в низькочастотних областях. PSNR-HVS-M [94] є ще однією модифікацією PSNR, що враховує CSF і міжкоефіцієнтне контрастне маскування базисних функцій ДКП. Розрахунок PSNR-HVS та PSNR-HVS-M передбачає декілька ключових коригувань традиційного PSNR:

1) зважування на основі частоти: алгоритм спочатку застосовує блочне ДКП, щоб розділити зображення на частотні компоненти. У кожному блоці DCT 8x8 низькочастотним коефіцієнтам надається більша вага;

2) коригування середнього зсуву та контрасту: перед ДКП розрахунок PSNR-HVS застосовує зсув середнього значення та розтягування контрасту до оригінальних і спотворених зображень, щоб стандартизувати умови та усунути глобальні зсуви інтенсивності;

3) для PSNR-HVS-M враховується також маскувальний ефект.

PSNR-HVS та PSNR-HVS-M краще корелюють із суб'єктивними оцінками людини, ніж традиційний PSNR, що робить їх корисним у застосуваннях, де якість сприйняття має першочергове значення, наприклад, стиснення зображень і відео. PSNR-HVS та PSNR-HVS-M вимірюються у дБ (більше значення означає вищу якість зображення).

UQI є універсальним індексом якості зображення (UQI) [95], його розроблено шляхом моделювання будь-якого спотворення зображення як комбінації трьох факторів: втрати кореляції, спотворення яскравості та спотворення контрасту.

SSIM – це метрика структурної подібності [95], яка є модифікованою версією UQI. Вона працює в просторовій області та виконує три типи порівнянь між еталонними та спотвореними зображеннями: порівняння яскравості, порівняння

контракту та порівняння структури. Ці три компоненти поєднуються відповідно до функції комбінації та є відносно незалежними.

Багатомасштабний індекс структурної подібності (MS-SSIM) [96] — це розширений показник оцінки якості зображення, призначений для вимірювання сприйманої подібності між двома зображеннями шляхом моделювання характеристик HVS. На відміну від традиційних одномасштабних методів, таких як індекс структурної схожості (SSIM) [97], MS-SSIM оцінює зображення в кількох просторових роздільних здатностях, що краще узгоджується з багатомасштабним сприйняттям візуальної інформації HVS.

Дослідження показали, що MS-SSIM добре корелює із MOS, загальною суб'єктивною метрикою якості, що робить MS-SSIM потужним інструментом для об'єктивної оцінки якості в завданнях, таких як стиснення та передача зображень. Значення MS-SSIM варіюються від 0 до 1, де останнє відповідає ідеальній якості.

Основний принцип роботи метрики FSIM [98] полягає в тому, що людський зір сприймає зображення на основі їхніх характерних ознак низького рівня. Використовуються два типи характеристик: конгруентність фази та величина градієнта. Конгруентність фаз використовується для зважування внеску кожної точки в подібність спотвореного та еталонного зображень. Значення FSIM варіюються від 0 до 1, де останнє відповідає ідеальній якості.

MDSI - це індекс схожості середнього відхилення [99], що обчислюється за допомогою комбінованої подібності градієнта, схожості кольоровості і об'єднання відхилень. Величина градієнта використовується для вимірювання структурних спотворень, а характеристики кольоровості використовуються для вимірювань колірних спотворень, потім карти подібності об'єднуються, щоб створити карту подібності градієнта та кольоровості. Згідно з MDSI, зображення ідеальної якості оцінюється нульовим показником якості, оскільки в його карті подібності немає змін.

Треба додати, що існує безліч метрик візуальної якості, але наведені вище є одними з найкращих з точки зору кореляції з MOS. Також вони є одними з найкращих з точки зору роботи із зашумленими зображеннями. Нещодавно було

проведено дослідження, що підтверджує даний факт для ситуації, коли зображення спотворено шумами різної природи [100]. База даних TID2013 [101] містить 25 тестових кольорових зображень з 24 типами та 5 рівнями спотворень. Було розглянуто дві підмножини типів спотворень «Noise» та «Actual» (табл. 1.1).

Таблиця 1.1 – Значення коефіцієнту кореляції Спірмена для 22 метрик візуальної якості

Метрика	Всі типи завад	Набір Noise	Набір Actual	Набори Noise та Actual
MDSI	0,8897	0,9275	0,9387	0,9374
CVSSI	0,8090	0,9248	0,9350	0,9341
MCSD	0,8045	0,9224	0,9326	0,9323
PSNRHA	0,8198	0,9230	0,9388	0,9322
GMSD	0,8004	0,9211	0,9314	0,9318
PSNRHMA _m	0,8541	0,9221	0,9387	0,9315
PSIM	0,8926	0,9189	0,9309	0,9303
PSNRHAY	0,7794	0,9184	0,9272	0,9275
PSNRHVS	0,6536	0,9172	0,9257	0,9263
PSNRHMA	0,8137	0,9151	0,9343	0,9250
IGM	0,8023	0,9099	0,9220	0,9227
PSNRHMA _y	0,7570	0,9107	0,9209	0,9226
VSI	0,8967	0,9101	0,9258	0,9218
SR-SIM	0,8076	0,9070	0,9211	0,9206
HaarPSI	0,8730	0,9063	0,9168	0,9190
ADM	0,7861	0,9113	0,9201	0,9189
PSNRHVSM	0,6246	0,9061	0,9175	0,9188
FSIM _c	0,8510	0,9022	0,9150	0,9164
ADD_GSIM	0,8310	0,9023	0,9151	0,9159
IQM2	0,7955	0,8995	0,9103	0,9122
ADD_SSIM	0,8023	0,9008	0,9119	0,9120
FSIM	0,8011	0,8969	0,9108	0,9117

Ці підмножини включали по декілька типів шумів з переліку: адитивний гаусівський шум, адитивний шум у колірних компонентах, просторово корельований шум, замаскований шум, високочастотний шум, імпульсний шум, шум квантування, розмиття за Гаусом, зменшення шуму зображення, мультиплікативний гаусів шум, стиснення із втратами шумових зображень. Аналіз даних, наведених у табл. 1.1, свідчить про наступне. По-перше, декілька елементарних метрик забезпечують приблизно однакову ефективність характеристики візуальної якості триканальних зображень з типами спотворень, які мають місце в зображеннях ДЗ. По-друге, ці елементарні метрики належать до різних груп (сімейств), таких як метрики на основі PSNR (як PSNRHA і PSNRHVSM) і метрики на основі SSIM (як ADD SIM і FSIM). Між тим, існують і елементарні метрики, які не належать до цих сімейств. Це означає, що існує багато ефективних елементарних метрик, хоча жодна з них не є досконалою. При цьому деякі метрики візуальної якості, наприклад, SSIM можуть бути гіршими за PSNR [100]. Приклад лінеаризації наведено на рис. 1.2.

Лінеаризація MOS часто використовується для порівняння ефективності метрики [102]. Діаграма розсіювання для кожної елементарної метрики показує пари значення метрики (після лінеаризації) проти MOS для кожного розглянутого тестового зображення, типу спотворення та рівня. І одним з висновків є той факт, що PSNR має значне розсіювання навколо ідеальної прямої кореляції в порівнянні із метриками візуальної якості. Крім того, очевидним є факт високої кореляції найкращих метрик візуальної якості між собою.

1.4 Шуми на зображеннях

Придушення шумів на зображенні, як важливий крок до покращення якості зображення, завжди було актуальною темою для досліджень. Оскільки шум, межі та текстура є високочастотними компонентами, їх часто важко розрізнити в процесі усунення шумів. З численних типів шуму, що переважає в різних типах зображень, адитивний білий шум Гауса (АБГШ), імпульсний шум (для якого найбільш

відомою є модель Сіль і Перець), шум квантування, шум Пуассона та спекл-шум є найбільш часто обговорюваними шумами в літературі [103].

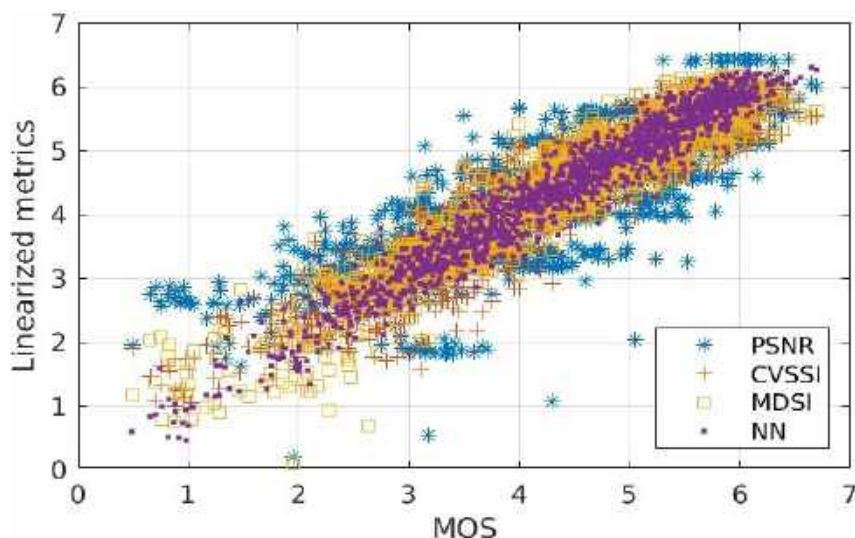


Рисунок 1.2 – Діаграма розсіювання, що ілюструє кореляцію чотирьох метрик (після лінеаризації) проти справжнього MOS (діапазон 0-7) для підмножини бази даних TID2013, підмножини Noise та Actual (NN – метрика на основі нейромережі)

АБГШ в основному виникає в аналогових схемах під час отримання та передачі зображення. Поширеність інших типів шумів, таких як шум квантування, імпульсний шум, спекл-шум і шум Пуассона, виникає в основному через дефекти виробництва, бітові помилки та підрахунок фотонів під час отримання зображення [104]. Загалом, як відновити значущу інформацію із зашумлених зображень у процесі фільтрації та отримати зображення високої якості є життєво важливою проблемою. Незважаючи на те, що за останні два десятиліття було досягнуто значного прогресу в придушенні шумів на зображенні [105], [106], попит на високоефективні методи усунення шумів у зображенні все ще актуальний у практичних застосуваннях. Основна причина полягає в тому, що з математичної точки зору природа зменшення шуму зображення є зворотною проблемою, а її рішення не є єдиним. Тому ключовим питанням є те, як змодельовати зображення і отримати ефективне рішення за апріорних обмежень.

Шуми, які є на зображеннях, що формуються системами дистанційного зондування Землі або системами медичної візуалізації, залежать від способу формування зображень, діапазону хвиль і завад, викликаних погодними умовами та іншими зовнішніми факторами. Для систем ДЗ, що працюють в оптичному діапазоні хвиль (і сусідніх з ним діапазонах), таких як система ДЗ авіаційного базування Aviris [107], шум має гаусову адитивну та квазіпуассонівську компоненти. Для систем ДЗ, що працюють у радіодіапазоні хвиль, типовими є такі типи шумів [108]:

- адитивний шум, може бути просторово некорельованим або просторово корельованим, розподіл гаусів або негаусів;
- сигнально-залежний шум, розподіл гаусів, релеєвський та експоненційний;
- змішані варіанти.

Для радіолокаційних зображень, що формуються системами ДЗ, типовим є превалювання мультиплікативних завад над адитивними [109] і часто адитивною шумовою складовою при фільтрації зображення нехтують, зосередивши увагу на придушенні мультиплікативної складової. Для систем з радіолокатором бічного огляду характерний гаусів (нормальний) розподіл. Відносна дисперсія мультиплікативного шуму σ_μ^2 для зображень, що формуються такими системами, зазвичай має порядок від 0,003 до 0,05. Наприклад, зображення міліметрового діапазону, що наведено в Додатку Е, на рис. Е.7, спотворено мультиплікативним шумом з $\sigma_\mu^2=0,005$.

Для систем ДЗ з РСА мають місце релеєвський, однобічний експоненціальний або Г-розподіл [110]. Релеєвський розподіл, як і однобічний експоненційний розподіл, типові для однопоглядових РСА. В Додатку Е, на рис. Е.8 наведено приклад такого зображення, сформованого системою дистанційного зондування TerraSAR-X [111].

Г-розподіл має місце при багатопоглядовому формуванні РСА-зображень [112]. У міру збільшення кількості поглядів Г-розподіл поступово наближається до

нормального. Відносні дисперсії σ_μ^2 для Γ -розподілу зазвичай мають порядок від 0,05 до 0,27.

Найбільш складною для зображень систем ДЗ є придушення шуму для однопоглядових зображень РСА, що зумовлено найбільшою відносною дисперсією мультиплікативного шуму для таких зображень. Слід зазначити, що шум на зображеннях систем ДЗ та медичних систем, як правило, є просторово-корельованим [113]. Тому для його ефективного придушення потрібна наявність ефективних методів сліпої оцінки як типу і дисперсії шуму, так і його спектра.

Шум у мультиспектральних і гіперспектральних системах ДЗ моделюється як сума сигнально-залежного (СЗШ) та сигнально-незалежного (СНШ) компонентів в k -тому каналі, що описується рівнянням:

$$I_{ij}^n(k) = n_{ij}^{C3} \left(I_{ij}(k) \right) + n_{ij}^{CH}(\sigma_a^2), \quad (1.13)$$

де i та j – координати пікселей на зображенні, $i = 1, \dots, I$, $j = 1, \dots, J$, $k = 1, \dots, K$, I, J – визначають розмір зображень, K – кількість каналів; $I_{ij}(k)$ – ij -й піксель істинного неспотвореного шумом зображення в k -му каналі; n_{ij}^{C3} – сигнально-залежна шумова компонента в ij -й пікселі (дисперсія залежить від $I_{ij}(k)$); n_{ij}^{CH} – сигнально-незалежна шумова компонента в ij -й пікселі (дисперсія адитивного шуму не залежить від $I_{ij}(k)$ та дорівнює σ_a^2).

Локальна дисперсія цього шуму зазвичай визначається як [114]:

$$\sigma_{ij}^n(k) = \sigma_{a\ ij}^2(k) + p(k)I_{ij}(k), \quad (1.14)$$

де $\sigma_{a\ ij}^2(k)$ – дисперсія адитивної сигнально-незалежної компоненти (зазвичай має гаусів розподіл); $p(k)$ – параметр сигнально-залежної компоненти в k -му каналі.

Дослідження шуму в сучасних мультиспектральних системах ДЗЗ [115] демонструють домінування адитивного СНШ у деяких каналах, хоча СЗШ також

заслуговує на увагу і не може бути повністю знехтуваний. У деяких каналах мультиспектральних систем допускається ігнорування сигнально-залежного компонента.

Шуми радіолокаційних зображень мають іншу природу, зокрема виникає явище, відоме як спекл. Це явище спричинене когерентним додаванням (інтерференцією) сигналів, відбитих від великої кількості елементарних відбивачів, які потрапляють в межі одного елемента роздільної здатності радіолокаційної системи. Результируючий сигнал $Ae^{j\varphi} = \sum_{k=1}^N A_k e^{j\varphi_k}$, де A_k і φ_k є амплітудою та фазою сигналу, а N — кількість відбивачів, є сумою сигналів зворотного розсіювання з випадковими фазами та амплітудами. В такій формі радіолокаційні зображення представлені у вигляді комплексних значень.

Дослідження спекл-шуму в радіолокаційних системах показали, що цей шум є мультиплікативним: його дисперсія залежить від інтенсивності сигналу [116]. Модель спекл-шуму описується рівнянням:

$$I_{ij}^c = I_{ij} \mu_{ij}, \quad (1.15)$$

де I_{ij}^c — спотворене спекл-шумом зображення; μ_{ij} — випадкова мультиплікативна компонента з математичним очікуванням $m_\mu=1$ і дисперсією σ_μ^2 .

Для зменшення впливу спекл-шуму застосовується некогерентне накопичення (multilooking), що базується на усередненні незалежних зображень сцени (поглядів). Такий підхід дозволяє зменшити спекл-шум, хоча й погіршує просторову роздільну здатність. Некогерентне усереднення може здійснюватись по азимуту або діапазону. Це підвищує радіометричну точність і покращує інтерпретацію радіолокаційних зображень [117].

Амплітуда A спекл-шуму для однопоглядових зображень підпорядковується закону розподілу Релея:

$$P_A(A) = \frac{2A}{\sigma} \exp\left(-\frac{A^2}{\sigma}\right), A \geq 0, \quad (1.16)$$

де σ визначає СКВ шуму. Інтенсивність шуму A^2 має експоненційний розподіл. Для багатопоглядових зображень, спекл-шум описується гамма-розподілом:

$$P_I(I) = \frac{L^L I^{L-1} e^{-LI}}{\Gamma(L)}, \quad (1.17)$$

де L — кількість поглядів, $I > 0$.

За великої кількості поглядів ($L > 10$) розподіл наближається до гаусового. Спектральний аналіз підтверджує просторово-корельовану природу шуму, оскільки більша частина енергії зосереджена в низькочастотному спектрі. Це враховується при фільтрації та тематичній обробці даних, що значно впливає на ефективність аналізу.

Приклади зображень, спотворених СЗШ, наведено в Додатку Е, на рис. Е.9-Е.13.

Вивчення моделі шуму є дуже важливою частиною обробки зображень. Без попереднього знання моделі шуму не можемо розробити та виконати ефективні дії з придушення шумів. Отже, тут було розглянуто та представлено різні моделі шуму, доступні в цифрових зображеннях.

В Додатку Е, на рис. Е.14 наведено приклади РСА зображень системи Sentinel-1 і відповідних двовимірних ДКП-спектрів для вертикальної та горизонтальної поляризацій.

1.5 Методи фільтрації зображень

Шум є одним із основних факторів, які більшою чи меншою мірою погіршують якість усіх зображень, отриманих різними типами систем візуалізації [118], [119]. Нелінійна фільтрація зображень є однією з найбільш досліджуваних областей цифрової обробки зображень. Інтенсивне вивчення цієї тематики ведеться

вже понад 30 років. Вагомий внесок у розвиток теорії нелінійної фільтрації зробили такі вчені, як Д.С. Лі, В.П. Ярославський, Я.Т. Астола, К.О. Єгіазарян, В.В. Лукін, Е.П. Сімонселлі, Х. Портілья, А.Н. Венецанопулос, А. Фої та інші.

Бувають практичні ситуації, коли інтенсивність шуму невисока і немає необхідності у видаленні шуму. Прикладами є кольорові (цифрові фото) зображення, отримані в умовах хорошого освітлення [118], [120], або більшість компонентних зображень багато- та гіперспектральних даних ДЗ в піддіапазонах з великим відношенням сигнал/шум [121]. Однак існує досить багато практичних випадків, коли придушення шуму є бажаним для покращення візуальної якості зображення чи відео [120], [122], [123], для забезпечення кращої класифікації даних дистанційного зондування [124] тощо. Зокрема, необхідність придушення шумів є очевидним для цифрових зображень і відео, отриманих в умовах поганого освітлення [120], [125], [126], для деяких піддіапазонів гіперспектральних і мультиспектральних даних [121], для радіолокаційних зображень тощо.

На сьогоднішній день існує велика кількість розроблених методів придушення шумів на зображеннях. Проблема полягає в тому, що багато з них вимагають інтенсивних обчислень, наприклад, ті методи, які базуються на нелокальному підході, який інтенсивно вивчається та вдосконалюється в останні роки [103], [125], [127], [128]. При цьому часто буває бажано виконувати придушення шуму ефективно, просто і одночасно достатньо швидко. Наприклад, це потрібно для придушення шумів у зображеннях і відео в цифрових камерах, у попередній обробці зображень на борту космічних носіїв, де обмежені обчислювальні та енергетичні засоби. Тоді швидкодія фільтрації становить головний інтерес, хоча інші характеристики ефективності фільтра також важливі. Операції, які використовуються для усунення шумів, часто мають бути стандартними, простими, швидкими та/або добре узгоджуватися з іншими операціями, що використовуються на початкових етапах обробки зображення, як оцінка параметрів шуму та/або стиснення. Це значно зменшує кількість класів застосовних методів придушення шумів, що робить необхідним пошук компромісних рішень на основі фільтрів у ковзному вікні [77], [121] і методів

усунення шумів на основі простих і ефективних ортогональних перетворень [129], [130].

На основі теорій і методів, прийнятих алгоритмами придушення шумів, їх можна загалом класифікувати за такими категоріями:

- 1) фільтрація у просторовій області (у ковзному вікні), яка залежить від просторової кореляції значень пікселів,
- 2) фільтрація в області ортогональних перетворень, яка використовує енергетичну компактність зображення,
- 3) методи на основі розрідженого представлення, які базуються на словниках вивченого представлення,
- 4) методи на основі низького рангу, які використовують адаптивні словники
- 5) методи на основі глибокого навчання, які використовують глибокі згорткові нейронні мережі з врахуванням особливостей шуму.

Простий лінійний фільтр [131] здійснює усереднення значень пікселів у межах фільтраційного вікна без урахування будь-якої додаткової інформації. Хоча цей метод добре пригнічує шум, він також надмірно згладжує деталі зображення.

Медіанний фільтр [132], як і його узагальнення — альфа-усічений фільтр, демонструє робастність, тобто здатність ефективно боротися з імпульсним шумом. Однак цей фільтр менш ефективний у збереженні дрібних деталей та текстур. Для покращення результатів збереження деталей були запропоновані його модифікації, наприклад, центрально-зважений медіанний фільтр.

Подальший розвиток медіанної фільтрації призвів до створення FIR-медіанних фільтрів [133], які враховують просторову інформацію у ковзному вікні. Вони краще зберігають деталі зображень, хоча дещо поступаються в ефективності придушення шуму на однорідних ділянках.

Фільтри Фроста [134] є експоненціально зваженими адаптивними фільтрами, що виконують придушення спекл-шуму в зображеннях за допомогою методу мінімального середньоквадратичного оцінювання. Техніка придушення шуму, адаптована фільтром Фроста для придушення спекл-шуму, присутнього в центральному пікселі (i, j) , у вікні обробки розміром $(2n+1) \times (2n+1)$. Ефективність

роботи фільтра Фроста значною мірою залежить від оптимального вибору коефіцієнта демпфування (або налаштування), оскільки менший сприяє кращому придушенню шуму, але не дозволяє фільтру задовільно зберегти межі, тоді як більший сприяє кращому збереженню меж, але не дозволяє фільтру ефективно придушувати шум. Отже, виконання придушення шуму в усьому зображенні за допомогою постійного значення коефіцієнта демпфування часто призводить до отримання неоптимальних результатів. В ідеалі він має змінюватися адаптивно відповідно до варіацій локальної статистики пікселів, що знаходяться у вікні обробки, щоб досягти кращих результатів, але класичний фільтр Фроста не дозволяє змінювати ці параметри адаптивно.

Сигма-фільтр (фільтр Лі) [75], [135] заснований на статистичних властивостях гаусового шуму, який зазвичай припускається присутнім у багатьох типах цифрових зображень. У розподілі Гауса більшість значень потрапляє в певний діапазон (наприклад ± 2 сигма) навколо середнього значення. Цей статистичний діапазон дозволяє сигма-фільтру ідентифікувати та виключати пікселі, які значно відрізняються від інтенсивності центрального пікселя, ймовірно, представляючи шум або інші особливості (наприклад, межі), а не аналогічні сусідні пікселі. Усереднюючи лише пікселі в цьому діапазоні, сигма-фільтр зберігає важливі структури зображення на відміну від традиційних фільтрів, які, як правило, розмивають межі. Сигма-фільтр є обчислювально ефективним і відносно простим у реалізації. Незважаючи на свою ефективність, сигма-фільтр має обмеження: він може бути не настільки ефективним для видалення висококонтрастних ізольованих шумових ділянок. Для зображень із негаусовим шумом ефективність сигма-фільтра може бути знижена, оскільки припущення щодо розподілу шуму можуть бути недотриманими.

Спекл-фільтрація є складною задачею в обробці РСА-зображень, до якої протягом останніх трьох десятиліть підходили багато дослідників. Ці піонерські роботи, представлені Лі [136], Фростом та ін. [137] і Куан та ін. [138] досягли успіхів у однопоглядній фільтрації РСА-зображень, що, ймовірно, є найважчим випадком для усунення шумів. Його складність пов'язана з вищезазначеною

природою та властивостями спеклу, а також з необхідністю задовольнити низку суперечливих вимог. Бажано придушити шуми в однорідних областях зображення, зберегти середні значення в цих областях і одночасно зберегти цінну інформацію у вигляді меж, об'єктів малого розміру (мішеней) і особливостей текстури. За останні двадцять років було запропоновано багато фільтрів для зменшення шуму РСА-зображень. Варто згадати декілька підходів. Локально адаптивні спекл-фільтри були розроблені в 90-х роках минулого століття і на початку цього століття [139]. Ці методи засновані на жорсткому або м'якому перемиканні між двома або декількома різними виходами фільтрів, приділяючи особливу увагу виявленню локальної неоднорідності та збереженню деталей. Практично в той же час стали популярними методи придушення шумів на основі вейвлетів [140]. Типовою особливістю видалення спеклу на основі вейвлетів є те, що зазвичай використовується триетапна процедура, де пряме гомоморфне перетворення логарифмічного типу застосовується до даних на першому етапі, щоб отримати зображення, спотворене адитивним шумом. Крім того, проблема вейвлет-фільтрів полягає в тому, що їм досить важко ретельно адаптуватися до просторової кореляції спеклів. Було запропоновано декілька методів, заснованих на відмінних від вейвлет-ортогональних перетвореннях (наприклад, ДКП) [141]. Ці методи більш ефективні для придушення спеклів в однорідних областях зображення, ніж вищезгадані фільтри Лі та Фроста. Однак збереження меж і деталей фільтрів на основі ДКП все ще потребує вдосконалення.

У сенсі збереження деталей нещодавно було запропоновано цікавий і потужний інструмент під назвою метод Level Set [142]. Спочатку підхід Level Set був розроблений для сегментації зображення. Очевидною перевагою цього підходу є те, що він здатний надійно виявляти та локалізувати неоднорідності на зображеннях. Цю властивість також можна використовувати для фільтрації РСА-зображень. Існують вражаючі приклади обробки реальних РСА-зображень, хоча підхід Level Set не був інтенсивно статистично перевірений і кількісно охарактеризований. Таким чином, існує потреба в тому, щоб оцінити цей підхід для обробки однопоглядових РСА-зображень та поєднати його з ДКП та іншою

фільтрацією на основі ортогональних перетворень. Застосування детектора на базі методів машинного навчання для отримання інформації про локальну активність зображення може призвести до покращення показників ДКП-фільтра. Тому завданням є також розробка комбінованого методу локально-адаптивної фільтрації однопоглядових РСА-зображень.

Люди, які мають справу з фільтрацією зображень, знають, що часто важко передбачити, чи стане якість зображення кращою після придушення шумів чи ні порівняно з якістю оригінального (зашумленого) зображення, яке підлягає обробці. Такий прогноз, якщо він достатньо надійний і точний, може заощадити зусилля та час у ланцюжку обробки зображень, особливо якщо є необхідність, щоб цей ланцюжок працював автоматично. Така потреба може існувати в цифрових камерах, у системах дистанційного зондування, особливо якщо вони працюють автономно, наприклад, на борту супутника, і якщо кількість зображень (або компонентів зображення), що підлягають обробці, є великою, як це відбувається в мульти- і гіперспектральних системах дистанційного зондування. Підхід до прогнозування для випадку адитивного шуму було, вперше запропоновано в роботі [5]. Але є необхідність показати, що той самий підхід, відповідним чином модифікований, також може бути застосовний до ситуацій, коли зображення спотворені залежним від сигналу шумом з апіорно відомими або точно оціненими параметрами.

Проривом у теорії нелінійної фільтрації стало використання ортогональних перетворень, таких як ДКП. Це дозволило розділяти інформаційні й шумові складові, значно підвищуючи ефективність фільтрації, зокрема для текстурованих ділянок. Найбільш ефективним на сьогодні є фільтр VM3D [78], який застосовує тривимірне ДКП. Цей метод знаходить схожі ділянки у зображенні, групує їх у тривимірні блоки та виконує над ними декореляцію. Завдяки цьому VM3D забезпечує найкращі результати як у придушенні шуму, так і у збереженні інформаційних деталей. Розвиток фільтрації спрямований на врахування більш складних характеристик зображень, наприклад, схожості між ділянками різних

зображень, що дозволяє створювати бази (кодові книги) фрагментів без шуму для покращення результатів обробки.

Дослідження показують, що фільтрація на основі ДКП є однією з найефективніших за критеріями ступеня придушення шуму та швидкодії. Алгоритм роботи такого ДКП-фільтра можна описати так. Фільтрація зазвичай проводиться у квадратних блоках (наприклад, 8x8 або 16x16 пікселів, що забезпечує прийнятну швидкість внаслідок можливості використовувати швидкі алгоритми). У кожному блоці виконується пряме ДКП, застосовується порогова обробка і виконується зворотне ДКП. В результаті для всіх елементів блоку одержують первинно відфільтровані значення. Блоки можуть вибиратися без перекриття, з частковим перекриттям та з повним перекриттям. Дослідження показують ефективність застосування обробки з повним перекриттям блоків. Для кожного пікселя фільтровані значення, отримані у різних блоках, яким належить цей піксель, усереднюються.

Втім, в теорії та практиці фільтрації зображень залишаються відкриті питання [128], [143]. Одне з них полягає в тому, що шум найчастіше вважається адитивним, хоча його слід розглядати як сигнал-залежний для більшості вищезгаданих типів зображень [144], [145], [146], [147]. Інша проблема полягає в тому, що ефективність фільтрації залежить від багатьох факторів, включаючи властивості зображення, яке підлягає обробці, інтенсивність і тип шуму, доступність апріорної інформації про характеристики шуму, тип фільтра та його параметри тощо.

Останнім часом згорткові нейронні мережі з'явилися як потужний інструмент для придушення шумів у зображенні [148], [149], [150]. Однак ці ЗНМ ефективні для придушення шуму лише тоді, коли вони навчені певним характеристикам. Наприклад, модель DRUNet [148] навчена пригнічувати АБГШ і вимагає двовимірної карти значень СКВ шуму як входні дані. Ця карта має бути оцінена [151] окремим методом, наприклад SDNet, або бути апріорно відомою. Використання цієї карти дозволяє DRUNet ефективно пригнічувати шум із просторово-змінним СКВ на зображенні. Однак DRUNet не здатний ефективно

пригнічувати просторово корельований шум. Деякі інші CNN, такі як VNet [149], не вимагають карти шуму як вхідних даних, а замість цього оцінюють її автоматично. Однак це відбувається за рахунок обмеження спектром і розподілом шуму, для якого мережа була навчена. Загалом, основним недоліком НМ є задовільна узагальнююча спроможність щодо даних, на яких було цю мережу навчено. Якщо розподіл нових даних буде відрізнятися від тренувальних даних, результат роботи НМ може бути непередбачуваним. Крім того, переваги НМ-методів фільтрації у порівнянні з найкращими традиційними зазвичай мають місце при нереально високих інтенсивностях завад.

Значний інтерес викликають методи, що використовують спектрально-просторові моделі, оскільки вони дозволяють одночасно враховувати спектральні та геометричні характеристики зображень. Багатоканальна фільтрація активно інтегрується у задачі дистанційного зондування Землі, де вона використовується для аналізу мультиспектральних і гіперспектральних даних, а також для покращення точності прогнозування екологічних і кліматичних змін [152].

Підсумовуючи, можна сказати, що особлива увага має бути приділена методам придушення складних (сигнально-залежних, просторово-корельованих) шумів. При цьому, має бути забезпечена прийнятна швидкодія, а також бути враховано візуальну якість. В цих випадках необхідно аналізувати декілька варіантів просторово-корельованих шумів (або використовувати корельовані шуми з характеристиками, що є притаманні конкретним популярним системам). Деякі тестові зображення наведено в Додатку Е, на рис. Е.15, Е.16. Також слід додати, що для деяких задач відсутні тестові зображення, тому доводиться створювати їх власноруч.

Дослідження зосереджено на методах на основі ДКП, тому що вони забезпечують просте керування своїми параметрами, а також забезпечують прийнятну швидкість.

1.6 Методи стиснення зображень із втратами

Як відомо, існують методи стиснення зображень без втрат і з втратами [153]. Межі, досяжні за допомогою стиснення без втрат, здаються практично досягнутими. Коефіцієнт стиснення (КС) для існуючих методів зрідка сягає 5 навіть для стиснення гіперспектральних даних, коли міжканальна кореляція використовується повною мірою [154]. Однак часто потрібні більші значення КС. Тому необхідно застосувати стиснення отриманих даних ДЗ із втратами. Основною особливістю стиснення з втратами є те, що воно вносить спотворення в зображення ДЗ, за рахунок чого зображення може деградувати. Стиснення з втратами може бути корисним лише за умови, що внесені спотворення не мають достатньо негативного впливу на цілі, для досягнення яких призначені отримані дані ДЗ (класифікація місцевості та/або оцінка параметрів, виявлення конкретного об'єкта тощо). Одне з припущень полягає в тому, що внесені спотворення повинні бути такого ж рівня або меншими, ніж погіршення через шум у вихідних даних [155]. Таким чином, шумові характеристики повинні бути прийняті до уваги і тому вони повинні бути відомі заздалегідь або попередньо оцінені [156], [157], [158]. Це також означає, що необхідно мати можливість контролювати внесені спотворення та/або забезпечити бажаний рівень спотворень. Більше того, часто це має відбуватися автоматично, наприклад, у режимі бортової компресії [8]. Дещо інше припущення можливе, якщо стиснуті зображення підлягають візуальному огляду та аналізу. Тоді внесені спотворення повинні бути такими, щоб не погіршувати візуальну якість зображення [159]. Тоді необхідно враховувати як специфічні властивості зображень, наприклад варіації їх динамічного діапазону [124], [156], [160], так і особливості системи зору людини (СЗЛ). Нарешті, ще одне припущення полягає в тому, що внесені спотворення повинні бути такими, щоб вони не мали (помітного) негативного впливу на точність класифікації або виконання інших операцій обробки даних ДЗ на завершальних етапах. Зазначимо, що зниження точності класифікації пов'язане з метриками, що характеризують внесені спотворення [161].

Стиснення зображень із втратами — це техніка, яка використовується для зменшення розміру файлу цифрових зображень шляхом відкидання частини вихідних даних, що призводить до отримання меншого зображення з більшою ефективністю зберігання, яке наближається до оригіналу, але не є ідентичним. Ця форма стиснення широко використовується в програмах обробки цифрових зображень, щоб збалансувати якість зображення та розмір файлу, особливо там, де пам'ять і пропускна здатність обмежені, наприклад у веб-контенті та мобільних додатках, дистанційному зондуванні та медичній діагностиці та ін.

Стиснення зображень із втратами базується на концепції перцептивної надлишковості — частини зображення, до яких зорова система людини (HVS) менш чутлива, можуть бути стиснуті з меншою точністю без помітного впливу на якість сприйняття. Цей підхід використовує обмеження людського зору, який більш чутливий до змін яскравості, ніж до тонких відмінностей у кольорі, а також менш чутливий до високочастотних деталей (таких як дрібні деталі та текстури).

Основною перевагою стиснення зображень із втратами є те, що воно значно (у рази чи у десятки разів) зменшує розмір файлу, забезпечуючи швидшу передачу та зменшуючи потреби в сховищі. Наприклад, стиснення стандарту JPEG часто може зменшити розмір файлу зображення більше, ніж на 90%, що робить його популярним для онлайн-зображень, де швидкість завантаження та пропускна здатність є ключовими факторами. Однак це відбувається ціною безповоротної втрати даних. Основним обмеженням стиснення з втратами є артефакти — візуальні спотворення, що виникають під час стиснення. Зображення JPEG, наприклад, можуть мати блоковий ефект, дзвін або розмитість, особливо при вищих ступенях стиснення. Артефакти стають особливо помітними на зображеннях із високим контрастом або дрібними деталями та можуть погіршити якість зображення до неприйняттого рівня. Останні розробки у стисненні зображень із втратами включають стиснення на основі вейвлетів (використовується у таких форматах, як JPEG 2000) і методи на основі трансформації в глибокому навчанні [162]. Вейвлет-стиснення, наприклад, дозволяє аналізувати дані з різною роздільною здатністю, пропонуючи кращу якість за нижчих бітрейтів і зменшуючи

видимі артефакти. Методи, керовані штучним інтелектом, які використовують НМ для стиснення зображень на основі вивчених шаблонів, є перспективними у забезпеченні високих коефіцієнтів стиснення з мінімальною втратою сприйняття [163]. Але однією з ключових проблем є неясність процедури керування КС для забезпечення необхідної якості.

Стиснення з втратами здатне забезпечити значно більші значення КС, але за рахунок внесених спотворень. Завжди є проблема досягти відповідного компромісу між якістю стисненого зображення (що характеризується багатьма різними способами) і КС [153], [164], [165]. Для цього є декілька причин. По-перше, стиснення, що забезпечує фіксований КС, який часто використовується на практиці [166], [167], призводить до стиснених зображень, якість яких може змінюватися в широких межах [168]. Для заданого КС для зображень з простішими структурами внесені спотворення менші, а якість стисненого зображення вища, тоді як для зображень із більш складною структурою спотворення більші, і, отже, якість зображення може бути незадовільною, оскільки деяка важлива інформація може бути неминуче втрачена, що є небажаним. Звичайно, деякі покращення можуть бути досягнуті завдяки використанню кращого кодера, адаптації до вмісту зображення [169], використанню міжканальної кореляції шляхом тривимірного стиснення [166], [167] або у деякі інші способи. Однак позитивний ефект може бути обмеженим, і/або можуть існувати деякі обмеження, наприклад, необхідність застосування стандарту стиснення зображення [153], [167]. Таким чином, є необхідність в аналізі або розробці підходів, які передбачають стиснення з втратами багатоканальних зображень ДЗ з певним контролем внесених спотворень у поєднанні з одночасним бажанням забезпечити більший КС. Пояснимо переваги такого підходу, коли він може бути доцільним і які умови для його реалізації. Нагадаємо, що дані ДЗ, стиснуті із втратами, можуть бути корисними, якщо: (а) внесені втрати не призводять до суттєвого зниження ефективності вирішення кінцевих завдань обробки зображень ДЗ, таких як сегментація, виявлення об'єктів, спектральне розмішування, класифікація, оцінка параметрів [82], [170], [171]; (б) спотворення через стиснення з втратами не з'являються самі як артефакти, що

призводять до небажаних негативних ефектів під час вирішення вищезгаданого кінцевого завдання.

Як було зазначено вище, ще одною важливою проблемою є присутність шуму на зображеннях ДЗ та медичних зображеннях. Цей нюанс також треба враховувати при розробці чи модифікації методів стиснення із втратами. Більше того, бажано враховувати природу шумів, щоб забезпечити найкраще стиснення з точки зору візуальної якості [102].

Аналогічно фільтрації, будемо розглядати одноканальні та багатоканальні зображення. Основну увагу буде приділено методам стиснення на основі ортогональних перетворень. Це пов'язано з найбільш прийнятним компромісом між якістю та ступенем стиснення, швидкістю стиснення та простотою забезпечення необхідної якості [102]. Крім того, методи на основі ортогональних перетворень є одними з найбільш ефективних в плані кривих швидкості-спотворення (КШС) та простими в плані варіювання характеристиками. Кожен кодер має параметр, що контролює стиснення (ПКС), для якого зрозуміла поведінка метрик якості [172].

Розглянемо найбільш використовувані кодери. Найбільш відомим і використовуваним методом стиснення зображень із втратами є JPEG (Joint Photographic Experts Group) [173], який використовує ДКП для розбиття зображення на частотні компоненти.

JPEG2000 є іншим стандартом стиснення зображень [87], розробленим як вдосконалена альтернатива класичному JPEG. Основою JPEG2000 є використання дискретного вейвлетного перетворення, яке забезпечує багатомасштабне подання зображення, що дозволяє ефективно розділяти низько- та високочастотні компоненти. Стандарт використовує арифметичне кодування на основі алгоритму EBCOT (Embedded Block Coding with Optimal Truncation), який забезпечує прогресивне відновлення зображення. JPEG2000 підтримує гнучке управління якістю через багаторівневу структуру кодового потоку, що дозволяє масштабувати якість зображення залежно від пропускну здатності каналу або доступних ресурсів. У порівнянні з JPEG, стандарт JPEG2000 має кращу стійкість до

артефактів стиснення, зокрема немає блокового ефекту, завдяки обробці даних у частотній області. Стандарт також надає можливості для обробки багатоспектральних та гіперспектральних даних завдяки підтримці багатоканальних зображень. JPEG2000 широко використовується у задачах дистанційного зондування (наприклад, для зображень супутників Sentinel-2), де потрібне ефективне стиснення великих обсягів даних із мінімальними втратами якості. У медичній візуалізації стандарт забезпечує високу якість реконструкції зображень для таких форматів, як DICOM, зберігаючи важливі діагностичні деталі.

Алгоритм Set Partitioning in Hierarchical Trees (SPIHT) — це високоефективна техніка стиснення зображень на основі вейвлетів, розроблена Аміром Саїдом і Вільямом А. Перлманом у 1996 році [174]. SPIHT широко оцінюється завдяки своїй ефективності стиснення, прогресивній передачі, а також високоякісній реконструкції, особливо в програмах, де потрібна висока точність, наприклад, медичні зображення та супутникові зображення.

Кодер AGU [175] — це метод стиснення зображень на основі ДКП, який базується на традиційному підході JPEG, але вдосконалює його, усуваючи такі обмеження, як блокові артефакти та якість стиснення. Цей підхід використовує три ключові вдосконалення: використання більших блоків 32x32 пікселів замість блоків 8x8, оптимізовану систему кодування в бітовій площині та постобробку для зменшення артефактів. Більші розміри блоків у більшості випадків підвищують ефективність стиснення, захоплюючи більше даних зображення на блок, що призводить до меншої кількості блокових артефактів і забезпечує вищу якість зображення в розпакованому зображенні. Кодер AGU вирішує проблему деблокінгу шляхом включення фільтрації на основі ДКП як етапу постобробки. Крім того, існує версія AGU-M, що також враховує ефект маскування.

Розширений кодер на основі дискретного косинусного перетворення (ADCT) — це алгоритм стиснення зображень [176], призначений для досягнення високоякісного стиснення з нижчою обчислювальною складністю, ніж інші вдосконалені кодери зображень, такі як JPEG2000 і AGU. Розроблений як удосконалення традиційних методів ДКП, ADCT інтегрує вдосконалені схеми

розділення, контекстне кодування та спрощені методи кодування коефіцієнтів для покращення ефективності стиснення та якості зображення. ADCT використовує ДКП, яке перетворює числове представлення зображення в частотні компоненти, дозволяючи виділяти та стискати важливу візуальну інформацію, відкидаючи зайві дані. Однак замість використання фіксованого розміру блоку (наприклад, блоки 8x8 у стандартному JPEG) ADCT використовує оптимізовану схему розділення зі змінними розмірами блоків. Розміри блоків можуть варіюватися від 8x8 до 256x256 пікселів, що дозволяє адаптуватися як до дрібних деталей, так і до більших однорідних областей зображення. Крім того, існує розроблена ADCT-M [176], що також враховує ефект маскування.

Кодек BPG (Better Portable Graphics) — це метод стиснення зображень, представлений Фабрісом Беллардом у 2014 році [177], [178]. Його основна мета — забезпечити вищий коефіцієнт стиснення, ніж JPEG, із збереженням найвищої якості зображення. Формат BPG досягає цього завдяки використанню прогресу сучасних відеокодеків, зокрема на основі стандарту HEVC (високоєфективне відеокодування), також відомого як H.265 [179]. На відміну від JPEG, який зазвичай обмежується 8-бітною глибиною кольору, BPG підтримує до 14 біт на канал, що дозволяє захоплювати ширший діапазон кольорів і градієнтів. Як правило, BPG забезпечує значно кращу якість зображення, ніж JPEG, за того самого розміру файлу. Хоча JPEG бореться з артефактами стиснення на низьких бітрейтах, BPG зберігає чіткість і деталізацію, що робить його придатним для використання в Інтернеті, де пропускна здатність обмежена.

High Efficiency Image File Format (HEIF) — це передовий формат зображень, розроблений Moving Picture Experts Group (MPEG) і завершений у 2015 році [180]. Він створений як ефективний контейнер для зображень і послідовностей зображень, здатний значно зменшити розміри файлів, зберігаючи високу якість. HEIF зазвичай асоціюється з HEIC (High Efficiency Image Coding), оскільки він часто використовує HEVC (High Efficiency Video Coding) для стиснення. Однак HEIF є контейнерним форматом, тобто він може зберігати зображення, стиснуті різними кодеками, хоча HEVC є найпоширенішим.

AVIF (формат файлу зображення AV1) — це розширений формат стиснення зображень, який походить від відеокодека AV1 (AOMedia Video 1), розробленого Alliance for Open Media (AOM) [181]. Розроблений як для високої ефективності, так і для якості, AVIF надає безкоштовну альтернативу таким форматам, як JPEG, PNG і HEIC, з відкритим вихідним кодом і забезпечує високу ефективність у підтримці якості зображення на низьких бітрейтах. AVIF отримав популярність завдяки своєму потенціалу для забезпечення менших розмірів файлів і чудової якості зображень, особливо у веб-додатках, фотографіях і ескізах відео. Подібно до інших сучасних кодерів зображень, AVIF використовує блоковий підхід кодування, розбиваючи зображення на одиниці, які можна незалежно обробляти та стискати. AVIF підтримує широкі колірні гами з глибиною кольору до 12 біт на канал, що робить його придатним для додатків високої точності.

Два приклади поведінки КШС [182] представлені в Додатку Е, на рис. Е.17-Е.18 для традиційних тестових зображень у відтінках сірого Peppers і Baboon. Перший має просту структуру (багато квазіоднорідних областей), тоді як другий має складну структуру (є майже повністю текстурним). Залежності представлені для коефіцієнта стиснення (КС) від 8 до 64 (іноді до 32), тобто в типовому діапазоні значень КС. Розглядаються три методи стиснення: JPEG з рівномірним квантуванням; SPIHT; та модифікація кодера ADCT під назвою ADCT-M, призначена для забезпечення високої якості зображення за рахунок нерівномірного квантування коефіцієнтів ДКП [168].

Відповідно до теорії всі залежності монотонно спадають. Але між ними є достатні відмінності, які в загальних рисах можна сформулювати так:

1) значення PSNR трохи кращі для SPIHT для КС приблизно на 10; для більшого КС, найкращі значення PSNR надає ADCT-M; JPEG – найгірший;

2) найкращі значення PSNR-HVS-M забезпечує ADCT-M, перевага порівняно з іншими кодерами значна, декілька дБ; це показує, що достатня перевага у візуальній якості стиснених зображень може бути забезпечена сучасними методами стиснення, такими як ADCT-M [168].

Важливим спостереженням є те, що для $KC=8$ значення PSNR становлять близько 37 дБ для тестового зображення Peppers і близько 28 дБ для тестового зображення Baboon, що показує, що MSE спотворень значно відрізняються. Деградації, очевидно, більші для зображення зі складнішою структурою. Це відбувається для всіх трьох розглянутих кодерів. Це має місце для всіх кодерів, і це проблема не тільки при стисненні зображення, але і при стисненні відео з фіксованим бітрейтом [168].

Щоб вирішити цю проблему, можна застосувати ітераційне стиснення/декомпресію з контролем якості зображення після кожного кроку ітерації та подальшим налаштуванням параметру, контролюючого стиснення (PKC).

Переваги цієї процедури полягають у наступному [168]:

- 1) вона застосовна для будь-якої метрики якості за умови, що залежність цієї метрики від PKC є монотонною (ця умова зазвичай справедлива, хоча її слід перевірити заздалегідь для досліджуваного кодера та відповідної метрики);
- 2) зазвичай можна досягти високої точності надання бажаного значення використовуваної метрики; зазвичай можна забезпечити бажаний PSNR або PSNR-HVS-M з похибкою менше 0,2 дБ, і це підходить для практичного застосування.

Водночас ітераційна процедура має ряд недоліків, а саме:

- 1) апіорі невідомо, скільки кроків ітерації знадобиться, щоб здійснити стиснення даного зображення з бажаною якістю згідно з обраною метрикою даним кодером; в [168] показано, що можливі 10 кроків ітерації (а інколи навіть більше);
- 2) параметри ітераційної процедури (початкове значення PKC та його крок, точність надання метрики) мають бути обґрунтовано задані; це може у середньому зменшити кількість кроків ітерації;
- 3) потрібен великий час для завершення ітерацій, якщо використовувана техніка стиснення/декомпресії недостатньо швидка; наприклад, кодер ADCT-M виконує оптимізацію схеми розбиття та стиснення зображень блоків контентно-адаптивного розміру і це потребує певного часу.

Деякі пропозиції щодо того, як зменшити кількість ітерацій, можна знайти в роботі [168]. Зокрема, початкове значення ПКС може бути встановлено на основі попередньо зібраної статистики. Нижче наведемо один приклад, запозичений з [168]. В Додатку Е, на рис. Е.19 представлено усереднені залежності КК від метрики якості (PSNR-HVS-M у розглянутому випадку) для двох кодерів на базі ДКП, AGU [175] і ADCT [183]. Як видно, КК близько 17 для кодера AGU забезпечує компресію в середньому з PSNR-HVS-M близько 41 дБ, тобто з практично непомітними спотвореннями. Однак значення PSNR-HVS-M для окремих зображень коливаються в досить широких межах від 39 дБ до 46 дБ. Тим часом зазвичай бажано забезпечити PSNR-HVS-M з похибкою менше 0,5 дБ. Таким чином, навіть якщо початковий КК встановлено належним чином, потрібні ітерації.

В Додатку Е, на рис. Е.20 представлено усереднені залежності КК від PSNR-HVS-M для двох кодерів на базі ДКП AGU-M та ADCT-M, які є модифікаціями AGU та ADCT відповідно, які враховують особливості HVS шляхом квантування більш високі просторові частоти з більшими кроками квантування (ці кроки визначаються заданим КМ і спеціальними матрицями). Для цих кодерів проблема однакова. Наприклад, щоб забезпечити бажаний PSNR-HVS-M=41 дБ, потрібно використовувати КМ близько 10. Але, фактично, встановлення фіксованого КМ=10 для кодера AGU-M забезпечує PSNR-HVS-M в межах від 40 до 44 дБ. Таким чином, необхідне додаткове коригування КМ.

Ітераційні процедури з рекомендованим кроком зміни потребують кількох кроків для забезпечення бажаного PSNR-HVS-M в обох розглянутих прикладах. Однак це не повністю вирішує проблему, і все одно бажано мати одноетапну або двоетапну процедуру стиснення зображення з бажаною якістю.

Величезна кількість даних сьогодні надається за допомогою існуючих датчиків ДЗ, як космічних, так і бортових [153], [184]. Обсяг даних є особливо великим, якщо зображення є гіперспектральними (тобто мають сотні зображень із піддіапазонами) та/або мають високу роздільну здатність. Зауважимо, що обидві тенденції (створення та використання багатоканальних систем, а також отримання даних високої роздільної здатності) характерні для останніх років. Обсяг

отриманих даних додатково збільшується за рахунок більш частих спостережень за зондованим рельєфом [184] – звичною практикою стало проведення моніторингу території досить часто, наприклад, щотижня. Отримані дані ДЗ необхідно передати, зберегти та/або розповсюдити. Для кожної з цих операцій може бути бажаним стиснення даних [153], [154], [164]. У той же час є декілька причин, які можуть перешкодити ефективному виконанню цих операцій. Щодо передачі даних: пропускна здатність каналу зв'язку, який використовується для передачі даних, може бути обмежена, час для передачі може бути обмежений, час і потужність для стиснення також можуть бути обмежені [153], [164]. Те саме стосується розповсюдження даних, хоча обмеження зазвичай менш суворі порівняно з передачею даних по низхідній лінії зв'язку. Пам'ять для зберігання даних ДЗ також може бути проблемою, незважаючи на швидкий розвиток нових засобів в останні роки [184]. Тому часто потрібно стискати зображення ДЗ [90], [154].

Таким чином, необхідно контролювати внесені спотворення для всіх вищезгаданих стратегій. Тут під «контрольованим» маємо на увазі декілька аспектів. По-перше, спотворення мають бути виміряні, оцінені або спрогнозовані, щоб переконатися, що вони не перевищують допустиме порогове значення відповідно до певної метрики (критерію) [185]. По-друге, внесені спотворення можна точно виміряти, лише якщо стиснення та декомпресія вже зроблені. Потім, якщо потрібно змінити рівень спотворення, потрібно змінити параметри кодера та виконати розрахунок метрики після наступної ітерації стиснення/декомпресії [186]. Це часто непрактично, особливо на борту. Тоді доцільніше говорити про оцінку або прогнозування спотворень без компресії та декомпресії, але з наближеним забезпеченням бажаної якості стиснутих даних. Звичайно, КС також може бути важливим. Потім необхідно забезпечити відповідний компроміс між КС та внесеними спотвореннями. Зауважте, що КС також залежить від використовуваного кодера. У цьому сенсі варто враховувати міжканальні кореляції, властиві багатоканальним даним ДЗ, що можна зробити різними способами [187]. Можна застосовувати різні перетворення [188] або здійснювати різні групування складових зображень [189]. Значну увагу приділено стисненню

зображень із втратами з урахуванням типу шуму [190] та характеристик [191]. Стверджено та досліджено можливе існування оптимальної робочої точки (ООП) та її прогноз [159]. Розглянуто проблеми прогнозування КС та його забезпечення для кодерів на основі ДКП [192]. Водночас, проблеми прогнозування якості стисненого зображення та забезпечення бажаної якості ще досконально не проаналізовані. У цьому напрямку проведено певну роботу. Зокрема, було запропоновано підхід до прогнозування якості для вейвлет-компресії зображень дистанційного зондування [193]. Відоме прогнозування MSE внесених спотворень для JPEG [194]. Однак контроль і прогнозування значень метрик для більш просунутих кодерів, таких як AGU [175] і ADCT [183], які значно перевершують JPEG [168], не були розроблені до останніх років.

Однією з типових ситуацій є те, що потрібно забезпечити бажаний КС (наприклад, потрібно передати зображення заданого розміру через лінію зв'язку протягом заданого інтервалу часу) [87]. Це була лише одна з причин розробки стандарту JPEG2000 і подібних методів. Тут є дві проблеми. По-перше, для даного КС якість стисненого зображення може змінюватися в досить широких межах залежно від складності зображення. Наприклад, варіації метрики візуальної якості PSNR-HVS-M [94] можуть перевищувати 20 дБ для того самого КС і, очевидно, такі варіації є величезними. Інтервал може включати зображення з PSNR-HVS-M близько 40 дБ, коли спотворення практично непомітні, і з PSNR-HVS-M близько 20 дБ, коли якість стиснених зображень дратує. По-друге, в даний час існують методи стиснення зображень, які значно перевершують JPEG2000 за характеристиками швидкості/спотворення [168], і бажаний КС для них можна досить легко забезпечити завдяки останнім досягненням [22], [195]. Інша ситуація полягає в наступному. Припустимо, що потрібно досить швидко забезпечити бажану якість стисненого зображення в першу чергу, а забезпечити більший КС в другу чергу. Тоді відразу виникає декілька питань:

- 1) який кодер використовувати?
- 2) як встановити ПКС, щоб забезпечити бажане значення метрики, яка характеризує якість?

3) яку метрику використовувати та які її значення для конкретного кодера?

1.7 Напрями дослідження

Дослідження в даній роботі спрямовані на вдосконалення підходів до обробки, фільтрації та стиснення зображень у контексті задач, пов'язаних із медичною візуалізацією, дистанційним зондуванням та іншими сферами, де важливе забезпечення високої якості зображень та ефективності їх обробки.

Основними напрямками дослідження визначено декілька взаємопов'язаних задач, які мають на меті підвищення ефективності та точності обробки зображень у різних сферах застосування.

По-перше, важливим завданням є створення нових методів виявлення текстур та неоднорідностей. Для цього пропонується аналіз локальних параметрів, що дозволяє ідентифікувати межі об'єктів та виділяти текстурні ділянки навіть у випадках, коли зображення значно зашумлені. Одним із рішень є використання відносної локальної дисперсії, квазірозмаху та коефіцієнтів косинусного дискретного перетворення, що дає можливість підвищити точність сегментації. Окрему увагу приділено застосуванню сучасних класифікаторів – зокрема, машин опорних векторів та нейронних мереж. Такі підходи здатні зменшити кількість хибних результатів, підвищити рівень достовірності класифікації текстур і водночас оптимізувати обчислювальні затрати.

Другою задачею визначено придушення шумів, насамперед спекл-шуму, що характерний для зображень радіолокаційних систем та медичної візуалізації. Для цього пропонуються адаптивні методи фільтрації, здатні враховувати як частотні, так і локальні характеристики сигналу. Це дозволить зменшити інтенсивність шумів, одночасно зберігаючи важливі структурні елементи, такі як межі органів або деталей об'єктів. Додатково планується впровадження інноваційних алгоритмів, серед яких методи на основі рівневих множин (Level Set) та вейвлет-перетворень. Вони орієнтовані на досягнення балансу між ефективним придушенням шуму та збереженням дрібних деталей.

Третій напрям пов'язаний із прогнозуванням ефективності фільтрації. Йдеться про створення моделей, здатних оцінити майбутні покращення якості ще до початку обчислень. Зокрема, планується аналіз змін метрик PSNR і PSNR-HVS-M з метою раціонального використання ресурсів та пришвидшення процесів обробки. Крім того, моделі повинні бути універсальними та придатними до різних типів шумів – адитивних, мультиплікативних і залежних від сигналу. Це забезпечить широку практичну придатність розробок.

Четверте завдання полягає у стисненні зображень із втратами при збереженні бажаної якості. Передбачається розробка двоетапних алгоритмів, які дадуть змогу мінімізувати втрати інформації при забезпеченні прийняттого рівня візуального сприйняття. Особливе значення це має у медичних застосуваннях, де навіть незначна втрата деталей може вплинути на діагностику. Планується створення методів, здатних адаптуватися як до простих зображень, так і до даних із високою текстурною насиченістю.

П'ятий напрямок присвячений дослідженню впливу шумів на процеси обробки. Йдеться про просторово корельовані шуми, що виникають, наприклад, у рентгенівських знімках або даних дистанційного зондування. Їхній вплив на класифікацію та інші етапи обробки може бути критичним. У цьому контексті важливим є вдосконалення алгоритмів стиснення, які враховують специфіку шумових структур, що дозволить підвищити ефективність у реальних умовах.

Шостий напрямок стосується покращення роботи з медичними та мультиспектральними даними. Запропоновано рекомендації щодо оптимального стиснення рентгенівських і стоматологічних знімків при збереженні діагностично важливої інформації. Також значну увагу приділено поканальному аналізу мультиспектральних зображень, що дає змогу оптимально налаштовувати параметри обробки для кожного спектрального каналу та зменшувати вплив високочастотних перешкод.

Усі зазначені завдання спрямовані на вирішення актуальних проблем, що виникають у медичній діагностиці, дистанційному зондуванні та інших сферах, де

якість візуальної інформації має вирішальне значення. Результати досліджень створюють основу для нових технологічних рішень, спрямованих на автоматизацію, підвищення точності та ефективності аналізу даних.

Додатково передбачається довести низку положень щодо прискорення обробки. Зокрема, планується показати необхідність використання більш складних моделей шуму під час проєктування методів. Також буде проаналізовано нові тенденції у фільтрації та стисненні зображень, потенціал яких поки що недостатньо реалізований. Нарешті, передбачається продемонструвати, що вибір конкретної стратегії обробки залежить від практичних обмежень і пріоритетів завдання.

Основні наукові результати, наведені в першому розділі, опубліковано в роботах автора [17], [19], [31], [34], [39], [46], зазначених в анотації.

РОЗДІЛ 2 ІНТЕЛЕКТУАЛЬНА ФІЛЬТРАЦІЯ ЗОБРАЖЕНЬ

В даному розділі розглянуто питання, що стосуються фільтрації зображень. В першу чергу розглянуто питання детектування текстур на спотворених завадами зображеннях за допомогою комбінування декількох параметрів локальної активності. Крім цього, розглянуто адаптивну фільтрацію зображень, спотворених сигнально-залежними та просторово-корельованими завадами, за допомогою алгоритмів машинного навчання. В останньому підрозділі розглянуто питання оцінки ефективності фільтрації для різних типів зображень. Основні результати автора, присвячені даним аспектам, відображено в публікаціях [1], [2], [5], [6], [7], [9], [10], [11], [13], [14], [15], [16], [17], [19].

2.1 Методи детектування та збереження текстур

2.1.1 Виявлення текстурних ділянок SVM-класифікатором на зображеннях, спотворених завадами

Текстурні ознаки широко використовують при сегментації зображень [196], класифікації матеріалів [197] та типів поверхонь на зображеннях дистанційного зондування [198], для пошуку зображень за подібністю у великих базах даних [72], комп'ютерній графіці і т.ін. Це пов'язано з тим, що текстури є одним з характерних елементів на зображеннях [119], [199]. При цьому існують і тією чи іншою мірою вирішені різні завдання виявлення текстурних ділянок. Наприклад, часто бажано виявити текстурну ділянку чи ділянки, для яких є еталон чи набір текстурних ознак [200], [201]. Однак є ситуації, коли необхідно виявити та визначити межі текстурних ділянок за відсутності апріорних відомостей про ознаки текстур [198], [202]. Таке завдання, зокрема, виникає у разі потреби оцінити текстурні ознаки, щоб потім використовувати їх при класифікації об'єктів [202], [203].

Одна з проблем при застосуванні зазначених видів обробки та аналізу полягає в тому, що вихідні зображення можуть бути спотворені завадами (адитивними,

сигнально-залежними, мультиплікативними залежно від типу зображень [79], [204], [205], [206], [207], [208], [209]). Присутність завад ускладнює виявлення текстурних ділянок [204] та видобування (оцінювання) текстурних ознак, знижує візуальну якість зображень загалом та текстурних ділянок зокрема [79]. У зв'язку з цим часто застосовують фільтрацію зображень для придушення завад. Однак придушення завад на текстурних ділянках є проблематичним для більшості відомих фільтрів, включаючи найсучасніші та найефективніші [128]. Навіть фільтр BM3D [78], який на даний момент є одним з найефективніших при придушенні адитивного білого гаусового шуму, у певних випадках лише трохи придушує завади та покращує візуальну якість текстурних зображень [79], [128], трохи перевищуючи простіші фільтри, наприклад, фільтр на основі ДКП [130], який показав себе одним із найкращих у плані збереження текстури [79], [204]. Тут слід зазначити таке. Для багатьох методів фільтрації при обробці текстурних ділянок використовують локально-адаптивний підхід [204], [205], [206], [208], [209], відповідно до якого для однорідних та неоднорідних (текстурних, околів меж та малорозмірних об'єктів) ділянок використовують різні фільтри або фільтри з різними параметрами. Такий підхід зазвичай приносить успіх, але для текстур та околів меж доцільно застосовувати різні фільтри [204]. Використовувані в [204] для вирішення такої задачі два локальні параметри (локальна дисперсія і квазірозмах, нормовані з урахуванням апріорно відомих параметрів завад), а також пост-обробка бінарних полів цих параметрів не завжди справляються з поставленою метою. Враховуючи, що статистика коефіцієнтів ДКП в блоках здатна характеризувати текстуру [79], ідея полягає в тому, що спільне використання трьох або більше параметрів, що описують текстурні ознаки з різних сторін, здатне поліпшити виявлення текстурних ділянок. Нижче розглядається один з можливих шляхів спільної обробки - на основі машини опорних векторів (МОВ, SVM) [210], що показала себе в останні роки одним з найбільш ефективних класифікаторів [211]. Мета цього підрозділу – розробити класифікатор «текстура – не текстура», здатний працювати за досить інтенсивних завад і проаналізувати його властивості. Розглядається випадок АБГШ.

Припустимо, що класифікація виконується попиксельно, тобто рішення про належність даного (розглянутого) пікселя класу «Текстура» або «Не текстура» виконується на основі застосування заздалегідь навченого класифікатора до набору вхідних ознак, розрахованих для ковзного вікна. Також припустимо, що навчання класифікатора виконується заздалегідь, offline. Для формування навчальної вибірки для класифікатора використано різні набори значень кількох локальних параметрів. Під локальністю мається на увазі те, що це параметри обчислюються у ковзних вікнах невеликого розміру: в даному випадку отримані та розглянуті результати для ковзних вікон розміру 5x5, 7x7 та 9x9 пікселів. Розглянуті локальні параметри обрані та застосовані, виходячи з таких міркувань:

- різні локальні параметри проявляють себе з неоднаковою ефективністю при детектуванні текстурних ділянок [204];

- використовувані параметри базуються на декількох різних ознаках та принципах детектування меж (або неоднорідностей), що дозволяє класифікатору бути більш універсальним по відношенню до властивостей даних, що обробляються.

Набір параметрів визначено експериментально і включає наступні параметри:

1) нормована (відносна) локальна дисперсія (ВЛД):

$$\delta_{ij} = \frac{\sigma_{ij}^2}{\sigma_{add}^2}, \quad (2.1)$$

де σ_{ij}^2 – локальна дисперсія, розрахована у ковзному вікні, а σ_{add}^2 – дисперсія адитивного шуму, присутнього на зображенні (припускаємо, що шум має відому або точно оцінену дисперсію). Цей параметр чутливий до слабкоконтрастних неоднорідних ділянок і має значно більші значення в ковзному вікні, що потрапив на межу або текстуру, ніж у вікні, що потрапив на однорідну ділянку;

2) квазірозмах (КР) (для випадку присутності адитивної завади з урахуванням нормування):

$$QR_{ij} = \frac{(I_{ij}^{(q)} - I_{ij}^{(p)})}{\sigma_{add}}, \quad (2.2)$$

де $I_{ij}^{(q)}$ та $I_{ij}^{(p)}$ – відповідно q -а та p -а порядкові статистики ранжованої вибірки. Значення параметрів q і p вибираються емпірично залежно від характеру завад, що присутні на зображенні. У дослідженні ці значення були встановлені рівними 23 і 3, 45 і 5, 72 і 10, відповідно (тобто $q \approx 0,9N$ і $p \approx 0,1N$, де N – загальна кількість пікселів, що належать ковзному вікну, $q+p=N+1$). Даний параметр веде себе подібно до відносної локальної дисперсії, робастно характеризуючи масштаб даних у ковзному вікні;

– детектор на основі ДКП, що базується на наступному принципі. У ковзному вікні обчислюється ДКП. Після цього розраховується кількість коефіцієнтів ДКП, що перевищили поріг, який у аналізованих дослідах був обраний рівним $2\sigma_{add}$. Даний параметр (може бути нормований на кількість пікселів у блоці N) працює в спектральній області та враховує той момент, що для текстур зазвичай характерна наявність достатньої кількості відносно явно виражених просторових спектральних компонентів.

Вибір значень порогів або способів нормування для локальних параметрів, що розраховуються, пов'язаний з типом завад, присутніх на оброблюваному зображенні. Випадок адитивних завад розглядається як найпростіший, але для сигнально-залежних чи мультиплікативних завад нескладно змінити нормування.

В якості навчального було використано зображення Barbara (рис. 2.4, а), спотворене адитивним гаусовим шумом з нульовим середнім та середньоквадратичним відхиленням (СКВ), рівним 5. Особливістю цього тестового зображення є те, що для нього відносно легко можна в інтерактивному режимі виділити області, які людина вважає текстурними, тобто сформувати набори пікселів, які відповідають класам «Текстура» та «Не текстура».

Підхід, який передбачає об'єднання кількох різних детекторів, виникає з таких міркувань. Параметри, що використовуються окремо, часто не здатні

ефективно «розділити» класи, що розглядаються («текстура» і «не текстура», в даному випадку). Для демонстрації цього наведемо декілька прикладів гістограм значень на виході трьох елементарних детекторів, отриманих окремо на текстурних і нетекстурних ділянках зображення (Додаток Ж, рис. Ж.1-Ж.3).

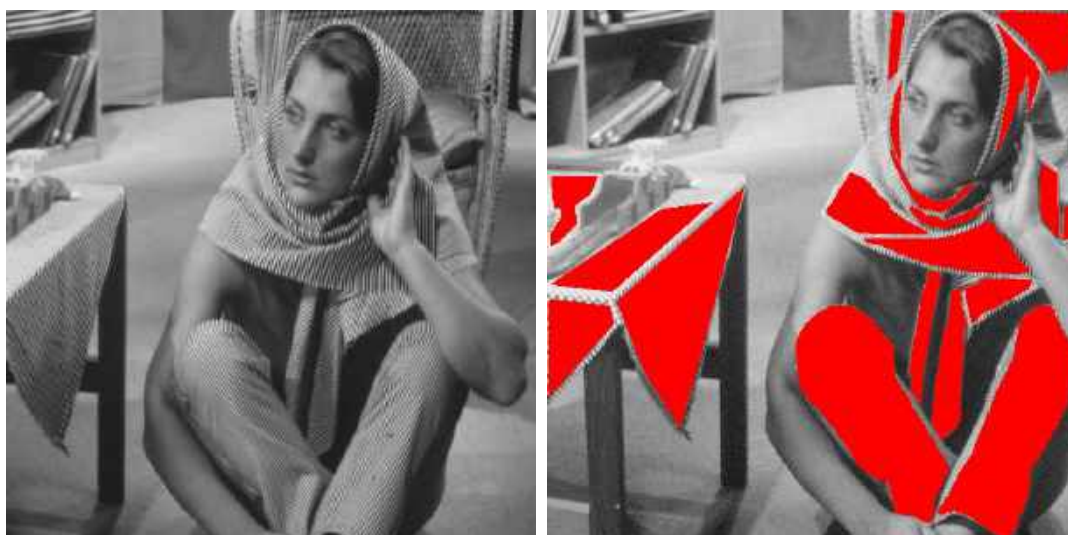
З урахуванням різної поведінки локальних параметрів для аналізованих класів має сенс спробувати об'єднати декілька найбільш ефективних елементарних детекторів у припущенні, що результуюча система (класифікатор) матиме більшу чутливість і специфічність.

Використовувався класифікатор на базі МОВ. Цей класифікатор вимагає попереднього навчання на заздалегідь розміченому наборі даних. Тому для кожної точки вхідного зображення з координатами $[x; y]$ формувалася вектор вхідних ознак F , що складається з точок з такими ж координатами $[x; y]$ у матрицях локальних параметрів. Цьому вектору F зіставлялася ознака S , що визначається за формулою:

$$S = \begin{cases} -1, \text{ нетекстура} \\ 1, \text{ текстура} \end{cases}, \quad (2.3)$$

При цьому вектор F був представлений у трьох варіантах: $F1=[ВЛД \text{ КР}]$, $F2=[ДКП]$, $F3=[ВЛД \text{ КР} \text{ ДКП}]$, де, наприклад, точкам вектора $F1$ відповідали точки локальних параметрів ОЛД і КР. Відповідно були позначені і класифікатори. Наприклад, $F1\_5 \times 5$ означає використання вектора властивостей $F1$ для ковзного вікна 5×5 пікселів.

Як зазначалося, зображення Barbara (рис. 2.1, а) було попередньо вручну розмічено (рис. 2.1, б). Червоний колір відповідав текстурним ділянкам. Інші ділянки вважалися нетекстурними. Розмір навчальної вибірки дорівнював 1000 елементів (1000 пар $F-S$), а розмір двох тестових вибірок дорівнює 500 елементів. Тестові вибірки ніяк не були задіяні в процесі навчання класифікатора і використовувалися для оцінки ефективності класифікатора та його прогностичних властивостей.



а

б

Рисунок 2.1 – Зображення Barbara: оригінал (а); спотворене з СКВ = 5 (б), червоним позначені області текстур

Для аналізу ефективності класифікатора було використано згаданий у розділі 1 ROC-аналіз [91]. Ключовими показниками ROC-аналізу є чутливість (ймовірність правильного виявлення текстур, TPR), специфічність (ймовірність правильного виявлення нетекстур, SPC), точність (ймовірність коректної класифікації текстур і нетекстур, ACC) і коефіцієнт кореляції Меттью (MCC) . При цьому розглядалися результати лише тестових вибірок, бо для навчальної вибірки значення ймовірностей прагнули до 100%. Нижче наведена таблиця 2.1 порівняльного аналізу показників для отриманих класифікаторів.

Аналіз значень ACC у разі обчислення локальних параметрів у вікні 5x5 та їх обробці класифікатором призводить до наступних висновків. Значення точності класифікатора F1_5x5 (з урахуванням ВЛД і КР) становлять 79% для обох тестових наборів. Це на 2-3% гірше ніж для найпростішого класифікатора F2_5x5, який працює на основі ДКП. Отже, класифікатор на основі ДКП дуже ефективний навіть сам по собі. Разом з тим, використання всіх трьох локальних параметрів у класифікаторі F3_5x5 призводить до збільшення точності класифікатора до 83-85%. При цьому величина помилки першого роду (1-TPR) становить 7-9%, а

величина помилки другого роду (1-SPC) - 14%, тобто текстури виявляються краще (надійніше), ніж не текстури.

Таблиця 2.1 – Порівняльний аналіз класифікаторів на основі зображення Barbara, спотвореного шумом з СКВ=5

Класифікатор	Тестовий набір №1			
	TPR	SPC	ACC	MCC
F1_5x5	0,91	0,67	0,79	0,59
F2_5x5	0,84	0,80	0,82	0,65
F3_5x5	0,93	0,76	0,85	0,70
F1_7x7	0,85	0,74	0,80	0,59
F2_7x7	0,84	0,75	0,80	0,59
F3_7x7	0,85	0,82	0,84	0,68
F1_9x9	0,86	0,72	0,79	0,59
F2_9x9	0,83	0,81	0,82	0,64
F3_9x9	0,82	0,86	0,84	0,67
	Тестовий набір №2			
F1_5x5	0,90	0,68	0,79	0,60
F2_5x5	0,87	0,75	0,80	0,62
F3_5x5	0,91	0,76	0,83	0,68
F1_7x7	0,86	0,72	0,79	0,59
F2_7x7	0,88	0,72	0,79	0,60
F3_7x7	0,85	0,82	0,83	0,66
F1_9x9	0,87	0,75	0,81	0,62
F2_9x9	0,81	0,80	0,80	0,60
F3_9x9	0,80	0,86	0,83	0,65

Результати аналізу для інших розмірів ковзних вікон (7x7 та 9x9 пікселів) подібні до результатів аналізу для вікон 5x5. Знову значення ACC виявляються більшими для класифікатора F3, що використовує всі три ознаки (локальні параметри). Значення ACC приблизно однакові для всіх трьох аналізованих

розмірів ковзних вікон, але значення TPR і MCC для великих вікон гірше. Тому, судячи з отриманих даних, не варто використовувати занадто великі вікна.

Для будь-якого класифікатора важливо перевірити, як він працює для даних, для яких навчання не проводилося. Нижче представлені результати роботи класифікаторів для зображень Barbara для шуму зі СКВ=5 і СКВ=10 (випадок представляє інтерес, оскільки рівень шуму помітно інший) і Baboon для шуму з СКВ=5 (випадок представляє інтерес, оскільки характеристики текстури зовсім інші) та СКВ=10 (інші і текстура, і рівень завад). Аналіз результатів на рис. 2.5 для F1_5x5 показує, що паралельно з хорошим виявленням текстур (рис. 2.5, а і 2.5, в), класифікатор помилково «вважає» текстурами і різноманітні неоднорідності – межі та їх околиці, малорозмірні об'єкти та їх околиці. Це й не дивно: адже ВЛД та КР – це насамперед детектори меж [204].

Цікава особливість спостерігається при виявленні текстурних ділянок для зображень з СКВ = 10 (рис. 2.2, б і 2.2, г). Очевидно, що ділянки зі слабкою (малоконтрастною) текстурою вже не виявляються (зверніть увагу на частину скатертини на столі, що звисає на рис. 2.2, б). На рис. 2.3 наведено результати застосування класифікатора F3_5x5. Цей класифікатор помітно рідше розпізнає межі об'єктів (ніжка столу на рис. 2.3, а та 2.3, б) як текстуру (порівняно з класифікатором F1_5x5, рис. 2.2, а та 2.2, б). Проблеми залишаються з виявленням малоконтрастної текстури. Разом з тим розроблений класифікатор, будучи навченим для одного типу текстури, добре працює і для інших текстур (див. результати для другого тестового зображення Baboon). Для тестування працездатності отриманого класифікатора було використано чотири тестових зображення, що є повністю текстурні зображення (рис. 2.4 а, б, в, г, зображення взяті з електронного каталогу зразків текстур) та результат роботи класифікатора для цих зображень на базі F3_5x5 (рис. 2.5). Таким чином, ідеальний класифікатор повинен забезпечувати значення TPR для кожного із зображень рівним 1. До вихідних зображень був штучно доданий шум зі СКВ відповідно 5 і 10.

Зважаючи на відсутність попередньо підготовленої карти текстур для зображень передбачалося, що відсоток текстур становить 100%. Тому показник

TPR визначався як відношення коректно класифікованих пікселів до кількості пікселів (розмір зображень – 512x512 пікселів).

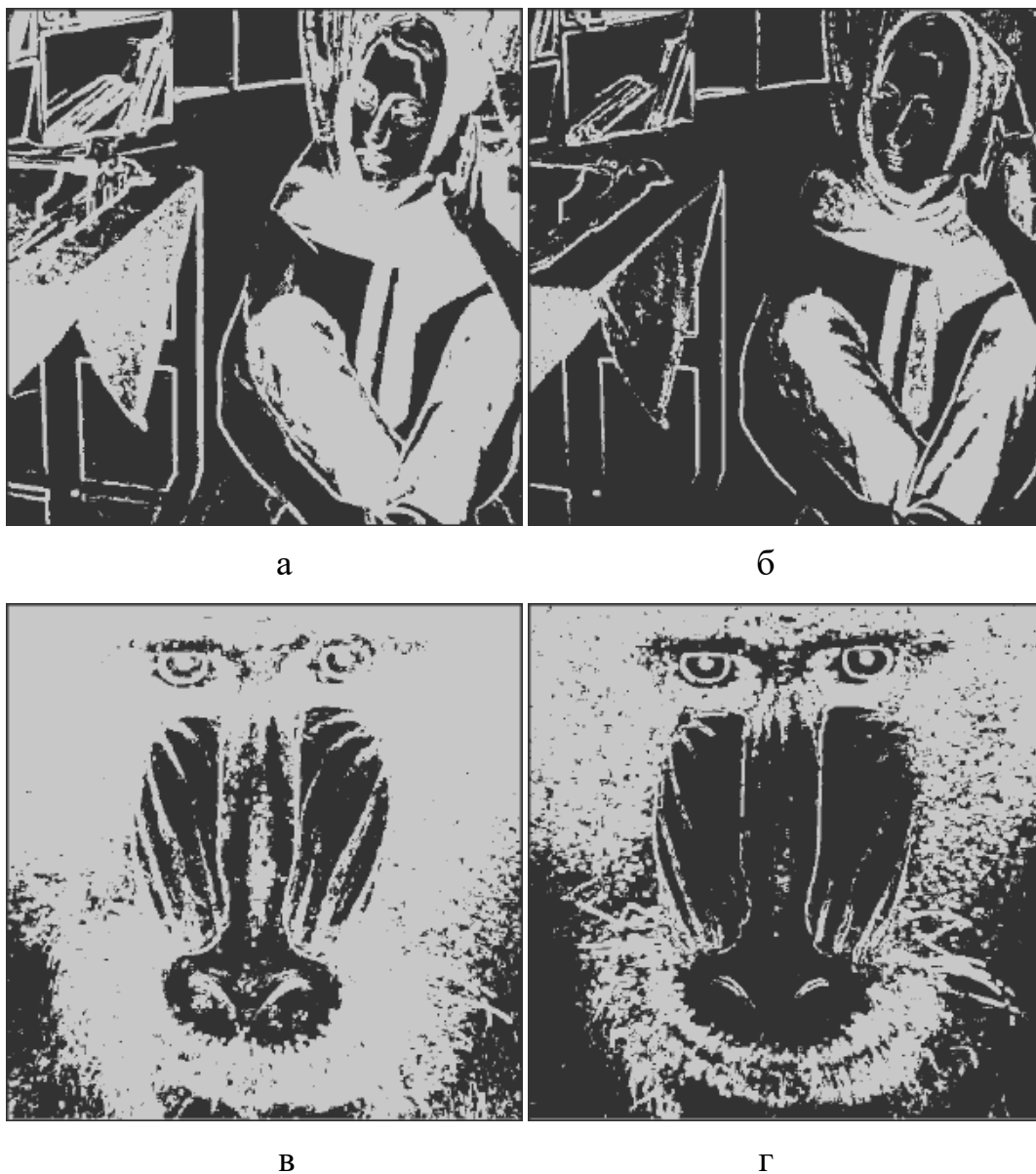


Рисунок 2.2 – Класифікатор F1_5x5: Barbara, СКВ = 5 (а); Barbara, СКВ = 10 (б); Baboon, СКВ = 5 (в); Baboon, СКВ = 10 (г)

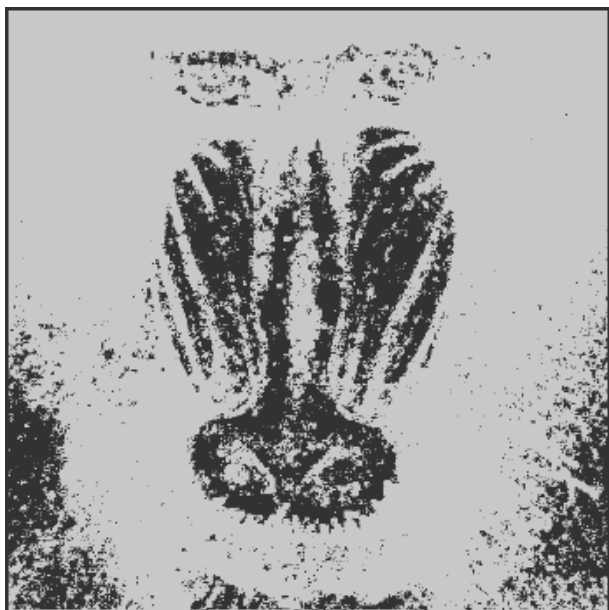
Показано результати для вікна 5×5, бо виходячи з аналізу результатів зображення Barbara (див. вище), саме цей розмір вікна дає найменшу помилку класифікації серед представлених розмірів вікон. Аналіз даних таблиці 2.2 показує, що класифікатор, загалом, працює навіть краще, ніж для тестових даних зображення Barbara.



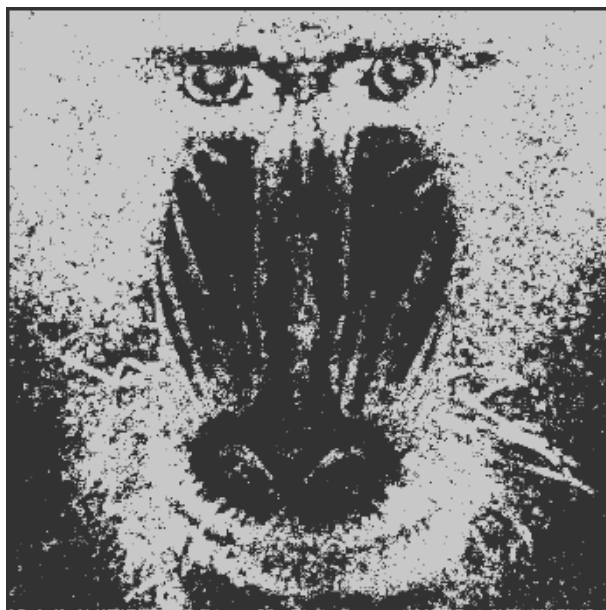
а



б



в



г

Рисунок 2.3 – Класифікатор F3_5x5: Barbara, СКВ = 5(а); Barbara, СКВ = 10 (б);
Baboon, СКВ = 5 (в); Baboon, СКВ = 10 (г)

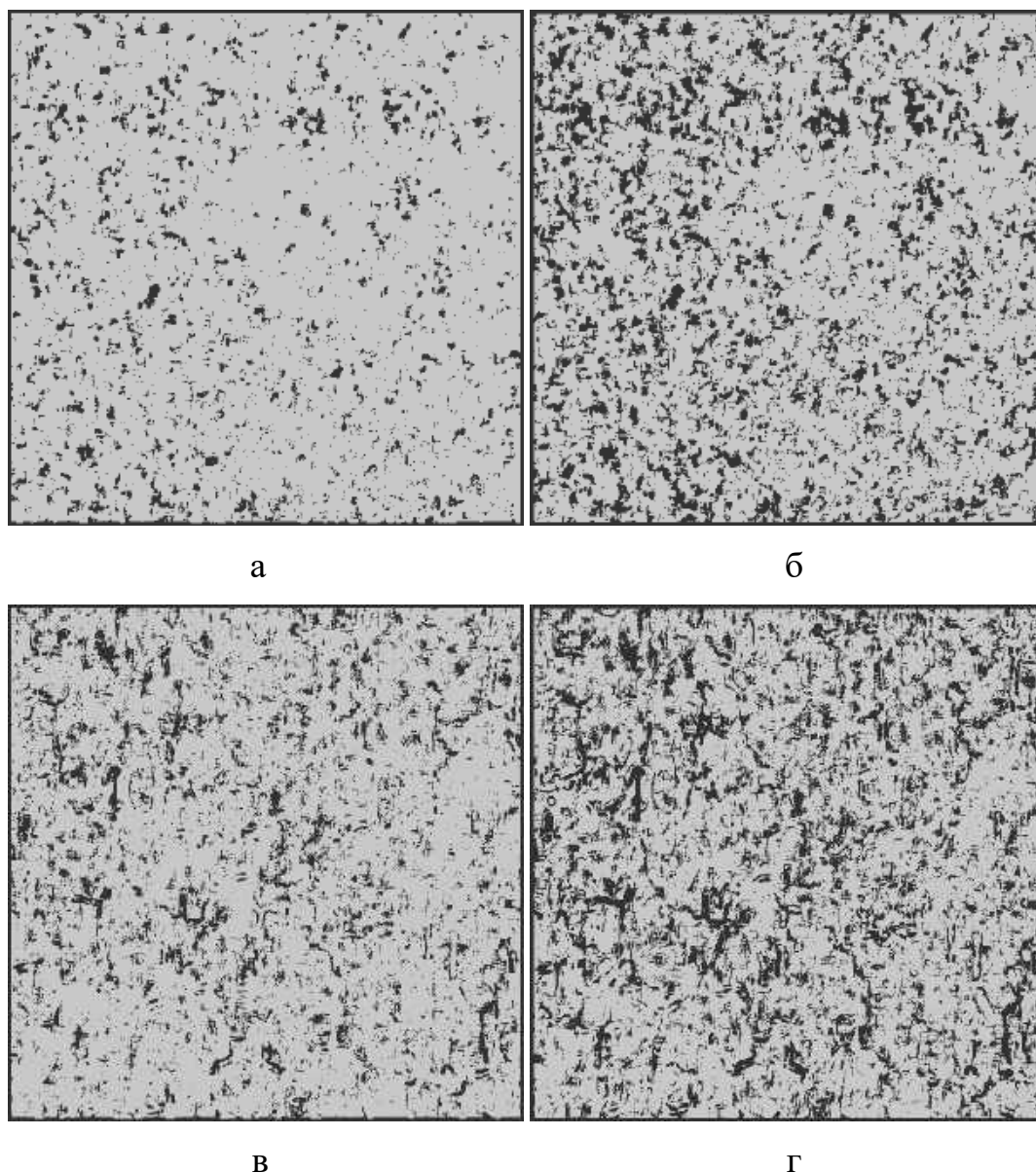


Рисунок 2.5 – Класифікатор F1_5x5 для повністю текстурних зображень

Спільне використання трьох вхідних параметрів на вході класифікатора призводить, як правило, до більших значень TPR, ніж при використанні одного або двох вхідних параметрів. Цей ефект має місце і при СКВ, що дорівнює 10 (див. дані в табл. 2.3).

Крім того, аналіз даних у таблиці 2.3 показує, що класифікатор при СКВ=10 працює помітно гірше, ніж для тестових даних. Цьому є логічне пояснення. Для навчання були використані дані, спотворені шумом зі СКВ=5, тому зображення, спотворені шумом з іншим СКВ, що значно відрізняється від присутнього в навчальних даних, будуть класифіковані гірше.

Таблиця 2.3 – Аналіз класифікаторів з використанням тестових зображень (рис. 2.4), спотворених шумом зі СКВ=10

Класифікатор	СКВ = 10			
	TPR, а	TPR, б	TPR, в	TPR, г
F1_5x5	0,94	0,57	0,71	0,61
F2_5x5	0,89	0,52	0,68	0,53
F3_5x5	0,94	0,62	0,78	0,67

2.1.2 Виявлення текстурних ділянок нейронною мережею на зображеннях, спотворених завадами

Подальшим вдосконаленням запропонованої методики є застосування простої НМ (персептрона), а також додавання ще двох локальних параметрів до переліку тих, що було застосовано у попередньому підрозділі:

1) простий детектор текстур (ДТ) [212]. Основний принцип детектування визначення «щільності» текстури шляхом обчислення пікових значень інтенсивності пікселів і віднесення кількості цих пікових значень до загальної кількості пікселів у вікні. Даний параметр не враховує статистики завад на зображенні і є, по суті, повністю геометричним параметром;

2) детектор Харріса (ДХ) [213] - ще один геометричний параметр, результуюче значення якого для ковзного вікна залежить від відношення середніх значень елементів, що знаходяться в різних частинах вікна.

ДХ раніше добре проявив себе при виявленні меж, і його залучення, ймовірно, може розрізнити наявність у вікні меж або текстур. ДП використовувався в [212] саме для виявлення текстурних ділянок, проте за відсутності завад.

З метою поєднання різнорідних параметрів була використана НМ, що складається з 3-х шарів. Кількість нейронів у вхідному шарі відповідає кількості вхідних параметрів, тобто було різним в кожному випадку (розглядаються випадки від 2 до 4 вхідних параметрів). Кількість нейронів у прихованому шарі: на 2 більше,

ніж у вхідному (відповідно до емпіричних правил навчання нейронних мереж). Вихідний шар містить 2 нейрони, які відповідають розпізнаваним класам.

Для навчання НМ було використано зображення Barbara (рис. 2.1), спотворене адитивним гаусовим шумом з нульовим середнім і СКВ, рівним 5.

З усього набору вхідних параметрів було складено всі можливі комбінації з двох, трьох та чотирьох параметрів. Насамперед усі підготовлені та навчені нейронні мережі були протестовані на зображенні Barbara. Деякі результати трьох вхідних параметрів наведено в таблиці 2.4.

Таблиця 2.4 – Результати перевірки нейронних мереж на зображенні Barbara, спотворених шумом зі СКВ=5

Вхідні параметри			AUC
ВЛД	КР	ДХ	0,431
ВЛД	КР	ДКП	0,835
ВЛД	КР	ДТ	0,813
ВЛД	ДХ	ДКП	0,696
ВЛД	ДХ	ДТ	0,841
ВЛД	ДКП	ДТ	0,732
КР	ДХ	ДКП	0,811
КР	ДХ	ДТ	0,822
КР	ДКП	ДТ	0,802
ДХ	ДКП	ДТ	0,641

Очевидно, що ефективності детекторів на основі навчених НМ істотно відрізняються. Є варіанти наборів вхідних параметрів, які явно поступаються іншим ефективності. Їхні значення AUC лежать у межах 0,43..0,696, але при цьому є і комбінації, що забезпечують значення AUC вище 0,8. У зв'язку з цим залишимо для подальшого дослідження декілька варіантів класифікатора на основі НМ, які забезпечують приблизно однакові результати виявлення текстурних ділянок на використаному тестовому зображенні. Аналогічно для комбінації з двох та чотирьох вхідних параметрів для подальшого аналізу були залишені лише

найкращі комбінації, які при верифікації на зображенні Barbara забезпечували значення AUC вище 0,8. Для того, щоб мати можливість перевірити ефективність застосування навченої нейронної мережі, було створено додаткові тестові зображення, що містять явно виражені текстурні та однорідні ділянки (рис. 2.6). Текстури на цих зображеннях мають різну міру зернистості, а також контрастності та регулярності.

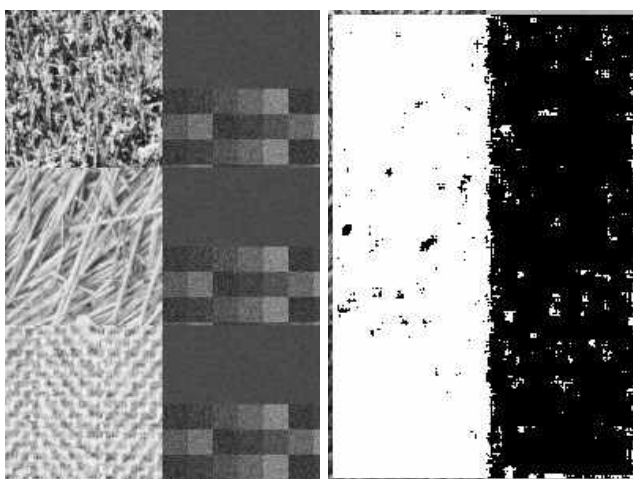


Рисунок 2.6 – Класифікатор на базі НМ для тестових зображень

Нижче наведені таблиці щодо порівняльного аналізу показників для залишених для аналізу найкращих класифікаторів з урахуванням навчених НМ. У таблицях наведено набори параметрів, які були вхідними для нейронної мережі та результати оцінки їх ефективності для чотирьох тестових зображень.

Аналізуючи результати, наведені у таблицях, можна зробити такі висновки. Як однозначні «лідери» можна виділити дві нейронні мережі: з 4-ма вхідними параметрами (КР, ДХ, ДТ і ДКП) і з 3-ма вхідними параметрами (КР, ДХ і ДКП). Зазначимо, що в обох наборах немає ВЛД. Зазначені два набори вхідних параметрів забезпечують дуже добрі результати на всіх чотирьох тестових зображеннях. Слід зазначити, що ефективність всіх нейромережевих детекторів нижче при обробці зображення «Кора» (рис. 2.4, б). Це, мабуть, пов'язані з тим, що це зображення явно виділяється з набору характерним розміром і структурою текстури. Нейронні мережі були навчені значно меншого характерного розміру текстури.

Таблиця 2.5 – Порівняльний аналіз класифікаторів для зображення «трава», спотвореного шумом зі СКВ=5

Вхідні параметри				AUC
КР	ДХ	ДТ	ДКП	0,993
ДХ	ДТ			0,955
КР	ДХ	ДКП		0,990
КР	ДХ	ДТ		0,623
ОЛД	ДХ	ДТ		0,950
ОЛД	КР	ДТ		0,761
ОЛД	КР	ДКП		0,763

Таблиця 2.6 – Порівняльний аналіз класифікаторів для зображення «кора», спотвореного шумом зі СКО=5

Вхідні параметри				AUC
КР	ДХ	ДТ	ДКП	0,967
ДХ	ПД			0,908
КР	ДХ	ДКП		0,972
КР	ДХ	ДТ		0,659
ОЛД	ДХ	ДТ		0,409
ОЛД	КР	ДТ		0,885
ОЛД	КР	ДКП		0,731

Таблиця 2.7 – Порівняльний аналіз класифікаторів для зображення «сіно», спотвореного шумом зі СКО=5

Вхідні параметри				AUC
КР	ДХ	ДТ	ДКП	0,988
ДХ	ПД			0,975
КР	ДХ	ДКП		0,981
КР	ДХ	ДТ		0,864
ОЛД	ДХ	ДТ		0,536
ОЛД	КР	ДТ		0,912
ОЛД	КР	ДКП		0,897

Таблиця 2.8 – Порівняльний аналіз класифікаторів для зображення «тканина», спотвореного шумом зі СКО=5

Вхідні параметри				AUC
КР	ДХ	ДТ	ДКП	0,981
ДХ	ПД			0,985
КР	ДХ	ДКП		0,975
КР	ДХ	ДТ		0,921
ОЛД	ДХ	ДТ		0,814
ОЛД	КР	ДТ		0,944
ОЛД	КР	ДКП		0,929

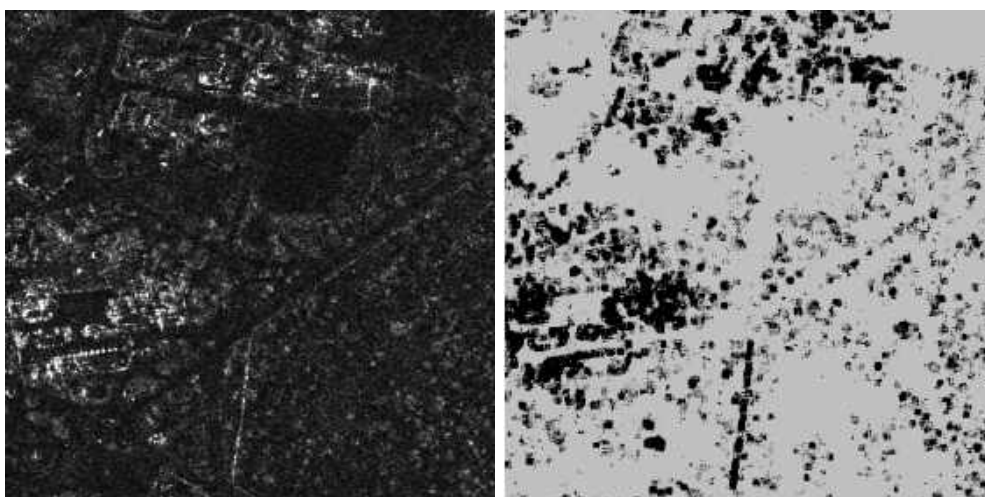


Рисунок 2.7 – Однопоглядове зображення РСА та його класифікаційна карта

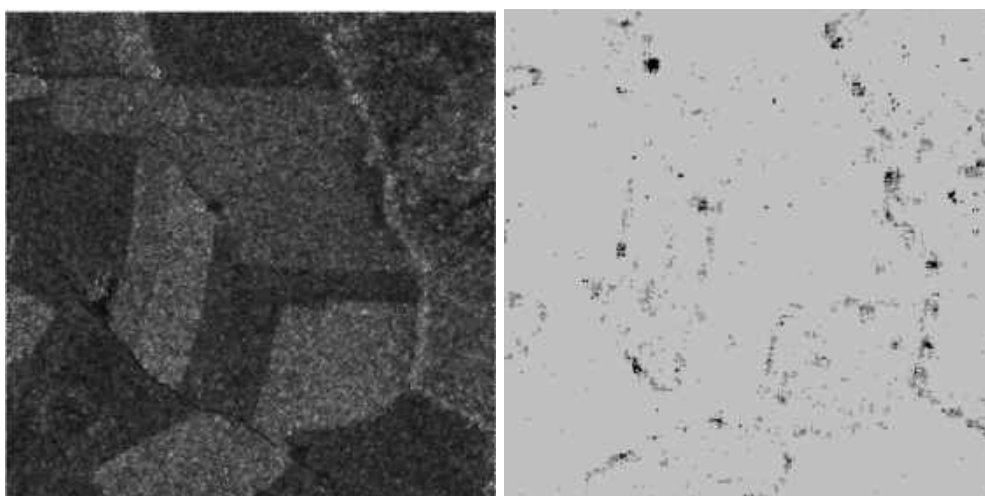


Рисунок 2.8 – Однопоглядове зображення РСА та його класифікаційна карта

Однак, незважаючи на це, ефективність найкращих нейронних мереж залишається досить високою (значення AUC на рівні 0,96, 0,97 і вище). Отриманий детектор на базі НМ було протестовано на реальних зображеннях ДЗЗ (рис. 2.7 та 2.8). Як видно, найбільш ефективними виявилися НМ, для яких параметри, що задіяні, базуються на введених в розгляд геометричних ознаках (ДХ, ДТ). Така особливість є характерною для детектування саме текстурних ділянок, тому що для меж та малорозмірних об'єктів найкраще проявили себе статистичні параметри.

Як було зазначено вище, шум у реальних зображеннях (наприклад, отриманих за допомогою РСА) може бути складнішим, ніж АБГШ. Тоді може бути корисною певна попередня обробка. Наприклад, попереднє усунення шумів РСА зображення за допомогою локального статистичного фільтра Лі [214] може зробити залишковий шум майже гаусівським, хоча така попередня обробка була використана до того, як вхідні параметри були розраховані з урахуванням мультиплікативної природи шуму. Класифікація зображень системи ДЗЗ Terra SAR-X за допомогою найбільш ефективної отриманої НМ представлена на рис. 2.7.

Нарешті, на рис. 2.8 показано РСА-зображення, яке має дуже просту структуру (це сільськогосподарська зона з великими однорідними регіонами та краями між ними). Досить багато пікселів, які належать до країв, виявляються та класифікуються правильно. Однак у виявлених контурах країв є розриви. Це свідчить про проблеми з виявленням і класифікацією, якщо контрасти низькі та/або інтенсивність шуму висока (як це зазвичай відбувається на однопоглядових зображеннях РСА).

2.2 Адаптивна фільтрація зображень за допомогою алгоритмів машинного навчання в умовах сигнально-залежних та просторово-корельованих завад

2.2.1 Аналіз ефективності комбінованого придушення спеклу в однопоглядових зображеннях РСА

Висока роздільна здатність може бути забезпечена, якщо РСА працює в режимі одного погляду.

Однією з проблем у розробці нових методів придушення спеклу РСА-зображень, а також оцінки ефективності для них є відсутність загальноприйнятих тестових (безшумних) і змодельованих (зашумлених) РСА-зображень. Подібна проблема була актуальною для дослідників, що займаються оптичною обробкою зображень. Вони вирішили цю проблему, прийнявши стандартний набір зображень у градаціях сірого та кольорових зображень, який включає еталонні зображення (наприклад, Lena, Baboon, Barbara, Peppers тощо). Тим часом співтовариство обробників зображень РСА все ще використовує наступні три практики. Перша практика моделює спекл для стандартних оптичних зображень, таких як Lena, Boat тощо [215]. Другий генерує зображення із спеклом на основі оптичних зображень природних сцен, отриманих з бортових платформ [215]. Нарешті, третя практика застосовує та аналізує реальні зображення до та після видалення спеклу [215], [216].

Ці недоліки стимулювали розробку більш складних моделей зображень РСА [217], [218]. Вони здатні моделювати просторово корельований спекл [217] та інші більш складні явища [219], [220], [221] на одно- та багатопоглядових зображеннях РСА [218].

Метою цього підрозділу є проведення кількісного аналізу ефективності та порівняння широкого набору сучасних методів придушення спеклу. Особливості цього дослідження полягають у використанні більш адекватних моделей однопоглядових РСА-зображень, які враховують просторову кореляцію шуму, а також у використанні трьох кількісних критеріїв, а саме інтегральної

середньоквадратичної похибки (MSE), розрахованої для всього зображення, локальної MSE, визначеної для локально активних областей тестових зображень, і метрика MSSIM [96], яка здатна адекватно характеризувати візуальну якість оригінальних (зашумлених) і очищених зображень. Особливість полягає в тому, що зони неоднорідності виявляються за допомогою підходу Level Set [142], який нещодавно продемонстрував свою ефективність при обробці PCA та інших типів зображень, що спотворені шумом [6].

Увагу сфокусовано на фільтрації однопоглядних PCA-зображень. Це найскладніший випадок через найвищу інтенсивність спеклу. Існує два варіанти моделювання тестових зображень PCA із шумом. В обох випадках шумова модель шуму виражається формулою (1.15). В першому випадку розглядатимемо амплітудні дані, для яких відносна дисперсія спеклу $\sigma_{\mu}^2 = 0,273$ та які відповідають Релеєвському розподіленню. В другому випадку інтенсивність розглянутих зображень відповідає негативному експоненціальному розподіленню з дисперсією $\sigma_{\mu}^2 = 1$. У дослідженні, не втрачаючи загальності, вивчається перший випадок (вихідні та відфільтровані зображення для двох випадків можна легко перетворити з одного представлення на інше простими гомоморфними перетвореннями [73]).

Особливістю моделі (1.15) у дослідженні є те, що спекл моделюється як просторово корельований мультиплікативний шум. Просторові кореляційні властивості спекла моделюються таким чином, що вони практично такі ж, як і для однопоглядних зображень PCA, створених космічною системою TerraSAR-X [222]. Безшумне тестове зображення ДЗЗ FR01 представлено на рис. 2.9, а, тоді як отримана версія з шумом показана на рис. 2.9, б. Безшумні тестові зображення мають 8-бітний растровий формат, тоді як зображення з шумом представлені як 16-бітні 2D-масиви даних, щоб уникнути ефектів відсікання. Відфільтровані зображення також були представлені в такому ж форматі, щоб зменшити вплив помилок округлення на оцінки критеріїв ефективності фільтрації.

Вихідна (інтегральна) MSE розраховується як:

$$MSE_{out} = \sum_{i=1}^{I_{IM}} \sum_{j=1}^{J_{IM}} (I_{ij}^f - I_{ij}^{true})^2 / (I_{IM} J_{IM} - 1), \quad (2.4)$$

де I_{ij}^f – значення відфільтрованого зображення у точці з координатами i,j ; I_{IM}, J_{IM} – це розміри зображення (в цьому дослідженні вони дорівнюють 512).

На додаток до інтегрального MSE будемо використовувати локальну MSE_{het} , яка визначається для неоднорідних ділянок зображення як:

$$MSE_{het} = \sum_{i,j \in G_{het}} (I_{ij}^f - I_{ij}^{true})^2 / (|N_{G_{het}}| - 1), \quad (2.5)$$

де G_{het} визначає пікселі зображення, що відносяться до неоднорідних ділянок; $N_{G_{het}}$ визначає кількість таких пікселів.



а

б

Рисунок 2.9 – Зображення FR01: оригінал (а), спотворено шумом зображення (б)

Ці неоднорідні області були виявлені за допомогою методу Level Set, який визначає місцезнаходження об'єктів малого розміру та розділяє їхні межі. Цей підхід базується на рівнянні Гамільтона-Якобі [223]. Його здатність виявляти неоднорідності на однопоглядових РСА-зображеннях була вивчена для змодельованих і реальних даних у [6]. Розглянемо детальніше метод Level Set.

Метод Level Set є технікою активного контуру, передній рух якого керується:

$$\psi^{(k+1)} = \psi^{(k)} - \Delta t \psi_t |\nabla \psi^{(k)}|, \quad (2.6)$$

де $\psi^{(k)}$ визначає рівень поверхні на k -й ітерації; $\psi^{(0)}$ – значення на першій ітерації (зазвичай, $\psi^{(0)} \in [-1; 1]$); Δt – одиничний часовий крок; $|\nabla(\cdot)|$ – величина градієнта початкової поверхні. Термін ψ_t – функція еволюції (або швидкості), яка визначає рух фронту. Було адаптовано закон еволюції набору рівнів, запропонований у [142], для виявлення малорозмірних цілей. Функція розповсюдження базується на локально оціненому коефіцієнті варіації (КВ) і пороговому значенні однорідності, введеному в [224] як

$$T = 1 + \frac{\sqrt{1+2\sigma_\mu^2}}{\sqrt{2M^2}} \sigma_\mu, \quad (2.7)$$

де M – розмір вікна; σ_μ – теоретичне СКВ спекла. Модель визначення цілі можна представити як [142]:

$$\psi_t = (\hat{\sigma}_\mu - T) - \varepsilon k, \quad (2.8)$$

де εk – швидкість викривлення, описана в [225]; $\hat{\sigma}_\mu$ – локальна оцінка σ_μ . Остаточні результати сегментації задаються точками перетину нуля $\psi^{(k)}$.

В Додатку Ж, на рис. Ж.4 показано деякі результати застосування цього методу до змодельованого РСА-зображення. Це зображення моделює великі однорідні області різного значення, одну велику текстурну область, декілька протяжних об'єктів (як огорожі на природних зображеннях). Крім того, було спеціально додано п'ятнадцять об'єктів малого розміру в однорідні області з різною середньою інтенсивністю.

Не дивно, що було виявлено три об'єкти, які представляють найбільший контраст щодо навколишнього фону (див. верхню центральну частину зображення). Інші об'єкти невеликого розміру не були виявлені, і їх неможливо виявити візуально, оскільки вони мають менший контраст. Дивно, але була виявлена велика текстурна область. Таким чином, Level Set метод може бути корисним для обробки зображень РСА.

Ці результати попереднього аналізу показують, що, подібно до інших методів виявлення, ефективність методу Level Set залежить від контрасту об'єкта. Тут, навпаки, припускаємо співвідношення $C = I_{obj}/I_{bg}$, де I_{obj} – справжнє значення зображення, яке приймається однаковим для всіх пікселів, які належать малому об'єкту, а I_{bg} позначає справжнє значення для навколишнього фону. Експерименти показали, що для $C \leq 1,6$ підхід на основі Level Set не виявляє об'єкти, головним чином, через вплив спеклу. Таким чином, доцільно дослідити вплив контрастності об'єкта на надійність виявлення об'єктів. Крім того, можна очікувати, що ймовірність правильно виявити об'єкт P_{cd} залежить від кількості пікселів, які йому належать.

Для оцінки ефективності алгоритму було проведено тести з іншим зашумленим зображенням. В Додатку Ж, на рис. Ж.5, а показано спеціально створене безшумне зображення з досить великою кількістю об'єктів малого розміру, які використовуються для тестування виявлення об'єктів на основі набору рівнів. Об'єкти представляють різні контрасти щодо однорідного фону, і C змінюється від 1,5 до 5,0. Об'єкти мають різну кількість пікселів N_{ss0} . Спекл було змодельовано для зображення в Додатку Ж, на рис. Ж.5, а, і, крім того, до нього було застосовано наш метод виявлення. В Додатку Ж, на рис. Ж.5, б показано результати виявлення об'єктів.

Візуальний аналіз карти показує, що об'єкт виявляється з більшою ймовірністю, коли його контраст або параметр N_{ss0} збільшуються (наприклад, $N_{ss0} > 9$).

Ще одне важливе спостереження полягає в тому, що об'єкти з більшим Nss0 виявляються цілком задовільно. Тести з рис. Ж.5, а і Ж.5, б (Додаток Ж) чітко ілюструють цю тенденцію.

Ще одна перевага підходу Level Set полягає в тому, що для змодельованого зображення не спостерігалось помилкових виявлень, хоча невеликі об'єкти з низьким C були пропущені. Зазвичай досягти відповідного компромісу між ймовірністю правильного виявлення та частотою помилкових тривог є важким завданням шляхом встановлення відповідного порогу в алгоритмах виявлення цілі [73]. Метод Level Set досить добре справляється з цим завданням, не створюючи помилково виявлені цілі, а також забезпечує правильне виявлення об'єктів із досить низьким контрастом відносно навколишнього фону.

Таблиця 2.9 – Ймовірність правильного визначення для різних контрастів об'єктів

Контраст об'єкту	Ймовірність правильного визначення
1,50	0,000
2,00	0,014
2,25	0,028
2,50	0,070
3,00	0,380
3,50	0,760
4,00	0,850
4,50	0,920
5,00	0,950

Що стосується експериментів із реальними зображеннями PCA, було протестовано запропонований алгоритм на зображеннях TerraSAR-X. На сайті <https://intelligence.airbus.com/> є декілька зображень, представлених у вигляді комплексних значень. Використовуючи їх, сформувано амплітудні однопоглядові зображення розміром 512x512 пікселів для регіону в Індонезії для даних

поляризації HH і VV, як це показано в Додатку Ж, рис. Ж.6. Ці зображення характеризуються досить високою взаємною кореляцією (схожістю). Коефіцієнт крос-кореляції для оригінальних (зашумлених) зображень дорівнює 0,63. Якщо зображення обробляються ефективним спекл-фільтром, коефіцієнт крос-кореляції зростає і становить близько 0,8 [226]. Для цих зображень PCA, метод Level Set створює контури навколо невеликих об'єктів або інших неоднорідностей. В Додатку Ж, рис. Ж.6, а і Ж.6, б показують ці контури білим кольором. Результати для двох розглянутих зображень досить схожі.

Після виявлення контурів області пропонується виконати постобробку для отримання бінарної карти. На рис. 2.10, а показано приклад цієї карти для тестового зображення FR01. Справді, невеликі об'єкти, межі та їх околиці виявляються добре.

Третім розглянутим показником є MSSIM, визначений для всього тестового зображення. Причини його використання полягають у тому, що це один із найкращих показників візуальної якості для зображень у відтінках сірого [227]. Отримані дані представлені в таблиці 2.10 для обох тестових зображень. Як відомо, ефективність цих фільтрів залежить від розміру ковзного вікна. Тому було застосовано декілька розмірів вікон, щоб визначити оптимальний відповідно до використовуваних критеріїв. Було використано вдосконалену реалізацію фільтра Лі [116]. Варто відзначити, що кращу ефективність придушення шумів в однорідних областях для обох фільтрів забезпечують більші розміри ковзного вікна. Згідно з аналізом даних, оптимальний розмір ковзного вікна для фільтра Лі становить 7x7 пікселів, що зазвичай рекомендується для практичного застосування.

При цьому розмір вікна 5x5 пікселів забезпечує краще збереження країв і деталей. Щодо фільтра Фроста, то вікно 17x17 пікселів забезпечує найменшу (найкращу) MSE_{out} , тоді як найкращу візуальну якість зображення забезпечує вікно 13x13 пікселів.

Загалом результати для фільтра Фроста (наприклад, з ковзним вікном 13x13 пікселів) є значно кращими, ніж для локального статистичного фільтра Лі. Крім того, фільтр Фроста [134] працює краще в порівнянні з іншими, більш складними фільтрами, як це буде видно з подальшого аналізу.

Іншим методом усунення шумів, який використовується для порівняння, є інтегро-диференціальний фільтр [228]. Насправді цей фільтр також залежить від розміру ковзного вікна, і результати представлені для вікон від 5×5 до 15×15 (позначаються як Vozel 5, ..., Vozel 15). Отримані дані представлені в таблиці 2.11. Як видно, мінімальна MSE_{het} спостерігалася для вікна 7×7 , тоді як оптимальна MSE_{out} та найкраща візуальна якість були забезпечені для ковзного вікна 9×9 .

Таблиця 2.11 також відображає дані для фільтра BM3D [78], що містить перетворення, що стабілізують дисперсію [216]. Цей метод усунення шумів призвів до MSE_{out} та MSE_{het} , значно більших, ніж для фільтра Фроста з оптимальними параметрами, а метрика MSSIM також має гірші значення.

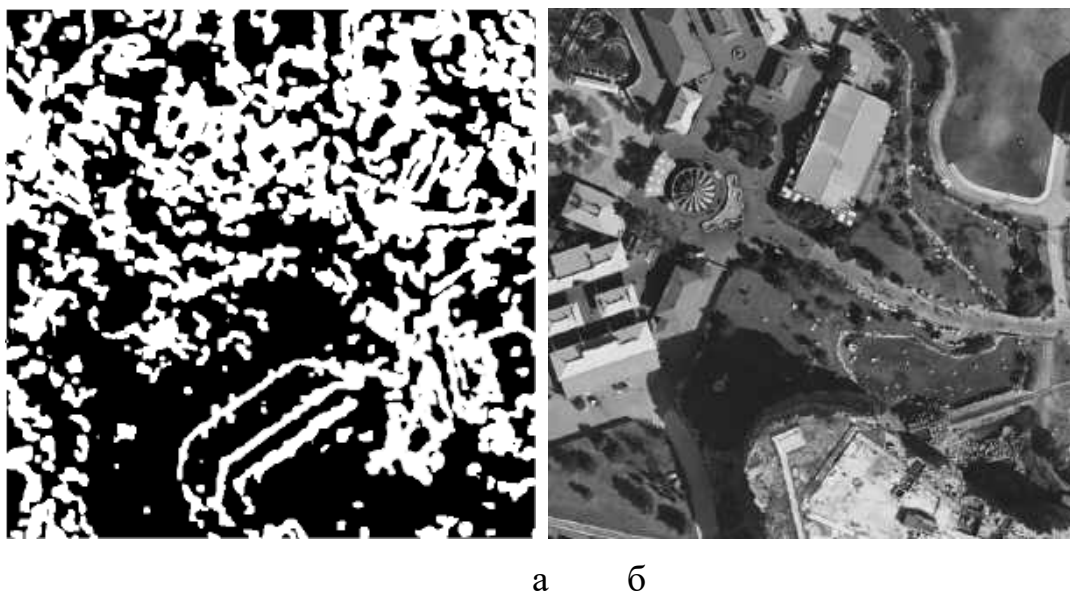


Рисунок 2.10 – Бінарна карта для зображення FR01 (а) та оригінал зображення FR02 (б).

Нижче наведено можливі причини, чому фільтр BM3D не працює достатньо добре для розглянутого застосування. По-перше, шум в зображеннях після прямого гомоморфного (стабілізуючого дисперсію) перетворення не є білим і не є гаусовим. Це ускладнює завдання пошуку подібних блоків [229]. По-друге, знайдені подібні блоки обробляються методом усунення шумів на основі ДКП, не адаптованого до просторового спектру шуму. Це додатково погіршує ефективність придушення

шуму. Тоді можна очікувати, що ефективність фільтрації на основі BM3D може бути покращена, якщо ці недоліки усунути та застосувати відповідні модифікації.

Таблиця 2.10 – Ефективність локальних статистичних фільтрів Лі та Фроста

Розмір ковзного вікна	FR01			FR02		
	MSE _{out}	MSE _{het}	MSSIM	MSE _{out}	MSE _{het}	MSSIM
Lee 5x5	620	783	0,778	640	831	0,766
Lee 7x7	567	843	0,780	598	988	0,767
Lee 9x9	589	958	0,761	625	1176	0,747
Frost 5x5	630	729	0,777	643	712	0,764
Frost 9x9	479	663	0,810	503	703	0,799
Frost13x13	460	670	0,817	485	722	0,806
Frost17x17	459	677	0,816	484	734	0,805

Метод придушення спеклу на основі Level Set включає алгоритм виявлення цілей, який враховує міру однорідності/гетерогенності для розрізнення об'єктів малого розміру, оточених великими однорідними областями. Потім метод спекл-фільтрації виконується блоками над однорідними областями та, крім того, зберігає краї та деталі. У таблиці 2.12 представлена оцінка ефективності цієї схеми фільтрації. Насправді його ефективність залежить від кількості ітерацій, і більша кількість ітерацій дає кращі результати в сенсі MSE_{out} та MSE_{het}. Проте отримані результати значно гірші, ніж для фільтра Фроста та методики з оптимальними параметрами [228].

Розглянемо фільтрацію на основі ДКП. Цей підхід до придушення спеклу пропонує багато варіантів. Тут вказано той, який розглядається в запропонованому підході.

Припускаємо обробку блоків, що повністю перекриваються, і адаптуємо цей процес фільтрації до просторового ДКП-спектру шуму шляхом застосування наступного комбінованого порогового значення [6], [230]:

$$D_t(m, n, k, l) = \begin{cases} D(m, n, k, l), \text{ якщо } |D(m, n, k, l)| \geq \beta \sigma_\mu I_{mn}^{mean} \sqrt{W_{kl}} \\ D^3(m, n, k, l) / (\beta \sigma_\mu I_{mn}^{mean} \sqrt{W_{kl}})^2 \text{ в інших випадках} \end{cases}, \quad (2.9)$$

де $D(m, n, k, l)$ – kl -елемент ДКП-перетворення блоку у лівому верхньому куті з координатами m, n . I_{mn}^{mean} визначає середнє значення для mn -го блоку; W_{kl} визначає нормалізований ДКП-спектр спеклу; k та l – індекси в ДКП області; β – параметр із рекомендованим значенням 4.5. Типовий розмір блоку становить 8×8 пікселів.

Таблиця 2.11 – Ефективність інтегро-диференціального фільтра та BM3D з VST.

Розмір ковзного вікна	FR01			FR02		
	MSE _{out}	MSE _{het}	MSSIM	MSE _{out}	MSE _{het}	MSSIM
Vozel 3	732	813	0,764	736	763	0,751
Vozel 5	510	651	0,807	526	520	0,795
Vozel 7	452	627	0,821	477	672	0,810
Vozel 9	438	634	0,824	467	692	0,812
Vozel 11	439	652	0,821	469	714	0,809
Vozel 13	446	672	0,816	477	735	0,803
Vozel 15	457	693	0,809	486	754	0,796
BM3D with VST	746	872	0,756	762	809	0,743

Крім того, оцінюємо дані, отримані для кожного блоку, і декілька різних значень параметра. Результати представлені в таблиці 2.13, і з їх аналізу можна зробити наступні висновки:

1) існують мінімуми як для MSE_{out} (всього зображення), так і для MSE_{het} (визначається для пікселів, які належать карті);

2) ці мінімуми спостерігаються для різних β , мінімальна MSE_{het} спостерігається для менших β (приблизно 4,6), оскільки менші β у фільтрації на основі ДКП забезпечують краще збереження меж і деталей; існують також

максимальні значення для метрики MSSIM, що спостерігаються практично для того самого, що й мінімальне значення MSE_{het} ;

3) значення найкращих критеріїв майже такі ж, як для фільтра Фроста з оптимальними налаштуваннями параметрів і трохи гірші, ніж для методу фільтрації [228].

Таблиця 2.12 – Ефективність придушення спеклу на основі Level Set.

Кількість ітерацій	FR01			FR02		
	MSE_{out}	MSE_{het}	MSSIM	MSE_{out}	MSE_{het}	MSSIM
1	1521	1519	0,710	1523	1377	0,696
2	952	1087	0,756	974	1101	0,743
3	765	957	0,773	793	1015	0,762
4	694	920	0,777	724	985	0,765
5	668	913	0,773	696	987	0,762

Таблиця 2.13 – Ефективність придушення спеклу на основі ДКП у поєднанні з частотно-залежним порогом і блоками пікселів 8x8

β	FR01			FR02		
	MSE_{out}	MSE_{het}	MSSIM	MSE_{out}	MSE_{het}	MSSIM
3,8	547	729	0,810	562	734	0,801
4,0	519	704	0,813	537	715	0,804
4,2	499	688	0,815	518	703	0,806
4,4	485	678	0,816	504	698	0,807
4,6	475	673	0,816	495	698	0,808
4,8	468	672	0,816	489	701	0,807
5,0	464	674	0,816	486	708	0,807
5,2	462	678	0,815	484	717	0,806
5,4	462	684	0,814	484	728	0,805
5,6	463	691	0,813	486	741	0,803

На основі аналізу результатів, наведених в [231] для інтенсивного шуму, можна очікувати, що кращі результати для розглянутого застосування можуть

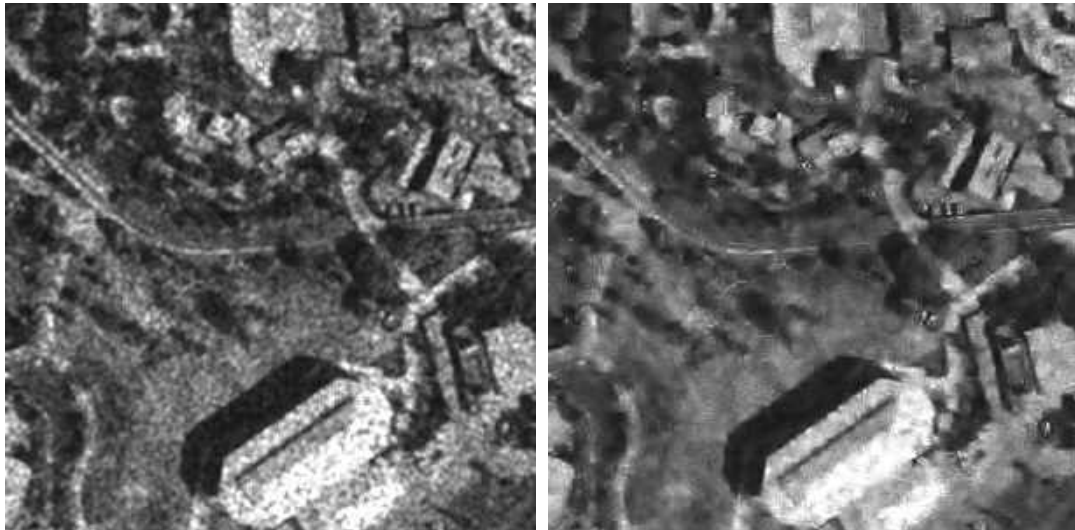
забезпечити фільтри на основі ДКП з блоками 16×16 пікселів. Щоб перевірити цю гіпотезу, було отримано дані моделювання для цієї версії придушення спеклу на основі ДКП, які відображені в таблиці 2.14. Хоча отримані значення досліджуваних критеріїв є задовільними, вони гірші, ніж ті, що представлені для фільтрів Фроста, методу Vozel і ДКП з блоками 8×8 з оптимальними параметрами. Основна проблема полягає у збереженні меж/деталей для обробки зображень у блоках 16×16 .

Дослідження також проводилися для жорсткого частотно-залежного порогу з розмірами блоків 8×8 і 16×16 пікселів. Різниця між ними обома полягає в тому, що оптимальні значення β становлять приблизно 3 відповідно до MSE_{out} (мінімум) і приблизно 2,8 відповідно до мінімуму MSE_{het} і максимуму $MSSIM$.

Варто зазначити, що MSE_{het} для всіх досліджуваних підходів придушення спеклу завжди більша, ніж MSE_{out} для всього зображення. Це пояснюється тим, що будь-який фільтр краще видаляє шум в однорідних областях зображення, ніж в неоднорідних [137], [214]. Крім того, області неоднорідності зазвичай відповідають областям пікселів з великим локальним середнім і, отже, великим локальним (вхідним) MSE через мультиплікативну природу спеклів.

На рис. 2.11 наведено два приклади відфільтрованих зображень. Як видно, фільтру на основі ДКП вдалося значно зменшити шум і забезпечити краще збереження меж/деталей, хоча його все ще варто вдосконалити.

Підсумовуючи результати моделювання, наведені вище, можна очікувати, що подальше покращення ефективності придушення спеклу може бути забезпечено завдяки кращому збереженню дрібних деталей (мішеней) і меж, тобто кращій обробці однопоглядових РСА-зображень у неоднорідних областях. Зауважимо, що площа, яка відповідає таким неоднорідним регіонам, досить велика. Для тестового зображення з шумом на рис. 2.9, 6 відсоток таких пікселів становить близько 44% на карті, представлений на рис. 2.10, а (білі пікселі). Припускаємо, що гетерогенні області можуть бути виявлені методом Level Set. Тоді завдання полягає в тому, щоб досягти якомога меншого локального MSE_{het} .



а

б

Рисунок 2.11 – Оброблені зображення FR01: а – оброблене фільтром Лі 7х7, б – для фільтра на основі ДКП (комбіноване частотно-залежне порогове значення $\beta=4,9$)

Покращення для збереження меж/деталей було запропоновано раніше в [6], і його ефективність була продемонстрована якісно, але не кількісно. Таким чином, нижче зосереджуємося на критеріях MSEhet і MSSIM для кількох версій комбінованих (локально адаптивних) методів придушення спеклу. Припущення полягає в тому, що фільтр на основі ДКП блоках 8х8 у поєднанні з частотно-залежним пороговим значенням $\beta=4,9$ підходить для обробки однорідних областей зображення. Таким чином, повинні перевірити, який фільтр більше підходить для застосування до виявлених неоднорідних областей.

У [6] запропоновано застосовувати фільтрацію на основі ДКП у поєднанні з частотно-незалежним жорстким порогом:

$$D_t(m, n, k, l) = \begin{cases} D(m, n, k, l), & \text{якщо } |D(m, n, k, l)| \geq \beta \sigma_\mu I_{mn}^{med} \\ 0 & \text{в інших випадках} \end{cases}, \quad (2.10)$$

де I_{mn}^{med} визначає медіану значень для mn -го блоку. Насправді, існує два основних підходи. По-перше, адаптація до спекл-просторового спектру використовується в однорідних областях завдяки (2.10). По-друге, застосовується адаптація до

локального вмісту за рахунок виявлення різнорідних областей методом Level Set і налаштування кращого збереження деталей та меж для цих областей.

Було перевірено цю локально-адаптивну версію, і надані результати представлені в таблиці 2.15. Як видно, отримані значення MSEout, MSEhet і MSSIM кращі, ніж для будь-якого фільтра на основі ДКП, дослідженого вище для обох тестових зображень.

Інший варіант був запропонований у роботі [6] і застосовує 8x8-блочне визначення порогу вейвлета Хаара у виявлених неоднорідних областях. Спочатку було проаналізовано ефективність самого вейвлета Хаара. Встановлено, що приблизно однакові (оптимальні) результати отримують, якщо жорсткий поріг встановити $T(m, n) = \beta \sigma_{\mu} I_{mn}^{\text{mean}}, \beta \approx 3,5$ або $T(m, n) = \beta \sigma_{\mu} I_{mn}^{\text{med}}, \beta \approx 4,0$.

Отримані дані моделювання для останнього варіанту наведені в таблиці 2.15. Результати трохи кращі, ніж для першої версії адаптивного фільтра.

Оскільки фільтр Фроста дав дуже хороші результати, він надихнув на ідею, що зменшення шуму на основі ДКП із частотно-залежним порогом можна поєднати з фільтром Фроста. Відповідний вихід локально адаптивного фільтра може бути надано як:

$$I_{ad}(i, j) = \begin{cases} a_1 I_{DCT}(i, j) + a_2 I_{Fr}(i, j), & \text{якщо } M(i, j) = 0 \\ a_3 I_{DCT}(i, j) + a_4 I_{Fr}(i, j), & \text{в інших випадках} \end{cases} \quad (2.11)$$

де для невід'ємних ваг $a_1 + a_2 = 1$; $a_3 + a_4 = 1$; $a_1 \geq a_3$. $I_{DCT}(i, j)$, $I_{Fr}(i, j)$ – вихідні значення ij -пікселя для фільтрів ДКП та Фроста відповідно. $M(i, j) = 0$ визначає значення карти локальної активності для локально пасивних ділянок (де неоднорідності виявляються методом Level Set). Крім того, було використано інший механізм для вдосконалення схеми придушення спеклу. Припускаємо, що виходи двох фільтрів корельовані, хоча коефіцієнт кореляції не дорівнює одиниці. Тоді можна забезпечити деяке часткове усереднення залишкового шуму.

Було досліджено комбіновану фільтрацію для блоку 8x8 ДКП із комбінованою пороговою обробкою для $\beta \approx 4,9$ та фільтром Frost 9x9. Отримані

результати ($a_1=0,6$, $a_3=0,4$) представлені в нижньому рядку таблиці Ж.1, що наведено у Додатку Ж (позначено як ДКП+Frost). Хоча є деякі покращення порівняно з базовими фільтрами ДКП і Frost, переваги недостатні. Причина в тому, що залишковий шум сильно корелює для цих фільтрів.

Подібно до (2.11), поєднали усунення шумів на основі ДКП з фільтром, представленим у [137]. Єдина відмінність полягає в тому, що в цьому випадку $a_1=a_2=0,5$, $a_3=0,3$, $a_4=0,7$. Результати представлені в нижньому рядку таблиці Ж.2, Додаток Ж. Загалом ці результати кращі, ніж для будь-якого іншого розглянутого методу усунення шумів, розглянутого в цьому дослідженні.

Зауважимо, що для всіх локально-адаптивних фільтрів, описаних тут, відповідність першому або другому типу області визначається положенням центрального пікселя блоку на карті країв. Візуально перевібивши результати, можна побачити, що фільтр зберіг деталі в неоднорідних областях для змодельованих зображень. Для цього застосували фільтр на основі ДКП у поєднанні з частотно-залежним порогом $\beta=4,9$ для однорідних областей і фільтр на основі ДКП з частотно-незалежним порогом або $T(m, n) = \beta \sigma_{\mu} I_{mn}^{med}$ для виявлених неоднорідних областей.

По-перше, використали $\beta=4,0$ згідно з таблицею Ж.2 у Додатку Ж. Вихідне зображення представлено в Додатку Ж, рис. Ж.7, а, і можна спостерігати краще збереження деталей та меж порівняно з неадаптивним варіантом (див. вихідне зображення на рис. 2.11, б). Можна очікувати, що різкість меж може бути ще кращою, якщо β менше 4 для фільтра, застосованого до неоднорідних областей. В Додатку Ж, рисунок Ж.7, б ілюструє випадок налаштування $\beta=3,0$, що призводить до ще кращого збереження країв/деталей, але артефакти стають більш помітними.

В Додатку Ж, на рис. Ж.8 показано вихідні зображення для двох інших адаптивних версій. Вихідне зображення, отримане для фільтра ДКП+Хаар, представлено в Додатку Ж, на рис. Ж.8, а. Важливою особливістю є те, що багато висококонтрастних невеликих об'єктів зберігаються дуже добре, якщо вони були виявлені за допомогою методу Level Set. Вихідні дані для комбінації усунення шумів на основі ДКП і фільтра, представленого в [228], показані в Додатку Ж, на

рис. Ж.8, б. Фільтр забезпечує краще придушення шумів, але межі та дрібні деталі менш чіткі.

2.2.2 Використання нейронних мереж для адаптивної фільтрації РСА-зображень та аналіз її ефективності

Як показано вище, завдання ефективної фільтрації може бути вирішено [232]:

- 1) надійним виявленням областей локальної активності зображення;
- 2) застосуванням фільтрів, що ефективно придушують шум на однорідних ділянках та більш «м'яких» алгоритмів згладжування на локально-неоднорідних ділянках (в околах малорозмірних об'єктів та меж, а також на текстурях).

Розробці таких методів присвячено велику кількість робіт [232]. Ефективність підходу залежить від наступних трьох факторів: вибору фільтра для придушення завад на однорідних ділянках, надійності виявлення локальних неоднорідностей та якості збереження деталей алгоритмом, вибраним для досягнення цієї мети. Зазначимо, що багато методів вирішення цих трьох часткових завдань або не враховують характерні особливості спекл-шуму, описані вище, або мають недостатню швидкодію та ефективність.

Метою цього дослідження є розробка комбінованого методу локально-адаптивної фільтрації однопоглядових РСА-зображень за допомогою нейронної мережі.

У методі застосовується НМ прямого поширення [233], навчена "з вчителем" на тестових зображеннях методом зворотного поширення помилки. При цьому як вхідні параметри для НМ застосовуються не самі елементи зображення, а вектори вихідних значень щодо простих (елементарних) детекторів меж, що працюють у ковзному вікні розміром 5x5 пікселів. В якості елементарних детекторів меж були відібрані ВЛД, КР, ДХ та детектор на основі ДКП. Слід сказати, що дані детектори меж були обрані найбільш ефективні серед безлічі розглянутих.

Для фрагмента тестового зображення одержана карта меж, що представлена на рис. 2.12. Алгоритм отримання меж представлений на рис. 2.13.

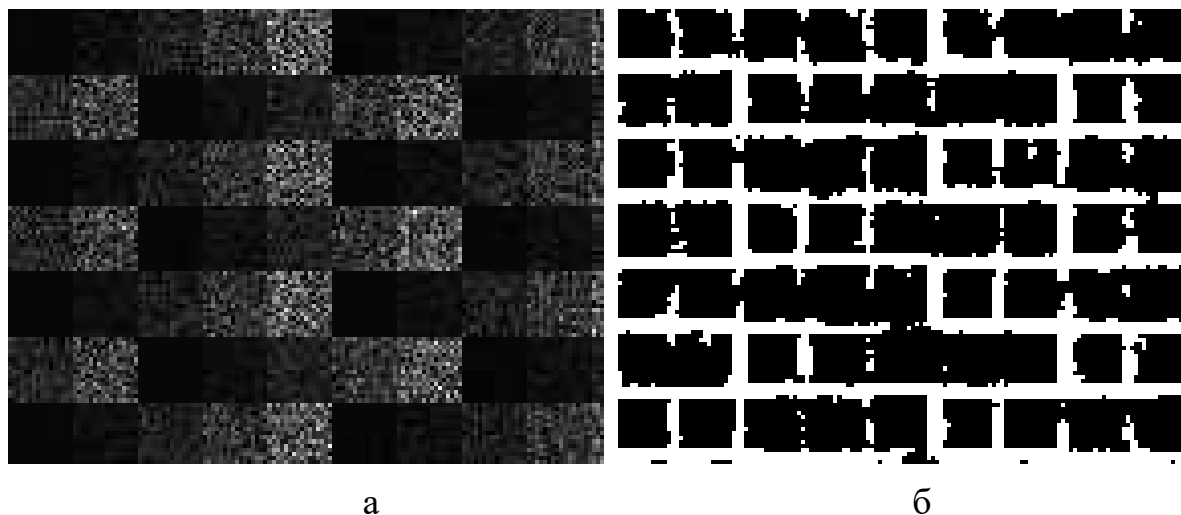


Рисунок 2.12 – Фрагмент тестового зображення (а) та отримана карта меж для фрагмента тестового зображення (б)

Зазначимо, що на реальних РСА-зображеннях спекл-шум має просторову корельованість, а нейромережа навчалася на тестових зображеннях з некорельованим спекл-шумом. Тому для усунення впливу корельованості шуму для реальних зображень проводилося попереднє проріджування обробленого зображення фільтром Лі в два рази для кожної з осей.

Після цього до них застосовувався нейромережевий детектор і таким чином були отримані первинні карти локальної активності. Після цього отримана бінарна карта перетворюється на вихідний масштаб простим розмноженням її елементів (значення одного пікселя первинної карти присвоюється чотирьом пікселям квадрата 2x2 елемента). Саме такі карти використовуються надалі при локально-адаптивній фільтрації з жорстким перемиканням та аналізі її ефективності.

Пропонований нижче метод обробки заснований на локально-адаптивному підході та полягає у виборі різних алгоритмів фільтрації (або різних параметрів одного і того ж фільтра) для різних областей зображення, що визначаються бінарною картою локальної активності. Локально-активні ділянки зображення детектуються за допомогою описаного вище нейромережевого детектора меж.

Приклад карти локальної активності, отриманої з його допомогою, показаний на рис. 2.10, а, де ці ділянки показані білим кольором.

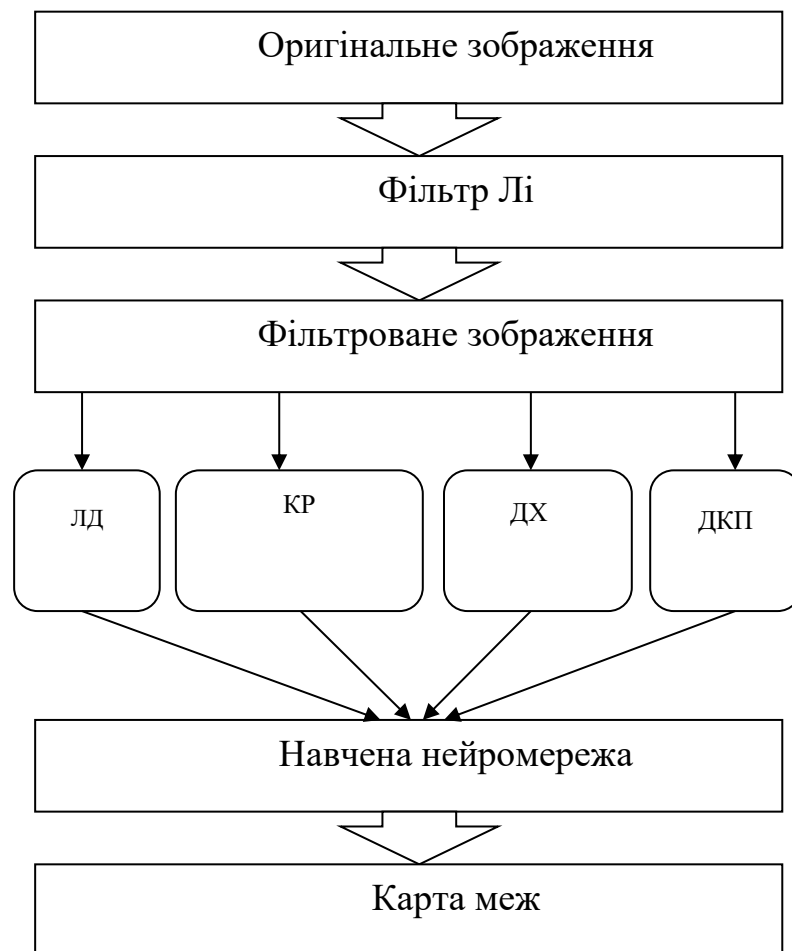


Рисунок 2.13 – Алгоритм отримання бінарних карт

На ділянках, зазначених детектором як межа або малорозмірний об'єкт, пропонується застосовувати ДКП-фільтрацію, що забезпечує краще збереження меж та деталей. Конкретні варіанти описані нижче. У свою чергу, на ділянках, які відповідають однорідним ділянкам, застосовується алгоритм фільтрації. При цьому значення β має бути достатньо великим, щоб забезпечувати ефективне придушення завад. Зазначимо, що для обробки зображень на локально-активних ділянках можливе застосування інших фільтрів.

Результати початкового етапу досліджень, отримані для зображення FR01 (рис. 2.9), представлені на рис. 2.14. Добре видно, що на виході фільтра BM3D (застосованого після ВСП логарифмічного типу зі зворотним ВСП та корекцією

середнього після фільтрації) мають місце артефакти (рис. 2.14, а), хоча межі збережені добре. У свою чергу, фільтр Лі з розміром ковзного вікна 7x7 (рис. 2.14, б) недостатньо придушує спекл-шум.

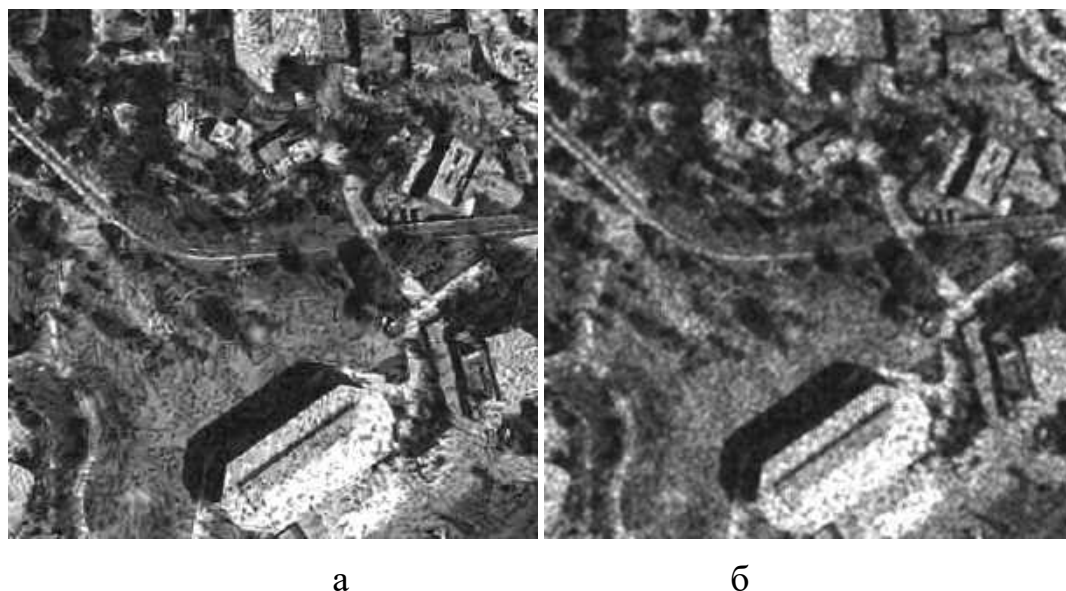


Рисунок 2.14 – Зображення FR01: на виході фільтра BM3D (а); відфільтроване фільтром Лі з ковзним вікном 7x7 пікселів (б)

Дані для порівняльного аналізу чотирьох тестових зображень, FR01-FR04 (рис. 2.15, оригінали зображень FR01 та FR02 наведені вище), наведені у таблицях 2.14-2.17. У таблицях 2.14 та 2.15 для тестових зображень FR02 та FR03, які не аналізувалися раніше, наведено докладніші дані числового моделювання (для різних значень параметрів аналізованих алгоритмів фільтрації). Фільтр на базі нейронної мережі позначено НМ.

Аналіз даних, наведених у табл. 2.14 та 2.15, показує наступне. Значення β , при яких спостерігаються мінімальні значення метрик для ДКП-фільтра з блоками 8x8 на основі (2.10), не збігаються. Для забезпечення мінімуму метрики MSE_{out} , яка, перш за все, характеризує ступінь придушення завад, необхідно використовувати значення β порядку 5,2. Мінімальні значення критерію MSE_{het} , що характеризує здатність фільтра зберігати деталі, мають місце при β порядку 4,2. Максимальні значення метрики $MSSIM$, що відповідає за візуальну якість,

спостерігаються при β порядку 4,6. Найкращим компромісним варіантом для неадаптивного ДКП-фільтра на основі (2.10) вважатимемо установку $\beta=4,8$.

Тут важливий той момент, що якщо розглядати кожну з метрик окремо, то положення оптимумів за параметром β для всіх тестових зображень приблизно збігаються. Цей факт дозволяє рекомендувати, яке значення β використовувати для обробки однорідних ділянок (наприклад, $\beta=5,2$), а яке – на неоднорідних (наприклад, $\beta=4,4$).

Другий висновок – ДКП-фільтр на основі (2.10) з обробкою в блоках 16×16 пікселів не призводить до більш ефективної обробки порівняно з обробкою в блоках 8×8 пікселів. Тому недоцільно використовувати такий алгоритм обробки ні окремо, ні в рамках локально адаптивного підходу для обробки, наприклад, на однорідних ділянках. Зазначимо, що оптимальне значення β приблизно дорівнює 4,4, якщо прагнути компромісу між аналізованими метриками.

Очевидно, що фільтр Лі (згідно з результатами досліджень найкращий розмір ковзного вікна – 7×7 пікселів) помітно поступається по ефективності обом ДКП-фільтрам з правильно встановленими значеннями β . Основна причина – невисока ефективність обробки на локально-неоднорідних ділянках, що виявляється у великих значеннях MSE_{het} . Фільтр BM3D з ВСП також недостатньо ефективний насамперед на однорідних ділянках, оскільки не враховує просторову корельованість завад.

Можна також порівняти дані, отримані для ДКП-фільтрів на основі (2.10), з даними для ДКП-фільтра з жорстким частотно-залежним порогом $T(m, n) = \beta \sigma_{\mu} I_{mn}^{mean}$, наведеними в вище в розділі 2.2.1.



а

б

Рисунок 2.15 – Оригінальні зображення: а – FR03, б – FR04

Таблиця 2.14 – Порівняльний аналіз методів фільтрації, зображення FR02

Фільтр, параметри	MSE _{out}	MSE _{het}	MSSIM
ДКП, 8x8, $\beta = 4,0$	537	715	0,804
ДКП, 8x8, $\beta = 4,4$	504	698	0,807
ДКП, 8x8, $\beta = 4,8$	489	701	0,807
ДКП, 8x8, $\beta = 5,2$	484	717	0,806
ДКП, 16x16, $\beta = 4,0$	547	755	0,798
ДКП, 16x16, $\beta = 4,4$	520	766	0,798
ДКП, 16x16, $\beta = 4,8$	513	796	0,793
Лі, 7x7	598	988	0,767
BM3D, ВСП	762	809	0,743
Frost_9x9	503	703	0,799
Frost_13x13	485	722	0,806
Frost_17x17	484	734	0,805
НМ 1	487	713	0,809
НМ 2	470	699	0,814

Для тестового зображення FR01 при рекомендованому $\beta = 3,0$ були отримані значення $MSE_{out}=443$, $MSE_{het}=672$ і $MSSIM=0,823$. Таким чином, значення MSE_{out} та $MSSIM$ трохи кращі, а значення MSE_{het} збігаються (див. табл. 2.16). Для тестового зображення FR04 були отримані $MSE_{out}=421$, $MSE_{het}=681$ і $MSSIM=0,769$, тобто значення MSE_{out} трохи краще, значення MSE_{het} трохи гірше, а значення $MSSIM$ збігаються.

Отже, варіанти неадаптивної ДКП-фільтрації з жорстким порогом і комбінованим порогом (2.10) забезпечують приблизно однакову ефективність обробки.

Цікаві результати були отримані для фільтра Фроста. Зі збільшенням розміру ковзного вікна підвищується ефективність придушення завад на однорідних ділянках, але погіршується збереження меж і деталей. При розмірі ковзного вікна 13×13 пікселів значення метрик майже такі самі, як і для кращих ДКП-фільтрів.

Таблиця 2.15 – Порівняльний аналіз методів фільтрації, зображення FR03

Фільтр, параметри	MSE_{out}	MSE_{het}	$MSSIM$
ДКП, 8×8 , $\beta = 4,0$	553	791	0,761
ДКП, 8×8 , $\beta = 4,4$	519	779	0,765
ДКП, 8×8 , $\beta = 4,8$	503	786	0,765
ДКП, 8×8 , $\beta = 5,2$	497	805	0,763
ДКП, 16×16 , $\beta = 4,0$	554	809	0,754
ДКП, 16×16 , $\beta = 4,4$	528	825	0,752
ДКП, 16×16 , $\beta = 4,8$	520	859	0,745
Лі, 7×7	597	1035	0,725
BM3D, ВСП	812	882	0,685
Frost_9x9	520	768	0,754
Frost_13x13	497	786	0,763
Frost_17x17	492	797	0,763
НМ 1	500	789	0,767
НМ 2	483	777	0,771

Таблиця 2.16 – Порівняльний аналіз методів фільтрації, зображення FR01

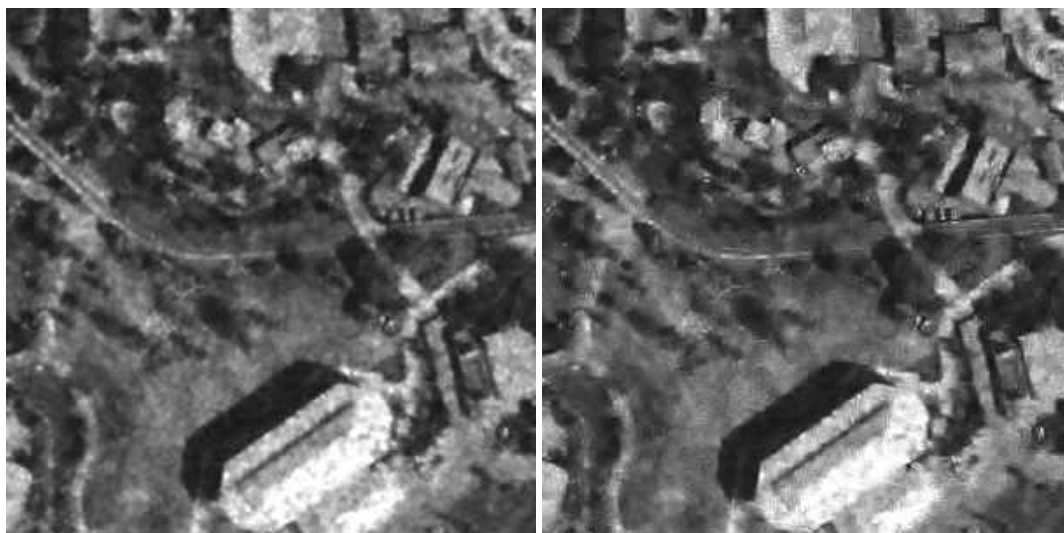
Фільтр, параметри	MSE _{out}	MSE _{het}	MSSIM
ДКП, 8x8, $\beta = 4,8$	469	672	0,817
ДКП, 16x16, $\beta = 4,4$	498	712	0,809
Лі, 7x7	566	843	0,780
BM3D, ВСП	746	872	0,757
Frost_13x13	460	670	0,816
НМ 1	466	680	0,823
НМ 2	455	655	0,826

Таблиця 2.17 – Порівняльний аналіз методів фільтрації, зображення FR04

Фільтр, параметри	MSE _{out}	MSE _{het}	MSSIM
ДКП, 8x8, $\beta = 4,8$	441	673	0,771
ДКП, 16x16, $\beta = 4,4$	454	721	0,763
Лі, 7x7	513	854	0,736
BM3D, ВСП	761	838	0,688
Frost_13x13	430	662	0,769
НМ 1	438	681	0,772
НМ 2	420	674	0,775

Питання полягає в тому, що на даний момент не ясно, яким чином враховувати при виборі розміру ковзного вікна просторовий спектр спеклу. Зображення на виході фільтра Фроста наведено на рис. 2.16, а. Як видно з порівняння зображень на рис. 2.16, б і 2.16, в, вихідна якість для обох ДКП-фільтрів висока, у другому випадку дещо вища чіткість контрастних меж та малорозмірних об'єктів. Водночас забезпечення вищої чіткості, як і раніше, бажане. Очевидно, що досягнутий гарний вихідний ефект та компроміс властивостей.

На підставі отриманих даних було вирішено розглянути два варіанти локально-адаптивних ДКП-фільтрів. Загальний алгоритм їх роботи описується виразом:



а

б



в

Рисунок 2.16 – Зображення FR01: на виході фільтра Фроста 13x13 (а); на виході фільтра ДКП з фіксованим $\beta=4,8$, 8x8 (б); на виході фільтра НМ 2 (в)

$$D_t(m, n, k, l) = \begin{cases} D_{het}(m, n, k, l), & \text{якщо } M(m+3, n+3) = 1 \\ D_t(m, n, k, l) & \text{якщо } M(m+3, n+3) = 0 \end{cases}, \quad (2.12)$$

де $M(m, n)$ – значення карти локальної активності у m -м пікселі; $D_{het}(m, n, k, l)$ – значення коефіцієнта після порогової обробки, що використовується для локально-активних блоків; $D_t(m, n, k, l)$ – значення коефіцієнта згідно з (2.10).

В першому випадку (НМ 1 у таблицях 2.14-2.17) використовується алгоритм:

$$D(m, n, k, l) = D^3(m, n, k, l) / (\beta_1 \sigma_\mu I_{mn}^{med} \sqrt{W_{kl}})^2, \quad (2.13)$$

де в I_{mn}^{med} – значення медіани у nm -му блоці.

При цьому враховується, що медіанне значення в блоці зазвичай менше, ніж середнє, особливо для околів малорозмірних деталей з позитивними контрастами щодо навколишнього фону. Дослідження, проведені для чотирьох тестових зображень, показали, що розумною практичною рекомендацією є вибір $\beta_1=5,2$.

Для другого варіанта (НМ 2) пропонується розраховувати поріг як $T_{nm} = \beta_2 \sigma_\mu I_{mn}^{mean}$ і використовувати жорстку порогову обробку виду:

$$D_t(m, n, k, l) = \begin{cases} D(m, n, k, l), & \text{якщо } |D(m, n, k, l)| \geq \beta_2 \sigma_\mu I_{mn}^{mean} \\ 0 & \text{в інших випадках} \end{cases}, \quad (2.14)$$

В результаті досліджень для аналізованих тестових зображень визначено значення, що рекомендується для цього варіанту, $\beta_2=4,4$.

Зазначимо, однак, що така рекомендація справедлива лише для спекла з розглянутими у цій роботі просторово-кореляційними характеристиками. На однорідних ділянках для обох варіантів пропонується використовувати ДКП-фільтр із порогом (2.10) з розміром блоку 8×8 $\beta=5,2$. Як впливає з аналізу даних для запропонованих варіантів локально-адаптивних фільтрів (НМ 1 та НМ 2, нижні два рядки у всіх таблицях), вони або перевершують всі інші згідно з усіма метриками, або трохи поступаються якомусь з інших фільтрів по одній з метрик. Для другого запропонованого варіанта отримано помітно найкращі результати.

2.3 Прогнозування ефективності фільтрації зображень

2.3.1 Прогнозування ефективності фільтрації сигнально-залежного шуму на оптичних зображеннях та одноканальних зображеннях ДЗЗ

Бувають практичні ситуації, коли зменшення шуму зображення дає суттєвий ефект. Покращення, що характеризується, наприклад, збільшенням PSNR приблизно на 5...12 дБ. У той же час існують практичні ситуації (низький рівень шуму та/або текстурний характер зображення, що підлягає фільтрації), коли придушення шуму не дає практичної користі або фільтрація навіть погіршує візуальну якість обробленого зображення порівняно з оригінальним [128], [143]. Тоді фільтрацію можна пропустити (уникнути). Постає практичне питання, як автоматично прийняти рішення фільтрувати зображення чи ні.

Останнім часом було зроблено декілька кроків вперед для відповіді на ці питання [5], [10], [234]. Основна ідея прогнозування ефективності фільтрації була вперше висунута в роботі [5] для випадку адитивного шуму.

Робота [10] стосується прогнозування для зображень, спотворених більш складним сигнально-залежним шумом, який має адитивну та квазіпуассонівську складові. Інші відмінні риси цього дослідження полягають в наступному. Як параметри, що описують ефективність фільтрації, були використані покращення метрик PSNR і PSNR-HVS-M. Більше зображень, ніж у [5], було використано для отримання діаграм розсіювання. Однак результати були отримані лише для однієї (хоча і важливої) моделі сигнально-залежного шуму.

Ще одна практична ситуація розглядається в роботі [234]. Знову ж таки, була розглянута модель адитивного шуму. Але кількість тестових зображень, які використовуються для отримання діаграми розсіювання, значно зросла. Крім того, у прогнозі використовувалися два і більше статистичних параметрів. Завдяки цьому було забезпечено більш точний прогноз. Було також показано, що спосіб визначення статистичного параметра впливає на прогноз. Беручи до уваги ці нещодавні результати, перша мета полягає в тому, щоб покращити ефективність прогнозу для випадку сигнально-залежного шуму. Друга мета полягає в тому, щоб вивчити, наскільки загальним є прогноз, тобто чи можна його використовувати для інших моделей сигнально-залежного шуму.

Нижче представлено підходи того, як можна здійснити прогноз ефективності придушення шуму. У цьому сенсі буде нагадано основні вимоги до прогнозування.

По-перше, потрібно мати якийсь параметр, здатний адекватно характеризувати ефективність придушення шуму. Більшість дослідників використовують стандартні критерії ефективності фільтрації як вихідну MSE (після фільтрації) для заданої інтенсивності шуму, який можна охарактеризувати дисперсією шуму для випадку адитивного шуму. Однак шум може бути різної інтенсивності та не бути адитивним. Таким чином, перша пропозиція полягала у використанні співвідношення:

$$MSE_{out}/\sigma_{eq}^2 = \sum_{i=1}^I \sum_{j=1}^J (I_{ij}^{true} - I_{ij}^f)^2 / \sum_{i=1}^I \sum_{j=1}^J (I_{ij}^{true} - I_{ij}^n)^2, \quad (2.15)$$

де σ_{eq}^2 – еквівалентна дисперсія шуму; I_{ij}^{true} , I_{ij}^f , I_{ij}^n – істинне, відфільтроване та вихідне (зашумлене) значення зображення в ij -му пікселі відповідно; I та J позначають розмір зображення.

Зауважимо, що коефіцієнт (2.15) характеризує ефективність фільтрації для різних типів шуму та різної інтенсивності шуму. Якщо це співвідношення значно менше одиниці, фільтрація є ефективною.

Подібним чином також можна охарактеризувати ефективність придушення шуму за допомогою:

$$\begin{aligned} IPSNR &= PSNR_{out} - PSNR_{inp}, \\ PSNR_{out} &= 10 \log_{10} (\sum_{i=1}^I \sum_{j=1}^J (I_{ij}^{true} - I_{ij}^f)^2 / IJ), \\ PSNR_{inp} &= 10 \log_{10} (\sum_{i=1}^I \sum_{j=1}^J (I_{ij}^{true} - I_{ij}^n)^2 / IJ), \end{aligned} \quad (2.16)$$

яке визначається у дБ. Значення IPSNR близько нуля (менше 1 дБ) вказує на те, що ефективність фільтрації низька і немає практичної потреби її виконувати.

Крім того, можливе також [10], [234] використання:

$$IPSNR - H = PSNR - HVS - M_{out} - PSNR - HVS - M_{inp}, \quad (2.17)$$

де наведені метрики вимірюються так само в дБ. Якщо $IPSNR - N$ менше 1 дБ, то фільтрація так само не має сенсу.

Спільний аналіз (2.17) з (2.16) або з (2.15) дозволяє отримати повне уявлення про потенціал фільтрації, якщо прогноз є достатньо точним.

Тут підходимо до другої вимоги – прогноз повинен бути достатньо точним; інакше прогнозування є марним. Щоб зрозуміти цю вимогу, повинні згадати деталі того, як виконується прогноз. Передбачається, що апріорно отримано залежність параметра, що характеризує ефективність фільтрації, від параметра, що характеризує статистику зображення та шуму. Приклади таких залежностей, які необхідно описати аналітично, представлені на рис. 2.17 червоною (стандартні оптичні дані) і синьою (типові дані дистанційного зондування) кривими, підігнаними до точок діаграми розсіювання. Як параметр, що характеризує властивості зображення та шуму, застосована ймовірність $P_{2\sigma}$. Це середня ймовірність того, що амплітуди коефіцієнтів ДКП у блоках 8×8 пікселів не перевищують значення $2\sigma(n, m)$ для блоку з верхнім лівим кутом у nm -му пікселі аналізованого зображення. Поріг $2\sigma(n, m)$ розраховується індивідуально для кожного блоку з урахуванням апріорі відомої залежності локальної дисперсії від локального середнього (для блоку) у випадку сигнал-залежного шуму ($\sigma(n, m) = \sigma_{add}$ для випадку адитивного шуму, де σ_{add} позначає стандартне відхилення шуму). Пізніше покажемо, як можна визначити $\sigma(n, m)$.

Особлива вимога до залежності полягає в тому, щоб вона була монотонною (функція $y = ae^{bx}$ використовувалася для підгонки до невід’ємних значень параметрів a і b , $IPSNR$ служить як y і $P_{2\sigma}$ як x для графіків на рис. 2.17).

Ще одна вимога – гарна підгонка кривих. Говорячи «гарна», маємо на увазі дві речі. По-перше, підгонку слід виконувати за допомогою відповідного типу функцій, забезпечуючи високі значення параметрів, що описують якість підгонки. По-друге, відстань точок діаграми розсіювання від підігнаної кривої (отриманої залежності) повинна бути відповідно малою (наприклад, менше 1 дБ).

Тут варто згадати, як були отримані діаграми розсіювання. Є набір тестових зображень і набір шумових характеристик. Кожне тестове зображення, спотворене

шумом із заданим варіантом характеристик, фільтрується та визначаються параметри (2.16), (2.17) та/або (2.18). Також розраховується параметр $P_{2\sigma}$. Потім на діаграмі розсіювання отримується одна точка. Отже, завдання на попередньому етапі наступні. По-перше, потрібно отримати відповідну кількість точок діаграми розсіювання, де їх горизонтальна вісь (що відповідає) покрита досить щільно. По-друге, підгонка повинна бути виконана досить якісно. Нижче показано деякі аспекти вибору тестових зображень, вибору параметрів шуму та підгонки кривої до діаграм розсіювання. Треба зауважити, що такі діаграми розсіювання можна отримати і підгонку можна виконати досить просто (див. рис. 2.17).

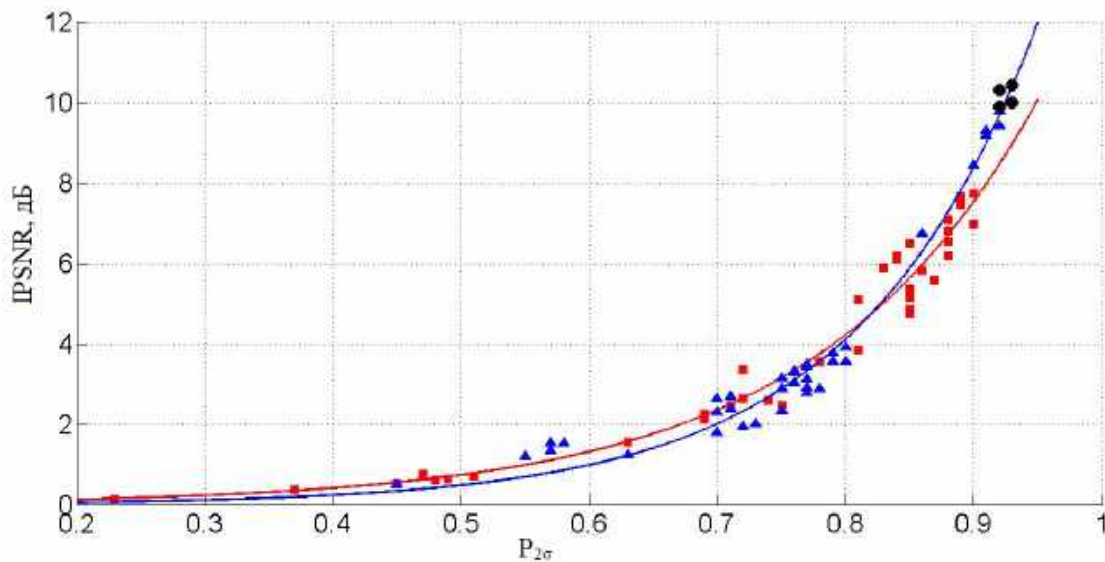


Рисунок 2.17 – Приклад залежностей, отриманих підгонкою кривої до діаграм розсіювання (червоні квадрати для оптичних даних, сині трикутники для даних дистанційного зондування, чорні точки для даних, спотворених спеклом)

Припустимо, тепер маємо підігнану криву. Потім для даного зображення необхідно обчислити параметр $P_{2\sigma}$ (або якийсь інший, див. далі), підставити його в отриману залежність і розрахувати один або декілька параметрів, що характеризують ефективність фільтрації. З розуміння цього методу випливає остання вимога до прогнозування. Він має бути достатньо швидким, зокрема швидшим за саму фільтрацію. Цю вимогу може задовольнити параметр $P_{2\sigma}$. ДКП у

8x8 піксельних блоках є стандартною операцією в обробці зображень і вона швидка [235]. Крім того, фактично немає необхідності обчислювати локальні оцінки параметра у всіх можливих положеннях блоків перед усередненням цих локальних оцінок. Останні дослідження [10] показали, що для аналізованого зображення достатньо мати близько 500 локальних оцінок. Завдяки цьому прогнозування можна виконувати значно швидше, ніж стандартне видалення шумів на основі ДКП, і у багато разів швидше, ніж фільтрація BM3D.

Нарешті, може зацікавити загальність розглянутого підходу до прогнозування ефективності фільтрації. Представлено підхід для двох типів фільтрів. Перший — стандартний фільтр на основі ДКП. Другим фільтром є BM3D [78], який вважається одним з найкращих [128] у видаленні АБГШ. З одного боку, обидва ці фільтри засновані на ДКП. Це є причиною того, що існує жорстка залежність між ефективністю фільтрації і статистикою коефіцієнтів ДКП в блоках. З іншого боку треба зауважити, що ці фільтри є одними з кращих [128]. Таким чином, оцінюючи ефективність фільтрації для них, оцінюємо практичну межу ефективності фільтрації, досягну на даний момент.

Пам'ятаймо, що поріг для фільтрації з жорстким порогом встановлюється як $T(n, m) = \beta \sqrt{f(\bar{I}(n, m))} = \beta \hat{\sigma}(n, m)$, де β параметр, який зазвичай дорівнює 2,7, а локальне середнє значення $\bar{I}(n, m)$ слугує локальною оцінкою дійсного значення зображення; $f(*)$ це апіорно відома або попередньо оцінена залежність локальної дисперсії від локального середнього (дійсне зображення). Вираз $\hat{\sigma}(n, m) = \sqrt{f(\bar{I}(n, m))}$ використовується для визначення локальних оцінок $P_{2\sigma}$ в блоках.

З вищезазначеного можна очікувати, що точність прогнозу залежить від багатьох факторів, таких як:

- 1) вибраний набір тестових зображень та їх властивості;
- 2) використовуваний набір шумових характеристик;
- 3) вибраний тип функції підгонки та її параметри (наприклад, порядок полінома);
- 4) використаний метод підгонки (стандартний LMSE або якийсь стійкий).

Нижче приділимо увагу кожному з цих факторів. Почнемо з тестових зображень. З самого початку було зрозуміло [5], що набір зображень повинен включати зображення низької, середньої та високої складності, оскільки будь-який тип із них можна зустріти на практиці. Типовим рішенням у таких випадках є використання стандартних тестових зображень, тому використали набір зображень, який включав добре відомі зображення: Baboon, Barbara, Goldhill, Lena та Peppers. Додали до них зображення Grass, представлене на рис. 2.18. Зображення Lena та Peppers мають низьку складність, а Baboon і Grass – зображення високої складності (дуже текстурні). Цей набір тестових зображень був названий «оптичним».



Рисунок 2.18 – Оригінал зображення Grass

Інший набір тестових зображень також містив шість зображень. Це тестові зображення FR01, FR02, FR03 і FR04 і досить відомі тестові зображення Frisco і Diego. Зображення FR01, FR02, FR03 і FR04 мають середню складність, зображення Diego містить багато дрібних деталей і досить складне, зображення Frisco досить просте. Далі цей набір буде називатися «дистанційне зондування».

Пояснимо, чому використали два набори. Першою метою було проаналізувати, як набір зображень впливає на підгонку кривої. Другою метою було дослідити, чи впливає вміст зображень на результати фільтрації. Нагадаємо,

що люди часто думають, що якщо результати були отримані для типових тестових зображень, їх не можна екстраполювати на дані дистанційного зондування або на якийсь інший конкретний тип зображень.

Точки розсіювання, отримані для тестового набору «Оптичний», представлені червоними квадратами на рис. 2.17. У свою чергу, синіми трикутниками позначені точки розсіювання, отримані для набору «Дистанційне зондування». Криві, окремо підігнані до частин «Оптичний» і «Дистанційне зондування» точок діаграми розсіювання, майже збігаються. Точки для обох наборів групуються разом. Таким чином, можемо констатувати, що вміст зображення (те, що зображено) суттєво не впливає на ефективність фільтрації. У свою чергу, на ефективність усунення шумів значно більше впливає складність зображення та інтенсивність шуму.

Тут хотіли б нагадати використану модель сигнально-залежного шуму, яка впливає з формули (1.13). Було припущено, що:

$$I_{ij}^n = I_{ij}^{true} + I_{ij}^{SI} + I_{ij}^{SD}, \quad (2.18)$$

де I_{ij}^{SI} та I_{ij}^{SD} визначають СНШ та СЗШ у ij -пікселі відповідно. Обидва компонента вважаються з нульовим середнім та відносяться до білого некорельованого гаусового шуму. Дисперсія цього шуму визначається як $\sigma_{ij}^2 = \sigma_0^2 + kI_{ij}^{true}$, де σ_0^2 визначає дисперсію СНШ, а k – коефіцієнт СЗШ. Щоб отримати точки розсіювання, було використано шість комбінацій (наборів) σ_0^2 і k . Вони представлені в таблиці 2.18. Ці комбінації вибрано для охоплення широкого спектру можливих практичних ситуацій. Наприклад, варіант 1 стосується домінуючої компоненти SI та інтенсивного шуму (еквівалентна дисперсія σ_{eq}^2 близько 125 і залежить від тестового зображення). Варіант 3 є прикладом переважної компоненти СЗШ (σ_{eq}^2 близько 105 і, таким чином, шум також інтенсивний). Навпаки, варіант 6 є прикладом переважającego адитивного (СНШ) шуму (σ_{eq}^2 близько 15 і, отже, шум не інтенсивний), його навряд чи можна

помітити на відповідних зашумлених зображеннях [236]. Інші комбінації відображають проміжні ситуації.

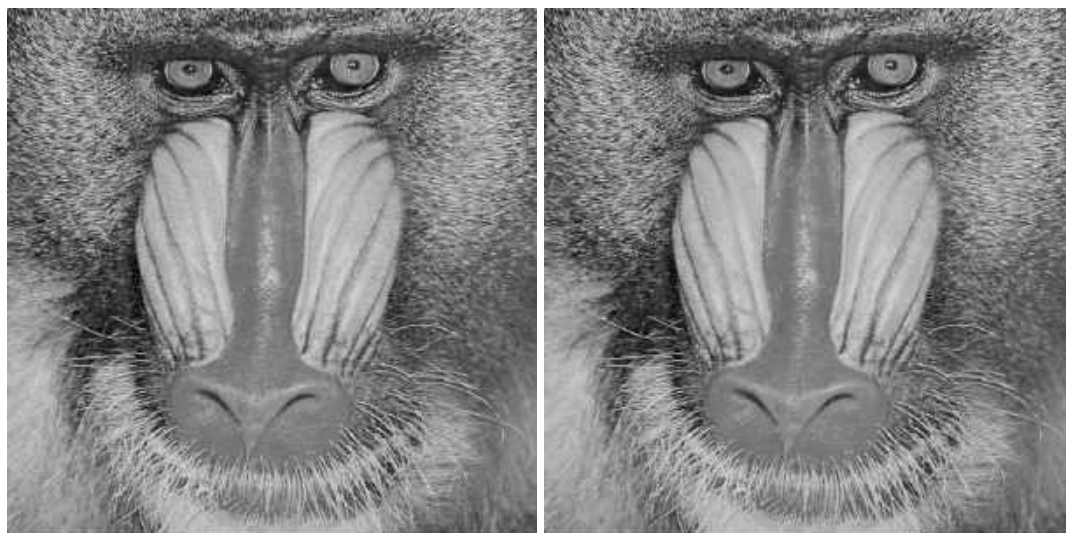
Тепер наведемо два приклади фільтрації. На рис. 2.19, а показано зашумлене зображення Baboon, спотворене сигнально-залежним шумом варіанту 6. Його відфільтрована версія представлена на рис. 2.19, б. Як бачимо, шум практично невидимий на вихідному зображенні з шумом, а відфільтроване зображення виглядає так само, як оригінальне. IPSNR для цього випадку становить 0,76 дБ, IPSNR-Н дорівнює 0,28 дБ і PSNR-HVS-Minr дорівнює 44,1 дБ.

Це показує, що фільтрація марна з двох причин. По-перше, IPSNR та IPSNR-Н обидва малі, і це вказує на те, що фільтрація неефективна. По-друге, PSNR-HVS-Minr перевищує 40 дБ, що відповідає практичній непомітності шумів та інших спотворень на зображеннях [182], [236].

Зауважте, що для розглянутого випадку (зображення Baboon, $\sigma_0^2=9$, $k=0,05$) і, згідно з підігнаними кривими на рис. 2.17, прогноз добре узгоджується з спостережуваним. Треба підкреслити, що для найбільш текстурних зображень (Grass, Baboon, Diego) спостерігалися досить малі значення і для сигнально-залежного шуму варіанту 6 (також для варіантів 4 і 5 для зображення Grass). Це узгоджується зі знаннями та практичним досвідом багатьох дослідників – найважче отримати користь від фільтрації для високотекстурних зображень і якщо шум не інтенсивним. Для $P_{2\sigma}=0,47$ обидві криві прогнозування дають IPSNR близько 0,6 дБ, тобто приблизно значення, отримане для розглянутого окремого випадку.

Таблиця 2.18 – Набори параметрів для сигнал-залежного шуму моделі (2.19)

Параметр	1	2	3	4	5	6
σ_0^2	100	25	9	36	36	9
k	0.2	0.6	0.8	0.4	0.1	0.05

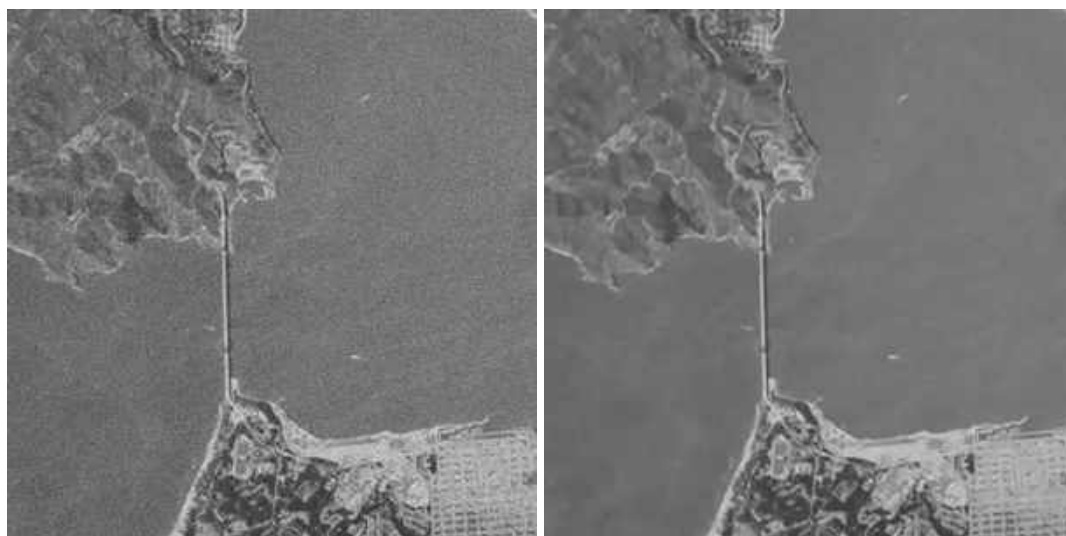


а

б

Рисунок 2.19 – Зображення Baboon: зашумлене, варіант 6 (а); відфільтроване (б)

Розглянемо протилежну практичну ситуацію. На рис. 2.20, а представлено зашумлене зображення Frisco, де використовується варіант 1 шуму ($\sigma_0^2=100$, $k=0,2$). Шум є інтенсивним, і це видно на цьому зображенні (PSNR-HVS-Minp=29,8 дБ). Відфільтроване зображення наведено на рис. 2.20, б.



а

б

Рисунок 2.20 – Зображення Frisco: а - зашумлене (варіант 1), б – відфільтроване.

Ефект фільтрації очевидний і це підтверджується $IPSNR=9,8$ дБ і $IPSNR-H=5,4$ дБ, тобто ці критерії є досить великими. Зверніть увагу, що в цьому випадку $P_{2\sigma}=0,92$. Це означає, що більшість коефіцієнтів DCT у блоках досить малі, і вони видаляються (присвоюються нульові значення) на етапі фільтрації.

Завдяки цьому фільтрація ефективна. Високу ефективність фільтрації можна передбачити з апроксимаційних кривих, представлених на рис. 2.17. Для обох кривих прогнозування дають близько 9...10 дБ, тобто приблизно значення, отримане для розглянутого окремого випадку.

Ці результати є прикладами того, що запропонований підхід до прогнозування ефективності зменшення шумів може працювати достатньо добре. Однак досі розглядався лише один окремий випадок сигнально-залежного шуму. Тому вирішено проаналізувати ще один важливий випадок сигнально-залежного шуму, а саме чистий мультиплікативний шум і, точніше, спекл-шум. Моделювання відбувалося з зображеннями FR01-FR04, спотвореними спеклом. Отримані результати частково показано на рис. 2.17 чорними крапками на діаграмі розсіювання. У цьому випадку $\hat{\sigma}(n, m) = \bar{I}(n, m)\sigma_{\mu}$ та всі розрахункові значення $P_{2\sigma}$ становлять приблизно 0,92. Це пояснюється тим, що шум дуже інтенсивний (це властивість спеклу). Отримані значення $IPSNR$ становлять приблизно 10 дБ і вони узгоджуються з прогнозованими значеннями, принаймні, для кривої, підігнаної за заданими значеннями «Дистанційного зондування». Ці результати показують, що запропонований підхід до прогнозування є досить загальним і може бути застосований для різних типів сигнально-залежного шуму за умови, що функція $f(*)$ апріорі відома або точно попередньо оцінена.

Параметр, який використовувався для прогнозування ефективності зменшення шуму вище та в попередніх публікаціях [5], [10], [234], був обраний емпірично без глибокої теоретичної основи для цього. Проте навіть цей статистичний параметр дає досить хороший прогноз. Результати, представлені в [10], показують, що параметр R^2 для підігнаної кривої $IPSNR(P_{2\sigma})$ становить близько 0,96 (подібні значення R^2 спостерігалися для $IPSNR(P_{2\sigma})$ у дещо інших

наближеннях). Це, загалом, відповідність хорошій якості, хоча її подальше вдосконалення є бажаним і можливим (шляхом ускладнення прогнозування).

Отримані наближення $\text{IPSNR-H}(P_{2\sigma})$ менш точні. Параметр R^2 для підігнаної кривої становить приблизно 0,92 (приблизно такі ж значення спостерігалися в [5] для адитивного шуму, а менші значення були отримані в [234] для великого набору тестових зображень). Це означає, що $\text{IPSNR-H}(P_{2\sigma})$ спрогнозувати важче, ніж $\text{IPSNR}(P_{2\sigma})$, хоча існують способи подолати цей недолік за допомогою більш складного прогнозування [234].

Варто зазначити, що в роботах [5] і [10], поряд з статистичним параметром $P_{2\sigma}$, було проаналізовано статистичний параметр $P_{2.7\sigma}$. Це середня ймовірність того, що амплітуди коефіцієнтів ДКП у блоках перевищують $2,7\hat{\sigma}(n, m) = 2,7\sqrt{f(\bar{I}(n, m))}$ для сигнально-залежного шуму або $2,7\sigma_{\text{add}}$ для чистого адитивного шуму. Аналіз точності підгонки показав, що залежності $\text{IPSNR}(P_{2.7\sigma})$ і $\text{IPSNR-H}(P_{2.7\sigma})$ підігнані дещо гірше (з меншим R^2), ніж відповідні залежності для $P_{2\sigma}$. Це привело до ідеї про те, що для розрахунку статистичних параметрів для прогнозування можна використовувати менші порогові значення. Зокрема, було показано в [234] для випадку адитивного шуму, що корисним може бути параметр $P_{0.5\sigma}$. Цей параметр розраховується як середня ймовірність того, що абсолютні значення коефіцієнтів ДКП в блоках не перевищують $0,5\hat{\sigma}(n, m) = 0,5\sqrt{f(\bar{I}(n, m))}$ для розглянутого випадку сигнально-залежного шуму.

Функції типу $y=ae^{bx}$ були використані для підгонки з невід'ємними значеннями параметрів a і b , де IPSNR або IPSNR-H слугують як y і $P_{0.5\sigma}$ як x . Відповідна діаграма розсіювання зображена на рис. 2.21. Точки діаграми розсіювання для набору «Оптичний» показані червоними квадратами, а відповідна підігнана крива також показана червоним кольором. Точки для набору «Дистанційне зондування» позначені синіми трикутниками, а підібрана крива має такий же колір. Видно, що обидві криві практично збігаються. Щоб підтвердити це, у таблиці 2.19 подано інформацію про підгонку.

На додаток до параметрів a і b апроксимаційних кривих і параметра R^2 , розглянутих вище, також представляємо середню квадратичну похибку прогнозу (в дБ). Параметри a і b для підігнаних кривих майже однакові. Але головний висновок полягає в тому, що підгонка IPSNR на основі $P_{0.5\sigma}$ краща (точніша), ніж на основі $P_{2\sigma}$ - значення R^2 досягають 0,99. RMSE також невелика, і аналіз їх значень показує, що IPSNR можна передбачити з похибкою менше 1 дБ. Це задовільна точність прогнозу на практиці. Згідно з представленими результатами, не варто проводити фільтрацію, якщо $P_{0.5\sigma}$ менше 0,15. Скатерограми для IPSNR- $H(P_{0.5\sigma})$ зображені на рис. 2.22. Всі позначення співпадають із наведеними раніше.

Підігнані криві мають схожий характер – вони зростають при зростанні $P_{0.5\sigma}$. Параметри підгонки представлені в таблиці 2.20. Значення a і b досить близькі, але не можна стверджувати, що криві збігаються. Основна різниця в прогнозі для підігнаних кривих має місце для великих $P_{0.5\sigma}$, коли IPSNR- H також є великими. Тоді ця різниця не така проблематична, оскільки рішення про фільтрацію зображення приймається в будь-якому випадку.

Очевидним позитивним моментом є те, що R^2 збільшився порівняно зі своїми значеннями для наближень на основі $P_{2\sigma}$ та $P_{2.7\sigma}$. Це означає, що $P_{0.5\sigma}$ варто використовувати для забезпечення більшої точності прогнозу. Тим не менш, точність прогнозу для IPSNR- H гірша, ніж для IPSNR. Аналіз підігнаних кривих показує, що не варто виконувати фільтрацію з метою підвищення візуальної якості фільтрованих зображень, якщо вона мала, наприклад, менше 0,2. Зауважте також, що IPSNR- H завжди менше, ніж IPSNR. Це ще раз показує, що метрика PSNR (і, відповідно, IPSNR) недостатньо адекватна для характеристики візуальної якості зображення.

Тепер перевіримо, чи добре працюють отримані наближення для моделі (2.20) сигнально-залежного шуму. Точки діаграми розсіювання, отримані для випадку чистого мультиплікативного шуму, позначені чорними крапками на рис. 2.21 і 2.22.

Як видно, усі чотири точки близькі до обох підігнаних кривих на рис. 2.21. Точки знаходяться між підігнаними кривими на рис. 2.22. Таким чином, можемо

стверджувати, що підхід прогнозування все ще є досить загальним, якщо прогнозування виконується з використанням вибраного набору тестових зображень та їхніх властивостей.

На початку підрозділу було згадано, що підгонка (і прогноз) залежить від вибраного набору тестових зображень та їхніх властивостей. Як було показано, різні набори тестових зображень можуть дати трохи різні результати підгонки, хоча спостережувана різниця була невеликою. Щоб дати практичні рекомендації, можна об'єднати результати для наборів Оптичне та Дистанційне зондування.

Для використаного типу апроксимуючих функцій отримані наступні вирази:

$$IPSNR = 0.162 \exp(11,36P_{0.5\sigma}), \quad (2.19)$$

$$IPSNR - H = 0.077 \exp(11,96P_{0.5\sigma}). \quad (2.20)$$

Значення R^2 для них дорівнюють 0,983 і 0,914 відповідно. Отже, якість прогнозування досить висока і краща для метрики IPSNR.

Однією з можливостей вибору тестових зображень для отримання діаграм розсіювання є використання більшої кількості тестових зображень. У [234] було використано 34 тестових зображення. Тоді отримані діаграми розсіювання є досить щільними, і це добре для підгонки. Однак, основною вимогою до вибору зображень для набору є надання точок з різними значеннями аргументів, які охоплюють весь діапазон можливих варіацій аргументів. Очевидно, що обраний тип функції впливає на підгонку. Хоча результати з використаною функцією $y = ae^{bx}$ досить хороші, потенційно можна знайти кращі наближення. Було перевірено можливість використання для апроксимації кубічного полінома для графіків розсіювання на рис. 2.21 і 2.22 (для обох наборів разом). Отримані значення R^2 дорівнюють 0,983 і 0,921 відповідно. Для підгонки було використано стандартні інструменти Matlab.

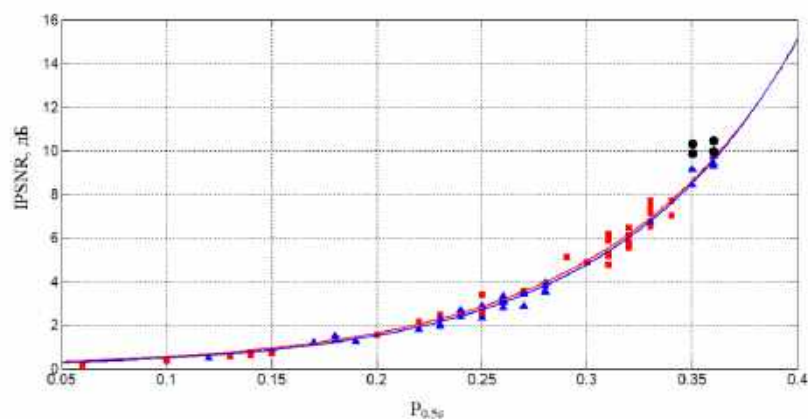


Рисунок 2.21 – Діаграми розсіювання $IPSNR(P_{0.5\sigma})$ та підігнані криві

Таблиця 2.19 – Параметри підгонки $IPSNR(P_{0.5\sigma})$

Набір зображень	a	b	R^2	RMSE
Оптичний	0,1754	11,13	0,976	0,39
Дистанційне зондування	0,1523	11,50	0,990	0,26

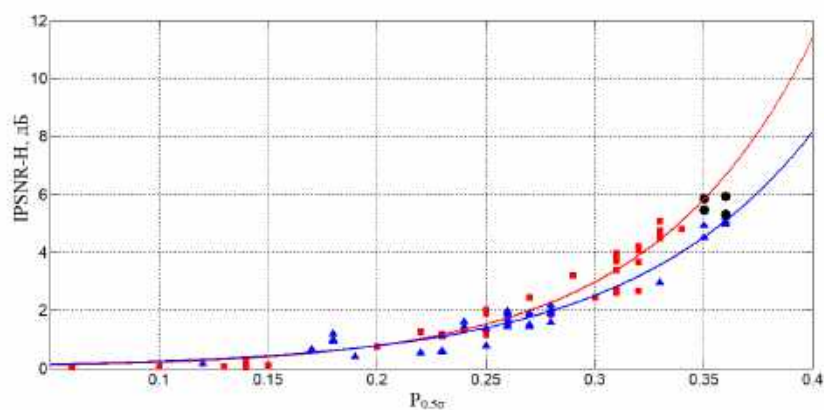


Рисунок 2.22 – Діаграми розсіювання $IPSNR-H(P_{0.5\sigma})$ та підігнані криві

Таблиця 2.20 – Параметри підгонки $IPSNR-H(P_{0.5\sigma})$

Набір зображень	a	b	R^2	RMSE
Оптичний	0.053	13.4	0.940	0.43
Дистанційне зондування	0.073	11.80	0.947	0.32

Нарешті, було застосовано метод найменшої середньоквадратичної похибки [237] для підгонки кривої. Цей метод добре працює, якщо в даних відсутні викиди. Візуальний аналіз діаграм розсіювання на рис. 2.17, 2.21 і 2.22 показує, що немає очевидних викидів.

2.3.2 Прогнозування ефективності фільтрації для попередньої обробки гіперспектральних зображень

Метою цього підрозділу є застосування розробленого раніше методу прогнозування до реальних даних і аналіз отриманих результатів з практичної точки зору. Для цього використовуємо групу каналів із двох гіперспектральних зображень сенсора Nuregon.

Слід розглядати випадки переважаючих сигнально-незалежних і сигнально-залежних компонентів, і інтенсивність шуму повинна змінюватися в широких межах. Тоді діаграма розсіювання виглядає так, як показано на рис. 2.23, а, і на неї можна підігнати криву та проаналізувати точність такої підгонки. Зрозуміло, що різні критерії якості можна передбачити з різною точністю. Вище було показано, що IPSNR зазвичай прогнозується точніше, ніж IPSNR-H.

Якщо є тільки один вхідний параметр, то доцільно застосовувати ймовірність $P_{0.5\sigma}$, що абсолютні значення коефіцієнтів ДКП не перевищують $0.5\hat{\sigma}(n, m)$ [234]. Діаграма розсіювання представлена на рис. 2.23, б. Аналіз як діаграм розсіювання, так і підігнаних кривих показує, що існують суворі монотонні залежності між вихідними та вхідними параметрами, які дозволяють робити висновки на основі прогнозованих даних. Припустимо, що IPSNR має бути більше 1 дБ, щоб вважати фільтрацію корисною. Тоді, грубо кажучи, можна не виконувати фільтрацію, якщо $P_{2\sigma}$ менше 0,6 або $P_{0.5\sigma}$ менше 0,2.

В обох випадках (рис. 2.23) в якості моделі прогнозування найбільш прийнятною була функція підгонки експоненти $y=ae^{bx}$. Зокрема, для підігнаної кривої на рис. 2.23, а $a=0,00797$ і $b=7,62$, тоді як для кривої на рис. 2.23, б параметри

становлять $a=0,11$ і $b=12,53$. Таким чином, створено метод прогнозування, і зараз стоїть завдання перевірити його на реальних даних.

Перш за все, необхідно мати деяку уяву про властивості шуму в гіперспектральних зображеннях. Для n -го каналу було використано модель (2.19).

На рис. 2.24 представлені оцінки цих параметрів, отримані методом [144] для наборів даних Hyperion (взято з <http://eros.usgs.gov/>). Є дві групи каналів з індексами 1...12, 62...77 і 231...240, які не використовуються в аналізі через дуже погану якість зображень у них. Через це оцінки параметрів шуму для них не представлені на графіках.

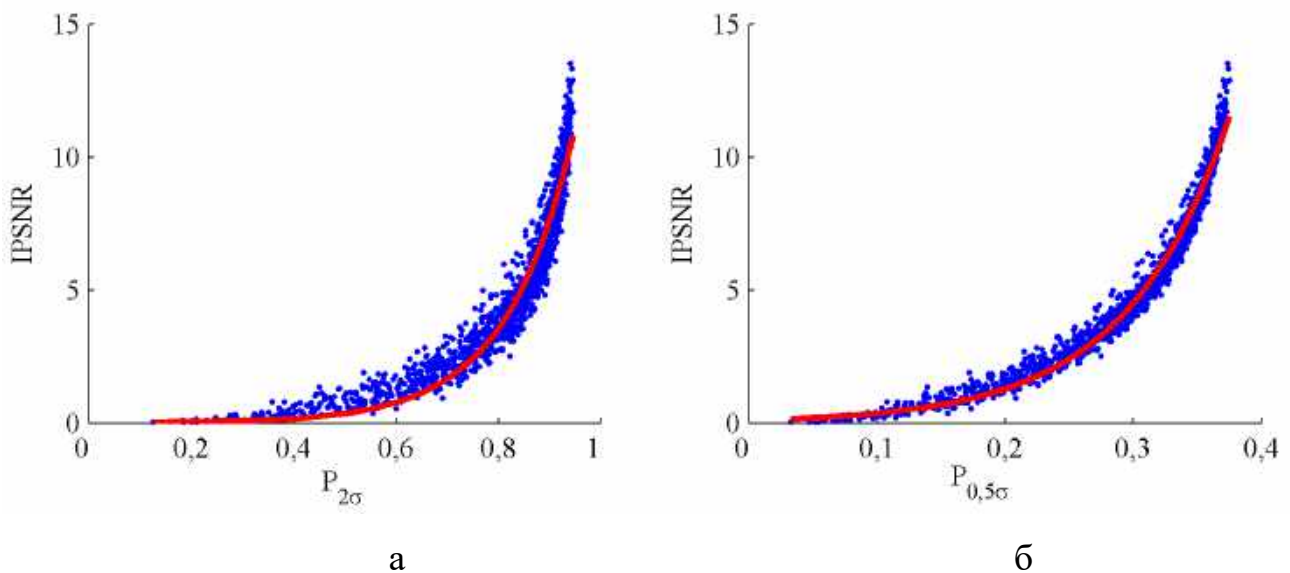


Рисунок 2.23 – Скатерограма IPSNR: від $P_{2\sigma}$ (а); від $P_{0.5\sigma}$ (б)

Оціночний аналіз показує наступне. По-перше, для всіх трьох наборів даних залежності оцінок від індексу каналу поведуться однаково. По-друге, є більш інтенсивний шум (більші оцінки $\sigma_0^2(n)$, так і $k(n)$) для зображень на краях діапазонів сенсора, наприклад, для піддіапазонів з індексами 13...15 та 59...61. Шум добре видно для візуалізованих зображень для цих піддіапазонів. При цьому в середині цього діапазону шум на один-два порядки менш інтенсивний. Насправді шум має таку інтенсивність (точніше, PSNR настільки ж високий), що шум практично не видно на візуалізованих зображеннях піддіапазону з індексами

близько 35...40. По-третє, можна оцінити внесок шумового компонента СЗШ шляхом обчислення еквівалентної дисперсії шуму (згідно 2.18) для нього як:

$$\hat{\sigma}_{eqSD}^2(n) = \sum_{i=1}^{I_{Im}} \sum_{j=1}^{J_{Im}} \frac{k(n)I_t(i,j,n)}{I_{Im}J_{Im}} \approx \sum_{i=1}^{I_{Im}} \sum_{j=1}^{J_{Im}} \frac{\hat{k}(n)I^n(i,j,n)}{I_{Im}J_{Im}} = \hat{k}(n)I_{mean}(n), \quad (2.21)$$

де $I_{mean}(n)$ визначає середнє значення зображення у n -му каналі. Тоді цю еквівалентну дисперсію можна порівняти з $\hat{\sigma}_0^2(n)$.

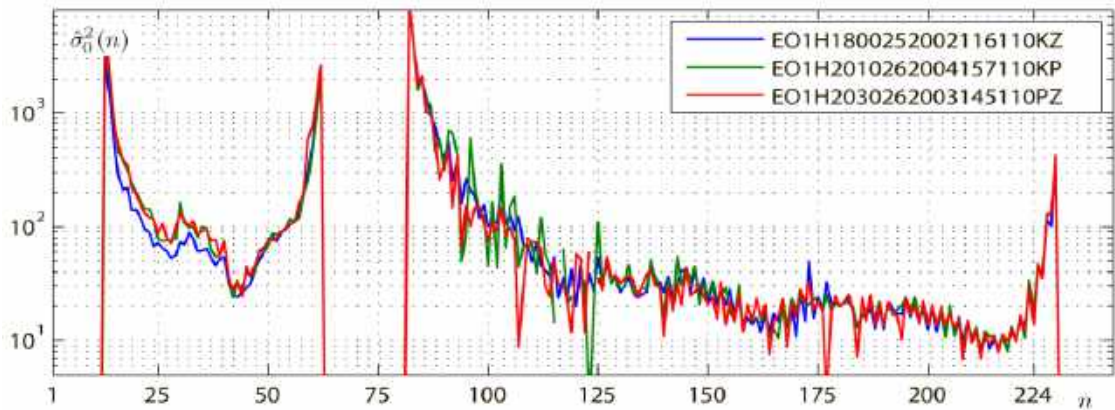
Такі порівняння були виконані в [156] і було показано, що для більшості каналів $\hat{\sigma}_{eqSD}^2(n)$ має такий самий порядок, як $\hat{\sigma}_0^2(n)$ або більше. Це означає, що при обробці та аналізі гіперспектральних даних слід враховувати СЗШ характер шуму. Винятком можуть бути декілька каналів (див. графіки на рис. 2.24), де оцінки $\hat{k}(n)$ негативні. Це переважно відбувається через недосконалість методу оцінювання та для каналів з дуже високим PSNR.

В аналізі було розглянуто канали з індексами 13...22. Як впливає з аналізу графіків на рис. 2.24, для них спостерігається тенденція до зниження інтенсивності шуму.

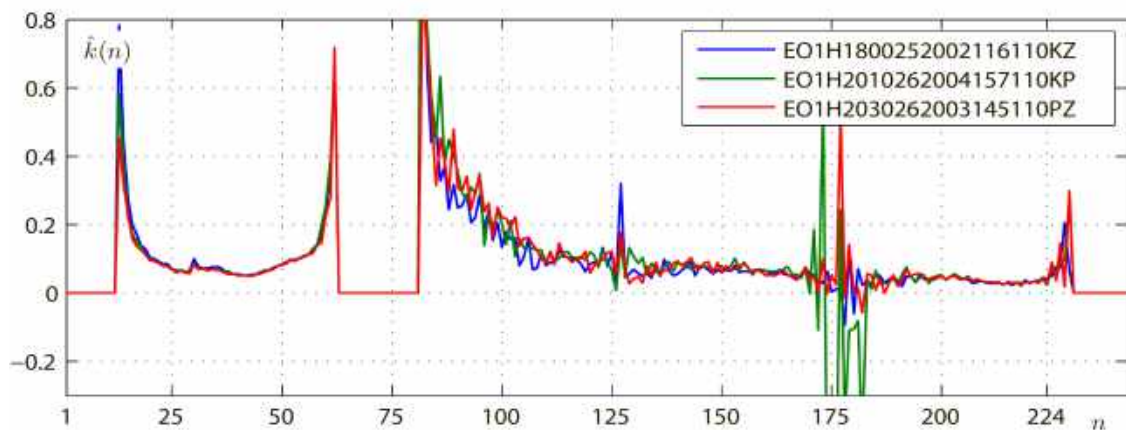
Ця тенденція підтверджується числовими даними, представленими в таблицях Ж.3 і Ж.4, Додаток Ж. При цьому динамічний діапазон для цих зображень залишається практично незмінним. Таким чином, спостерігається тенденція до збільшення вхідного PSNR. Через це найбільші значення $P_{2\sigma}$ спостерігаються для 13-го каналу (нагадаємо, що $P_{2\sigma}$ є великим, тобто наближається до максимально можливого значення 0,95, для зображень простої структури та шуму високої інтенсивності).

Відповідно до прогнозної залежності на рис. 2.23, а, IPSNR та IPSNR-H зростають, якщо $P_{2\sigma}$ більше. Через це максимальний очікуваний позитивний ефект від фільтрації прогнозується для 13-го каналу зображення. Цей ефект досить великий – прогнозований IPSNR (на основі $P_{2\sigma}$) досягає 5,6 дБ для першого набору даних і 9,06 дБ для другого набору даних. Таке покращення повинно бути очевидно, якщо візуально порівняти вхідні та вихідні зображення. Це можливо

зробити. В Додатку Ж, на рис. Ж.9 представлено вихідне (зашумлене) і вихідне (відфільтроване) зображення в 13-му каналі набору даних EO1H1800252002116110KZ. Хоча шум видно на вихідному зображенні (особливо в однорідних областях зображення), він ефективно придушується на вихідному зображенні разом із хорошим збереженням важливих деталей і меж.



а



б

Рисунок 2.24 – Оцінки $\hat{\sigma}_0^2(n)$ та $\hat{k}(n)$ для піддіапазонів трьох наборів даних Hyperion

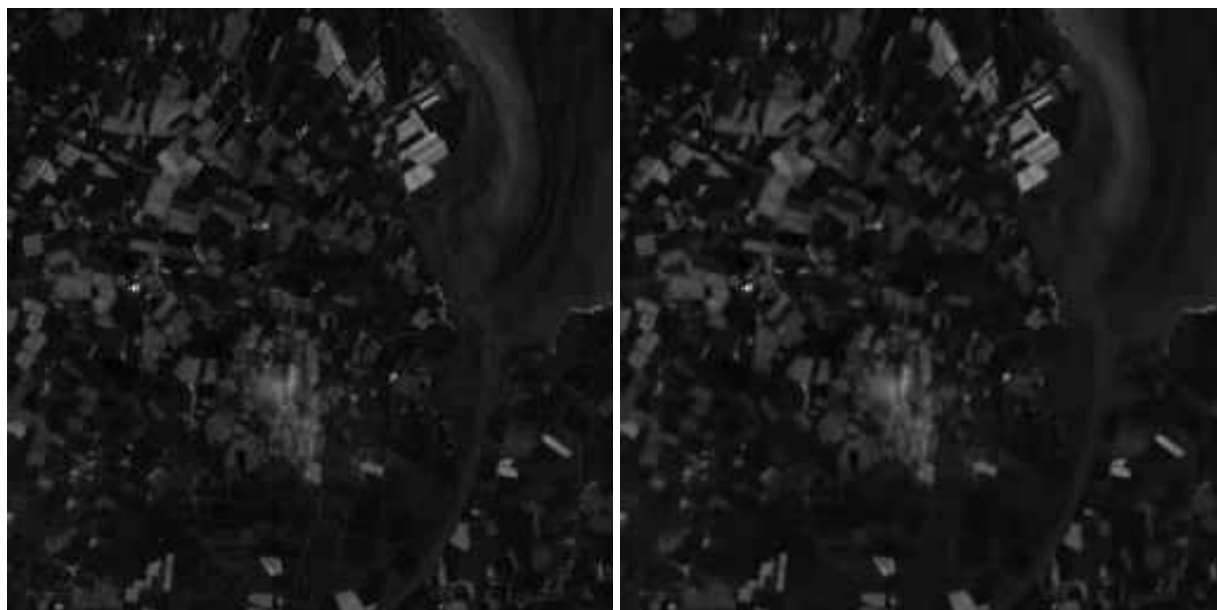
У той же час існує також багато каналів, для яких IPSNR, прогнозований на основі $P_{2\sigma}$, становить близько 1 дБ або менше. Водночас, прогнозовані значення IPSNR-Н (див. дані у крайніх правих колонках таблиць Ж.3 і Ж.4, Додаток Ж) навіть менше 1 дБ. Це означає, що покращення за розглянутими показниками завдяки фільтрації є незначними.

Це добре видно в Додатку Ж, на рис. Ж.10, де представлені вхідні та вихідні зображення для 22-го каналу. Шум на вихідному зображенні майже не помітний, а оброблене (відфільтроване) зображення виглядає практично так само, як і вхідне. Таким чином, немає необхідності фільтрувати це зображення.

Три інші спостереження, які впливають з аналізу даних у таблицях Ж.3 і Ж.4, Додаток Ж, є наступними. По-перше, усі три прогнози (на основі $P_{2\sigma}$, $P_{2.7\sigma}$ [10] і $P_{0.5\sigma}$) досить близькі. Це означає, що фактично всі три вхідні параметри можна використовувати на практиці. По-друге, прогнозований IPSNR-Н завжди менший, ніж IPSNR (обидва виражаються в дБ і, таким чином, можуть бути порівняні). Це є типовим для випадків придушення шуму, де IPSNR приблизно дорівнює 1...3 дБ не гарантує покращення якості зображення. По-третє, покращення залежить від складності зображення. Розглянутий фрагмент набору даних EO1H1800252002116110KZ (Додаток Ж, рис. Ж.9) є менш складним, ніж аналізований фрагмент іншого набору даних, оскільки в ньому присутні досить великі однорідні області зображення. Проаналізований фрагмент для набору даних EO1H2010262004157110KP представлений на рис. 2.25, де показані зображення в 13-му каналі.

Він має менші однорідні області та більше країв і деталей. Завдяки цьому шум менш помітний і його можна помітити, в основному, в частинах зображення зі світлішими кольорами (вища середня інтенсивність).

Відфільтроване зображення EO1H2010262004157110KP показано на рис. 2.25, б, і видно, що шум придушується, а корисна інформація зберігається.



а

б

Рисунок 2.25 – Вхідні (а) і вихідні (б) зображення EO1H2010262004157110KP в 13-му каналі даних Hyperion

Також проведено візуальний аналіз вхідних і вихідних зображень для інших каналів. Починаючи з n близько 18 стає важко знайти відмінності. Це означає, що, з одного боку, фільтрація не вносить спотворень, і це є цінною перевагою фільтра на основі DCT (наприклад, стандартний медіанний фільтр у таких випадках вносить видимі спотворення). З іншого боку, це означає, що немає необхідності застосовувати фільтрацію, і її можна пропустити, зберігши ресурси системи обробки даних ДЗ. Це може прискорити обробку гіперспектральних даних, де, фактично, можна пропустити фільтрацію для приблизно 70...80% каналів. Це може бути особливо важливо для бортової обробки даних.

Припустимо, що правило для уникнення фільтрації таке: пропустити фільтрацію даного компонентного зображення, якщо $P_{2\sigma} < 0,6$ (див. рис. 2.23, а). Тоді немає необхідності виконувати фільтрацію зображення для каналів з $n > 18$ підмножини EO1H2010262004157110KP і $n > 20$ для підмножини EO1H1800252002116110KZ (див. дані в таблицях).

Висновки до розділу 2

Цей розділ присвячено питанням підвищення ефективності інтелектуальної фільтрації зображень дистанційного зондування Землі та медичної діагностики в умовах присутності складних завад. У ході дослідження було показано, що найкращий результат для класифікації текстур на зображеннях з шумом дає одночасне використання локальних параметрів відносної локальної дисперсії, квазірозмаху та параметра на основі косинусного дискретного перетворення. Трохи кращі результати отримані при розмірі вікна, в якому обчислювалися параметри, що дорівнює 5×5 . Наступним етапом дослідження може стати удосконалення алгоритму навчання класифікатора або використання інших потужних класифікаторів (таких як Decision Trees або AdaBoost). Крім цього, можна дослідити агрегування вхідних параметрів класифікатора в ковзному вікні, що, ймовірно, може додатково покращити якість класифікації.

Порівнянні результати демонструє нейромережа на основі трьох параметрів: квазірозмаху, детектора Харріса та детектора на основі ДКП. Крім того, нейронна мережа, навчена лише за двома параметрами (детектором Харріса і простим детектором текстур) не надто відстає за ефективністю від вищезгаданих, проте вимагає менше обчислювальної потужності. Тому, коли першорядне значення має простота обробки, слід застосовувати саме цей підхід.

Виявлено, що при великому характерному розмірі текстур (великозернистих текстур) ефективність детекторів дещо падає. виправити цей недолік можна за рахунок навчання на кількох варіантах текстури, а також збільшенні розміру ковзного вікна до 7×7 пікселів.

Широкий набір методів придушення спеклу перевірено для кількох тестових зображень, спотворених просторово-корельованим спеклом. Показано, що основною проблемою є забезпечення якісного збереження меж та деталей. Щоб подолати її, запропоновано застосувати метод Level Set для виявлення неоднорідностей на РСА-зображеннях і, крім того, для детектування фрагментів зображення, для яких збереження деталей є першочерговим завданням. Тоді фільтр

на основі ДКП із частотно-залежним порогом забезпечує ефективне придушення спеклу в областях зображення, визнаних однорідними. А для кращого збереження деталей запропоновані декілька підходів, які також протестовані для реальних однопоглядових зображень РСА. Серед найкращих методів досягнення цієї мети - вейвлет-фільтр Хаара, адаптований до мультиплікативного шуму, та інтегро-диференціальний метод. Про це свідчать як кількісні дані, так і деякі наочні приклади, що представлені в роботі.

Запропоновано модифікований метод локально-адаптивної фільтрації однопоглядових РСА-зображень, що ґрунтується на ДКП-фільтрації з частотно-залежним порогом і використовує НМ детектор меж для детектування ділянок локальної активності зображення. Показано, що запропонований метод перевершує багато існуючих алгоритмів фільтрації за кількома показниками якості (MSEout, MSEhet, MSSIM). Наведено результати застосування запропонованого методу до зображень зі спекл-шумом, ідентичним за характеристиками реальних даних.

Показано, що ефективність усунення шумів при фільтрації на основі ДКП зображень, спотворених сигнально-залежним шумом, можна прогнозувати за умови, що характеристики сигнально-залежного шуму апіорі відомі або попередньо оцінені з високою точністю. Можна прогнозувати такі параметри, як покращення PSNR або покращення PSNR-HVS-M після фільтрації. Таке прогнозування є швидким і на його основі можна прийняти надійні рішення, чи варто проводити придушення спеклу ще перед початком фільтрації зображення, що дозволяє зберегти час та ресурси системи обробки даних.

Доведено, що підхід до прогнозування є досить загальним у тому сенсі, що він розроблений для сигнально-залежного шуму, який містить сигнально-незалежні та сигнально-залежні компоненти, тож і для мультиплікативного шуму. Крім того, продемонстровано та проаналізовано декілька способів покращення точності прогнозування. Один із них полягає у використанні статистичного параметра $P_{0.5\sigma}$ замість раніше емпірично запропонованих параметрів. Інші способи покращення прогнозу стосуються стратегій формування діаграми розсіювання та кращої підгонки кривої.

Описано та перевірено метод прогнозування ефективності придушення шуму для реальних багатоканальних зображень, спотворених сигнально-залежним шумом. Показано, що показники IPSNR та IPSNR-H можна передбачити досить добре. На підставі такого прогнозу можна прийняти рішення щодо доцільності застосування фільтрації до компонентних зображень. Показано, що в гіперспектральних даних є досить багато каналів, для яких доцільно використовувати фільтрацію, оскільки завдяки цьому можна забезпечити значне покращення якості зображення.

Основні наукові результати, наведені у другому розділі, опубліковано в роботах автора [1], [2], [5], [6], [7], [9], [10], [11], [13], [14], [15], [16], [17], [19], зазначених в анотації.

РОЗДІЛ 3 РОЗРОБКА ТА ВДОСКОНАЛЕННЯ АЛГОРИТМІВ СТИСНЕННЯ НЕЗАШУМЛЕНИХ ЗОБРАЖЕНЬ

В даному розділі розглянуті питання, що стосуються двоетапних процедур забезпечення бажаної якості зображень. Крім того, розглядаються питання стиснення зображень із забезпеченням візуально непомітних спотворень. Окрему увагу приділено адаптивному стисненню багатоканальних зображень дистанційного зондування Землі. Також розглянуто питання дивних зображень дистанційного зондування і їхніх властивостей. Основні результати автора, присвячені даним аспектам, відображено в публікаціях [34], [35], [40], [42], [44], [45], [46], [47], [48], [49], [50], [63], [65], [71].

3.1 Двоетапні процедури забезпечення необхідної якості

Щодня створюється величезна кількість зображень і відео, які потрібно зберігати, обробляти та/або передавати через канали зв'язку [238], [239], [240]. Завдяки цьому стиснення зображень і відео стало стандартною операцією в ланцюжку обробки зображень [87], [240], [241]. Існують методи стиснення без втрат і стиснення з втратами, де останні стають все більш корисними, оскільки коефіцієнт стиснення (КС), що забезпечується стисненням без втрат, часто не відповідає вимогам [87], [242]. Тим часом, досі існує багато невирішених питань щодо стиснення зображень із втратами, включаючи те, який кодер використовувати (обрати), як забезпечити бажаний КС (для деяких кодерів [87], [168]), як забезпечити швидке стиснення та декомпресію [243] тощо. Але одним із найважливіших завдань є забезпечення бажаної якості стиснутих зображень, оскільки в іншому випадку стиснення з втратами може втратити сенс [87], [168], [241], [242], [243]. Зауважте, що стиснення зображень із втратами — це область, де метрики якості з еталоном (МЯЕ), тобто метрики, розраховані для пари стиснутих і оригінальних зображень, в основному використовуються для контролю внесених втрат і для характеристики якості стисненого зображення. Часто можливо, що

потрібно стиснути візуально без втрат [182], [244]. Тоді поріг невидимості спотворень повинен бути відомий заздалегідь. Завдання, яке необхідно вирішити на етапі стиснення, полягає в тому, щоб отримати значення використовуваної метрики не гірше цього порогу або максимально близько до нього [245].

Як бачимо, необхідно вирішити декілька завдань. По-перше, слід вибрати хорошу метрику. По-друге, необхідно знати вищезазначені порогові значення. По-третє, має бути доступна процедура, здатна забезпечити ці порогові значення, і вона має бути достатньо швидкою та точною. На щастя, перші дві задачі останнім часом частково вирішено [18]. Щодо третього завдання існують два основні підходи [34], [37], [168], [194], [244]. Відповідно до одного з них [168], [244] пропонується застосовувати ітераційну процедуру. Виконується багаторазове стиснення/декомпресія, де кожен раз використана МЯЕ обчислюється та порівнюється з бажаним значенням (порогом), потім ПКС змінюється, щоб наблизитися до бажаного значення, і стиснення/декомпресія виконується знову [168]. За допомогою певних засобів кількість кроків можна зменшити у два-три рази [168], але це все одно кількість ітерацій залишається апріорі невідомою.

Метою цього підрозділу є розробка та аналіз гібридного рішення, яке називаємо двоетапним підходом. Він включає в себе попереднє стиснення/декомпресію з обчисленням використаної МЯЕ. Потім, використовуючи отриману раніше усереднену КШС, визначається, як необхідно змінити ПКС. Після цього виконується остаточне стиснення. Розглянуто та обговорено особливості запропонованої процедури, її переваги та недоліки.

Для початку порівняємо кодери AGU та JPEG2000. У таблиці 3.1 наведено значення PSNR для AGU та JPEG2000 для п'яти стандартних тестових зображень і чотирьох значень КС.

Як можна побачити, AGU забезпечує трохи кращі значення PSNR порівняно з JPEG2000 для всіх тестових зображень для всіх розглянутих КС. Паралельно дані в таблиці 3.1 чітко показують одну проблему при стисненні зображень із втратами – для даного КС якість стиснених зображень значною мірою залежить від властивостей зображення (вмісту, складності).

Таблиця 3.1 – Значення PSNR (дБ) для різних тестових зображень і КС

Тестове зображення	КС=8		КС=16		КС=32		КС=64	
	JPEG2000	AGU	JPEG2000	AGU	JPEG2000	AGU	JPEG2000	AGU
Lenna	40,33	40,52	37,27	37,46	34,15	4,51	31,02	31,50
Barbara	38,07	39,26	32,87	34,65	28,89	30,77	25,87	27,55
Baboon	29,11	29,70	25,57	26,12	23,18	23,69	21,68	22,01
Goldhill	36,54	37,03	33,24	33,65	30,53	31,09	28,49	28,97
Peppers	38,17	38,33	35,80	35,55	33,54	33,32	30,79	30,90

Наприклад, для КС=8 PSNR змінюється від 38...41 дБ (невидимі спотворення) для зображень досить простої структури (Lenna, Peppers) до 29...30 дБ (видимі спотворення) для тестового зображення Baboon, яке є високотекстурним. Так само, для КС=32, PSNR змінюється від $\approx 34,5$ дБ (майже непомітні спотворення) для Lenna до ≈ 23 дБ (дратівливі спотворення) для Baboon. Насправді межі цих варіацій можуть бути ще ширшими. Тож зрозуміло, що КС або ПКС слід налаштовувати індивідуально (адаптивним способом) для кожного зображення, яке має бути стиснутим з бажаною якістю. Крім того, з аналізу даних табл. 3.1 можна припустити, що КШС є монотонними. Продемонструємо, що це так у переважній більшості практичних випадків. На рис. 3.1 (взято з [168]) представлені залежності метрики візуальної якості PSNR-HVS-M від КС для тестового зображення Pole, яке має досить просту структуру. Залежності представлені для дев'яти різних кодерів, включаючи добре відомі JPEG (версія з нерівномірним квантуванням ДКП-коефіцієнтів) і JPEG2000 (версія без оптимізації візуальної якості), SPIHT, який є вільно доступною модифікацією JPEG2000, кодери на основі ДКП AGU, AGU-M, ADCT, ADCT-M, а також сучасні методи стиснення зображень WebP і H.265 (його версія для фотографій).

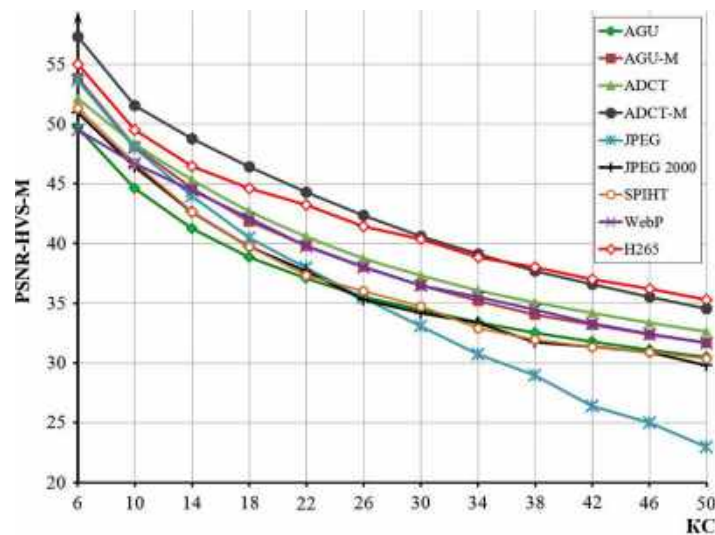


Рисунок 3.1 – Залежності PSNR-HVS-M від КС для тестового зображення Pole, стиснутого 9 різними методами стиснення

Аналіз залежностей на рис. 3.1 дозволяє зробити декілька висновків. По-перше, видно, що всі залежності монотонно спадають, і причини цього зрозумілі – більший КС пов'язаний з більшим рівнем спотворень. По-друге, модифікації кодера, призначені для забезпечення високої візуальної якості (ADCT-M, H.265), дійсно працюють найкращим чином. JPEG добре працює для малих значень КС, але його ефективність знижується найшвидше, якщо КС збільшується. По-третє, усі залежності є «гладкими» функціями, і ця властивість вже використовувалася на завершальних етапах ітераційної процедури [168], де лінійна інтерполяція останніх результатів використовується для розрахунку остаточного ПКС. Залежності на рис. 3.1. є функціями КС (такий спосіб представлення пояснюється необхідністю порівнювати дані для різних кодерів, для яких ПКС є різними; зокрема, для JPEG2000 як ПКС використовується bpr). Для кодерів на основі ДКП параметри стиснення змінюються кроком квантування (КК) або коефіцієнтом масштабування (КМ) [168], [175], [182]. Отже, цікаво побачити та проаналізувати КШС для цих аргументів (КК або КМ). Приклад таких залежностей наведено на рис. 3.2. для дев'яти стандартних тестових зображень різної складності в градаціях сірого.

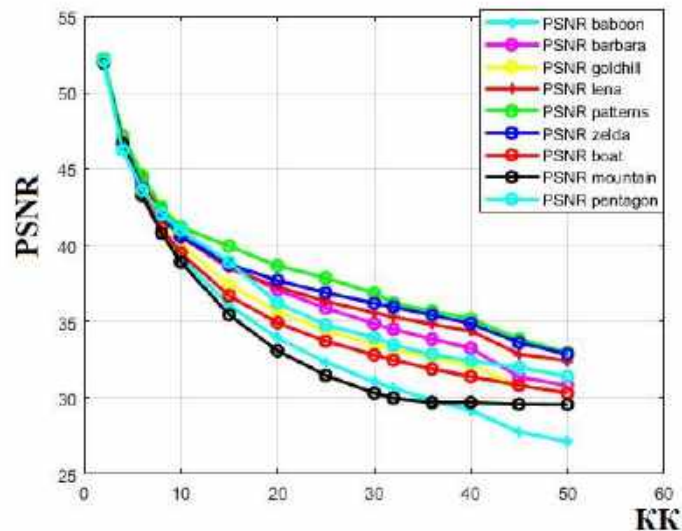


Рисунок 3.2 – Залежності PSNR від КК для дев'яти тестових зображень, стиснених AGU

Тестові зображення Lena та Zelda є «найпростішими», а тестові зображення Baboon і Mountain є найскладнішими. Тестове зображення Patterns є штучним та має специфічну структуру (див. <http://ponomarenko.info/agu.htm>). Аналіз отриманих залежностей показує наступне. По-перше, криві розташовані дуже компактним чином для малих КК (скажімо, $KK < 10$). Однак вони розходяться для більших КК і, як можна очікувати, поведінка кривої залежить від складності зображення, що стискається. Це призводить до того, що для забезпечення заданого PSNR для різних зображень потрібні досить різні значення КК. Наприклад, щоб забезпечити $PSNR \approx 35$ дБ (поріг невидимості спотворень [245]), потрібно встановити КК близько 14 для тестового зображення Mountain і близько 40 для тестових зображень Patterns і Zelda. Таким чином, знову приходимо до висновку, що ПКС (КК в розглянутому випадку) потрібно встановлювати індивідуально для кожного зображення, яке необхідно стиснути. По-друге, майже всі криві йдуть «майже паралельно» для КК в межах від 15 до 50, тобто в основній області стиснення. Така властивість, як правило, буде використовуватися нами в запропонованій двоетапній процедурі.

Подивіться на рис. 3.3. На ньому показана залежність для конкретного тестового зображення (в розглянутому випадку – Boat). Звичайно, вона

відрізняється від усередненої залежності, але ці залежності дуже схожі і для малих діапазонів варіації КК вони йдуть «майже паралельно». Ці спостереження дозволяють запропонувати наступну процедуру. Першим кроком є стиснення та розпакування зображення за допомогою початкового ПКС ($КК_1$ для AGU та ADCT). Цей початковий $ПКС_{init}$ можна визначити з усередненої залежності використаної метрики від використаного ПКС (як PSNR від КК, представленою на рис. 3.3) для даного кодера та для M_{des} .

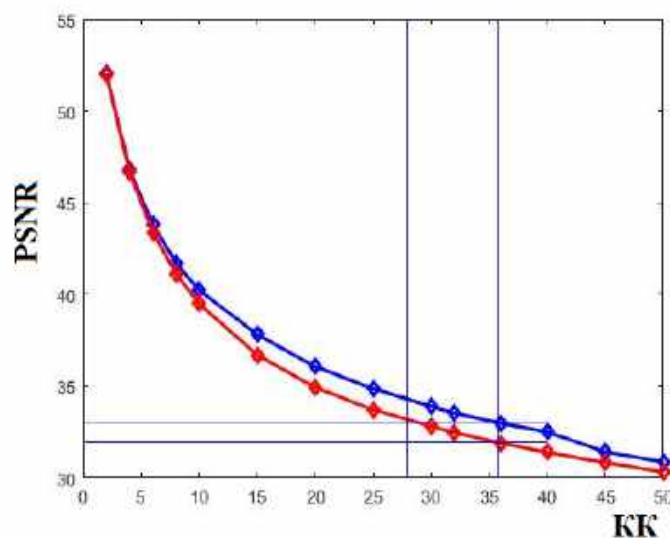


Рисунок 3.3 – Усереднена КШС (PSNR від КК, синій колір) і КШС для тестового зображення Boat (позначено червоним кольором)

Такі усереднені залежності (див. приклади в [168]) можна отримати заздалегідь в автономному режимі для великого набору тестових зображень. Крім того, слід визначити похідну метрики $dM/dПКС=M'$ для початкового значення ПКС. Потім обчислюється M_{init} , створений стисненням за допомогою $ПКС_{init}$. Звичайно, для конкретного зображення, яке стискається за допомогою M_{des} , значення M_{des} і M_{init} відрізняються. Тоді, припускаючи лінійну залежність M від ПКС для розглянутої частини залежності, можна отримати оцінку $ПКС_{des}$ як:

$$ПКС_{des} = ПКС_{init} + (M_{des} - M_{init}) / M' \quad (3.1)$$

Іншими словами, на другому кроці PKC_{des} оцінюється за допомогою лінійної апроксимації усередненого КШС. Після цього здійснюється стиснення з використанням PKC_{des} . Приклад показано на рис. 3.3. Припустимо, що потрібно забезпечити $PSNR_{des}=33$ дБ. Тоді, відповідно до усередненої кривої (синій), KK_{init} потрібно встановити рівним 35,74. Але для тестового зображення Boar це призводить до $PSNR_{init}$ близько 31,92 дБ, тобто менше, ніж бажано. Потім KK коригується та встановлюється рівним 27,95. Після стиснення цього зображення за допомогою цього KK $PSNR$ виявляється рівним 33,17 дБ, тобто значно ближче до 33 дБ, ніж спочатку і залишкова похибка є задовільно малою.

Наведений вище приклад стосується конкретного тестового зображення та конкретного $PSNR_{des}$. Тому потрібно провести статистичний експеримент, щоб проаналізувати, який позитивний (або, можливо, негативний) ефект спостерігається від другого кроку. Такі експерименти були проведені, і нижче наведено деякі результати. Щоб охарактеризувати точність, визначимо значення дисперсії (тут і далі по ансамблю тестових зображень) для першого та другого етапів для наданих метрик (у цьому випадку $PSNR_{init}$ та $PSNR_{prov}$). В Додатку К, у таблиці К.1 також представлені значення ΔKK і рекомендовані KK_{rec} , які використовуються на другому кроці. Як можна побачити, $PSNR_{init}$ після першого кроку змінюється в досить широких межах навколо бажаного $PSNR_{des}=34$ дБ. Для зображень простих структур вони більші за $PSNR_{des}$, тоді як для зображень складних структур ситуація протилежна.

Завдяки цьому ΔKK може бути як позитивним, так і негативним. У результаті KK_{rec} змінюється в широких межах від 14,73 для тестового зображення Baboon до понад 40 для тестових зображень Zelda та Patterns. Завдяки другому кроку $PSNR_{prov}$ стає ближчим до $PSNR_{des}$, ніж $PSNR_{init}$ для восьми з дев'яти тестових зображень. Як наслідок, дисперсія після двоетапного стиснення є значно меншою, ніж після одного кроку. Однак є одна проблема. Навіть після двох кроків дисперсія є досить великою, і похибки в деяких випадках перевищують 1 дБ, що є небажаним.

Це трапляється для зображень складної структури, для яких $PSNR_{prov}$ виявляється значно більшим за $PSNR_{des}$. Причини цього наступні. По-перше,

реальні значення похідних для відповідних КШС можуть істотно відрізнятися від M' для усередненої кривої. По-друге, лінійна апроксимація інколи може бути надто грубою і вона не працює належним чином (ΔKK для зображень складної структури приблизно вдвічі менше, ніж KK_{init}). Тепер проаналізуємо дані для $PSNR_{des}=37$ дБ. Вони наведені в Додатку К, у таблиці К.2. Видно, що завдяки другому кроку результати покращилися для всіх дев'яти тестових зображень – $PSNR_{prov}$ ближчі до $PSNR_{des}$, ніж $PSNR_{init}$. Дисперсія також значно зменшилася. Тільки для двох зображень складної структури похибки забезпечення бажаного $PSNR$ перевищують 1 дБ. Проте позитивним моментом є те, що $PSNR_{prov}$ перевищує $PSNR_{des}$ практично для всіх тестових зображень. Це означає, що завдання забезпечити якість зображення не гірше бажаного, можна вважати виконаним. Нарешті, в Додатку К, в таблиці К.3 представлені отримані дані для $PSNR_{des}=40$ дБ.

Завдяки другому етапу результати стали кращими для всіх дев'яти зображень, а похибки надання необхідного $PSNR$ є прийнятними для всіх тестових зображень, за винятком Baboon та Mountain, для яких надана якість значно вища за бажану. Отже, поставлена задача досягнута для випадку забезпечення $PSNR_{des}=40$ дБ.

Підсумовуючи представлені результати, можна констатувати наступне. Запропонована двоетапна процедура працює досить добре, якщо потрібно забезпечити $PSNR$ більше 37 дБ, тобто якщо введені спотворення повинні бути невидимими. Зауважимо, що подібні результати та практично такі ж висновки отримано для метрики $PSNR$ для кодера ADCT [168]. При цьому похибки забезпечення бажаного $PSNR$ виявляються занадто великими для $PSNR_{des}$ менше 35...37 дБ. Причини цього були пояснені вище.

Зосередимося тепер на розгляді трьох популярних метрик оцінки візуальної якості, а саме $PSNR$, $PSNR-HVS$ та $PSNR-HVS-M$. Щоб більш повно оцінити ефект двоетапного підходу, використали всі три метрики одночасно та порівняли їх (нагадаємо, що всі вони виражаються в дБ і обчислюються, у певному сенсі, аналогічно).

Щоб проілюструвати залежність показників якості від КК, було використано зображення Goldhill як приклад. Результати показані в Додатку К, на рис. К.1.

Як відомо [94], PSNR і PSNR-HVS рівні, якщо спотворення подібні до АБГШ. Аналіз кривих PSNR(KK) і PSNR-HVS(KK) показує, що для малих КК (менше 8) спотворення дійсно схожі на АБГШ. Однак для більшого КК спотворення на низьких просторових частотах є більшими, і це несприятливий факт. Щодо PSNR і PSNR-HVS, PSNR-HVS-M, то PSNR-HVS-M більше, якщо є ефект маскування. Це має місце при малих і середніх значеннях КК. Але для $KK > 40$ ефект маскування зникає і спотворення стають більш «низькочастотними». Також відомо, що внесені спотворення практично непомітні, якщо PSNR становить близько 35 дБ або більше, а PSNR-HVS-M становить близько 40 дБ або більше. Як можна побачити, це відбувається для зображення Goldhill, якщо КК становить приблизно 18.

Щоб проілюструвати роль та особливості метрик у візуальній оцінці якості стисненого зображення, було проведено порівняльний експеримент. Було отримано три зображення для тестового зображення Goldhill для трьох значень КК, а саме $KK=30, 35$ і 40 . Стиснені зображення показані на рис. 3.4, а також наведено значення метрик. Порівняння двох зображень нижче (для $KK=30$ і 35) показує, що значення різниці для відповідних показників становлять близько 1 дБ (різниця найменша для PSNR і найбільша для PSNR-HVS-M). Для цього випадку можна помітити невеликі відмінності в стиснених зображеннях, зокрема, для фрагментів, позначених зеленими прямокутниками.

Порівнюючи зображення для $KK=30$ і $KK=40$ (рис. 3.4, а і рис. 3.4, в), спостерігаємо більші відмінності в значеннях показників (вони складають близько 1,36 дБ для PSNR, близько 1,86 дБ для PSNR-HVS і близько 2,45 дБ для PSNR-HVS-M).

Можна помітити наявні відмінності. Тоді можна зробити висновок, що якщо PSNR (або PSNR-HVS або PSNR-HVS-M) змінюється більше, ніж на 0,5...1 дБ, можна помітити зміни у візуальній якості. Таким чином, ці відмінності можна використовувати як поріг точності. Коли похибка забезпечення бажаної якості менше або приблизно дорівнює 0,5 дБ, різницю помітити практично неможливо.

Коли похибка приблизно дорівнює 1 дБ, різницю легше побачити. Таким чином, повинні забезпечити бажану якість з похибками менше 1 дБ.



а

б

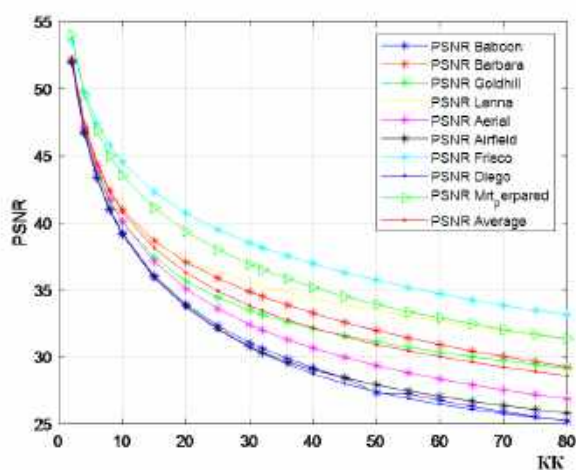


в

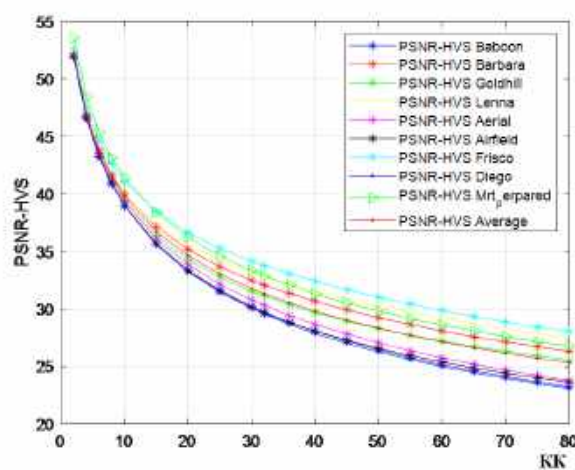
Рисунок 3.4 – Порівняння зображень для КК=30 (а), КК=35 (б) і КК=40 (в)

Щоб перевірити універсальність двоетапного методу, описаного вище, деякі зображення були змінені. Фактично, деякі з тестових зображень більш типові для дистанційного зондування та медичних зображень. При отриманні середньої кривої спотворення необхідно спочатку встановити КК на ряд значень для кожного зображення та отримати відповідне значення метрики візуальної якості. Потім усереднюють дані для кожної точки КК і, таким чином, отримують середню криву

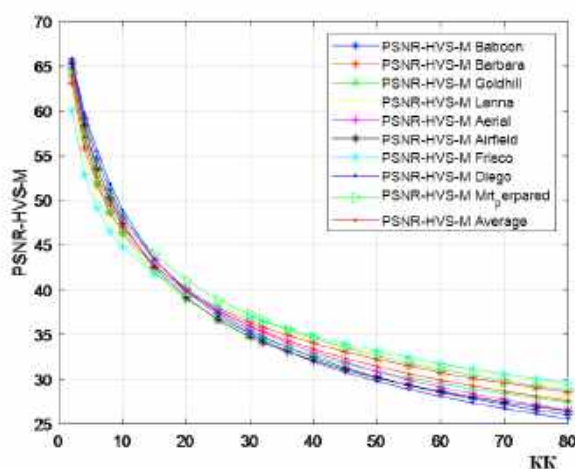
спотворення. Таким чином було отримано середні криві для трьох показників візуальної якості. Результати показані на рис. 3.5. Можна виявити, що незалежно від PSNR, PSNR-HVS або PSNR-HVS-M тенденція зміни кривої спотворень дуже схожа. На основі цієї характеристики метод, запропонований у статті [246], спочатку використовує середню криву спотворень (отриману заздалегідь в автономному режимі), щоб отримати початкове значення ПКС і виконати перше стиснення. Значення метрики, отримане в результаті першого стиснення, буде близьким до бажаного значення, але похибка невизначена (різні зображення мають різні похибки). Звичайно, це також можна спостерігати з рис. 3.5. Потім виконується друге стиснення.



а



б



в

Рисунок 3.5 – Залежність метрики від КК: PSNR (а), PSNR-HVS (б) і PSNR-HVS-M (в)

Корекція параметра для другого кроку стиснення пов'язана з помилкою, отриманою в результаті першого стиснення. Для корекції параметрів можна використовувати формулу (3.1). Методи розрахунку для двох інших показників візуальної якості однакові та виконуються за допомогою наступних формул:

$$KK_{des} = KK_{init} + (PSNR-HVS_{des} - PSNR-HVS_{init}) / M', \quad (3.2)$$

$$KK_{des} = KK_{init} + (PSNR-HVS-M_{des} - PSNR-HVS-M_{init}) / M', \quad (3.3)$$

де KK_{init} – початкове значення KK (приблизне значення з середньої кривої спотворення); $PSNR-HVS_{init}$ та $PSNR-HVS-M_{init}$ – значення метрик, отримане після першого етапу стиснення; $PSNR-HVS_{des}$ та $PSNR-HVS-M_{des}$ – бажане значення метрик; M' – похідна середньої кривої спотворення, що відповідає KK_{init} .

Таким чином, ідея полягає в тому, що після лінійної корекції, значення KK_{des} , ближче до бажаного значення, можна отримати для кожного зображення. Виправляючи KK один раз, можна досягти значного підвищення точності на другому етапі стиснення.

Щоб краще зрозуміти, як працює цей метод, у експериментах було встановлено бажані значення показників рівними 30, 35 і 40 дБ. Вони були проведені для МЯЕ PSNR, PSNR-HVS і PSNR-HVS-M. Детальні дані для одного з набору тестів наведено нижче у таблицях 3.2 та 3.3.

Розраховано дисперсії після першого та другого стиснення відповідно. З таблиці 3.2 видно, що при використанні двоетапного підходу дисперсія суттєво зменшилась, а дані більш стабільні; похибка між бажаною візуальною якістю ($PSNR-HVS_{des}$) і забезпеченою візуальною якістю ($PSNR-HVS_{prov}$) не перевищує 0,605 дБ, що є достатньо хорошою точністю. Порівняно з аналізом результатів PSNR у роботі [246] для HVS-метрики забезпечена точність практично не залежить від складності стиснутого зображення. Для PSNR і двох метрик на основі HVS для трьох бажаних рівнів візуальної якості було проведено дев'ять наборів експериментів, а статистичну інформацію експериментальних даних показано в таблиці 3.3.

Тут M_{des} — бажане значення розглянутої метрики візуальної якості, VAR_1 — дисперсія метрики візуальної якості для дев'яти тестових зображень, отриманих після першого етапу стиснення, VAR_2 — дисперсія метрики візуальної якості, отримана після другого (коригувального) кроку стиснення, $MAX_{\Delta final}$ є максимальною похибкою між M_{des} і наданими значеннями M_{pro} .

Таблиця 3.2 – Статистика та параметри забезпечення PSNR-HVS_{des}=30 дБ

Тестове зображення	KK_{init}	PSNR-HVS _{init}	ΔKK	KK_{rec}	PSNR-HVS _{prov}
Goldhill	38,9283	29,8194	-1,0614	37,8669	30,0071
Baboon	38,9283	28,2555	-9,7143	29,214	30,4306
Barbara	38,9283	30,8412	4,6843	43,6126	30,1307
Lenna	38,9283	31,7084	9,5133	48,4416	30,4254
Aerial	38,9283	28,8464	-6,4239	32,5044	30,1949
Airfield	38,9283	28,2629	-9,6731	29,2552	30,3089
Frisco	38,9283	32,6046	14,5038	53,4321	30,605
Diego	38,9283	28,0752	-10,718	28,21	30,5263
Mrt_prepared	38,9283	31,5009	8,3578	47,2861	30,1604
Дисперсія	2,9739		0,0401		

Таблиця 3.3 – Статистичні дані, отримані для дев'яти тестових зображень

Метрика	M_{des}	VAR_1	VAR_2	$MAX_{\Delta final}$
PSNR	40	4,4876	0,4714	1,2125
PSNR	35	7,1179	1,6761	3,6067
PSNR	30	9,1183	9,0517	5,2102
PSNR-HVS	40	0,7906	0,0408	0,3583
PSNR-HVS	35	1,6575	0,0369	0,6017
PSNR-HVS	30	2,9739	0,0401	0,605
PSNR-HVS-M	40	0,3675	0,0108	0,257
PSNR-HVS-M	35	1,0119	0,1041	0,720
PSNR-HVS-M	30	2,0312	0,0280	0,465

З даних у таблиці 3.3 можна побачити, що двоетапний підхід загалом покращив точність надання значення M_{des} для показників PSNR-HVS та PSNR-HVS-M. Після двох кроків дисперсія зменшилася приблизно до 1/10-1/70 однокрокової дисперсії, що вказує на те, що корекція КК зіграла позитивну роль. Максимальна дисперсія VAR_2 PSNR-HVS-M становить 0,1, що вказує на те, що залишкові похибки невеликі, а результати відносно стабільні. При цьому максимальна похибка становить 0,605 дБ (для PSNR-HVS) і 0,720 дБ (для PSNR-HVS-M). Згадуючи попередні результати, можна зробити висновок, що такі помилки є прийнятними. Можна також вважати, що запропонований метод здатний забезпечити якість зображення, близьку до бажаного значення, практично без помилок (різниць) для показників візуальної якості PSNR-HVS та PSNR-HVS-M.

Вище розглянуті кодери на основі ДКП. Втім, виникає питання, чи можна застосовувати двоетапний підхід для кодерів на основі малохвильових перетворень. В якості прикладу розглянемо SPIHT, характеристики якого близькі до характеристик JPEG2000. Зазначимо, що і для цих кодерів задача забезпечення бажаної якості є актуальною. Нижче розглядаємо МЯЕ PSNR та PSNR-HVS-M.

У таблиці 3.4 представлено значення PSNR і PSNR-HVS-M для SPIHT для п'яти типових тестових зображень і чотирьох значень КС. Дані в таблиці 3.4 чітко демонструють одну проблему при стисненні SPIHT – для даного КС якість стислих зображень дуже значною мірою залежить від властивостей зображення (вмісту, складності). Наприклад, для КС=8 PSNR змінюється від 40...44 дБ (невидимі спотворення) для зображень досить простої структури (Lenna, Mrt_prepared) до 30 дБ (чітко видимі спотворення) для тестового зображення Baboon, яке має текстурний характер. Подібним чином, для КС=16 значення PSNR змінюються від ≈ 38 дБ (невидимі спотворення) для Mrt_prepared до ≈ 26 дБ (дратівливі спотворення) для Baboon. Насправді межі цих варіацій можуть бути ще ширшими.

На рис. 3.6. показано різницю у візуальній якості зображень різної складності, коли КС дорівнює 16.

Таблиця 3.4 – Значення метрик (дБ) для різних тестових зображень і КС

Зображення	КС=2		КС = 4		КС = 8		КС = 16	
	PSNR	PSNR-HVS-M	PSNR	PSNR-HVS-M	PSNR	PSNR-HVS-M	PSNR	PSNR-HVS-M
Lenna	55,80	67,26	45,14	52,59	40,46	44,90	37,24	39,53
Barbara	53,37	64,71	43,59	50,05	37,41	40,53	32,08	33,64
Baboon	46,06	57,23	34,96	41,05	30,16	32,62	25,96	26,70
Goldhill	52,39	64,63	42,03	48,22	36,61	40,16	33,13	33,90
Mrt_prepared	102,32	116,25	52,41	62,84	44,12	48,54	38,21	40,24



а

б

Рисунок 3.6 – Тестові зображення, стиснені з КС=16: а) зображення високої складності Baboon, PSNR=25,96 дБ, PSNR-HVS-M=26,70 дБ, видимі спотворення; б) зображення низької складності Mrt_prepared, PSNR=38,21 дБ, PSNR-HVS-M=40,24 дБ, невидимі спотворення

Для зображення високої складності Baboon спотворення більші. Значення PSNR і PSNR-HVS-M становлять 25,96 дБ і 26,97 дБ відповідно, що нижче порога помітності спотворень (~35 дБ), і викривлення можна чітко спостерігати – вони виглядають як змазування в областях, позначених червоними та жовтими прямокутниками. Для зображення низької складності Mrt_prepared стиснення вносить менше спотворень, а значення PSNR і PSNR-HVS-M становлять близько 38 дБ і 40 дБ відповідно, спотворення непросто виявити візуальним оглядом.

Зрозуміло, що для даної кількості біт на один піксель (БНП) або КС (КС і БНП суворо залежать одне від одного, де для 8-бітного представлення зображень у відтінках сірого $КС \approx 8 / \text{БНП}$ [50]) якість може бути дуже різною. Завдяки аналізу даних в таблиці 3.4 можна зробити висновок, що візуальна якість зображення залежить як від БНП, так і від складності зображення. Це означає, що часто на практиці стиснення з втратами КС або БНП потрібно встановлювати індивідуально (адаптивним способом) для кожного зображення, яке потрібно стиснути з бажаною якістю. Тому складним завданням для SPIHT є прогнозування параметрів і забезпечення бажаної якості зображення.

Нещодавно запропонований двоетапний метод стиснення [50] частково вирішує цю проблему. Двоетапний метод протестовано для великої кількості зображень, результати показують, що візуальна якість усіх стиснених зображень має монотонну залежність від ПКС (БНП), а тенденція зміни якості зображення така, що її можна вважати приблизно лінійною в обмеженому інтервалі значень БНП. Таким чином, можна очкувати, що параметри середньої кривої швидкості спотворення, отримані заздалегідь, можуть бути використані для прогнозування та забезпечення бажаної якості зображення.

Двоетапний метод SPIHT для забезпечення бажаної якості аналогічний описаному вище для кодера AGU, а саме встановлює початкове значення контрольного параметра (БНП), використовуючи осереднену КШС, отриману в автономному режимі (заздалегідь), а потім проводить перше стиснення/декомпресію розглянутого зображення. Потім обчислюється використана метрика візуальної якості Metr_1 , але її значення може суттєво відрізнятися від бажаного (Var_1 після першого кроку може становити близько 40 дБ, таким чином, точність дуже низька). Маючи оцінку якості, отриману на першому етапі, і нахил (похідну) осередненої КШС, значення ПКС, що відповідає бажаній візуальній якості, обчислюється з використанням лінійної апроксимації, щоб виконати другий крок стиснення. Після налаштування параметрів значення візуальної якості, отримане за допомогою другого стиснення, зазвичай наближається до бажаного значення. Завдяки другому кроку дисперсія

зменшується в декілька разів, але все одно може бути занадто великою, особливо для PSNR (похибка може досягати 7...8 дБ). Причина полягає в тому, що використовуємо нахил, що відповідає точці БНП на КШС. У порівнянні з раніше запропонованим ітераційним методом [168], двоетапний метод працює швидше, але все ще бажано зменшити час обчислення та точність.

Ідея полягає в наступному. По-перше, була запропонована дуже швидка процедура прогнозування значень метрик при стисненні зображень із втратами для методів на основі ДКП, зокрема AGU. Простота цієї процедури пояснюється тим, що статистика коефіцієнтів ДКП в обмеженій кількості блоків 8x8 пікселів отримується та дуже швидко аналізується для прогнозування ефективності кодера. По-друге, існує досить тісний зв'язок між характеристиками AGU та SPIHT, який раніше використовувався для стиснення із втратами зображень із шумом [247]. Пропонується використовувати цю властивість, але спочатку її слід ретельно проаналізувати для випадку стиснення зображень, що є практично безшумними.

Проаналізуємо КШС, що отримані для двох тестових зображень. Вони представлені на рис. 3.7 як залежності PSNR і PSNR-HVS-M від КС. Видно, що існує тісний зв'язок між залежностями для розглянутих кодерів. Результати показують, що AGU на основі ДКП досягає трохи кращих значень МЯЕ, ніж SPIHT для того ж самого КС. Але алгоритм корекції (алгоритм перерахунку параметрів AGU у відповідні параметри SPIHT) повинен базуватися на більшому обсязі експериментальних даних.

Для кращого розуміння на рис. 3.7 наведено середні криві залежності показників візуальної якості від КС для двох кодерів. У випадку, коли значення візуальної якості кількох зображень осереднені, криві для SPIHT і AGU теж дуже схожі, і відхилення в основному є фіксованим для кожного значення КС. Після розрахунку середнє відхилення PSNR (SPIHT відносно AGU) становить 0,6232 дБ, а середнє відхилення PSNR-HVS-M (SPIHT відносно AGU) становить 0,9337 дБ. Потім значення метрик візуальної якості для SPIHT можна обчислити як:

$$PSNR_{SPIHT} = PSNR_{AGU} - 0,6235, \quad (3.4)$$

$$PSNR - HVS - M_{SPIHT} = PSNR - HVS - M_{AGU} - 0,9337, \quad (3.5)$$

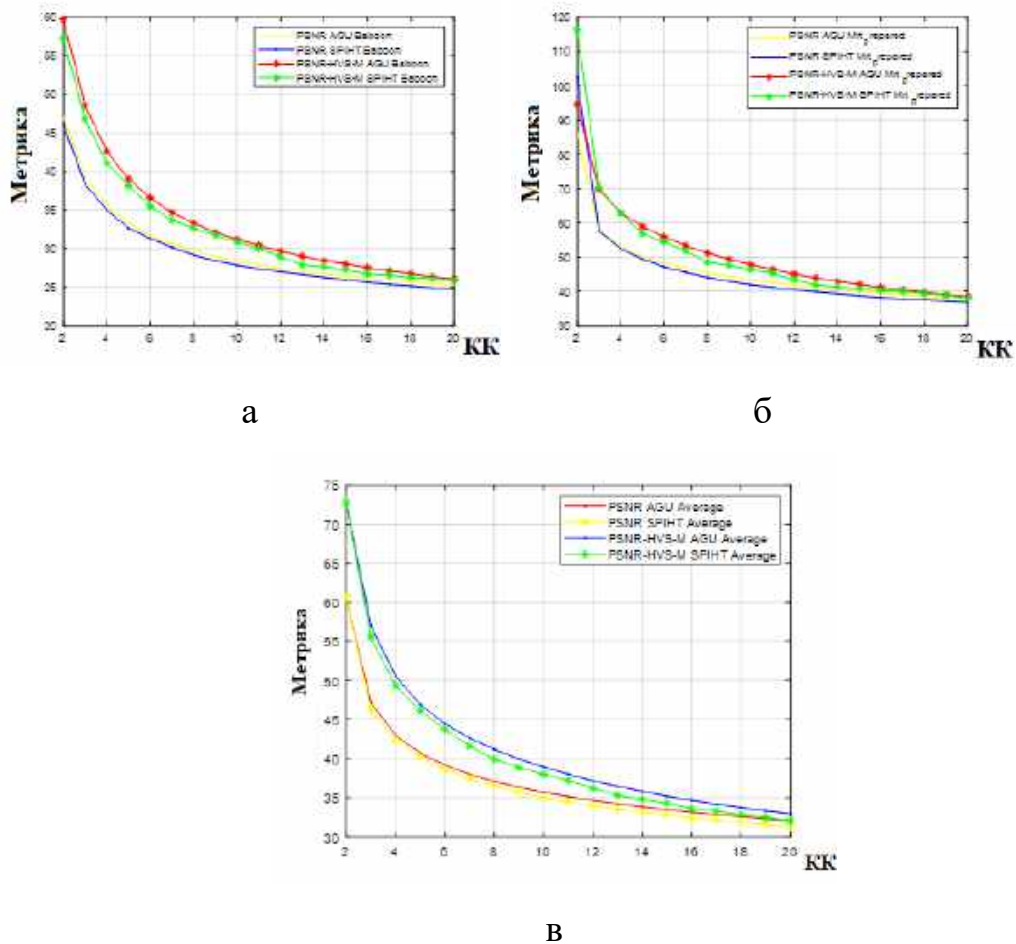


Рисунок 3.7 – Крива залежності тестових зображень: а) зображення високої складності Baboon; б) зображення низької складності MRT_prepared; в) середні залежності для AGU і SPIHT

Аналіз експериментальних результатів компресії SPIHT та AGU показує, що існує фіксований зв'язок між залежністю значень візуальної якості обох від КС. Оскільки прогнозування якості зображення після стиснення AGU було успішно досягнуто [248], стає можливим прогнозувати якість зображення для SPIHT за допомогою методу, що описано нижче.

Коротко розглянемо методику візуального прогнозування якості AGU. Теоретична основа полягає в тому, що існує пряма залежність між спотвореннями, внесеними через квантування коефіцієнта ДКП, і втратами в стиснених даних.

Будемо використовувати параметр, який позначає середню ймовірність того, що квантовані коефіцієнти ДКП у блоках 8x8 дорівнюють нулю [43], [166]:

$$P_{0q} = (\sum_{n=1}^{N_{bl}} N_n) / 64 N_{bl}, \quad (3.6)$$

де N_n визначає кількість коефіцієнтів ДКП, що мають бути обнулені (це відбувається, якщо величина коефіцієнта ДКП менша за $KK/2$) для n -го блоку; N_{bl} позначає кількість розглянутих блоків 8x8 пікселів, випадково вибраних у розглянутому зображенні. Вочевидь, що чим більше P_{0q} (більше коефіцієнтів, яким присвоєно нульові значення), тим більше КС. Функцію, яка апроксимує це співвідношення, можна надати як [166]:

$$KC = 0,9462 \exp(2,895 P_{0q}) + 1,045 \times 10^{-13} \exp(35,52 P_{0q}). \quad (3.7)$$

Це означає, що P_{0q} можна передбачити для даного КС або БНП і навпаки. У [248] запропоновано метод прогнозування PSNR кодера AGU. Цей метод обчислює коефіцієнти ДКП у блоках зображення 8x8 пікселів і квантує їх як:

$$D_q(n, k, l) = [D(n, k, l) / KK], k = 0, \dots, 7; l = 0, \dots, 7, \quad (3.8)$$

де $D(n, k, l)$, $k = 0, \dots, 7; l = 0, \dots, 7$ позначає набір коефіцієнтів ДКП для n -го блоку; $D(n, 0, 0)$ є постійним коефіцієнтом (DC), який пов'язаний з середнім в n -му блоці; $[\cdot]$ позначає округлення до найближчого цілого числа.

Як відомо, стандартний PSNR оцінюється як:

$$PSNR = 10 \log_{10}(MAX^2 / MSE), \quad MSE = \frac{1}{I_{lm} J_{lm}} \sum_{i=1}^{I_{lm}} \sum_{j=1}^{J_{lm}} (I_{ij}^c - I_{ij}^{or})^2, \quad (3.9)$$

де MAX визначає динамічний діапазон зображення; $I_{lm}J_{lm}$ визначає розмір зображення; I_{ij}^c та I_{ij}^{or} позначає значення ij -го пікселя стисненого та оригінального зображень відповідно. У [247], [248] було показано, що MSE для AGU можна швидко і досить точно спрогнозувати для заданого КК наступним чином. Розрахуємо різниці як:

$$\Delta D_q(n, k, l) = KK \times D_q(n, k, l) - D(n, k, l), k = 0, \dots, 7; l = 0, \dots, 7, \quad (3.10)$$

і тоді MSE для n -го блоку можна розрахувати за формулою (3.11):

$$MSE_n = \frac{1}{64} \sum_{k=0}^7 \sum_{l=0}^7 (\Delta D_q(n, k, l))^2. \quad (3.11)$$

Оцінку MSE для всього зображення можна розрахувати наступним чином:

$$MSE = \frac{1}{N} \sum_{n=1}^N MSE_n = \frac{1}{64N} \sum_{n=1}^N \sum_{k=0}^7 \sum_{l=0}^7 (\Delta D_q(n, k, l))^2, \quad (3.12)$$

і, після цього, метрику PSNR можна оцінити відповідно до (3.9). Стандартне відхилення похибок прогнозування менше 2 дБ, тобто краще, ніж для двоетапної процедури для SPIHT.

Коефіцієнти $D(n, k, l), k = 0, \dots, 7; l = 0, \dots, 7$ збираються для всіх розглянутих блоків і, таким чином, можна легко передбачити показники візуальної якості для будь-якого КК. Що стосується метрики візуальної якості PSNR-HVS-M, вона визначається як:

$$PSNR - HVS - M = 10 \log_{10} \left(\frac{255^2}{MSE_{HVS-M}} \right), \quad (3.13)$$

де MSE_{HVS-M} – специфічна MSE, яка враховує особливості HVS. Стаття [43] показує, як MSE_{HVS-M} можна швидко та точно оцінити та спрогнозувати для AGU.

Тож робочі характеристики AGU (KC , MSE , MSE_{HVS-M}) можуть бути швидко і точно спрогнозовані. Водночас відомі також залежності середнього PSNR та PSNR-HVS-M від KC для AGU та SPIHT (рис. 3.7, в, вирази (3.4) та (3.5)). Ці два аспекти забезпечують ідеї реалізації для прогнозування візуальної якості для SPIHT.

Є одна ключова проблема, яку потрібно вирішити. Контрольним параметром SPIHT є БНП, а ПКС AGU – KC . Щоб спрогнозувати якість стисненого зображення SPIHT для заданого БНП, потрібно визначити відповідний KC для AGU. Для цього давайте спочатку перетворимо БНП у KC як $KC=8/\text{БНП}$. Наступним етапом є перетворення KC в KK ; цей етап відносно складний, оскільки залежить від наявного зображення. Нагадаємо, що параметр P_{0q} відображає відсоток значень коефіцієнтів ДКП, обнулених після квантування. Відомий KC дозволяє визначити P_{0q} за допомогою (3.6). Звичайно, обернену функцію знайти нелегко, але P_{0q} можна отримати за допомогою покроково-лінійної інтерполяції.

Тоді потрібен алгоритм отримання KK , що відповідає визначеним P_{0q} і, відповідно, KC для розглянутого зображення. Один дуже простий варіант цього алгоритму показаний на рис. 3.8. Спочатку встановлюється початкове значення KK , наприклад, рівне 0, а потім поступово збільшується, поки відсоток коефіцієнтів ДКП з абсолютними значеннями меншими за $KK/2$ все ще менше, ніж P_{0q} . Коли це станеться, фіксуємо KK .

Четвертий етап у процесі прогнозування полягає у використанні методу для розрахунку значень PSNR і PSNR-HVS-M, що відповідають значенню KK , отриманому на третьому етапі. Після цього завершаємо весь процес прогнозування візуальної якості кодера стиснення SPIHT через AGU, використовуючи вирази (3.4) і (3.5).

У процесі прогнозування деякі обчислення використовуються для заміни фактичного процесу стиснення для досягнення мети прогнозування, і, звичайно, відхилення неминучі. Проведемо аналіз точності прогнозу. У експериментах було розглянуто два значення БНП, а саме 0,5 і 1.

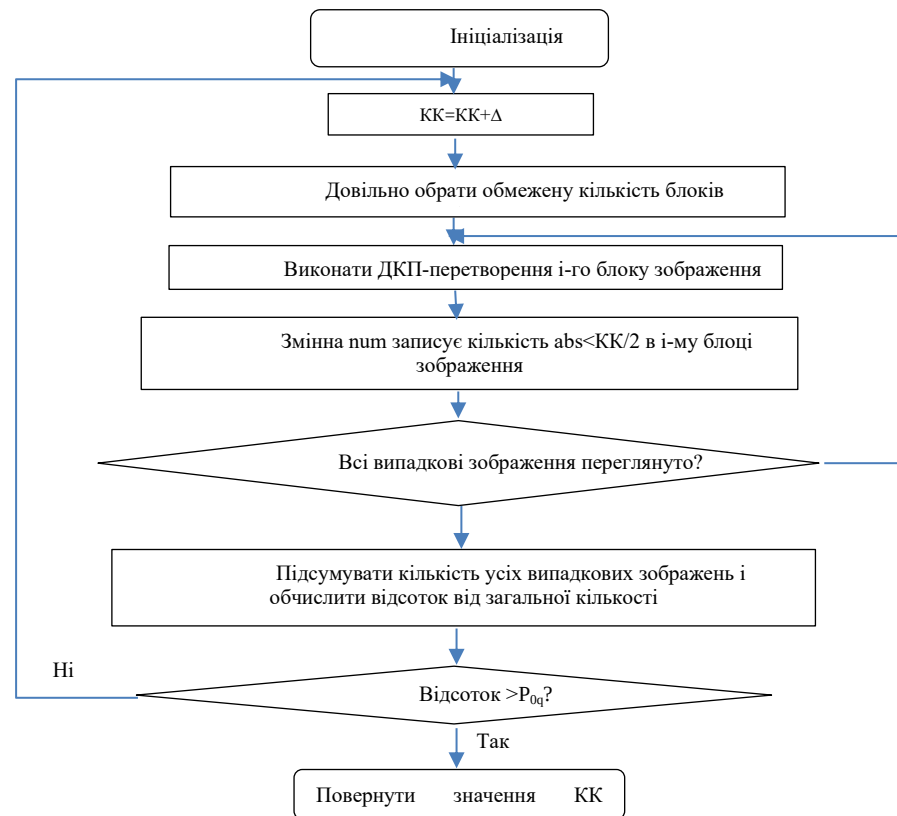


Рисунок 3.8 – Блок-схема для розрахунку KK

Щоб проаналізувати точність прогнозування, було використано два показники візуальної якості, PSNR і PSNR-HVS-M. В експерименті аналізувались 300, 500 і 1000 випадкових положень блоків. Порівняння показало, що точність даних для 500 і 1000 блоків не покращилася в порівнянні з 300 блоками, але витрати часу зросли через збільшення обсягу обчислень. Тому 300 вибрано як кількість випадкових блоків зображення для проведення експерименту. В експерименті було вибрано дев'ять часто використовуваних зображень (включаючи дуже складні текстурні зображення, помірно складні текстурні зображення та прості зображення). Експериментальні дані прогнозу PSNR представлені в таблиці 3.5, що відповідає даним прогнозу візуальної якості з БНП, що дорівнює 0,5 і 1 відповідно.

У цих таблицях $PSNR_{pre}$ представляє значення PSNR зображення, обробленого кодувальником SPIHT, розраховане програмою, а $PSNR_{real}$ є значенням PSNR зображення, фактично стисненого/розпакованого SPIHT.

Похибка – це відхилення між прогнозованим значенням і фактичним значенням, а час відноситься до часу, витраченого на процес прогнозування.

Таблиця 3.5 – Статистика даних PSNR

Зображення	БНП=0,5				БНП=1,0			
	PSNR _{pre}	PSNR _{real}	Помилка	Час, с	PSNR _{pre}	PSNR _{real}	Помилка	Час, с
Goldhill	33,6655	33,1294	-0,5361	0,0977	36,2791	36,6047	0,3256	0,0828
Baboon	26,1117	25,6281	-0,4836	0,1964	30,7143	29,8422	-0,8721	0,1227
Lenna	37,0936	37,2353	0,1417	0,0859	40,778	40,4585	-0,3195	0,0629
Barbara	35,2759	32,0839	-3,192	0,1097	39,0924	37,4122	-1,6802	0,0689
Aerial	28,6229	28,7333	0,1104	0,1832	32,7749	33,1914	0,4165	0,1079
Airfield	29,0245	27,4969	-1,5276	0,1558	31,9097	30,3014	-1,6083	0,1056
Frisco	42,8915	42,6874	-0,2041	0,0598	48,3571	47,7694	-0,5877	0,0638
Diego	26,5196	26,6317	0,1121	0,1912	29,4761	29,4905	0,0144	0,1250
Mrt_prepared	39,4486	38,2106	-1,238	0,0869	47,7793	44,1214	-3,6579	0,0578
CKB S	1,0274				1,2120			

Щоб статистично проаналізувати точність прогнозу, розрахунок СКВ було виконано для кожної групи похибок як:

$$S = \sqrt{\frac{1}{N-1} \sum_{i=1}^N (X_i - \bar{X})^2}, \quad (3.14)$$

З аналізу даних у таблиці 3.5 можна побачити, що точність прогнозування цього методу хороша, а максимальне відхилення становить близько 3,66 дБ. І загальний розподіл помилок розумний. Стандартні відхилення забезпечення бажаного PSNR становлять 1,0274 і 1,212 дБ відповідно. Таким чином, точність краща, ніж у двоетапного методу. Дані про PSNR-HVS-M наведено в таблиці 3.6. Тут PSNR_{pre} — це значення PSNR зображення, обробленого кодувальником SPIHT, розраховане програмою, а PSNR_{real} — це значення PSNR зображення, фактично стисненого/розпакованого SPIHT.

Таблиця 3.6 – Статистика даних PSNR-HVS-M

Зображення	БНП=0,5				БНП=1,0			
	PSNR-HVS-M _{pre}	PSNR-HVS-M _{real}	Помилка	Час, с	PSNR-HVS-M _{pre}	PSNR-HVS-M _{real}	Помилка	Час, с
Goldhill	36,2594	33,9034	-2,3560	0,2408	40,8164	40,1590	-0,6574	0,2105
Baboon	27,8017	26,6977	-1,1040	0,3336	35,0034	32,6220	-2,3814	0,2583
Lenna	39,0606	39,5323	0,4717	0,2125	44,1417	44,9048	0,7632	0,1926
Barbara	36,4551	33,6409	-2,8142	0,2305	42,6237	40,5339	-2,0898	0,2024
Aerial	30,2541	29,3158	-0,9383	0,3006	37,7715	36,6634	-1,1081	0,2786
Airfield	31,3978	29,1119	-2,2859	0,2706	36,7416	33,5935	-3,1481	0,2538
Frisco	42,9987	42,5198	-0,4789	0,2005	46,7146	48,9363	2,2217	0,1926
Diego	28,2534	27,4636	-0,7898	0,3229	33,7246	33,3702	-0,3544	0,2790
Mrt_prepared	42,7600	40,2362	-2,5238	0,1905	44,8965	48,5444	3,6479	0,1584
СКВ	1,0531				2,0946			

Відхилення PSNR-HVS-M, коли БНП дорівнює 0,5, трохи більше, ніж відхилення для БНП, що дорівнює 1. Максимальне відхилення становить 3,65 дБ, а стандартні відхилення вибірки становлять 1,0531 дБ і 2,0946 дБ відповідно. Вибіркове стандартне відхилення даних PSNR-HVS-M менше, ніж для PSNR, і помилка більш стабільна. У той же час, оскільки розрахунок PSNR-HVS-M включає перетворення ДКП та зворотне перетворення ДКП, споживаний час збільшується приблизно в 2 рази, ніж прогнозування PSNR. Зі статистичних даних PSNR і PSNR-HVS-M цей метод прогнозування може дуже добре прогнозувати якість стисненого зображення SPIHT на основі AGU, і точність була покращена. У таблиці 3.7 представлені дані для аналізу ефективності використання часу. Щоб більш об'єктивно зрозуміти часову ефективність методу прогнозування візуальної якості SPIHT, в таблиці наведено середній час прогнозування 9 тестових зображень використовується як коефіцієнт вимірювання в таблиці.

Таблиця 3.7 – Порівняння середнього споживання часу (для метрик PSNR і PSNR-HVS-M)

БНП	Час _{pre}	Час _{two-step}	Час _{spiht}
0,5	0,2558	0,8937	0,3914
1	0,2251	0,9220	0,3005

У таблиці 3.7 розглянуто два набори даних для двох значень БНП - 0,5 і 1. Прогнозування та розрахунок PSNR і PSNR-HVS-M виконувалося наступним чином. Час_{pre} означає час, необхідний для отримання значення візуальної якості за допомогою методу прогнозування; Час_{two-step} представляє витрати на перший крок стиснення/декомпресії за допомогою двоетапного методу та отримує значення візуальної якості за необхідний час; Час_{spiht} — це час, необхідний для SPIHT стислих зображень. Як видно з даних у таблиці 3.7, запропонований метод прогнозування та забезпечення якості значно покращив ефективність часу, оскільки він уникає етапу стиснення/декомпресії. Середній необхідний час становить 1/3 від двоетапного методу та менше часу, необхідного для фактичного процесу стиснення SPIHT.

Перейдемо до аналізу кодера BPG. З двома типовими значеннями 30 дБ і 35 дБ як бажаними значеннями дані кодерів стиснення BPG і SPIHT порівнювалися для трьох зображень, а експериментальні результати наведено в таблиці 3.8.

З порівняння трьох зображень можна побачити, що коли бажане значення метрики якості становить 35 дБ, коефіцієнти стиснення двох кодерів стиснення досить близькі. Тим часом, коли бажане значення становить 30 дБ, КС для BPG набагато вище, ніж у SPIHT (від 3 до 6 разів), особливо коли текстура зображення проста (як у зображеннях Goldhill і Lenna). Ефективність стиснення BPG вища, ніж у SPIHT, особливо в робочому діапазоні стиснення з втратами 30 дБ...35 дБ, що є типовим. Надзвичайно важливо забезпечити бажану якість візуального стиснення для цього діапазону, тоді як високий ступінь стиснення може бути досягнутий розглянутим кодером за умови відповідності вимогам візуальної якості. Ця проблема чітко демонструє переваги кодера BPG.

Таблиця 3.8 – Залежність КС від PSNR

Зображення	PSNR _{des} =35 дБ		PSNR _{des} =30 дБ	
	BPG	SPIHT	BPG	SPIHT
Goldhill	10,865	9,998	245,984	40,0037
Baboon	3,636	3,999	19,823	6,987
Lenna	27,285	26,665	245,834	80,0195

Дослідимо застосовність двоетапного методу для кодера BPG. Використовуємо 12 зображень як вибіркового набір зображень, включаючи стандартні тестові зображення та текстури. На рис. 3.9 наведено залежності PSNR та PSNR-HVS-M від Q.

Крива на рис. 3.9, а показує, що значення PSNR для $PSNR > 40$ дБ приблизно збігаються майже для всіх зображень (крім Frisco), а в інтервалі $30 \text{ дБ} < PSNR \leq 40 \text{ дБ}$ криві зображення майже паралельні, зокрема, для інтервалу $30 \text{ дБ} < PSNR < 35 \text{ дБ}$. Середня крива та кожна індивідуальна крива також приблизно паралельні. Загалом, ступінь розкиду між кривими менший, ніж для інших методів стиснення [44], [45], [50]. Тим не менш, значення PSNR для одного і того ж Q можуть значно відрізнятися, до 10 дБ. Це створює проблеми із забезпеченням бажаної якості. Залежності PSNR-HVS-M від Q демонструють ту саму тенденцію, але криві йдуть більш компактно, ніж залежності PSNR. Порівняно з кодером SPIHT можна сказати, що візуальна якість у стисненні за допомогою BPG менше залежить від вмісту зображення (або складності), а основним параметром прийняття рішення є ПКС Q. При цьому значення PSNR-HVS-M для одного і того ж Q можуть відрізнятися до 7 дБ. Таким чином, використання заданого Q може призвести до різної візуальної якості стиснутих зображень.

Середня КІШС для спотворення кодера BPG відображає залежність якості зображення від контрольного параметра Q. Ці криві були отримані з експериментальних статистичних даних набору зображень. Цей набір зображень містить зображення, які є стандартними для обробки зображень, зображення дистанційного зондування та вибрані типові зображення текстури (рис. 2.7).

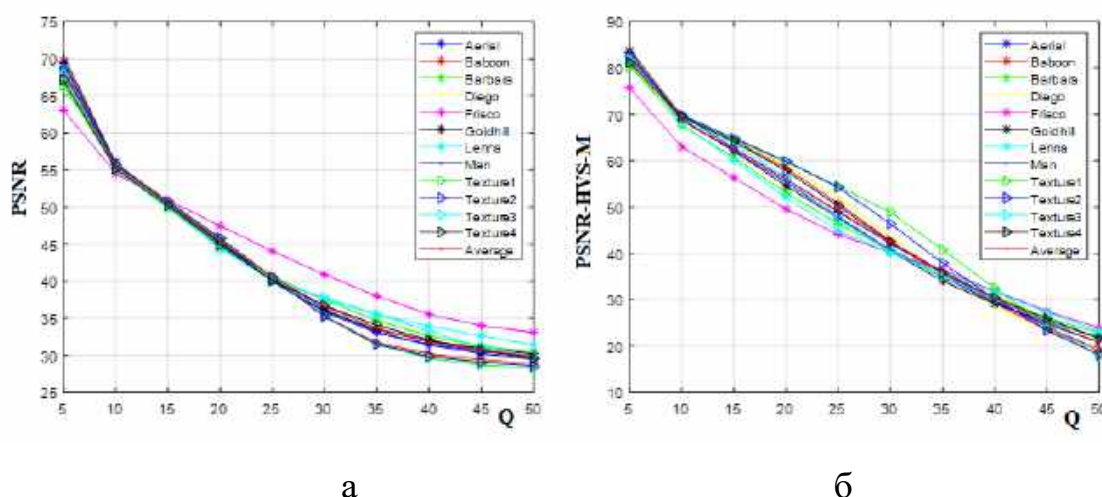


Рисунок 3.9 – Залежності PSNR (а) та PSNR-HVS-M (б) від Q

Видно, що типи зображень у наборі різноманітні, а складність зображень різна. Експериментальні дані цього набору зображень стали вихідними даними розробки двоетапного методу. Частина статистичних даних для залежностей PSNR-HVS-M від Q відображена в таблиці 3.9. На основі кривої середніх значень визначені середні значення похідної. На першому етапі стиснення необхідно розрахувати початкове значення ПКС. Основою розрахунку є те, що середнє значення візуальної якості, відповідне розрахунковому значенню, є достатньо близьким до бажаного значення.

Це розрахункове значення обчислюється аналогічно рівнянню (3.1):

$$Q_{init} = Q_{est} + \frac{PSNR-HVS-M_{des} - PSNR-HVS-M_{ave}}{M'}, \quad (3.15)$$

де Q_{est} – правий край інтервалу КІШС; $PSNR-HVS-M_{ave}$ – середнє значення PSNR-HVS-M, що відповідає Q_{est} . $PSNR-HVS-M_{des}$ – бажане значення, і M' є похідною, що відповідає Q_{est} . Розраховуючи у відповідності з даними у таблиці 3.9, можна зробити висновок, що початкове значення Q дорівнює 35.

Таблиця 3.9 – Залежність PSNR-HVS-M від Q (часткові дані)

Зображення	Q			
	5	30	35	40
Goldhill	83,310	40,499	34,071	29,122
Lenna	82,634	40,166	35,871	31,652
Середнє	81,909	42,431	36,095	30,278
Похідна	-2,649	-1,267	-1,163	-1,044

Припустимо, що треба забезпечити $PSNR-HVS-M_{des} = 35$ дБ для зображення Goldhill. Розраховуючи з даними таблиці 3.9, можна зробити висновок, що початкове значення Q дорівнює 35. Перше стиснення приводить до значення $PSNR-HVS-M_{init}$, що дорівнює 34,071 дБ. Це значення є досить близьким до бажаного значення, але точність все ще можна підвищити за допомогою другого кроку. На другому кроці Q коригується за допомогою рівняння:

$$Q_{des} = Q_{init} + \frac{PSNR-HVS-M_{des} - PSNR-HVS-M_{init}}{M'}. \quad (3.16)$$

Нове значення Q (дорівнює 34) використовується на другому кроці стиснення. Після другого кроку $PSNR-HVS-M_{prov}$ становить 35,394 дБ і є ближчим до бажаного значення.

На відміну від інших кодерів, ПКС Q у BPG можна встановлювати лише як ціле число. Чи впливає ця функція на вплив двоетапного методу на BPG? Необхідно провести більше експериментів. Щоб проаналізувати вплив двоетапного методу на BPG, використовуються три типові значення: 40 дБ, 35 дБ і 30 дБ для тестування, а також використовуються метрики PSNR і PSNR-HVS-M. Статистичні результати експерименту наведені в таблиці 3.10. Зі статистичних даних видно, що хоча ПКС для BPG може мати лише цілочисельні налаштування, дисперсія після стиснення на другому кроці (після корекції параметра) приблизно на порядок менша, ніж на першому кроці, а похибка після двоетапного стиснення прийнятна для практики.

Похибка для PSNR не перевищує 2,87 дБ, а для PSNR-HVS-M не перевищує 2,55 дБ. Можна побачити, що метод двоетапного стиснення підходить для кодерів BPG і може забезпечити достатнє підвищення точності. Щоб зрозуміти конкретні деталі даних, використовуємо бажане значення PSNR-HVS-M, яке становить 35 дБ для подальшого аналізу.

Таблиця 3.10 – Статистичні дані, отримані для дев'яти тестових зображень

Метрика	M_{des}	VAR_1	VAR_2	$MAX\Delta_{final}$
PSNR	40	1,227	0,108	1,032
PSNR	35	3,754	0,175	1,325
PSNR	30	1,856	0,781	2,87
PSNR-HVS-M	40	6,243	0,343	1,05
PSNR-HVS-M	35	3,087	0,916	2,545
PSNR-HVS-M	30	1,463	0,274	1,072

Параметр Q , показаний у таблиці 3.9, є параметром, який використовується для першого кроку стиснення. Він походить із осередненої КІШС і становить 35 для всіх зображень. PSNR-HVS- M_{init} після першого кроку стиснення різний для різних зображень, інтервал його зміни $[34,071 \sim 40,838]$ дБ, дисперсія 3,087, а максимальна похибка 5,838 дБ. Можна побачити, що хоча середній PSNR-HVS- M_{init} становить близько 35 дБ, існує великий діапазон його зміни. Водночас це значення використовується для зворотного зв'язку та розрахунку корекції Q для кожного окремого зображення (з використанням рівняння (3.16)). Використовуючи Q_{des} як параметр для другого кроку стиснення, значення PSNR-HVS-M, яке можна задати, звужується. Діапазон $[32,455; 36,097]$ (дБ), дисперсія — 0,916, а максимальна похибка — 2,545 дБ. Взагалі кажучи, незалежно від діапазону зміни або дисперсії, максимальна похибка, як правило, покращується після другого кроку стиснення. Зокрема, можна побачити, що в багатьох випадках уточнений параметр Q на другому кроці є таким самим, як і на першому кроці. У цьому випадку для

досягнення бажаної візуальної якості потрібен лише один крок. Другий крок стиснення можна пропустити, щоб покращити ефективність часу.

Таблиця 3.11 – Статистичні дані та параметри для PSNR-HVS- $M_{des}=35$ дБ

Test image	Q_{init}	PSNR-HVS- M_{init}	Q_{des}	PSNR-HVS- M_{prov}	КС
Aerial	35,000	35,683	35,000	35,683	5,997
Baboon	35,000	35,374	35,000	35,374	6,929
Barbara	35,000	35,665	35,000	35,665	16,755
Diego	35,000	35,447	35,000	35,447	7,827
Frisco	35,000	36,414	36,000	35,495	29,486
Goldhill	35,000	34,071	34,000	35,394	16,831
Lenna	35,000	35,871	35,000	35,871	31,028
Man	35,000	34,926	34,000	36,097	10,955
Texture 1	35,000	40,838	40,000	32,455	7,427
Texture 2	35,000	37,937	37,000	34,762	8,456
Texture 3	35,000	35,042	35,000	35,042	25,348
Texture 4	35,000	35,869	35,000	35,869	17,187
Дисперсія	3,087		0,916		

З аналізу даних у таблиці 3.11 впливають ще два цікаві аспекти:

1) найбільша помилка виникає для зображення Texture 1, а не для зображення Frisco. Нагадаємо, що найбільше відхилення від середньої кривої спотворень є для тестового зображення Frisco;

2) отриманий коефіцієнт стиснення для різних зображень сильно відрізняється. При однаковій візуальній якості КС змінюється від 6 до 31. Таким чином, двоетапний метод забезпечує BPG-стиснення з бажаною візуальною якістю, яка гарантує передумови для подальшої обробки зображення, переваги BPG можуть бути максимально використані, а ступінь стиснення можна отримати якомога більше.

Дослідження кодера ADCT розглянуто у Додатку А.

3.2 Стиснення зображень із забезпеченням візуально непомітних спотворень

Подібно до добре відомого JPEG, зрозуміло, що спотворення при стисненні зображень із втратами на основі ДКП викликані квантуванням коефіцієнтів ДКП [87], [248]. Якщо загальна кількість блоків 8×8 пікселів дорівнює N , тоді припустимо, що маємо коефіцієнти ДКП $D(n, k, l)$, $n=1, \dots, N$; $k=0, \dots, 7$; $l=0, \dots, 7$. Після квантування позначимо коефіцієнти як $D_q(n, k, l)$, $n=1, \dots, N$; $k=0, \dots, 7$; $l=0, \dots, 7$. Тоді середньоквадратична похибка буде мати вигляд:

$$MSE = \frac{1}{N} \sum_{n=1}^N MSE_n = \frac{1}{64N} \sum_{n=1}^N \sum_{k=0}^7 \sum_{l=0}^7 (\Delta D_q(n, k, l))^2, \quad (3.17)$$

де

$$D_q(n, k, l) = \left[\frac{D(n, k, l)}{KK} \right], k = 0, \dots, 7; l = 0, \dots, 7,$$

$$\Delta D_q(n, k, l) = KK D_q(n, k, l) - D(n, k, l), k = 0, \dots, 7; l = 0, \dots, 7,$$

Операція $[]$ означає округлення до найближчого цілого, а n – індекс блоку.

Зазвичай припускають, що помилки квантування мають рівномірний розподіл і, завдяки цьому, їх дисперсія становить $\Delta^2/12$, де Δ – розмір інтервалу. Оскільки у досліджуваному випадку $\Delta=KK$, можна очікувати, що MSE внесеної дисперсії (для рівномірного квантування коефіцієнтів ДКП) має становити близько $KK^2/12$. Однак більш детальний аналіз показує, що це не так.

Спочатку проаналізуємо залежність MSE внесених спотворень від KK для трьох тестових зображень: безшумних тестових зображень дистанційного зондування Frisco (що має просту структуру) та Airfield (що має більш складну структуру) та тестового зображення Frisco, спотвореного АБГШ. Залежності представлені на рис. 3.10 для кодера ADCT.

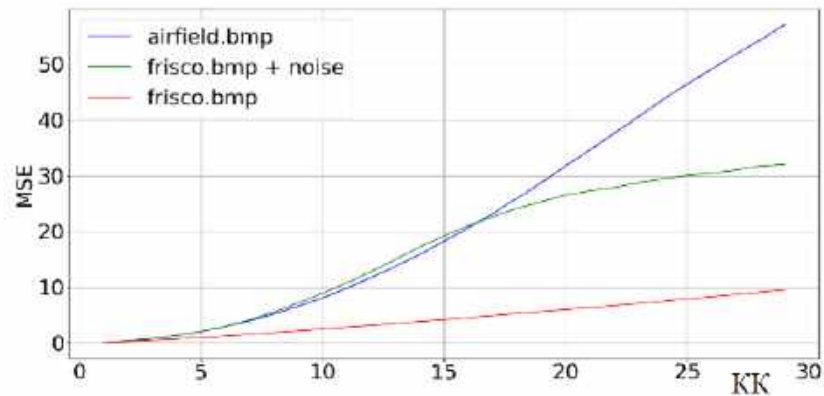


Рисунок 3.10 – Залежності для безшумного і зашумленого зображень Frisco і безшумного зображення Airfield

Перше спостереження полягає в тому, що існує достатня різниця в MSE (особливо для $KK > 5$) залежно від складності зображення. Друге спостереження полягає в тому, що наявність шуму змінює поведінку MSE порівняно з тим самим безшумним зображенням.

На рис. 3.11 наведено залежності КС від КК для тих же трьох зображень. Видно, що безшумне зображення з простою структурою стискається набагато краще, ніж два інших для того самого КК. Основна причина полягає в тому, що для безшумного зображення простої структури існує значно більше нульових значень коефіцієнтів ДКП після їх квантування.

Причина цих ефектів полягає в розподілі коефіцієнтів ДКП з ненульовою просторовою частотою. Розподіли представлені в Додатку К, на рис. К.2 в однаковому масштабі для всіх трьох зображень. Розподіли для безшумних зображень не є гаусовими, а розподіл для простіших за структурою зображень є значно вужчим. Наявність шуму (особливо якщо його дисперсія досить велика) робить розподіл ближчим до гаусового і ширшим. Такі відмінності в розподілах коефіцієнтів ДКП приводять до відмінностей розподілів помилок квантування. Якщо КК малий порівняно з «шириною» (масштабом) розподілу, то розподіл помилок квантування близький до рівномірного (прикладі наведені в Додатку К, на рис. К.3). Це трапляється для досить малого КК, зображень складної структури

та/або зображень із шумом. Тоді MSE введених спотворень має бути близьким до $KK^2/12$.

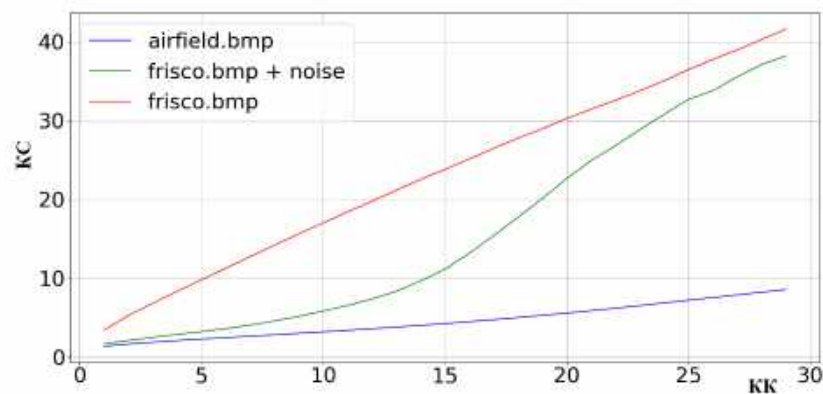


Рисунок 3.11 – Залежності КС від КК для безшумного і зашумленого зображень Frisco і безшумного зображень Airfield

У той же час, якщо розподіл коефіцієнтів АС ДКП є вузьким порівняно з КК, розподіл помилок квантування може значно відрізнятися від рівномірного. Приклади наведені в Додатку К, на рис. К.3, б та К.3, в. Це трапляється для зображень простої структури, які не зіпсовані шумом. Тоді очікується, що MSE введених помилок буде меншим, ніж $KK^2/12$. Таким чином, можна припустити, що MSE внесених спотворень якимось чином залежить від статистики коефіцієнтів ДКП, яка, у свою чергу, визначається властивостями зображення та шуму. Статистика коефіцієнтів АС ДКП, проаналізована вище, була зібрана в блоках пікселів 8×8 , але залежності аналізувалися для методів стиснення, які використовують блоки іншого розміру. Можливість такого аналізу була доведена в роботі [248], де було показано, що MSE, передбачена для заданої КК для набору 8×8 піксельних блоків, сильно корелює з MSE для кодерів AGU та ADCT. Крім того, набір блоків, у яких збирається статистика, може бути обмеженого розміру.

Іншими словами, немає необхідності збирати статистику коефіцієнтів АС ДКП для всіх можливих позицій блоків. Ці спостереження (див. [248] для більш детальної інформації) відкрили перспективи для забезпечення бажаного MSE або PSNR з високою швидкістю – дивись роботи [24], [25], [33].

Треба пояснити, у чому полягає ідея та що мається на увазі під прогнозуванням параметрів стиснення. Припустимо наступне. Маючи зображення, яке потрібно стиснути, і виконавши простий і швидкий аналіз його статистики, можна обчислити приблизне (прогнозоване) значення розглянутої метрики, що характеризує стиснення для даного кодера і ПКС, що використовується в цьому кодері. Як приклад – припустимо, що аналізуючи статистику коефіцієнтів ДКП у наборі блоків, можна розрахувати приблизну MSE для заданої КК для кодерів AGU або ADCT. Тоді стає можливим налаштувати ПКС (у нашому випадку КК) таким чином, щоб прогнозоване значення розглянутої метрики (наприклад, MSE) приблизно дорівнювало бажаному значенню цієї метрики (наприклад, $MSE_{des}=25$ для 8-ми бітових зображень, що приблизно відповідає порогу невидимості внесених спотворень[245]). Якщо ці припущення справедливі і відповідний алгоритм може бути реалізований, стає можливим здійснити стиснення з бажаною якістю без ітерацій.

Спроба прогнозувати PSNR для JPEG має декілька недоліків. По-перше, вона була розроблена тільки для JPEG. По-друге, передбачалося, що всі коефіцієнти ДКП повинні бути зібрані та збережені для оцінки масштабу розподілу (це не є серйозною проблемою, оскільки фактично оцінку масштабу можна зробити без збереження всіх коефіцієнтів ДКП). По-третє, припущення про те, що розподіл коефіцієнтів ДКП близький до лапласівського, невірне. Через це MSE і PSNR для даного КК прогножуються з помилками. Для PSNR похибки не перевищують 1 дБ, і це може бути прийнятним для практики. Оскільки всі коефіцієнти ДКП у 8x8 піксельних блоках повинні бути обчислені для прогнозування, то прогнозування займає майже стільки ж часу, скільки стиснення, і, фактично, передбачення є вбудованою частиною стиснення JPEG.

Для кодерів AGU і ADCT бажано виконувати прогноз швидше, ніж виконується сама компресія. У деяких роботах запропоновано два шляхи для цього [248], [249]. Перша ідея полягає в тому, що прогнозовану MSE для вдосконалених кодерів на основі ДКП (AGU, AGU-M, ADCT, ADCT-M) можна розрахувати як прогнозовану MSE для JPEG, помножену на відповідний коригувальний коефіцієнт

(значення цих коефіцієнтів близькі до одиниці). Друга ідея полягає в тому, що прогнозування можна зробити як наведено у виразі (для рівномірного квантування):

$$MSE_{pred} = \frac{1}{R} \sum_{r=1}^R MSE_r = \frac{1}{64R} \sum_{r=1}^R \sum_{k=0}^7 \sum_{l=0}^7 (\Delta D_q(r, k, l))^2, \quad (3.18)$$

де

$$\Delta D_q(n, k, l) = KK D_q(r, k, l) - D(r, k, l), k = 0, \dots, 7; l = 0, \dots, 7,$$

Ключовою відмінністю є те, що $R \ll N$. У [248], [249] показано, що R може бути в десятки разів меншим за N і часто достатньо мати 300...500 випадково вибраних блоків, щоб оцінити статистику коефіцієнтів ДКП з належною точністю.

Зауважимо, що (3.18) дозволяє прогнозувати MSE для заданої КК, але не дає прямої відповіді, як надати MSE_{des} . Одним із виходів може бути використання ітераційної процедури налаштування КК для отримання прогнозованого MSE, близького до MSE_{des} , з належною точністю. Це можуть бути ітераційні процедури, подібні до [168], які застосовуються для налаштування КК за допомогою блокової модифікації, але без стиснення/декомпресії зображення, тобто значно швидше. Похибка, що дорівнює 5% MSE_{des} , може вважатися належною точністю (правило зупинки).

Підсумовуючи цей опис, можна констатувати, що простота та досить висока швидкість налаштування КК для надання MSE_{des} забезпечується наступним. По-перше, двовимірне ДКП має бути виконано в блоках 8x8 пікселів, для яких існують численні ефективні апаратні та програмні реалізації [250]. По-друге, аналізується обмежена кількість блоків, менша за загальну кількість блоків 8x8 пікселів для даного зображення. По-третє, навіть якщо для знаходження потрібного КК треба 8...10 ітерацій, операції в (3.18) досить прості, швидкі та їх легко реалізувати.

При цьому цей етап можна додатково прискорити. Це можна зробити за допомогою підходу прогнозування, який спочатку був запропонований у [25], а потім розвинений у [24], [33]. Ідея полягає в наступному. Припустимо, що дана метрика може бути апроксимована як досить проста функція ПКС (наприклад, КК) і одного або двох параметрів, які описують розподіл коефіцієнтів АС DCT, де ці параметри можна обчислити дуже легко та швидко. Загальна формула для MSE_{pred} :

$$MSE_{pred} = (KK^2/12)f_0(x), \quad (3.19)$$

де $f_0(x)$ – функція параметрів X (або одного параметра).

Тоді завдання полягає в тому, щоб знайти цю функцію та вирішити, який параметр(и) використовувати як її аргумент(и). Завдання було вирішено з використанням діаграм розсіювання та регресії, а також попереднього досвіду [5]. Припустимо, що є набір тестових зображень, які стискаються розглянутим кодером з різними значеннями КК. Для кожного випадку стиснення вимірюється MSE внесених спотворень, визначається параметр(и) X і відома КК. Тоді кожен випадок можна представити як точку діаграми розсіювання, де вертикальна вісь відповідає $\hat{f}(x) = 12MSE/KK^2$, а горизонтальна вісь (або осі) – X . Приклад такої діаграми розсіювання, де X є єдиним параметром, показано в Додатку К, на рис. К.4 (кодер ADCT). Як параметр, що характеризує статистику коефіцієнтів ДКП, було використано P_0 – ймовірність того, що коефіцієнти ДКП після квантування стануть рівними нулю. Цей параметр також можна розглядати як ймовірність того, що абсолютні значення $D(r,k,l), r=0, \dots, R, k=0, \dots, 7; l=0, \dots, 7$ менші за $KK/2$. Зрозуміло, що P_0 можна дуже легко і швидко обчислити, і це робить прогнозування швидшим, ніж прогнозування на основі (3.18).

Параметр P_0 було обрано на основі попереднього досвіду [195], [251], де цей параметр використовувався для прогнозування КС (див. діаграму розсіювання в Додатку К, на рис. К.5). Точки розташовані дуже компактно, і це показує, що існує дуже суворя залежність між КС і P_0 , що дозволяє точно прогнозувати КС.

Давайте пояснимо, як P_0 описує властивості зображення. Для даного КК P_0 менший, якщо масштаб коефіцієнтів ДКП більший, тобто якщо складність зображення вища (і тоді КС менший, як видно в Додатку К, з рис. К.5, це добре узгоджується з графіками на рис. 3.11). Таким чином, це показує, що P_0 є параметром, який досить добре і адекватно характеризує складність зображення.

Тепер повернемося до діаграм розсіювання в Додатку К, на рис. К.4 (діаграми розсіювання на рис. К.4, а і К.4, б однакові, але підігнані криві різні). Діаграма розсіювання досить «компактна», і це вказує на те, що можемо досить добре апроксимувати функцію. Інтервал можливих аргументів (від 0 до 1) можна розділити на дві області (підінтервали). Перший підінтервал $P_0 < 0,6$, де спостерігається краща компактність точок діаграми розсіювання та оцінки $f_0(P_0)$ близькі до одиниці. Це означає, що для $P_0 < 0,6$ маємо $MSE_{pred} \approx (KK^2/12)$, і, отже, $KK_{des} = \sqrt{12MSE_{des}}$. Тоді для забезпечення непомітності спотворень, які зазвичай мають місце для MSE близько 20...25, необхідно використовувати КК близько 17 для кодерів AGU і ADCT. Пам'ятаючи про те, що для забезпечення еквівалентної кодерам AGU і ADCT якості, кодери AGU-M і ADCT-M повинні використовувати $KM_{des} \approx KK_{des}/1,6$. Забезпечити невидимість спотворень у цих кодерах можна встановленням $KM_{des} \approx 10,5$. Це досить добре узгоджується з емпіричними рекомендаціями, наведеними в [182]. Там було запропоновано встановити для AGU-M $KM_{des} \approx (I_{max} - I_{min})/20$, де I_{max}, I_{min} – максимальне і мінімальне значення пікселів зображення, що стискається. Для 8-бітних зображень $I_{max} - I_{min}$ зазвичай трохи менше 255 і, таким чином, отримуємо $KM_{des} \approx 12$. MSE у роботі [182] обчислюється як $(I_{max} - I_{min})^2/2000 \approx 30$, тобто трохи більше ніж MSE, яке рекомендуємо зараз (20...25).

Зверніть увагу, що якщо $P_0 < 0,6$, то КС < 6 (див. в Додатку К, на рис. К.5). Існує також підінтервал $P_0 \geq 0,6$ (див. в Додатку К, на рис. К.4), який відповідає більшому КС (див. в Додатку К, на рис. К.5) і більшим внесеним спотворенням. Точки діаграми розсіювання розташовані не так щільно, як для $P_0 < 0,6$. Проте, в будь-якому випадку, тенденції та властивості очевидні. По-перше, MSE внесених

спотворень менші або суттєво менші, ніж $KK^2/12$. Іншими словами, MSE зростає повільніше, ніж KK^2 , по мірі збільшення KK. Окремі приклади цього можна побачити на рис. 3.10. Оскільки для заданого KK ймовірність P_0 більша для простіших структурних зображень, це означає, що такі зображення можуть бути стиснуті з більшим КС, ніж складні за структурою зображення для тієї самої MSE внесених спотворень.

Важливим моментом є також формування діаграми розсіювання. При отриманні даних для діаграми розсіювання необхідно враховувати деякі рекомендації. По-перше, аргументи точок повинні покривати весь інтервал їх можливих значень. По-друге, дані повинні представляти різні ситуації, можливі на практиці. У цьому випадку це означає, що тестовий набір зображень повинен включати прості, помірно складні та складні зображення, а KK має варіюватися в дуже широких межах. Ці рекомендації були враховані при отриманні діаграм розсіювання в Додатку К, на рис. К.4 і К.5. Інше питання полягає у виборі функцій для підгонки. Декілька функцій можуть бути хорошими рішеннями. Зокрема, сума двох експонент (Додаток К, рис. К.4, а) і короткий ряд Фур'є (Додаток К, рис. К.4, б) дають схожі результати. В обох випадках R^2 становить близько 0,92 (це вважається хорошим результатом регресії), а RMSE становить 0,084 (що свідчить про досить високу точність прогнозу). Таким чином, отримана функція має вигляд:

$$f_0(P_0) = -0,007721 \exp(4,824P_0) + 1,112 \exp(-0,1455P_0), \quad (3.20)$$

Маючи цю апроксимацію, пропонуємо наступну процедуру забезпечення бажаного MSE_{des} :

- 1) починаємо з деякого KK_0 ; з огляду на те, що $KK_{des} = \sqrt{12MSE_{des}}$, рекомендується використовувати $KK_0 = \sqrt{12MSE_{des}}$; обчислити P_0 для цього KK_0 шляхом порівняння абсолютних значень отриманих коефіцієнтів АС ДКП з $KK_0/2$;
- 2) якщо $P_0 < 0,6$, використовувати KK_0 як KK_{des} та зупинити процедуру;

3) в іншому випадку збільшуємо KK і обчислюємо для нього P_0 доки $KK^2 f_0(P_0) \in [11MSE_{des}; 13MSE_{des}]$ (ширина інтервалу вибирається так, щоб забезпечити швидку конвергенцію і відповідну точність). Потім використовуємо KK як KK_{des} та зупиняємо процедуру. KK збільшується приблизно на $0,08KK_0$ на кожному наступному кроці.

Запропонована процедура зазвичай потребує не більше 3 кроків. Основний час обчислень на кожному кроці займає операція обчислення P_0 , яка є дуже простою та швидкою. Звичайно, стиснення та декомпресія не потрібні.

Вище приділили основну увагу прогнозуванню MSE для кодерів, де KK використовується як PKC . Це кодери AGU та $ADCT$, для яких діаграми розсіювання та основні залежності дуже схожі [25], [33]. Може здатися, що представлені результати важливі лише для метрики MSE , вищезгаданих кодерів і зображень у градаціях сірого. Однак це не так.

Перш за все, в [248] показано, що прогнозування MSE можна зробити на основі (3.18) для кодерів $AGU-M$ і $ADCT-M$. Єдина відмінність у цьому випадку полягає в тому, що використовується нерівномірне квантування, і заданий KM визначає окремі кроки квантування для кожної просторової частоти. Потім у кожному r -му блоці:

$$D_q(r, k, l) = [D(r, k, l)/KK(k, l)], k = 0, \dots, 7; l = 0, \dots, 7, \quad (3.21)$$

$$\Delta D_q(r, k, l) = KK(k, l) D_q(r, k, l) - D(r, k, l), k = 0, \dots, 7; l = 0, \dots, 7, r = 1, \dots, R,$$

Формули (3.21) повинні використовуватися при розрахунках MSE у формулі (3.18). Очевидно, що різниця не є істотною для швидкості обчислень. Корируючі коефіцієнти, наведені в [248] і приблизно рівні 0,9 для обох кодерів, повинні бути враховані при прогнозуванні MSE для $AGU-M$ і $ADCT-M$. Крім того, припускається, що прогнозування MSE за допомогою діаграм розсіювання та підгонки кривої можна зробити для кодерів $AGU-M$ і $ADCT-M$ подібно до AGU і $ADCT$. По-друге, оскільки MSE можна спрогнозувати, $PSNR$ також можна

спрогнозувати. Це можна зробити опосередковано, шляхом прогнозування MSE. Однак прогноз можна зробити і безпосередньо. Прогнозований PSNR можна представити як:

$$PSNR_{pred} = 10\log_{10}(255^2 / (\frac{KK^2}{12}) f_0(P_0)) = 58,92 - 20\log_{10}(KK) + f_1(P_0), \quad (3.22)$$

де $f_1(P_0) = -10\log_{10}(f_0(P_0))$. Діаграму розсіювання наведено на рис. 3.12. Відповідно до (3.22) та якщо $P_0 < 0,6$, існує майже лінійне падіння $PSNR_{pred}$ з $\log_{10}(KK)$. Вважаємо, що таке прогнозування також можливе для інших метрик.

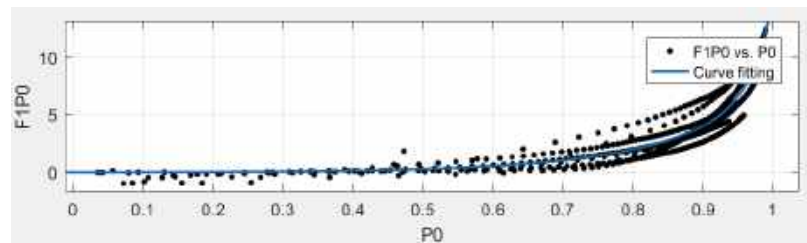


Рисунок 3.12 – Діаграма розсіювання $f_1(P_0)$ для зображень без шуму та підігнаної кривої, кодер ADCT

По-третє, наведені вище діаграми розсіювання були отримані для тестових зображень у градаціях сірого, які є типовими зображеннями оптичного або дистанційного зондування. Однак, дуже схожі діаграми розсіювання були отримані для медичних зображень сітківки (див. в Додатку К, на рис. К.6 і порівняйте діаграму розсіювання з діаграмами розсіювання в Додатку К, на рис. К.4).

В Додатку К, на рис. К.7 наведено результат стиснення кольорового зображення сітківки без видимих спотворень, тобто зберігається вся діагностично цінна інформація. Оскільки це зображення не має багато текстур і деталей, а також завдяки використанню 3D-версії кодера AGU, який використовує міжканальну кореляцію, досягнутий КС трохи перевищує 100. Це показує, наскільки великим може бути виграш в об'ємі медичних даних через їх розумне стиснення з втратами. Отже, підхід застосовний до різних типів зображень і є досить універсальним.

Нарешті, давайте проаналізуємо приклад для зображення природної сцени (тестове зображення Goldhill). В Додатку К, на рис. К.8 представлено оригінальне безшумне зображення та його стиснуті версії з $MSE=20, 40$ і 60 . Як видно, важко знайти візуальну різницю між зображеннями в Додатку К, на рис. К.8, а і К.8, б, тоді як спотворення помітні на рис. К.8, в і їх добре видно на рис. К.8, г. Це свідчить про доцільність використання розробленого підходу для різних застосувань.

Вже було згадано, що часто бажано спрогнозувати метрики візуальної якості. Щоб перевірити, чи можливо це, було отримано діаграму розсіювання для $MSE_{HVS-M}/(KK^2/12)$ від P_0 (Додаток К, рис. К.9). Як видно, це співвідношення становить приблизно $0,05$ для малих P_0 (це відбувається для малих KK та/або зображень складної структури), тобто PSNR-HVS-M приблизно на 13 дБ більше, ніж PSNR. Це означає, що внесені спотворення добре маскуються вмістом зображення і, швидше за все, їх неможливо помітити візуально. Різниця в PSNR-HVS-M і PSNR зменшується до $5...7$ дБ для $P_0 > 0,5$, тобто типових умов стиснення з втратами. Діаграма розсіювання та підібрана крива показують, що MSE_{HVS-M} можна добре спрогнозувати для заданого KK . Іншими словами, візуальні метрики також можна спрогнозувати за допомогою запропонованого підходу. Знову ж таки, сума двох експонент (саме цей випадок представлений в Додатку К, на рис. К.9) може слугувати апроксимаційною кривою з досить невеликою кількістю змінюваних параметрів.

Наведемо результати застосування запропонованого підходу до реальних гіперспектральних даних. Стиснення з втратами було застосовано до зображення набору даних сенсора Hyperion EO1H1800252002116110KZ. Сенсор Hyperion створює дані поганої якості (дуже шумні) в піддіапазонах з індексами $k=1,...,12$ і $k=58,...,76$. Зображення в цих піддіапазонах часто відкидаються під час аналізу, тому не стискали їх. Потім було порівняно два підходи до стиснення. Обидва передбачають покомпонентне стиснення. Перший був запропонований раніше [165]. Зображення стискаються після застосування перетворення, що стабілізує дисперсію (ПСД), яке враховує залежні від сигналу властивості шуму та перетворює цей шум на адитивний з дисперсією, приблизно рівною одиниці. Тоді

рекомендований $KK=3,5$ (це позначення використовується на рисунках нижче). Зворотне перетворення застосовується теж покомпонентно після декомпресії. Для запропонованого методу покомпонентні зображення були перетворені в інтервалі від 0 до 255. Потім для кожного з них застосовано кодер AGU з $KK=17$, що приблизно відповідає $PSNR_{des}=34,5$ дБ ($MSE_{des} \approx 24 \approx 17 \times 17 / 12$). Для відповідних даних використовується $KK=17$.

Отримані значення PSNR, розраховані між стисненими та оригінальними компонентними зображеннями, представлені в Додатку К, на рис. К.10. Як видно, PSNR для методу [165] у більшості піддіапазонів виявляється значно більшим, ніж встановлений $PSNR_{des}$. Лише в деяких піддіапазонах (з індексами 165...185), де вхідний PSNR є досить малим, визначені значення PSNR становлять близько 40 дБ (тобто внесені спотворення невидимі в декомпресованих зображеннях). Для запропонованого підходу PSNR для внесених втрат є значно меншим, але для всіх зображень піддіапазону PSNR все одно перевищує 35 дБ. Як впливає з аналізу даних в Додатку К, на рис. К.11, КС для всіх піддіапазонів перевищує 5 (більш детальне дослідження показує, що $P_0 > 0,6$ у всіх випадках). Таким чином, MSE менше, ніж $KK^2/12$, а наданий PSNR більший, ніж очікувалося. Основне спостереження щодо даних в Додатку К, на рис. К.11 полягає в тому, що КС для запропонованого методу в декілька разів більше, ніж для методу-прототипу майже для всіх піддіапазонів, за винятком діапазонів з малим вхідним PSNR. Таким чином, отримано істотний виграш в сенсі КС, а внесені спотворення залишилися непомітними.

Розглянемо тепер можливості тривимірного стиснення в групах. Є багато різних варіантів [165]. Було проаналізовано один із найпростіших, де компонентні зображення були перетворені до меж 8-бітного представлення, потім об'єднані в 4-канальні групи (позначені bl), а потім стиснені 3D-версією кодера AGU.

Після декомпресії зображення необхідно «розтягнути» компонентні зображення до початкових меж. Як і раніше, було використано $KK=17$. Для зручності порівняння отримані дані також представлені в Додатку К, на рис. К.10 та К.11, для 3D стиснення вони позначені як $KK=17$, $bl=4$.

Значення КС для 3D-випадку показані однаково для всіх компонентів однієї групи. Як видно, значення КС для тривимірного стиснення приблизно в два рази більші, ніж для запропонованого покомпонентного стиснення. Це очевидна перевага 3D-стиснення. Між тим, є також дуже цікаві спостереження, що впливають з аналізу даних для PSNR (Додаток К, рисунок К.10).

Як видно, існує багато піддіапазонів, для яких PSNR для 3D-стиснення є значно більшим (і внесені спотворення є значно меншими), ніж для покомпонентного стиснення. Значення PSNR майже однакові, якщо стискаються піддіапазони з малим вхідним значенням PSNR. Тут не наводяться приклади оригінальних і стиснених складових зображень, оскільки візуально вони ідентичні. Треба зауважити, що встановлення більшого $PSNR_{des}$ призводить до більшого PSNR внесених втрат і меншого КС для кожного компонентного зображення відповідно. Встановивши більший $PSNR_{des}$, можна гарантувати, що точність класифікації не погіршиться.

3.3 Адаптивне стиснення багатоканальних зображень

Розглянемо тепер вплив стиснення з втратами на класифікацію багатоканальних зображень ДЗ. Зокрема, зосередимося на випадку обмеженої кількості каналів, наприклад, кольорових, мультиполяризаційних або мультиспектральних зображень, через такі причини: (а) простіше продемонструвати ефекти, які мають місце на зображеннях через втрати стиснення та те, як ці ефекти впливають на класифікацію лише для випадку невеликої кількості компонентів зображення; (б) точна класифікація багатоканальних зображень з невеликою кількістю компонентів зазвичай є менш складним завданням, ніж класифікація гіперспектральних даних.

Повертаючись до переваг стиснення з втратами багатоканальних даних ДЗ, можна сказати наступне. По-перше, зображення, стиснуті з втратами, можуть бути класифіковані краще, ніж оригінальні зображення [252], [253]. Такий ефект зазвичай має місце, якщо оригінальні зображення зашумлені, а параметри кодера

налаштовані так, щоб ефект видалення шуму через стиснення з втратами був максимальним [165]. Однак, навіть якщо шум на зображеннях відсутній (або, точніше, пікові значення співвідношення сигнал/шум (PSNR) на вихідних зображеннях високі, а шум не видно на візуалізованих компонентних зображеннях), точність класифікації зображень може залишатися практично однаковою для відносно широкого діапазону варіації КС [252], [253], [254].

Ширина цього діапазону може бути різною. Вона залежить від наступних основних факторів:

- використовуваний класифікатор, наскільки він ефективний, як він навчений;
- використана техніка стиснення (кодер);
- наскільки добре розрізняються класи та які функції використовуються для їх опису;
- скільки класів присутні в багатоканальному зображенні;
- наскільки «складним» є розглянуте зображення (чи є воно високотекстурним, чи містить воно багато невеликих або протяжних об'єктів, який середній розмір фрагментів, що належать до кожного класу тощо).

Припустимо наступне:

- існують суворі залежності між спотвореннями через стиснення з втратами (які можна охарактеризувати певними метриками) та ТРСС для багатоканальних зображень ДЗ;
- маючи такі залежності, можна рекомендувати, які значення метрик необхідно надати, щоб забезпечити практично такий же ТРСС для стиснених даних, як і для вихідних даних;
- можна надати рекомендовані значення метрик, принаймні, для деяких сучасних методів стиснення на основі ДКП;
- порогові значення метрики можуть залежати від використовуваного класифікатора та властивостей багатоканального зображення, що підлягає стисненню з втратами;

– ефективність класифікатора також може залежати від того, як його навчено (наприклад, використання оригінального або стисненого багатоканального зображення).

Як було зазначено вище, існує багато підходів до класифікації багатоканальних зображень. Нижче розглянемо два з них. Перший базується на методі максимальної правдоподібності (ММП) [255], а другий спирається на навчання НМ [256]. В обох випадках вивчається попиксельна (pixel-wise) класифікація. Існує багато причин для використання піксельного підходу та саме цих класифікаторів: (а) для спрощення завдання класифікації та використання лише Q ознак ($q = 1, \dots, Q$), тобто значень певного багатоканального зображення в кожному пікселі; (б) показати проблеми попиксельної класифікації; (в) класифікатори на основі ММП та НМ вважаються одними з найкращих [256].

Зауважимо, що класифікатори даних ДЗ можуть бути навчені різними способами. Одним із варіантів є заздалегідь навчити класифікатор, використовуючи раніше отримані зображення, наприклад, нестиснуті, збережені для навчання чи інших цілей. Іншим варіантом є використання частини отриманого стисненого зображення для навчання класифікатора та його використання для класифікації всього зображення [257]. Можна очікувати, що результати класифікації будуть різними, а обидва варіанти мають як переваги, так і недоліки.

Функція класифікації (розпізнавання) для ММП заснована на обчисленні деякої метрики в просторі ознак:

$$\Phi = \begin{cases} 1, \rho(\vec{x}^*, \vec{x}_s) \geq tr \\ 0, \rho(\vec{x}^*, \vec{x}_s) < tr' \end{cases} \quad (3.23)$$

де $\vec{x}^*, \vec{x}_s$ є векторами ознак (атрибутів) для поточного та зразкового об'єктів (у цьому випадку пікселя зображення); $\rho(\blacksquare, \blacksquare)$ – використовувана метрика векторної подібності; tr позначає поріг прийняття рішення.

Як компоненти вектора ознак, загалом, можна розглядати як вихідні (піксельні (воксельні) значення багатоканального зображення), так і розраховані на їх основі «похідні» ознаки (наприклад, співвідношення цих значень).

Якщо $\Phi = 1$, то поточний об'єкт (піксель), визначений вектором ознак $\vec{x}^*$, відноситься до класу a_s .

Беручи до уваги стохастичний характер ознак $\vec{x}, \vec{x} \in R^c$, інформація про кожен клас об'єктів міститься в параметрах $\vec{\Theta}$ спільних ФЦЙ $f_c(\vec{x}; \vec{\Theta})$.

Ця інформація використовується для створення статистичних правил прийняття рішень

$$L = \frac{f_c(\vec{x}^*; \vec{\Theta} | a_2)}{f_c(\vec{x}^*; \vec{\Theta} | a_1)} \geq tr, \quad (3.24)$$

де L – відношення правдоподібності. Порогове значення tr визначається використаним статистичним критерієм; наприклад, байєсівський критерій забезпечує отримання оптимального класифікатора, коли доступна наступна інформація: ФЦЙ для всіх наборів шаблонів, ймовірності появи кожного класу та втрати, пов'язані з ймовірностями неправильної класифікації. У більшості практичних випадків для опису об'єкта використовується інформація статистичної вибірки, тобто статистичні оцінки ФЦЙ, отримані на етапі навчання, використовуються у відношенні правдоподібності L . Потім розмір вибірки, достовірність інформації про об'єкти спостереження та ефективність використання цієї інформації в правилі прийняття рішень в основному визначає якість прийнятих рішень.

Навчальні вибірки формуються за допомогою пікселів, що представляють заданий клас; зазвичай це деяка область (або області, набір пікселів) на зображенні, ідентифікована на основі деяких апріорних даних для зондованої місцевості. Основною вимогою до даних навчальних вибірок є їх репрезентативність — пікселі вибірки повинні відповідати одному класу на місцевості; такий клас повинен займати територію, яка досить добре представлена пікселями на зображенні з

заданою роздільною здатністю. При побудові багатоальтернативного правила прийняття рішень ($K > 2$) використовується критерій максимальної правдоподібності; поріг tr у рівнянні (3.24) приймається рівним 1. Критерій максимальної правдоподібності дає змогу усунути невизначеність рішення (коли жоден із класів не можна вважати кращим за інші), не вимагає знань априорних ймовірностей класів і функції втрат, дозволяє оцінити надійність рішень і може бути легко узагальнений на випадок багатьох класів. Відповідно до цього критерію вважається, що контрольна вибірка (виміряні значення ознак у поточному пікселі зображення) належить до класу $1 \leq u \leq K$, для якого функція правдоподібності (вірніше її оцінка, отримана на етапі навчання) є максимальним:

$$f_c(\vec{x}^*; \vec{\Theta} | a_u) = \max_{1 \leq k \leq K} \{f_c(\vec{x}^*; \vec{\Theta} | a_k)\} \Rightarrow \vec{x}^* \in a_u, \quad (3.25)$$

Якщо припущення щодо нормального закону розподілу вектора ознак вірні, метод максимальної правдоподібності рівняння (3.25) забезпечує оптимальне розпізнавання. Для підвищення точності апроксимації емпіричних негаусових розподілів ознак та врахування можливої наявності сильних кореляційних зв'язків доцільно застосовувати багатопараметричні закони розподілу на основі нелінійних перетворень нормальних випадкових величин. Ці перетворення включають перетворення Джонсона. Для опису класу спектральних ознак пропонується застосувати SB-розподіл Джонсона [258], [259]:

$$f(x) = \frac{\eta\lambda}{\sqrt{2\pi}(x-\varepsilon)(\varepsilon+\lambda-x)} e^{-\frac{1}{2}(\gamma+\eta \ln \frac{x-\varepsilon}{\varepsilon+\lambda-x})^2}, \quad (3.26)$$

де η та γ – два параметри форми, пов'язані з асиметрією та ексцесом ($\eta > 0$, $\lambda \in (-\infty; +\infty)$); ε і λ – два параметри масштабу випадкової величини x ($\varepsilon \in (-\infty; +\infty)$; $\varepsilon \leq x \leq \varepsilon + \lambda$).

Оцінки параметрів для ФЩЙ $\vec{\Theta} | a_k = \{\varepsilon_k, \lambda_k, \eta_k, \gamma_k\}$ можна знайти шляхом числової оптимізації функції втрат, що мінімізує інтегральну MSE представлення

емпіричного розподілу (гістограми) за допомогою теоретичної моделі в рівнянні (3.27).

Багатовимірний варіант рівняння (3.26) такий [258]:

$$f_c(\vec{x}|a) = (2\pi)^{-\frac{c}{2}} ||\Xi||^{-\frac{1}{2}} \prod_{v=1}^c \frac{\eta_v \lambda_v}{(x_v - \varepsilon_v)(\varepsilon_v + \lambda_v - x_v)} \times \\ \times \exp \left[-\frac{1}{2} \sum_{v,k=1}^c \Xi_{vk}^{-1} (\gamma_v + \eta_v \ln \frac{x_v - \varepsilon_v}{\varepsilon_v + \lambda_v - x_v}) (\gamma_k + \eta_k \ln \frac{x_k - \varepsilon_k}{\varepsilon_k + \lambda_k - x_k}) \right], \quad (3.27)$$

де c – розмір вектора ознак $\vec{x}$; Ξ – вибіркова кореляційна матриця.

Після процесу створення та оцінки навчальних вибірок пікселі зображення сортуються (пов'язуються) з класами на основі вирішального правила. У разі застосування попиксельної класифікації дані для кожного пікселя з координатами (i, j) обробляються окремо (незалежно). Вектор значення для поточного пікселя $\vec{x}_{ij}^*$ поміщається в математичні моделі цільового класу в рівнянні (3.24). Отримані результати $\hat{f}_c(\vec{x}_{ij}^*)|a_k, k = 1 \dots K$, порівнюють між собою та вибирають максимальну оцінку функції правдоподібності; його номер – це номер класу, до якого відноситься поточний піксель (i, j) .

На сьогоднішній день існує багато НМ і підходів до їх навчання та оптимізації [233]. Обрано просту, але ефективну НМ для застосування – повнозв'язаний персептрон, що працює у поєднанні з самоорганізованою картою для обробки отриманих вагових коефіцієнтів або ймовірностей приналежності пікселя до одного з класів. Деталі архітектури такі: вхідний рівень для вхідних даних триколірних компонентів, один прихований нейронний шар із 30 повністю підключеними нейронами (ця кількість показала однакову ефективність із більшою можливою кількістю нейронів прихованого шару). Для прихованого шару використовується функція активації гіперболічного тангенсу. Як функція глобального навчання використовується зворотне розповсюдження масштабованого спряженого градієнта. Вихідний шар використаної НМ є самоорганізованою картою, повністю пов'язаною з нейронами прихованого шару.

Він забезпечує відображення вихідних ймовірностей рішення на список класів у вигляді векторів. Для навчання НМ було взято 70% створеного набору, інші 30% даних використовуються для перевірки. Запропонований класифікатор простий у використанні та швидкий (що є важливим моментом для застосування класифікації до даних дистанційного зондування). Загальний процес навчання повторювався 100 разів з повною перестановкою набору даних, і отриманий класифікатор був застосований до тестових зображень.

Створюючи тестове багатоканальне зображення, врахували наступні типові властивості класів, які можуть мати місце в реальних зображення ДЗ:

властивість 1: статистичні характеристики ознак зрідка мають розподіли, близькі до гаусового; оскільки в цьому випадку використаними ознаками є значення кольорових компонентів (див. тестове кольорове зображення та його компоненти в Додатку К, на рис. К.12, а - К.12, г відповідно), то ця властивість відноситься до них;

властивість 2: об'єкти, що відносяться до певного класу, рідко бувають абсолютно однорідними (навіть водну поверхню зазвичай не можна вважати абсолютно однорідною); багато факторів призводять до різноманітності (варіацій) значень пікселів, як-то варіації фізичних і хімічних характеристик поверхні, що відбиває або випромінює, шум і т.д.;

властивість 3: об'єкти «перекриваються» (перетинаються) у просторі ознак; саме ця властивість зазвичай призводить до неправильної класифікації, особливо якщо таке перекривання має місце для багатьох ознак;

властивість 4: кількість пікселів, що належать до різних класів у кожному зображенні ДЗ, може сильно відрізнятись, може бути так, що відсоток пікселів, що належать до одного класу, становить менше 1%, тоді як для іншого класу можуть бути десятки відсотків відповідних пікселів;

властивість 5: в одному пікселі можуть бути об'єкти двох і навіть більше класів [233], [260], що вимагає застосування методів декомпозиції, але не розглядаємо такі (складні) ситуації в подальших дослідженнях.

Покажемо, що більшою чи меншою мірою включено ці властивості в тестове зображення (Додаток К, рисунок К.12). В Додатку К, таблиця К.4 містить дані про сім класів, гістограми для них та апроксимації за допомогою SB-розподілу Джонсона.

Клас 1 можна умовно трактувати як «Дороги», клас 2 – як сільськогосподарські поля (Поле-У), клас 3 – як сільськогосподарські поля іншого кольору (Поле-Р), клас 4 – як «Вода», клас 5 – як «Трава», клас 6 – як «Дерева», клас 7 – як якийсь дуже текстурний або неоднорідний клас, як міська зона чи виноградник («Текстура»).

Як бачимо, властивість 1 (негаусів розподіл) виконується практично для всіх окремих розподілів. Властивість 2 (внутрішня неоднорідність) має місце для всіх класів, особливо для класів 2 і 7. Властивість 3 (перетин ознак) має місце для багатьох класів — усі три ознаки для класів 1 і 2, усі три ознаки для класів 5 і 6; ознаки класу 7 практично перетинаються з усіма іншими класами (перетин ознак також можна побачити з порівняння зображень компонентів, де деякі класи важко розрізнити, наприклад, класи 2 і 4 у зеленому компоненті, рис. 2, с). Властивість 4 також спостерігається — є невеликий відсоток пікселів, що належать до класу 1, у той час як досить багато пікселів належать до класів 4, 5 і 7. На рис. 3.14, а представлена справжня карта класів, де клас 1 показано світло-коричневим, клас 2 — жовтим; 3-й клас — червоним, 4-й клас — синім; 5-й клас — світло-зеленим, 6-й клас — темно-зеленим, а 7-й клас — білим.

Класифікатор ММП, застосований до нестисненого зображення (Додаток К, рис. К.13, а), створює карту класифікації, наведену на рис. К.13, б. Як бачимо, помилкових класифікацій досить багато. Класи 4 і 6 класифікуються найкращим чином (відповідні імовірності правильної класифікації P_{44} і P_{66} дорівнюють 0,85 і 0,82), тоді як клас 5 визначається найгіршим чином ($P_{55} = 0,19$). Класи 1 і 7 розпізнаються погано (P_{11} і P_{77} дорівнюють відповідно 0,48 і 0,42).

Класи 2 і 3 характеризуються $P_{22} = 0,67$ і $P_{33} = 0,77$ відповідно. Досить низькі ймовірності класифікації можна пояснити кількома факторами. Деякі з них уже згадувалися вище — багато перетинів у просторі ознак. Ще одна можлива причина

полягає в тому, що класифікатор ММП базується на апроксимаціях розподілу, і ці апроксимації можуть бути не ідеальними (див. гістограми та їх апроксимації в таблиці 3.15). Досить багато пікселів помилково віднесено до класу 7 («Текстура»), для якого розподіли дуже широкі та перетинаються з розподілами для інших класів.

В Додатку К, на рис. К.12, де представлено зображення, стиснуті за умови $\text{PSNR-HVS-M} \approx 36$ дБ і $\text{PSNR-HVS-M} \approx 30$ дБ. У цих випадках значення КС становлять приблизно 4,5 і 8,9 відповідно (для покомпонентного стиснення). Для випадку $\text{PSNR-HVS-M} \approx 36$ дБ помітні внесені спотворення [245]. Вони здебільшого проявляються як згладжування варіацій у квазіоднорідних областях (див. фрагменти для класу 6 (темно-синій) для зображень в Додатку К, на рис. К.12, а,е). Ефекти такого придушення шуму або високочастотних коливань через стиснення з втратами відомі для стиснення з втратами [242], [261], і вони можуть навіть мати позитивний ефект для класифікації (за певних умов). Для випадку $\text{PSNR-HVS-M} \approx 30$ дБ спотворення через стиснення з втратами (див. стиснене зображення в Додатку К, на рис. К.12, е) проявляються у згладжуванні варіацій і розмитті меж/об'єктів (це добре видно для класу 1). Імовірності правильної класифікації для класів наведені в Додатку К, в таблиці К.5. Зверніть увагу, що класифікатор ММП був «навчений» з використанням наближень розподілу, отриманих для нестисненого триканального зображення. Як видно, ймовірності для окремих класів по-різному залежать від якості стисненого зображення та ступеня стиснення. P_{11} неухильно зменшується і стає близьким до нуля. Це відбувається з двох причин. По-перше, функції для класу 1 достатньою мірою перетинаються з ознаками для класу 2. По-друге, через стиснення з втратами розподіл ознак після стиснення відрізняється від розподілу ознак для нестисненого зображення, де ця різниця збільшується зі зменшенням PSNR-HVS-M і збільшенням КС. Це ілюструють розподіли, представлені в Додатку К, таблиця К.6 для класу 1 і, частково, в таблиці К.7 для класу 2.

Аналіз даних таблиці К.5 (Додаток К) також свідчить про наступне. Є деякі класи (Класи 3–6), для яких більший КС і менший PSNR-HVS-M (що відповідає більш достатньому рівню стиснення з втратами) призводить до зменшення

ймовірності правильної класифікації (проаналізуйте дані в стовпцях для P_{33} , P_{44} , P_{55} і P_{66}). Водночас є класи, для яких вірогідність правильної класифікації зростає — див. дані для P_{77} . Це пов'язано зі зменшенням кількості неправильних класифікацій в області текстури (див. карти в Додатку К, на рис. К.13, в,г). Існує також один клас (Клас 2), для якого P_{22} є найбільшим для PSNR-HVS-M приблизно 40 дБ (через ефект фільтрації шуму), але він зменшується для меншого PSNR-HVS-M (більший КС).

Одна ймовірна причина, чому класифікатор ММП, навчений для нестиснених зображень і застосований до стиснених, втрачає ефективність, полягає в тому, що розподіли ознак, спостережувані для стиснутих даних, відрізняються від тих, що спостерігаються для вихідних даних ДЗ і використовуються для навчання класифікатора. Тоді можна очікувати, що класифікатор ММП має бути навчений для стиснених зображень. Було проведено відповідне дослідження, результати представлені в табл. 3.12. Висновок, який випливає з аналізу даних у табл. 3.12, може здатися тривіальним — необхідно навчити класифікатор ММП для стиснених зображень з однаковими характеристиками стиснення.

Дотримання цього правила призводить до значного підвищення точності класифікації. Наприклад, якщо вихідні дані використовуються під час навчання, а потім класифікатор ММП застосовується до стиснутих даних, що забезпечує PSNR-HVS-M = 42 дБ, тоді TPCC $P_{cc} = 0,501$. У той же час, якщо в навчанні використовувалося стиснене зображення, що забезпечує PSNR-HVS-M = 42 дБ, P_{cc} радикально зростає і стає рівним 0,579. Ще більш дивовижні результати спостерігаються для випадків стиснення зображень, які відповідають PSNR-HVS-M, що дорівнює 36 і 30 дБ. P_{cc} для відповідного навчання досягає 0,597 і 0,527 відповідно. Зверніть увагу, що $P_{cc} = 0,627$ для оригінального зображення, класифікованого за допомогою навчання для вихідного зображення.

Тоді виявляється, що для зображень, стиснутих за допомогою PSNR-HVS-M, що дорівнює приблизно 40 дБ, спостерігається майже такий самий P_{cc} , якщо їх навчити для відповідних стиснутих зображень. Класифікаційні карти, отримані для цієї методики навчання, представлені на рис. 3.13, а,б.

Таблиця 3.12 –Загальна ймовірність правильної класифікації P_{cc} для різних зображень залежно від «навчальних» даних

Зображення, що використовуються для навчання		Загальна ймовірність правильної класифікації для стиснутих зображень		
		PSNR-HVS-M=42 дБ	PSNR-HVS-M=36 дБ	PSNR-HVS-M=30 дБ
0	Оригінал (нестиснений)	0,501	0,434	0,442
1	Стиснений із забезпеченням PSNR-HVS-M = 42 дБ	0,579	0,566	0,563
2	Стиснений із забезпеченням PSNR-HVS-M = 36 дБ	0,519	0,597	0,599
3	Стиснутий із забезпеченням PSNR-HVS-M = 30 дБ	0,408	0,446	0,527

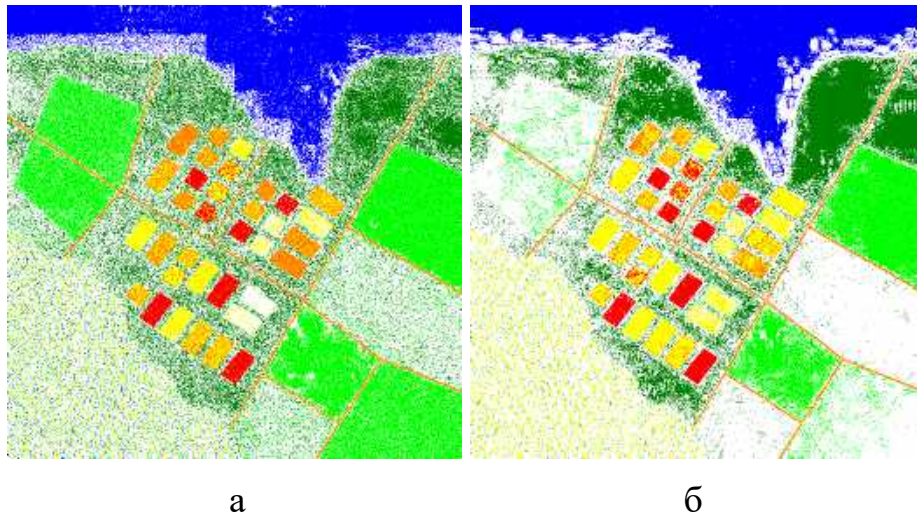


Рисунок 3.13 –Карти класифікації для стисненого зображення із забезпеченням PSNR-HVS-M = 36 дБ з використанням класифікатора ММП, навченого для того самого стиснутого зображення (а); стиснене зображення із забезпеченням PSNR-HVS-M = 30 дБ за допомогою класифікатора ММП, навченого для того самого стисненого зображення (б).

Їх можна порівняти з картами в Додатку К, рис. К.13, в,г відповідно. Це порівняння чітко демонструє, що навчання для відповідних стиснених даних є кращим. Багато класів розпізнаються помітно краще. Один цікавий момент полягає в тому, що багато неправильних класифікацій можуть з'являтися в околах гострих

меж у багатоканальних зображеннях (біля меж класу 4). Це пояснюється тим, що такі межі розмиваються стисненням із втратами, і з'являються пікселі з «проміжними значеннями», які можна назвати «неправильними класами». Тепер проаналізуємо дані для класифікації оригінального триканального зображення за допомогою навчання НМ, навченої для цього зображення. Отримана матриця плутанини представлена в таблиці 3.13. Відповідна карта наведена у Додатку Б на рис. Б.1, а. Як бачимо, добре розпізнаються декілька класів: Поле-Р (клас 3), Вода (клас 4), Деревя (Клас 6). Існує багато неправильних класифікацій для класів «Дороги» (з класом Поле-У, який має практично ті самі кольори) і «Текстура» (його пікселі можуть бути помилково пов'язані з класами «Вода», «Деревя», «Трава» та, рідко, Поле-У). Результати відносно погані для класу 2 і класу 5. Таким чином, навіть для цього, майже ідеального, випадку (класифіковане зображення те саме, що використовувалося для навчання, застосовано класифікатор на основі НМ, здатний працювати з негаусовим розподілом ознак), результати класифікації далекі від досконалості.

Таблиця 3.13 – Матриця плутанини для нестисненого зображення, класифікована НМ, навченою для цього зображення.

Клас	Ймовірність рішення						
	Дороги	Поле-У	Поле-Р	Вода	Трава	Деревя	Текстура
Дороги	0,356	0,645	0	0	0	0	0
Поле-У	0,109	0,608	0,002	0	0	0	0,281
Поле-Р	0	0,004	0,996	0	0	0	0
Вода	0	0	0	0,986	0	0	0,014
Трава	0	0	0	0	0,744	0,245	0,012
Деревя	0	0	0	0	0,072	0,923	0,005
Текстура	0	0,03	0	0,26	0,116	0,153	0,441

Результати для НМ-класифікатора наведені в Додатку Б. Висновки, що з них випливають, аналогічні тим, що наведені вище для ММП-класифікатора.

Підсумовуючи отримані результати та висновки, можна констатувати наступне:

- класифікація будь-яким класифікатором не є ідеальною, якщо ознаки перетинаються в просторі ознак та/або якщо використовуються не всі можливі ознаки;
- якщо класифікатор навчено для нестиснутих даних, а потім застосовано до стиснених зображень, тоді більше стиснення (більший КС, менший PSNR або PSNR-HVS-M) призводить загалом до погіршення класифікації; при цьому різниця в P_{cc} і ймовірності правильної класифікації для окремих класів може бути незначною; точність класифікації для деяких конкретних класів може навіть підвищитися завдяки придушенню шуму;
- варто навчати класифікатор (принаймні, класифікатори на основі MMM та НМ) для тих «умов», у яких він буде застосований; ці умови можна описати значеннями показників якості, які планується надати на етапі стиснення, або значенням КК, яке буде використано;
- різні класифікатори можуть давати достатньо різні класифікаційні карти для тих самих даних ДЗ; кількісні параметри, що характеризують класифікацію, також можуть значно відрізнятися; хорошим прикладом є дані для класу 6 — класифікатор ММП не розпізнає їх достатньо добре, тоді як НМ-класифікатор працює краще, а ситуація для класу 7 є протилежною;
- стиснення з втратами негативно впливає на класи, які в основному пов'язані з невеликими та протяжними об'єктами; причина в тому, що стиснення з втратами розмиває такі типи об'єктів і змінює розподіл ознак; у цьому сенсі використання метрик візуальної якості та встановлення бажаних значень для них на етапі стиснення може бути корисним, щоб уникнути значного погіршення точності класифікації для таких класів;
- для попиксельної класифікації неправильні класифікації зазвичай виглядають як ізольовані пікселі; якщо спотворення через стиснення є великими, неправильна класифікація може виглядати як групи пікселів, особливо в околах

висококонтрастних меж на зображеннях ДЗ; це небажано, і варто уникати таких ситуацій, які здебільшого трапляються, якщо спотворення через стиснення з втратами є видимими (якщо PSNR-HVS-M = 30 дБ і, рідше, якщо PSNR-HVS-M = 36 дБ) ;

– це наводить на думку, що варто проводити стиснення з втратами або без візуально помітних спотворень, або з спотвореннями, пов'язаними з частковим придушенням шуму в квазіоднорідних областях зображення; цим можна керувати, встановивши бажаний PSNR-HVS-M (≈ 40 дБ) або відповідний ПКС для використовуваного кодера (наприклад, КК для AGU).

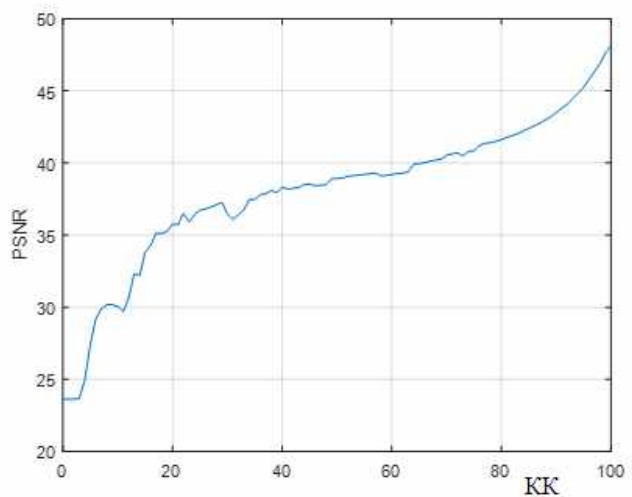
3.4 Аномальні (дивні) зображення в дистанційному зондуванні, їх властивості

Нещодавні дослідження [262], [263] продемонстрували, що КШС можуть не бути монотонними. Зображення, на яких це відбувається, було названо дивними. Дослідження вперше були проведені для зображень у градаціях сірого для кодера AGU [175], для якого було виявлено існування дивних зображень. Потім була виявлена можлива немонотонність КШС для JPEG і кольорових зображень. На самому початку здавалося, що немонотонність КШС може спостерігатися для зображень штучного походження, які містять великі області білого кольору. Однак пізніше було показано [262], що дивні зображення можуть мати природне походження, наприклад, кольорові зображення, отримані в умовах поганого освітлення (вночі). Немонотонність можна спостерігати не тільки для залежностей MSE або PSNR від ПКС, а й для метрик візуальної якості. Також встановлено, що не стільки кодер, скільки властивості зображення визначають, чи є дивним дане зображення чи ні. Як зазначено вище, розглядалися лише штучні та природні (звичайні) зображення сцени. Питання в тому, чи можуть зображення ДЗ бути дивними і які основні властивості таких зображень? Частково відповісти на ці питання є метою цього дослідження. Нижче показано, що дивні зображення ДЗ існують і вони мають специфічні властивості, пов'язані зі складністю зображення.

Варто навести приклади. Зображення (рис. 3.14, а) не є нетиповим, схожі зображення сцени можна знайти в базах зображень і на сайтах. Його особливість полягає в тому, що воно містить велику темну однорідну область. Як видно з рис. 3.14, б, залежність PSNR від КК не є монотонною. Вона має досить багато локальних мінімумів і максимумів. З аналізу випливає, що, наприклад, доцільніше стискати розглянуте зображення за допомогою КК=29, ніж КК=31, оскільки в першому випадку одночасно забезпечується більший КС і кращий PSNR (для КК=29, PSNR=37,26 дБ і КС=29,12, тоді як для КК=31, PSNR=36,08 дБ і КС=28,35). Як видно, різниця в PSNR перевищує 1 дБ і таку різницю можна візуально виявити при порівнянні двох стиснених зображень.



а



б

Рисунок 3.14 –Приклад дивного зображення (а) і КШС PSNR(КК) для JPEG (б) для цього зображення

Тоді виникає питання, яке зображення можна вважати дивним? Першою формальною відповіддю може бути наступне. Якщо припустити, що КШС монотонно спадає, то зображення є дивним, якщо існує принаймні одне i , для якого:

$$Metr(i - 1) < Metr(i) \cap Metr(i) > Metr(i + 1), \quad (3.28)$$

або, простіше, $Metr(i) > Metr(i - 1)$.

В свою чергу, якщо передбачається, що КШС монотонно зростає, то зображення можна вважати дивним, якщо існує принаймні одне i , для якого:

$$Metr(i - 1) > Metr(i) \cap Metr(i) < Metr(i + 1), \quad (3.29)$$

або, простіше, $Metr(i) < Metr(i - 1)$, тобто якщо існує хоча б один локальний мінімум.

Тут i – індекс масиву значень КШС, який використовується в аналізі. Наприклад, для JPEG він збігається зі значеннями КК, які є цілими числами від 1 до 100. Якщо ПКС є не КК, а іншим параметром, то виникає питання, яку кількість зразків I ($i=1, \dots, I$) треба аналізувати для даної КШС. Більший I дозволяє виконувати більш ретельний аналіз, але вимагає більшої кількості операцій стиснення зображення, декомпресії та метричної оцінки. Є кодери, для яких ПКС може приймати будь-яке значення в певному діапазоні. Прикладами є значення БНП для JPEG або SPIHT, які можуть змінюватися від 0 до 8 для зображень у відтінках сірого, представлених як 8-бітні 2D-дані. Для кодерів на основі ДКП, керованих КК, таких як, наприклад, кодер AGU [175], мінімальний КК прагне до 0, тоді як максимальне значення, як правило, не обмежене. Через це доцільно розглядати значення КК у такому діапазоні, щоб стиснуті зображення не були повністю спотворені (спотворення не мають занадто дратувати). Таким чином, проаналізували дані для КК від 1 до 100, використовуючи цілі значення (хоча, загалом, можна застосовувати не цілі значення КК).

Тут варто звернути увагу на наступне. Аналізуючи КШС, дослідники часто отримують їх, використовуючи розріджені значення ПКС, наприклад, встановлюючи КК рівним 5, 10, 15, ..., 100. У такому випадку існує велика ймовірність того, що КШС може здаватися монотонною, а дивність зображення не спостерігається (не виявлено).

Повернемося до визначення дивних зображень. Визначення може бути суворішим. Припустимо, що КШС зменшується. Тоді зображення можна вважати дивним, якщо виконується хоча б одне з умов:

$$\Delta M = \text{Metr}(i) - \text{Metr}(i - 1) > \delta. \quad (3.30)$$

Подібним чином, для збільшення КШС, зображення може розглядатися як дивне, якщо є принаймні один i , який:

$$\Delta M = \text{Metr}(i) - \text{Metr}(i - 1) < -\delta. \quad (3.31)$$

Тут δ — попередньо встановлений поріг, який показує, що несподіваний «стрибок» є значним. Наприклад, для метрики PSNR δ можна встановити приблизно 0,3-0,5 дБ. Дійсно, якщо знайдений ΔPSNR (3.30) для КШС PSNR(KK) дорівнює, скажімо, 0,01 дБ, це не надто проблематично на практиці.

Таким чином, можна розглянути дві практичні ситуації: 1) зображення формально дивне (ФД), але абсолютне значення найбільшого знайденого ΔM мале (умови (3.30) або (3.31) не виконуються); 2) зображення строго дивне (СД), це має місце, якщо виконуються умови (3.30) або (3.31).

Нагадаємо, що головний інтерес – зображення ДЗ. Через це вибрали чотири набори даних триканальних (кольорових) зображень із вільно доступної бази даних UC Merced Land Use Dataset (<http://weegee.vision.ucmerced.edu/datasets/landuse.html>). Кожен отриманий набір даних містить 31 зображення та набори даних мають назви Agricultural, Airplane, Beach, Dense Residential. В Додатку К, на рис. К.14 представлено по чотири маленькі копії типових зображень для кожного набору даних. У датасеті Beach знайдено досить багато (двадцять одне) СД-зображень, ще чотири формально дивні. У наборі даних Dense Residential не знайдено дивних зображень.

Навіть візуальний огляд зображень в Додатку К, на рис. К.14 пояснює, чому це так. Набір даних Agricultural містить декілька зображень із досить великими

однорідними областями (наприклад, крайнє праве зображення у відповідному рядку). Зображення в наборі Airplane і, особливо, в Dense Residential більш неоднорідні. Нарешті, зображення в наборі Beach зазвичай містять дві великі (квазі)однорідні області, які відповідають поверхні піску та води. Зауважте, що крайнє ліве зображення у відповідному рядку не є дивним, тоді як два крайніх правих є СД.

Вище навели опис властивостей деяких зображень ДЗ. При цьому складність зображення можна охарактеризувати кількісно. В роботі [264] представлено п'ять параметрів, які можуть характеризувати складність зображення та позначаються як ентропія (E), співвідношення країв (ER), контраст (C), кореляція (CO) та енергія (EN). Параметри E, ER та C мають менші значення для простіших структурних зображень, тоді як параметри CO та EN мають протилежні властивості [264]. Ідея (припущення) полягає в тому, що вищезазначені параметри можна співвіднести з ймовірністю того, що зображення буде дивним. Щоб перевірити цю ідею, було розраховано параметри для зображень у наборах даних. Для прикладу наведемо частину даних для п'ятнадцяти перших зображень у наборі Agricultural. Вони наведені в таблиці 3.14.

Як видно, параметри мають різні діапазони варіації. Межі для ентропії - від 5,54 до 7,22 (фактично від 5,1 до 7,5 для всіх 124 розглянутих зображень), ER змінюється в межах від 0,14 до 0,40 (від 0,07 до 0,41 для всіх зображень), C від 0,07 до 2,25 (від 0,02 до 2,25) для всіх зображень), CO змінюється від 0,11 до 0,95 (від 0,05 до 0,99 для всіх зображень), а EN змінюється від 0,06 до 0,31 (від 0,06 до 0,44 для всіх зображень). Спостерігаються такі основні властивості - дивні зображення (які зазвичай мають досить просту структуру та більший відсоток пікселів, які належать до однорідних областей) зазвичай характеризуються меншими значеннями E, ER та C, тоді як CO та EN для них мають більші значення. Однак ця залежність не є жорсткою. Розраховали ранговий коефіцієнт кореляції Спірмена (SROCC) між параметрами та числовим представленням дивності зображення (0 для звичайних зображень, 1 для ФД-зображень і 2 для СД-зображень) для трьох наборів даних (нагадаємо, що для набору даних Dense Residential немає СД і ФД

зображень, тому кореляцію для нього визначити неможливо). Результатом є те, що існує досить велика кореляція дивності та контрастності зображення (C), ентропії (E) та енергії (EN). Інші два параметри демонструють високу кореляцію (великі абсолютні значення SROCC) лише для окремих наборів даних – див. дані в таблиці 3.15.

Таблиця 3.14 – Приклад значень параметрів для набору зображень

Зображення	E	ER	C	CO	EN	Дивне чи ні
Img00	6,4324	0,3530	0,4345	0,6235	0,2194	ТАК/НІ
Img01	6,2656	0,3562	0,4256	0,5284	0,2287	НІ
Img02	6,3933	0,3537	0,4558	0,5644	0,2027	НІ
Img03	6,1998	0,3901	0,3953	0,5045	0,2571	НІ
Img04	6,8800	0,1361	0,3637	0,8594	0,1518	НІ
Img05	6,9877	0,3997	1,3106	0,3825	0,0856	НІ
Img06	6,8322	0,4031	1,2088	0,3352	0,1006	НІ
Img07	6,8351	0,3936	1,7582	0,1097	0,0901	ТАК /НІ
Img08	6,9941	0,3868	1,7872	0,2861	0,0720	НІ
Img09	7,2211	0,3958	2,2514	0,2974	0,0616	НІ
Img10	6,6044	0,1987	0,0767	0,9473	0,2759	ТАК
Img11	5,5384	0,1577	0,4204	0,1340	0,3065	ТАК /НІ
Img12	6,5364	0,1982	0,0699	0,9470	0,2930	ТАК
Img13	6,4552	0,3876	1,1089	0,1444	0,1278	НІ
Img14	6,9084	0,4041	1,6922	0,3034	0,0824	НІ

Аналіз демонструє наступне. Немає жодного параметра, що характеризує складність зображення, який дозволяє надійно виявляти дивні зображення до їх стиснення шляхом встановлення деякого порогу. Однак існує досить тісний зв'язок між параметрами, що описують складність зображення, і його дивністю, які потенційно можуть бути використані для виявлення.

Було побудовано дві залежності для дивного зображення Beach16. Вони наведені в Додатку К, на рис. К.15. КІШС PSNR(KK) (рис. К.15, а) використовувалася в попередньому аналізі для виявлення дивних зображень. Як

видно, ця КІШС має декілька локальних мінімумів і максимумів, що підтверджує дивність цього зображення. На рис. К.15, б представлено КІШС MDSI(KK), де MDSI є індексом подібності середнього відхилення [99]. Вона є однією із найкращих метрик візуальної якості, яка має властивість бути меншою для кращої візуальної якості.

Як видно, згідно з КІШС на рис. К.15, б, Beach16 є дивним зображенням також згідно з традиційним PSNR. Крім того, найбільші коливання MDSI спостерігаються для того ж інтервалу варіації KK, що і для PSNR – для KK близько 10.

Таблиця 3.15 – Коефіцієнти рангової кореляції Спірмена між дивністю зображення та параметрами, що характеризують складність зображення

Набір зображень	C	CO	ER	EN	E
Agricultural	-0,6057	0,4971	-0,3942	0,5849	-0,3274
Airplane	-0,4270	0,1586	0,0366	-0,0366	-0,1830
Beach	-0,7882	0,6660	-0,8004	0,6110	-0,6220
Всі зображення	-0,7154	0,3387	-0,1336	0,6159	-0,6786

Було показано [262], що зображення може бути дивним не лише якщо воно стиснене за допомогою JPEG, але й за допомогою інших кодерів. Через це було вирішено провести більш ретельне дослідження для кодера AGU [175].

Якщо зображення є триканальним, то AGU можна застосовувати покомпонентно або використовувати 3D-версію AGU, розроблену для стиснення багатоканальних (зокрема, гіперспектральних) зображень. Нижче використано перший варіант, оскільки цікаво, чим відрізняються компоненти триканальних зображень у сенсі їх дивності. Оскільки дослідження зосереджено на дивних зображеннях, то було розглянуто ті самі 31 зображення з набору даних Beach.

Для кожної складової визначено наступні параметри отриманих залежностей PSNR від KK: 1) кількість локальних максимумів; 2) значення метрики Nom фонового вмісту [265], яка є ще однією метрикою, що характеризує складність

зображення. Цей показник має більші значення для зображень простішої структури. Крім того, представлено кількість локальних максимумів для PSNR, визначених для компонента Y. Отримані дані представлені в таблиці 3.16.

Їх аналіз показує декілька цікавих тенденцій. По-перше, є недивні зображення, для яких усі КШС (для кожної з трьох компонентів і в сукупності) є монотонними (наприклад, beach00.tif). У той же час є також зображення, для яких деякі часткові КШС є монотонними, тоді як інші містять локальний максимум або локальні максимуми (див. дані для зображень beach06.tif і beach21.tif).

Нарешті, є зображення, для яких спостерігаються локальні максимуми для всіх компонентів (наприклад, зображення beach12.tif або beach16.tif).

По-друге, кількість максимумів зазвичай різна для КШС, отриманих для різних компонентів. Для компонента Y кількість локальних максимумів приблизно така ж, як і для КШС для компонентів R, G і B. Приклад чотирьох залежностей наведено на рис. 3.15, де видно декілька максимумів для всіх компонентів.

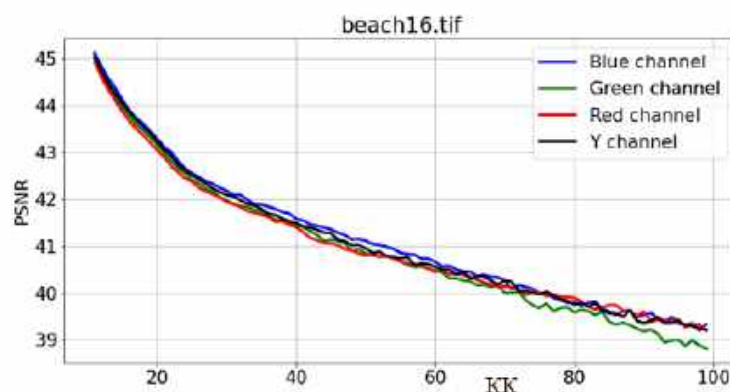


Рисунок 3.15 –Залежності PSNR від КК для кодера AGU, застосованого покомпонентно до зображення beach16

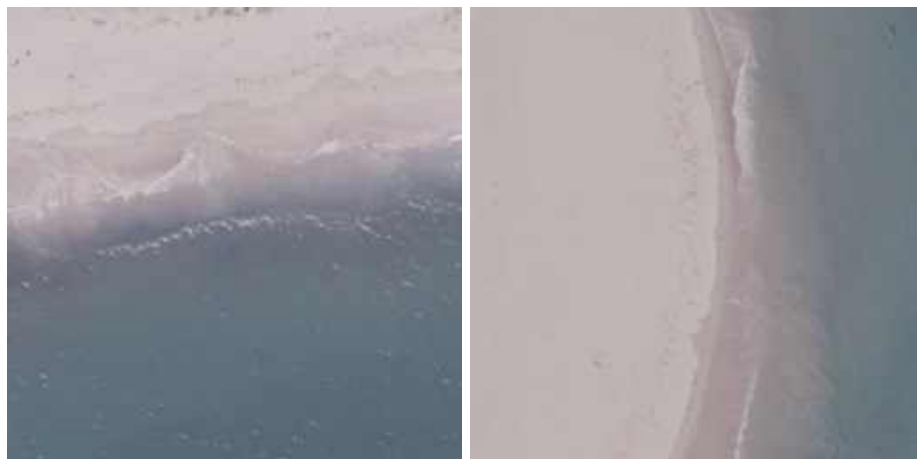
По-третє, дивні зображення, отримані для JPEG, здебільшого ті ж самі, як і для кодера AGU. У першу чергу це стосується зображень beach11-beach17. Зображення beach14 і beach16 показані на рис. 3.16. Як видно, обидва зображення містять в основному дві квазіоднорідні ділянки водної поверхні та пляжного піску.

Таблиця 3.16 – Статистичні характеристики для залежностей PSNR від значень метрики Ном та КК для кодера AGU, застосованих покомпонентно до триканальних зображень набору Beach

Зображення	Кількість локальних максимумів				НОМ			
	R	G	B	Y	R	G	B	Y
beach00.tif	0	0	0	0	0,166	0,176	0,173	0,174
beach01.tif	0	0	0	0	0,141	0,149	0,14	0,146
beach02.tif	0	0	0	0	0,106	0,122	0,088	0,117
beach03.tif	0	0	0	0	0,099	0,111	0,091	0,104
beach04.tif	0	0	0	0	0,097	0,105	0,074	0,095
beach05.tif	0	0	0	0	0,123	0,142	0,145	0,142
beach06.tif	0	1	1	1	0,2	0,187	0,163	0,176
beach07.tif	3	0	2	2	0,192	0,182	0,163	0,178
beach08.tif	1	2	2	3	0,158	0,155	0,146	0,152
beach09.tif	0	0	0	0	0,175	0,167	0,137	0,16
beach10.tif	0	0	0	0	0,181	0,173	0,146	0,166
beach11.tif	3	3	4	4	0,326	0,3	0,241	0,283
beach12.tif	4	2	2	2	0,303	0,284	0,214	0,259
beach13.tif	11	6	8	5	0,306	0,291	0,213	0,276
beach14.tif	5	8	11	11	0,305	0,276	0,232	0,261
beach15.tif	4	9	3	4	0,314	0,289	0,275	0,288
beach16.tif	15	18	15	17	0,394	0,359	0,337	0,361
beach17.tif	6	6	7	9	0,316	0,285	0,26	0,28
beach18.tif	15	14	14	13	0,313	0,286	0,252	0,276
beach19.tif	1	0	2	0	0,194	0,219	0,257	0,188
beach20.tif	0	1	2	1	0,246	0,253	0,294	0,232
beach21.tif	0	1	0	0	0,229	0,253	0,254	0,245
beach22.tif	0	0	2	0	0,245	0,255	0,266	0,245
beach23.tif	1	0	1	0	0,238	0,282	0,237	0,257

Продовження таблиці 3.16

Зображення	Кількість локальних максимумів				НОМ			
	R	G	B	Y	R	G	B	Y
beach24.tif	3	0	4	1	0,233	0,28	0,252	0,248
beach25.tif	3	2	4	1	0,232	0,217	0,198	0,211
beach26.tif	2	2	4	3	0,271	0,259	0,219	0,246
beach27.tif	0	0	0	0	0,246	0,235	0,225	0,233
beach28.tif	0	0	0	0	0,233	0,21	0,171	0,211
beach29.tif	1	0	0	1	0,209	0,21	0,18	0,201
beach30.tif	1	0	0	1	0,233	0,249	0,217	0,238



а

б

Рисунок 3.16 – Зображення beach14 (а) і beach16 (б), які є дивними для кодерів JPEG і AGU

По-четверте, значення метрики Ном зазвичай більші для зображень, визначених як дивні. Значення цього показника зазвичай досить близькі для всіх трьох компонентів. Це не дивно, оскільки компоненти кольорових зображень зазвичай сильно корельовані (схожі один на одного). Розрахували SROCC між незвичайністю зображення (було встановлено 0 зображенням, для яких КІС є

монотонною, і одиницю тим зображенням, які мають принаймні один локальний максимум) і значеннями Nom . Отримано наступні значення SROCC: 0,71 для R, 0,63 для G, 0,57 для B і 0,67 для Y компонента. Таким чином, кореляція знову висока, і параметр Nom виглядає досить інформативним. Проте все ще важко встановити певний поріг надійного розрізнення зображень на дивні та не дивні.

Висновки до розділу 3

Цей розділ присвячено питанням підвищення ефективності стиснення із втратами зображень дистанційного зондування Землі. Запропоновано двоетапну процедуру стиснення зображення з втратами для забезпечення бажаної якості стиснення зображення з втратами відповідно до PSNR (попередні експерименти показують, що підхід також працює для деяких інших показників). Його ефективність проаналізована для набору тестових зображень у відтінках сірого, і показано, що підхід працює достатньо добре для стиснення зображень із забезпеченням невидимих спотворень.

Результати доводять, що двоетапний метод забезпечує хорошу точність забезпечення бажаної якості зображення згідно з цими метриками, усі значення дисперсії та максимальні похибки кращі (менші), ніж для PSNR, а максимальні похибки не перевищують 0,72 дБ. Це можна вважати хорошим практичним досягненням.

Розглянуто проблему прогнозування та забезпечення необхідної якості зображень, стиснутих методом SPIHT. Показано, що для заданого БНП якість зображення може змінюватися в широких межах. Двоетапна процедура забезпечує кращу точність забезпечення якості, але її точність все ж може бути незадовільною. Використовуючи простоту виконання SPIHT і AGU і використовуючи раніше запропоновані підходи до прогнозування характеристик ефективності для AGU, запропоновано алгоритм прогнозування PSNR і PSNR-HVS-M. Це швидше і точніше, ніж двоетапна процедура.

Двоетапний метод для зображень у градаціях сірого виконано для кодера BPG на основі HEVC. Завдяки аналізу показників PSNR і PSNR-HVS-M показано, що цей метод забезпечує задовільну точність у досягненні бажаної якості.

Результати двоетапного методу стиснення для кодера ADCT проаналізовано за допомогою двох метрик візуальної якості, одна з яких враховує особливості зорового сприйняття людини. Для PSNR, коли ΔKK є негативним і його абсолютне значення більше, ніж $KK_{init}/2$, вводиться модифікація. Завдяки цьому рівень помилок після двох кроків значно зменшився. Швидкий контроль точності двоетапного методу стиснення в поєднанні з високою ефективністю стиснення ADCT заощаджує більше місця, часу та використання обчислювальних ресурсів, тим самим краще сприяючи екологічним обчисленням.

Розглянуто проблему швидкого стиснення зображень різного походження з втратами із забезпеченням бажаної якості (часто непомітності внесених спотворень). Показано, що існують досить чіткі залежності показників якості від KK і KM для кодерів із втратами на основі ДКП. Запропонований підхід є загальним і може бути застосований для різних типів зображень у медичній практиці, візуалістиці, ДЗЗ, робототехніці тощо. Крім того, продемонстровано, що можливе прогнозування багатьох показників візуальної якості. Також показано, що 3D-стиснення зображень, зібраних у групи, дає значно кращі результати. Наведено приклади для реальних даних гіперспектрального ДЗ.

Розглянуто завдання стиснення багатоканальних зображень із втратами шляхом введення різних рівнів спотворень для аналізу їх впливу на точність класифікації. Було досліджено два класифікатори, а саме ММП та НМ. Був використаний кодер AGU. Крім того, було використано два інших кодери на основі трансформації з коротким аналізом їх ефективності.

Було також показано, що основні залежності спостерігаються для різних кодерів (зокрема, методів стиснення на основі ДКП і вейвлетів) і для даних, що надаються різними мультиспектральними системами. В запропонованому підході до стиснення з втратами можна застосувати інші показники візуальної якості, окрім PSNR-HVS-M. Таким чином, підхід є досить загальним. Нарешті, надано практичні

рекомендації щодо встановлення параметрів кодера, щоб уникнути значного зниження точності класифікації. Коротко згадуються можливості вдосконалення класифікації та описуються напрямки майбутніх досліджень.

Аналіз монотонності КШС для чотирьох наборів даних типових зображень ДЗ було проведено для двох кодерів (JPEG та AGU). Показано, що існує певний відсоток зображень, для яких залежності PSNR від КК для JPEG і AGU не є монотонними. Найчастіше це трапляється із зображеннями набору даних Beach, для яких зображення здебільшого містять дві квазіоднорідні області, що відповідають пляжному піску та поверхні води. Також показано, що існує досить багато параметрів, призначених для характеристики складності зображення, які мають досить високу кореляцію з ймовірністю того, що зображення буде дивним.

Основні наукові результати, наведені у третьому розділі, опубліковано в роботах автора [34], [35], [40], [42], [44], [45], [46], [47], [48], [49], [50], [63], [65], [71], зазначених в анотації.

РОЗДІЛ 4 АЛГОРИТМИ СТИСНЕННЯ ЗАШУМЛЕНИХ ЗОБРАЖЕНЬ ТА ПРОГНОЗУВАННЯ ХАРАКТЕРИСТИК СТИСНЕННЯ

У цьому розділі розглядаються підходи до стиснення зображень, що спотворені шумом (зашумлених зображень) для кодерів на основі ДКП. Ці алгоритми ґрунтуються на прогнозуванні середньоквадратичних похибок для заданого кроку квантування з використанням простого аналізу складності зображення. Проаналізовано статистичні та просторово-спектральні характеристики спотворень, внесених кодерами на основі ДКП, і показано, що вони залежать від кроку квантування та локального вмісту. Показано, що MSE спотворень можна досить легко спрогнозувати. Ці прогнози є достатньо точними і можуть бути використані для належного встановлення кроку квантування. Також розглядається підхід до прогнозування візуальної якості. Продемонстровано, що можна оцінити MSE-HVSM в обмеженій кількості блоків 8x8 пікселів, а потім передбачити значення цієї метрики для всього зображення для розглянутого кодеру. Досліджено вплив вмісту зображення та кількості аналізованих блоків. Основні роботи автора, використовувані в цьому розділі, наведені у публікаціях [20], [25], [27], [32], [33], [37], [43], [55].

4.1 Аналіз характеристик внесених спотворень при стисненні зображень

Почнемо аналіз з кодера AGU і розглянемо три значення КК, що дорівнюють 5, 10 і 20, та чотири тестових зображення ДЗ FR01-FR04. Таблиця 4.1 містить такі параметри: БНП, які можна перерахувати в КС як $КС = 8/БНП$ (для зображень у відтінках сірого, представлених у вигляді двовимірних 8-бітних масивів даних), PSNR і PSNR-HVS-M. Як видно, випадок КК = 5 може бути пов'язаний зі стисненням майже з втратами, оскільки для всіх тестових зображень значення БНП трохи перевищують 3, значення PSNR є більшими за 35 дБ, а значення PSNR-HVS-M перевищують 55 дБ. Випадок КК = 20 стосується видимих спотворень, які не є дратівливими. Для КК = 20 БНП становить приблизно 1,4, тобто КС перевищує 5.

Оскільки значення PSNR трохи менші за 35 дБ, а значення PSNR-HVS-M становлять приблизно 39,7 дБ для всіх чотирьох тестових зображень, спотворення можна помітити під час візуального огляду. Випадок КК = 10 можна розглядати як проміжний: КС становить приблизно 3,5, PSNR – приблизно 39,7 дБ, а PSNR-HVS-M – приблизно 47,3 дБ, тобто спотворення непомітні. Наведені приклади добре узгоджуються з результатами роботи [168], де показано, що в середньому для AGU КК потрібно встановити рівним 16, щоб забезпечити непомітність внесених спотворень.

Таблиця 4.1 – Характеристики стиснення тестових зображень для AGU

ПКС	КК=5	КК=10	КК=20
FR01			
БНП	3,24	2,29	1,40
PSNR, дБ	45,06	39,74	34,75
PSNR-HVS-M, дБ	55,20	47,31	39,71
FR02			
БНП	3,28	2,34	1,46
PSNR, дБ	45,13	39,77	34,71
PSNR-HVS-M, дБ	54,98	47,46	39,87
FR03			
БНП	3,32	2,37	1,46
PSNR, дБ	44,97	39,53	34,45
PSNR-HVS-M, дБ	55,48	47,31	39,55
FR04			
БНП	3,09	2,19	1,35
PSNR, дБ	45,21	39,85	34,71
PSNR-HVS-M, дБ	54,28	47,11	39,63

З даних у таблиці 4.1 може здатися, що значення PSNR, PSNR-HVS-M і БНП (КС) для даного КК майже однакові, і це не проблема забезпечити бажану якість. Однак це не так (причина приблизного збігу характеристик ефективності для розглянутих чотирьох тестових зображень полягає в тому, що їх складність

приблизно однакова). У [168] показано, що, наприклад, для БНП = 1,6 значення PSNR-HVS-M змінюються в межах від 30 дБ до 53 дБ. У [266] також показано, що значення PSNR для одного і того ж КК можуть відрізнятися до 10 дБ (для КК близько 20). Коротко розглянемо попередні дані для ADCT. Вони представлені в таблиці 4.2, де надаються значення КС/БНП і MSE/PSNR для чотирьох тестових зображень і трьох значень КК. Як бачимо, основні тенденції збігаються:

- КС збільшується (БНП зменшується), якщо КК збільшується;
- MSE збільшується (PSNR зменшується), якщо КК стає більшим;
- MSE зростає майже пропорційно KK^2 , тому отримане значення для $KK = 5$

визначається як $MSE = (KK^2/12)$

Таблиця 4.2 – Характеристики стиснення тестових зображень для ADCT

Зображення	КК	КС	БНП	MSE	PSNR
FR01	5	2,19	3,01	2,08	44,95
FR02	5	2,65	3,03	2,03	45,06
FR03	5	2,53	3,18	2,10	44,91
FR04	5	2,74	2,93	1,98	45,16
FR01	10	3,97	2,02	7,34	39,47
FR02	10	3,87	2,08	7,13	39,60
FR03	10	3,71	2,17	7,71	39,26
FR04	10	3,98	2,02	7,03	39,66
FR01	20	6,78	1,18	23,03	34,51
FR02	20	6,46	1,24	23,17	34,48
FR03	20	6,37	1,26	25,38	34,09
FR04	20	6,73	1,19	23,92	34,34

Тим часом для $KK = 10$ реальні значення MSE вже помітно менше, ніж $KK^2/12$, а для $KK = 20$ ця різниця ще збільшується ($400/12 \approx 33,33$). Справжні MSE стають ≈ 24 для розглянутих тестових зображень. Досить ґрунтовний аналіз тенденцій, що спостерігаються, та їхніх причин проведено в роботі [34]. Відмінності в залежностях детально пояснюються в цій роботі (а також вище в

підрозділі 3.2), де показано, що розподіли коефіцієнтів ДКП суттєво відрізняються для розглянутих трьох зображень і параметрів цих розподілів (наприклад, відсоток коефіцієнтів ДКП, які стають нульовими після квантування) та знаходяться в досить тісному зв'язку з MSE і КС. У той же час представлені результати можна розглядати як основу ідеї про те, що MSE можна передбачити для даної КК, якщо зображення, яке потрібно стиснути, якимось швидко проаналізовано для визначення його основних властивостей.

Необхідність вивчення статистичних і просторово-спектральних властивостей спотворень, що вносяться різними кодерами, важлива з декількох причин. По-перше, у [155] було зазначено, що спотворення через стиснення з втратами мають бути обмеженої інтенсивності та мати просторово «рівномірний» розподіл, щоб уникнути артефактів (включаючи артефакти класифікації). По-друге, результати в [248] продемонстрували, що спотворення мають розподіл, близький до гаусового для відносно малих КС і КК для зображень складної структури, тоді як розподіл може бути негаусовим для зображень простих структур. Однак причини цього не пояснюються. Через це варто нагадати дані, нещодавно представлені в роботі [267], де було виконано візуальний та кількісний аналіз різницевих зображень. Зауважте, що різницеві зображення досить часто використовуються для аналізу стиснення зображень із. Нехай $\{I^{or}(i, j), i = 1, \dots, I_{Im}, j = 1, \dots, J_{Im}\}$ — вихідне зображення розміром $I_{Im} \times J_{Im}$ пікселів і $\{I^c(i, j), i = 1, \dots, I_{Im}, j = 1, \dots, J_{Im}\}$ представляють відповідне стиснене зображення. Обидва зображення є доступними, і зображення різниць можна обчислити як:

$$\{\Delta(i, j) = I^{or}(i, j) - I^c(i, j), i = 1, \dots, I_{im}, j = 1, \dots, J_{im}\}, \quad (4.1)$$

У більшості практичних ситуацій $\Delta(i, j)$ є цілими числами, які можуть бути негативними, позитивними та дорівнювати нулю. Для візуального представлення існує декілька способів представлення різницевих зображень, наприклад:

- як абсолютні значення різниць (включаючи попереднє збільшення);
- за допомогою п'єдесталу, наприклад, як:

$$\{\Delta_p(i, j) = I^{or}(i, j) - I^c(i, j) + 128, i = 1, \dots, I_{im}, j = 1, \dots, J_{im}\},$$

Один із прикладів Δ_p представлено на рис. 4.1 для тестового зображення FR01, стисненого за допомогою AGU з КК = 20. Флуктуації (спотворення) навряд чи можна помітити, оскільки PSNR приблизно дорівнює порогу видимості спотворень (трохи менше 35 дБ). Візуалізовані спотворення абсолютно не видно для КК = 5 і КК = 10. Через це різницеве зображення тут не показано. Оскільки очікується, що різниці (4.1) можуть бути нестационарним двовимірним процесом, його просторовий спектральний або кореляційний аналіз слід виконувати з обережністю. Один з методів такого аналізу, застосований для залежних від сигналу або просторово інваріантних даних, був запропонований в [268], [269]. Необхідно визначити моду локальних оцінок ексцесу, отриманих в області ДКП. Якщо мода менше 3,75, шум можна вважати просторово некорельованим. Отримані результати представлені в таблиці 4.3.

Як видно, спотворення можна вважати просторово некорельованими для всіх чотирьох тестових зображень і всіх трьох значень КК. Проте можна спостерігати тенденцію до збільшення моди (тобто до більшої просторової кореляції спотворень) при зростанні КК. Подібний аналіз було проведено для кодера ADCT. Результати практично однакові — спотворення можна вважати просторово некорельованими, принаймні, для $КК \leq 20$.

Проведемо статистичний аналіз по блоках, враховуючи можливу просторову нестационарність спотворень. Використаємо блоки розміром 8×8 пікселів, які не перекриваються, і обчислимо в кожному з них наступні два параметри (де позиція кожного блоку визначається координатами верхнього лівого кута n і m):

$$\sigma_{dist}^2 = \frac{1}{64} \sum_{i=m}^{m+7} \sum_{j=n}^{n+7} (\Delta_p(i, j) - \bar{\Delta}), \quad (4.2)$$

$$\sigma_{im}^2 = \frac{1}{64} \sum_{i=m}^{m+7} \sum_{j=n}^{n+7} (I^{or}(i, j) - \bar{I^{or}}), \quad (4.3)$$

$$\text{де } \bar{\Delta} = \frac{1}{64} \sum_{i=m}^{m+7} \sum_{j=n}^{n+7} \Delta_p(i, j) \text{ та } \overline{I^{or}} = \frac{1}{64} \sum_{i=m}^{m+7} \sum_{j=n}^{n+7} I^{or}(i, j).$$

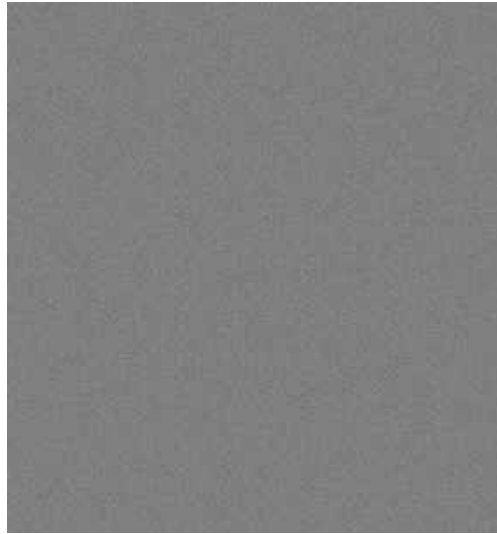


Рисунок 4.1 – Різницеве зображення Δ_p для стисненого зображення FR01 для $KK = 20$

Таблиця 4.3 – Мода локальних оцінок ексцесу

Зображення	KK = 5	KK = 10	KK = 20
FR01	2,6974	2,7507	2,8896
FR02	2,7189	2,7642	2,9065
FR03	2,7192	2,7446	2,8437
FR04	2,7573	2,7933	2,8510

Отримані дані можна представити у вигляді графіків розсіювання середньоквадратичної помилки (RMSE) спотворень σ_{dist} проти RMSE спотворень σ_{im} , що характеризує локальну (контентну) активність. Для тестового зображення FR01 отримані діаграми розсіювання наведені в Додатку Л, на рис. Л.1 для трьох значень KK. Їх аналіз показує наступне:

– можна спостерігати великі області σ_{im} , де σ_{dist} є, загалом, випадковими, але їх середнє значення є практично постійним; пізніше такі ділянки будуть називатися областю насичення;

– не дивно, що в таких областях середнє значення σ_{dist} приблизно дорівнює $(KK^2/12)0,5$ — приблизно 1,4 для $KK = 5$ (в Додатку Л, на рис. Л.1, а), приблизно 2,7 для $KK = 10$ (в Додатку Л, на рис. Л.1, б) і близько 5 для $KK = 20$ (в Додатку Л, на рис. Л.1, в);

– якщо σ_{im} досить малий, то спостерігається тенденція зменшення σ_{dist} (в середньому) зі зменшенням σ_{im} ; ця тенденція може спостерігатися, коли $\sigma_{im} \leq KK$.

Діаграми розсіювання, отримані для інших трьох тестових зображень, дуже схожі на ті, що представлені Додатку Л, на рис. Л.1. Щоб проілюструвати це, в Додатку Л, на рис. Л.2 наведені діаграми розсіювання для тестового зображення FR02. Як видно, висновки, які можна зробити з їхнього аналізу, такі ж, як наведено вище. Отримавши такі діаграми розсіювання, можна здійснити регресію, тобто підгонку кривих. Для цього використовується таке наближення:

$$\sigma_{dist}^{appr} = a \cdot \exp(b\sigma_{im}) + c, \quad (4.4)$$

де a , b і c — параметри функції, які потрібно оцінити. Для підгонки кривих та оцінки їх параметрів була використана функція `fminsearch` пакету MATLAB.

Параметри підігнаних кривих наведені в Додатку Л, в таблиці Л.1. Для одного й того ж KK значення параметрів підігнаних кривих для розглянутих тестових зображень дуже близькі. Особливо це стосується параметра c , який описує «рівень насичення».

Представлені діаграми розсіювання пояснюють, чому MSE спотворень, внесених стисненням із втратами в зображення простих структур, відрізняється від MSE внесених спотворень для зображень складної структури. Зображення складної структури здебільшого мають блоки, де σ_{im} достатньо великий через високу локальну активність, і тоді σ_{dist}^2 для таких блоків у середньому приблизно дорівнює $KK^2/12$. Тоді MSE для всього зображення також близьке до $KK^2/12$. Тим часом, коли KK збільшується, блоків у зоні насичення стає менше, і MSE для всього зображення виявляється меншим, ніж $KK^2/12$.

Представимо тепер графіки розсіювання σ_{dist} від σ_{im} для кодера ADCT. Вони наведені в Додатку Л, на рис. Л.3.

Діаграми розсіювання в Додатку Л, на рис. Л.3 можна порівняти з відповідними діаграмами розсіювання на рис. Л.1 та Л.2. Порівняння показує, що основні властивості діаграм розсіювання дуже схожі. Знову ж таки, спостерігаються зони насичення для $\sigma_{im} > KK$, де σ_{dist} приблизно дорівнює $(KK^2/12)^{0,5} = KK/3,47$. Коли $\sigma_{im} \leq KK$, тоді σ_{dist} зменшується зі зменшенням KK . Це зменшення можна пояснити наступним чином.

У квазіоднорідних областях зображення (блоках) існує великий відсоток коефіцієнтів ДКП, які менші за KK (особливо якщо KK достатньо великий). Помилки квантування для таких коефіцієнтів ДКП в середньому мають менші абсолютні значення порівняно з випадком, коли більшість коефіцієнтів ДКП мають абсолютні значення, більші за KK (див. розподіли помилок квантування в [194]). Оскільки помилки квантування в області ДКП менші, локальні MSE для відповідних блоків стиснених зображень також менші. Карта σ_{im} у блоках (збільшена в 5 разів для кращої візуалізації) для тестового зображення Frisco представлена на рис. 4.2, а, тоді як на рис. 4.2, б показана карта σ_{dist} в блоках (збільшена в 28 разів для кращої візуалізації) для того самого тестового зображення. Чітко видно, що σ_{dist} менше (пікселі темніші) в однорідних областях зображення, де значення σ_{im} менші (пікселі темніші).

Діаграми розсіювання для інших тестових зображень дуже схожі. Параметри кривих, отримані за допомогою апроксимації (4.4), представлені в таблиці 4.4. У порівнянні з відповідними значеннями в Додатку Л, в таблиці Л.1 вони дуже схожі. Єдина відмінність полягає в тому, що значення параметра c для кодера ADCT трохи більші. Вони приблизно дорівнюють $KK/3,47$ (1,44, 2,88 і 5,76 для KK , що дорівнює 5, 10 і 20 відповідно). Оскільки для трьох значень KK отримано велику кількість σ_{dist} і σ_{im} , що відповідають один одному, можна отримати сукупні діаграми розсіювання для обох розглянутих кодерів. Крім того, представлено нормалізовані графіки розсіювання залежності $12^{0,5} \sigma_{dist} / KK$ від σ_{im} / KK . Для кодера AGU

агрегована діаграма розсіювання представлена на рис. 4.3, а. Загальні властивості цієї діаграми розсіювання подібні до представлених раніше. Існує монотонно зростаюча частина підігнаної кривої, яка спостерігається для $\sigma_{im}/KK < 1$, і «частина насичення», де $12^{0,5}\sigma_{dist}/KK$ близьке до одиниці, що має місце для $\sigma_{im}/KK \geq 1$. Підігнана крива виражається як:

$$f\left(\frac{\sigma_{im}}{KK}\right) = -0,7381 \cdot \exp\left(-\frac{2,8526\sigma_{im}}{KK}\right) + 0,9685. \quad (4.5)$$

Зверніть увагу, що значення параметра c у (4.5) не дорівнює 1, але воно досить близьке до одиниці.

Аналогічно для кодера ADCT була отримана агрегована діаграма розсіювання. Вона представлена на рис. 4.3, б. Порівняння діаграм розсіювання на рис. 4.3 показує, що їх основні властивості досить близькі. Зверніть увагу, що велике співвідношення σ_{im}/KK відноситься до «дуже активних» локальних ділянок (блоків, де спостерігаються гострі краї або великоконтрастні об'єкти невеликого розміру) і досить малих значень KK .

Найбільші варіації $12^{0,5}\sigma_{dist}/KK$ щодо підігнаних кривих мають місце для $\sigma_{im}/KK \approx 1$. Такі ситуації спостерігаються для блоків, які відповідають низькоконтрастним межам, деталям і текстурам, які є найбільш типовими для зображень природної сцени.

Отримана підібрана крива представлена як:

$$f\left(\frac{\sigma_{im}}{KK}\right) = -0,9498 \cdot \exp\left(-\frac{4,0992\sigma_{im}}{KK}\right) + 0,9762. \quad (4.6)$$

У цьому випадку значення параметра c ще ближче до одиниці.

Отримавши вирази (4.5) і (4.6) для кодерів AGU і ADCT відповідно, можна передбачити σ_{dist} для кожного k -го блоку 8×8 пікселів даного зображення, використовуючи відповідну оцінку σ_{im} , отриману для цього блоку. Для даного блоку та KK маємо:

$$\sigma_{dist} \approx \left(\frac{KK}{12^{0,5}} \right) f \left(\frac{\sigma_{im}}{KK} \right). \quad (4.7)$$

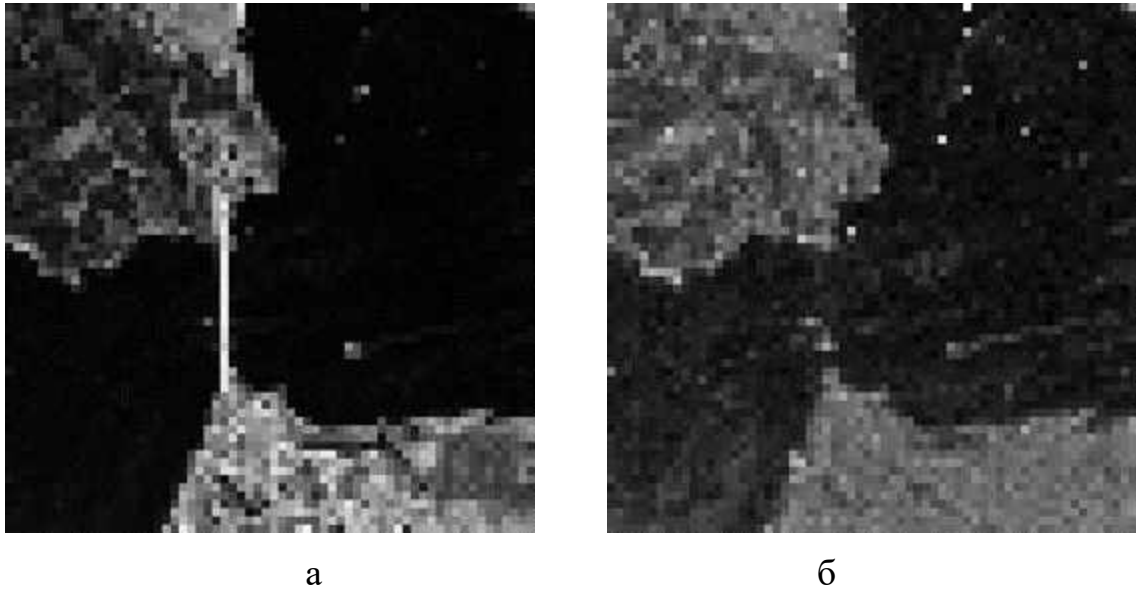


Рисунок 4.2 – Карта σ_{im} (а) та σ_{dist} (б) для тестового зображення Frisco

Припускаючи знання оцінок $\sigma_{dist}(k)$, $k = 1, \dots, K$, де K позначає загальну кількість блоків (питання, як блоки можуть бути розташовані та якою має бути їхня кількість, будуть обговорені пізніше), MSE для всього зображення можна передбачити як:

$$MSE \approx \sum_{k=1}^K \frac{\sigma_{dist}^2(k)}{K}, \quad (4.8)$$

або

$$MSE \approx \sum_{k=1}^K \frac{KK}{12} f^2(\sigma_{im}(k)/KK)/K, \quad (4.9)$$

де $\sigma_{im}(k)$ – значення СКВ в k -му блоці, яке визначається за формулою (4.3). Зауважте, що дуже легко обчислити всі $\sigma_{im}(k)$, $k = 1, \dots, K$ заздалегідь, якщо позиції блоків також відомі заздалегідь.

Таблиця 4.4 – Параметри підігнаних кривих для кодера ADCT

Зображення	KK=5	KK=10	KK=20
FR01	a = -2,42	a = -2,95	a = -4,69
	b = -1,44	b = -0,52	b = -0,20
	c = 1,45	c = 2,77	c = 5,25
FR02	a = -1,59	a = -2,97	a = -5,17
	b = -1,00	b = -0,50	b = -0,21
	c = 1,45	c = 2,77	c = 5,33
FR03	a = -2,86	a = -4,22	a = -4,96
	b = -1,44	b = -0,74	b = -0,22
	c = 1,45	c = 2,80	c = 5,40
FR04	a = -1,35	a = -2,86	a = -5,59
	b = -0,85	b = -0,43	b = -0,20
	c = 1,45	c = 2,84	c = 5,64

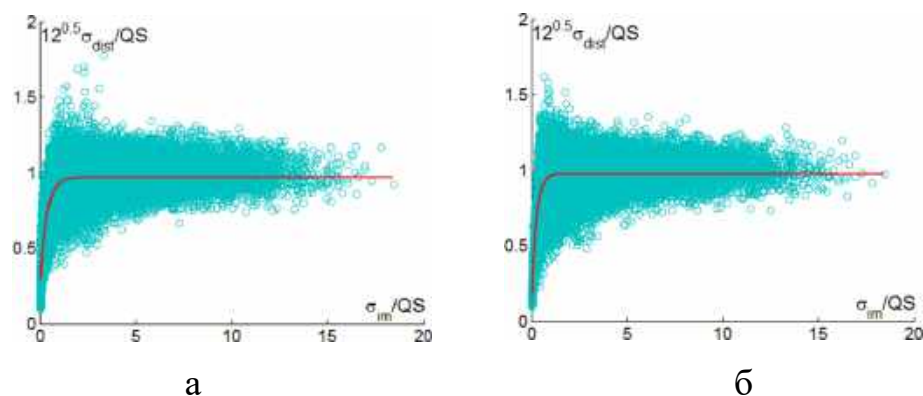


Рисунок 4.3 – Агрегована діаграма розсіювання для різних тестових зображень і значень KK, кодер AGU(а) та ADCT(б)

Проаналізуємо точність прогнозування MSE для внесених спотворень. В Додатку Л, в таблиці Л.2 представлені дані для чотирьох тестових зображень, використаних для формування діаграм розсіювання для кодера AGU. Як видно, справжні значення MSE близькі до відповідних прогнозованих (MSE_{pred}), відносна різниця вважається невеликою для практичних застосувань, оскільки вона не перевищує 10%. Дані для кодера ADCT для тих самих чотирьох зображень представлені в Додатку Л, в таблиці Л.3. Їхній аналіз показує, що справжні та

прогнозовані значення також досить близькі. Максимальна різниця не перевищує 8%, тобто 0,3 дБ щодо PSNR.

Однак більш цікавими є результати для зображень, які не використовувалися в навчанні (отримання підгоночних кривих для діаграм розсіювання). Такі перевіірочні дані були отримані для 16 інших тестових зображень різного походження. Серед них є декілька традиційних зображень, таких як: Lena, Baboon, Barbara, Boat, Goldhill, Peppers, Man. Є також високотекстурні зображення, такі як Grass та Bikes (див. зменшені копії в Додатку Л, на рис. Л.4). Декілька зображень дистанційного зондування, наприклад Frisco, Airfield, Lu01, San Diego, A13sm, Sent01, Sent02 також були використані в аналізі.

Мета використання зображень різного походження — продемонструвати, що запропонований підхід до прогнозування може бути застосовний у різних практичних ситуаціях. Деякі результати представлені в таблицях 4.5 і 4.6, де показано найбільш цікаві приклади.

Разом з цим, нижче представлені статистичні результати, що характеризують точність прогнозу. Аналіз даних таблиці 4.5 свідчить про наступне. По-перше, для $KK = 5$ значення MSE суттєво не відрізняються (варіюються від 1,55 до 2,14) і вони добре прогнозовані. По-друге, для $KK = 10$ різниця в значеннях MSE збільшується (вони змінюються від 4,2 для зображення простої структури Lu01 до 7,81 для зображення складної структури Grass), але прогноз все ще вважається досить точним.

Нарешті, для $KK = 20$ MSE може помітно відрізнитися залежно від складності зображення (від 10,21 для Lu01 до 30,92 для тестового зображення Grass).

Однак прогнозовані значення добре узгоджуються зі справжнім MSE, і це відбувається для всіх зображень, які використовуються в аналізі. Для всіх двадцяти зображень, використаних в аналізі, також було визначено зміщення та дисперсію для справжніх і прогнозованих значень.

Для $KK = 5$ вони дорівнюють -0,11 і 0,18; для $KK = 10$ — дорівнює 0,07 і 0,51; для $KK = 20$ — 2,77 і 8,66 відповідно. Зміщення та RMSE є меншими за 15%

середніх значень для кожного КК, і така точність є прийнятною для практики (фактично, PSNR прогнозується з максимальною похибкою менше ніж 1,5 дБ).

Таблиця 4.5 –Порівняння справжніх і прогнозованих значень MSE (кодер AGU) для зображень, використаних під час експериментальної перевірки

Зображення	КК	MSE	MSE _{pred}
A13sm	5	1,74	1,70
A13sm	10	5,53	4,87
A13sm	20	15,02	12,63
Grass	5	2,14	1,95
Grass	10	7,81	8,25
Grass	20	30,92	31,38
Man	5	1,89	1,81
Man	10	5,86	6,44
Man	20	16,54	20,51
Lu01	5	1,55	1,67
Lu01	10	4,20	5,33
Lu01	20	10,21	14,82
Sent01	5	2,04	1,94
Sent01	10	7,04	7,36
Sent01	20	22,57	25,37

Аналіз даних, представлених у таблиці 4.6, показує наступне. По-перше, якщо КК = 5, значення MSE коливаються в досить вузьких межах (від 1,69 до 2,14) і їх можна спрогнозувати досить добре. По-друге, при КК = 10 значення MSE змінюються від 4,61 для зображення Lu01, що має просту структуру, до 8,21 для зображення Grass зі складною структурою. Точність прогнозування висока для всіх зображень. По-третє, для КК = 20 MSE може змінюватися в широкому діапазоні. Прогнозовані значення досить близькі до відповідних справжніх, і, фактично, це має місце для всіх зображень, використаних у дослідженні. Що стосується зміщення та дисперсії, то вони такі: для КК = 5 вони дорівнюють -0,15 та 0,019; для КК = 10 — дорівнює 0,07 і 0,51; для КК = 20 — дорівнює 2,77 та 8,66 відповідно.

Таблиця 4.6 –Порівняння справжніх і прогнозованих значень MSE (кодер ADTC) для зображень, використаних під час експериментальної перевірки

Зображення	KK	MSE	MSE _{pred}
A13sm	5	1,84	1,86
A13sm	10	5,46	6,11
A13sm	20	14,85	16,37
Grass	5	2,14	1,99
Grass	10	8,21	7,94
Grass	20	31,92	31,62
Man	5	2,04	1,89
Man	10	6,33	6,85
Man	20	17,39	22,22
Lu01	5	1,69	1,78
Lu01	10	4,61	5,81
Lu01	20	11,25	15,97
Sent01	5	2,09	1,98
Sent01	10	7,39	7,70
Sent01	20	23,63	27,46

Зміщення та RMSE є меншими за 16% від середніх значень для кожного KK, і така точність є прийнятною для практики (фактично, PSNR прогнозується з максимальною похибкою менше 1,4 дБ). Порівняння даних у таблицях 4.5 і 4.6 показує, що значення MSE_{pred} для кодера ADCT зазвичай трохи більші, ніж для кодера AGU. Цей висновок також можна зробити з порівняння виразів (4.5) і (4.6), де в останньому випадку значення параметра c більше. Крім того, була також проаналізована точність прогнозування для методу [33] з використанням кодера ADCT для KK = 20. Мінімальний прогнозований MSE спостерігається для зображення Lu01 (11,63), максимальний (34,43) — для тестового зображення Grass. Статистичний аналіз показує, що похибка між прогнозованими та справжніми значеннями MSE дорівнює -1,44, а дисперсія дорівнює 7,11. Це означає, що точність прогнозування така ж, як і для методу, запропонованого вище.

Залишається декілька аспектів для обговорення. По-перше, усі наведені вище результати були отримані для неперекриваючих блоків, які повністю покривають область зображення. Таким чином, для зображень 512×512 пікселів, які використовуються в аналізі, під час обчислення $\sigma_{dist}(k)$, $k = 1, \dots, K$, задіяно 4096 блоків. Можна очікувати, що використання блоків, що перекриваються, може покращити передбачення. Тож було використано блоки, що повністю перекриваються, із загальною кількістю $(512 - 7) \times (512 - 7) = 255025$. Однак це не призвело до підвищення точності прогнозу. Причому відповідні значення MSE_{pred} для блоків, що не перекриваються, і блоків, що повністю перекриваються, відрізняються несуттєво — менше ніж на 0,04%. Наприклад, для тестового зображення FR04 і $KK = 20$ (кодер AGU) значення MSE_{pred} дорівнюють 22,240 і 22,264 для блоків, що не перекриваються, і повністю перекриваються, відповідно. Крім того, був вивчений випадок, коли 1000 блоків були розміщені випадковим чином. Отримані результати прогнозу практично не відрізняються від відповідних даних для блоків, що не перекриваються. Отже, є спосіб для прискорення прогнозу шляхом аналізу обмеженої кількості блоків.

Виходячи з розробленого методу прогнозування MSE, на практиці необхідний алгоритм для забезпечення бажаного значення MSE_{des} . Можна запропонувати декілька варіантів. Наприклад, як перший крок, розрахунок коригувального коефіцієнта як:

$$K_{cor} \approx \sum_{k=1}^K f^2(\sigma_{im}(k)/\sqrt{12MSE_{des}})/K \quad (4.10)$$

та

$$KK_1 = \sqrt{12MSE_{des}/K_{cor}}, \quad (4.11)$$

через це визначаємо

$$MSE_1 \approx \sum_{k=1}^K \frac{KK_1^2}{12} f^2(\sigma_{im}(k)/KK_1)/K. \quad (4.12)$$

Якщо MSE_1 досить близький до MSE_{des} (наприклад, якщо вони відрізняються не більше ніж на 10%), значення $KK_1 = \sqrt{12MSE_{des}/K_{cor}}$ може використовуватися для остаточного стиснення. Якщо ця умова не виконується, кінцевий КК може бути розрахований як:

$$KK_f = KK_1 \sqrt{MSE_{des}/MSE_1}. \quad (4.13)$$

Цей алгоритм був спочатку протестований для кодера AGU для $MSE_{des} = 20$ для всіх двадцяти зображень, використаних у попередніх дослідженнях. Отриманий MSE коливається від 11,87 (зображення Barbara) до 20,59 (зображення Grass), його зміщення і дисперсія дорівнюють 3,22 і 5,79 відповідно, тому цю точність можна вважати задовільною. Кінцевий КК змінюється від 16,04 для зображення Grass до 24,99 для зображення Frisco, тобто, як і очікувалося, кінцевий КК є найбільшим для зображень простої структури та найменшим для зображень складної структури. Як можна побачити, отримані MSE в середньому менші, ніж бажано, і це, в основному, відбувається для зображень простих структур. З одного боку, це може бути корисним на практиці, оскільки є певний «запас» якості лише для зображень, для яких спотворення помітніші. З іншого боку, можна ввести деяку додаткову корекцію для усунення зміщення. Запропонований алгоритм також протестовано для кодера ADCT для $MSE_{des} = 20$. Наданий MSE змінюється в межах від 12,06 для зображення Barbara до 21,16 для зображення SanDiego. Зміщення і дисперсія дорівнюють 2,86 і 6,46 відповідно. Таким чином, знову отримано зміщення та RMSE помилок менше 15%; таку точність можна вважати задовільною. Підсумковий КК коливається в межах від 16,04 для зображення Grass до 24,90 для зображення Frisco. Додаткова перевірка запропонованого алгоритму була проведена для 300 версій у градаціях сірого зображень 512×512 пікселів із бази даних безшумних зображень TAMPERE17 [270]. Отримані результати для $MSE_{des} = 20$ представлені у вигляді гістограм для кодерів AGU та ADCT в Додатку

Л, на рис. Л.5. У більшості випадків надані значення MSE досить близькі до $MSE_{des} = 20$. Хоча отримані значення MSE зміщені відносно бажаного MSE, а дисперсія кінцевого MSE є досить великою, важливою перевагою є те, що отримані значення MSE здебільшого менші за бажане MSE.

Найменші отримані MSE спостерігаються для зображень із простою та специфічною структурою, наприклад, представлених в Додатку Л, на рис. Л.6. Як бачимо, алгоритм визначення кінцевого КК обчислювально дуже простий. Це вимагає обчислення $\sum_{k=1}^K \frac{KK_1^2}{12} f^2(\sigma_{im}(k)/KK_1)/K$ двічі, де функції $f(\bullet)$ досить прості, обчислення σ_{im} , $k = 1, \dots, K$ та елементарні операції порівняння та арифметичні операції. Реалізовані на процесорі Intel Core i7 L620 2,00 ГГц з 8 ГБ оперативної пам'яті, вони займають $0,363 \pm 0,025$ секунди для блоків, що не перекриваються, для зображення розміром 512×512 пікселів. Нагадаємо, що компресія за допомогою AGU займає $1,020 \pm 0,040$ секунди (тоді як для декомпресії потрібно $1,735 \pm 0,083$ секунди). Таким чином, прогнозування є значно швидшим, ніж стиснення. Зауважте, що додатково можливе використання меншої кількості аналізованих блоків — якщо, наприклад, 1000 блоків, розміщених у випадковому порядку, використовуються замість 4096 блоків, що не перекриваються, прогнозовані MSE відрізняються менш ніж на 1%, отже, такий підхід підходить для практики.

У цьому випадку прогнозування може бути реалізовано на порядок швидше, ніж стиснення. Для кодера ADCT час прогнозування для блоків, що не перекриваються, становить $0,374 \pm 0,065$ секунди, тоді як для стиснення потрібно $3,080 \pm 0,125$ секунди, тобто прогнозування та визначення фінального КК відбувається значно швидше, ніж стиснення (декомпресія вимагає $1,998 \pm 0,123$ секунди). Звичайно, прогнозування можна додатково прискорити порівняно з випадком блоків, що не перекриваються, використовуючи меншу кількість блоків. Отримані результати мають ще один позитивний результат. Вони дозволяють запропонувати більш адекватну модель (порівняно з [271]) для моделювання

спотворень через стиснення з втратами у вигляді просторово некорельованого гаусового шуму з дисперсією, що залежить від локальної активності зображення.

Розглянемо тепер поведінку візуальних показників якості в залежності від КК. Отримані залежності для тестового зображення *Mrt_prepared* розміром 512x512 пікселів представлені на рис. 4.4, а. Аналіз дозволяє зробити такий висновок.

Для $КК < 7$ усі показники перевищують 40 дБ (PSNR-HVS-M перевищує 50 дБ), це означає, що внесені спотворення непомітні. У дослідженні [32] для малих КК введено наступне наближення:

$$MSE_{HVS-M} = 0,02896KK^{1,976}. \quad (4.13)$$

Інтервал більших значень КК (від 7 до приблизно 30, що відповідає значенням MSE_{HVS-M} від 1 до приблизно 30, рис. 4.5) відноситься до найважливіших практичних випадків. При зростанні КК всі показники зменшуються (рис. 4.4, а). PSNR-HVS-M стає значно меншим, ніж PSNR для того самого КК.

Це показує, що внесені спотворення не схожі на АБГШ, більше спотворень припадає на низькі просторові частоти. З подальшим збільшенням КК PSNR-HVS-M також стає меншим, ніж PSNR. Це показує, що ефект маскування стає незначним, оскільки спотворення стають занадто великими. Тим часом PSNR-HVS-M все ще більші, ніж PSNR-HVS. Щоб перевірити, чи поведінка залежностей є однаковою для інших зображень, було оброблено тестове стоматологічне зображення. Отримані залежності наведені на рис. 4.4, б. Їхнє порівняння з відповідними залежностями на рис. 4.4, а, показує, що вони дуже подібні. На рис. 4.5 показано діаграму розсіювання, для якої кожна точка відповідає визначеному для одного тестового зображення, стисненого за допомогою заданого КК.

Як видно, значення для одного і того ж КК можуть бути досить різними, особливо для $КК > 30$, де значення можуть відрізнятися в декілька разів. Це означає, що такий прогноз, як (4.13), може призвести до помилок для окремих зображень, які можуть бути занадто великими, і потрібне точніше прогнозування. Для

вирішення цього завдання було вирішено повернутися до ідей у роботах [34], [249], які передбачають використання статистики зображення в обмеженій кількості блоків.

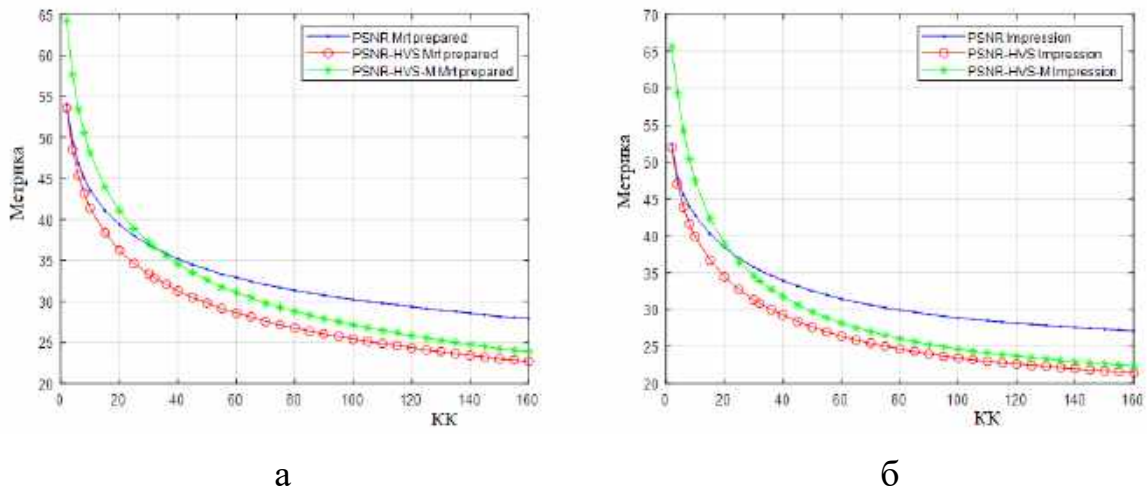


Рисунок 4.4 – Залежності метрик візуальної якості від КК для двох тестових зображень: Mrt_prepared (а) та Impression (б)

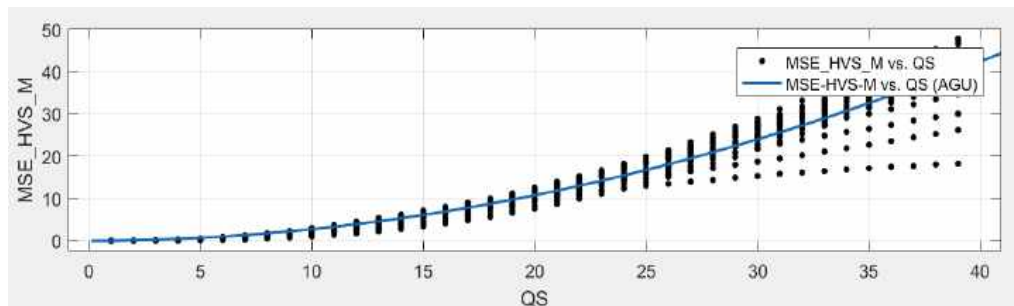


Рисунок 4.5 – Діаграма розсіювання MSE_{HVS-M} від КК для стиснення з втратами зображень у градаціях сірого для кодера AGU

Згадаймо, що тут MSE_{HVS-M} обчислюється як середнє значення локальних оцінок MSE_{HVS-M} у блоках (кількість таких локальних оцінок становить близько $PmJm/64$). Для традиційного підходу необхідно враховувати всі можливі позиції блоків і MSE_{HVS-M} повинен бути обчислений після стиснення зображення даним кодером.

У цьому випадку пропонується розрахувати оцінку MSE_{HVS-M} до стиснення з використанням обмеженої кількості блоків розміром 8×8 пікселів. У цього наміру є декілька обґрунтувань. По-перше, обчислення MSE_{HVS-M} базується на ДКП, а кодер AGU також базується на 2D ДКП (хоча розмір блоку не становить 8×8 пікселів). Через це очікуємо, що існує кореляція між MSE_{HVS-M} для кодера AGU і MSE_{HVS-M} , визначеним в обмеженій кількості блоків пікселів 8×8 . По-друге, алгоритм ДКП у блоках 8×8 є стандартною операцією, що використовується для стиснення зображень і відео. Тож, він має його швидкі апаратні та програмні реалізації. Тоді можна очікувати, що прогнозування може бути здійснено набагато швидше, ніж стиснення.

Отже, ідея полягає в наступному. Виберемо N блоків розміром 8×8 пікселів. Потім обчислимо 2D ДКП для кожного блоку з отриманням $D(k,l,n)$, де $k=0,\dots,7$ і $l=0,\dots,7$ – індекси просторової частоти, а $n=1,\dots,N$ – індекс блоку. Після цього можна провести квантування для заданого КК та визначити:

$$D_{dq}(k,l,n) = KK \left(\left\lceil \frac{D(k,l,n)}{KK} \right\rceil \right), \quad (4.14)$$

де $\lceil \cdot \rceil$ означає округлення до найближчого цілого числа.

Потім $MSE_{HVS-M}(n)$ обчислюється за допомогою таблиці ваг просторових частот і правила маскування (див. [245] для більш детальної інформації). Після цього розраховується середнє значення $MSE_{HVS-M \Sigma}$ для всіх блоків і прогнозується $MSE_{HVS-M AGU}$ для AGU (подробіці прогнозування наведені нижче).

Розглянемо тепер деякі попередні дані, які показують, що таке прогнозування можливе. В Додатку Л, на рис. Л.7 представлено два приклади діаграм розсіювання, отриманих для тестових зображень Airfield і Goldhill. Кольорові позначки стосуються оцінок, отриманих у блоках 8×8 пікселів, тоді як чорні позначки стосуються справжніх значень $MSE_{HVS-M AGU}$ для AGU. КК змінюється від 1 до 100 з кроком 1.

Аналіз наведених діаграм розсіювання та залежностей показує наступне:

1) $MSE_{HVS-M AGU}$ збільшується, коли KK стає більшим, залежність виглядає близькою до KK^2 , принаймні для малих KK ;

2) залежності $MSE_{HVS-M AGU}$ від KK здаються дуже подібними, хоча значення для зображення Airfield трохи більші, ніж для Goldhill для того самого KK (наприклад, розглянемо випадок $KK=80$);

3) значення $MSE_{HVS-M AGU}$, здається, здебільшого менші, ніж $MSE_{HVS-M \Sigma}$ для того самого KK ; тоді виникає ідея, що деякий фіксований коригувальний коефіцієнт або коригувальну функцію (залежну від KK) можна використовувати для перерахунку оцінки $MSE_{HVS-M \Sigma}$ до оцінки $MSE_{HVS-M AGU}$;

4) оцінки $MSE_{HVS-M \Sigma}$ поводяться досить компактно, де компактність оцінок, отриманих для більшого N , краща; це означає, що N , як і можна було очікувати, впливає на точність прогнозу і цей аспект має бути вивчений більш детально;

5) різноманітність оцінок $MSE_{HVS-M \Sigma}$ найбільша для $N=100$ (позначено зеленими трикутниками) і найменша для $N=1000$ (позначено фіолетовими прямокутниками).

Наведемо діаграми розсіювання для двох більш типових тестових зображень, Baboon і Barbara, де перше є високотекстурним, а друге є зображенням середньої складності (Додаток Л, рис. Л.8). Як бачимо, основні залежності однакові. Властивості даних для тестового зображення Barbara (Додаток Л, рис. Л.8, б) дуже подібні до властивостей для тестового зображення Goldhill (Додаток Л, рис. Л.7, б), тоді як властивості даних для тестового зображення Baboon (Додаток Л, рис. Л.8, а) є такими, що ближче до тестового зображення Airfield (Додаток Л, рис. Л.7, а), яке також є більш текстурним. З огляду на наведені вище залежності, було проведено наступний статистичний аналіз для розглянутої сукупності тестових зображень. Для кожного KK і кожного N було розраховано середнє значення та стандартне відхилення співвідношення $S = MSE_{HVS-M \Sigma} / MSE_{HVS-M AGU}$. Отримані дані представлені в Додатку Л, рис. Л.9. Вони трохи неочікувані, але цікаві та корисні. Існує інтервал KK від 1 до приблизно 7, де середнє значення S досить велике (значно більше одиниці) з очевидною тенденцією до швидкого зменшення. Значення стандартного відхилення також великі в цьому інтервалі. Ці ефекти не

можна побачити в Додатку Л, на рис. Л.7 і Л.8. Хоча значення $MSE_{HVS-M \Sigma}$ більші, ніж $MSE_{HVS-M AGU}$, вони все ще дуже малі. Якщо $KK > 7$, то співвідношення має стабільні значення, близькі до 1,34. Це співвідношення не залежить від кількості аналізованих блоків. Це означає, що прогнозований $MSE_{HVS-M AGU}$ можна легко розрахувати як:

$$MSE_{HVS-M pred} = MSE_{HVS-M \Sigma} / 1,34. \quad (4.15)$$

Отже, запропоновано алгоритм прогнозування $MSE_{HVS-M AGU}$. Але на практиці потрібен алгоритм, щоб забезпечити бажаний $MSE_{HVS-M des}$. Для цього пропонується зробити наступне:

- 1) обчислити початкове значення кроку квантування KK як $KK_{init} = (MSE_{HVS-M des} / 0,02896)^{\frac{1}{2}}$;
- 2) якщо $KK_{init} \leq 7$, використовувати цей KK_{init} як кінцевий крок квантування KK_{fin} для стиснення;
- 3) якщо $KK_{init} > 7$, то обчислити $MSE_{HVS-M \Sigma}$ з використанням обмеженої кількості блоків і перерахувати його до $MSE_{HVS-M AGU}$;
- 4) порівняти цей $MSE_{HVS-M AGU}$ з $MSE_{HVS-M des}$; якщо точність є задовільною (наприклад, $MSE_{HVS-M AGU}$ відрізняється від $MSE_{HVS-M des}$ менш ніж на $\epsilon\%$, тоді використовувати KK_{init} як KK_{fin} ; інакше визначити KK_{fin} як:

$$KK_{fin} = KK_{init} \left(\frac{MSE_{HVS-M des}}{MSE_{HVS-M AGU}} \right)^{\frac{1}{2}}, \quad (4.16)$$

де припускається, що $MSE_{HVS-M AGU}$ приблизно пропорційна KK^2 .

Запропонований алгоритм було застосовано до всіх дев'яти тестових зображень для двох значень $MSE_{HVS-M des}$, а саме рівного 4, що відповідає непомітності спотворень, і рівного 10, яке стосується видимих, але не дратівливих спотворень. Отримані дані наведено відповідно в таблицях 4.7 і 4.8. Як впливає з

аналізу даних у таблиці 4.7, $MSE_{HVS-M\Sigma}$ і $MSE_{HVS-M\ AGU}$, отримані для KK_{init} , здебільшого значно менші, ніж $MSE_{HVS-M\ des}$. Це означає, що апроксимація (4.13) не зовсім точна для малих КК. Таким чином, необхідна поправка типу (4.16). Після виправлення отриманий MSE_{prov} становить близько $MSE_{HVS-M\ des}$, деякі значення MSE_{prov} більші за $MSE_{HVS-M\ des}$ (це трапляється для зображень складної структури), а деякі менші (це має місце для зображень простих структур).

Зазначимо, що значення КС коливаються в дуже широких межах від 4...5 для зображень складної структури до майже 19 для простих структур.

Розглянемо тепер дані в таблиці 4.8. Знову ж таки, $MSE_{HVS-M\Sigma}$ і $MSE_{HVS-M\ AGU}$ для KK_{init} значно менші, ніж $MSE_{HVS-M\ des}$. Після корекції вони стають приблизно $MSE_{HVS-M\ des}$, хоча деякі з них виявляються занадто великими, як для тестового зображення Diego. Оскільки введено більше спотворень, значення КС зросли порівняно з даними в таблиці 4.7. Таким чином, точність надання $MSE_{HVS-M\ des}$ зросла порівняно з (4.13), але її все ще варто покращити. Це можна зробити, оцінивши складність стисненого зображення та врахувавши цей фактор.

Таблиця 4.7 – Дані для $MSE_{HVS-M\ des} = 4$

Зображення	KK_{init}	$MSE_{HVS-M\Sigma}$	$MSE_{HVS-M\ AGU}$	KK_{fin}	MSE_{prov}	КС
Aerial	12	2,014	1,503	20	6,195	5,75
Airfield	12	2,431	1,814	18	6,129	4,26
Baboon	12	2,041	1,523	19	5,769	4,50
Barbara	12	3,653	2,726	15	3,505	8,70
Diego	12	1,840	1,373	20	6,790	4,50
Frisco	12	4,314	3,220	13	3,482	18,5
Goldhill	12	2,744	2,047	17	5,648	8,60
Lena	12	3,648	2,722	15	4,254	12,5
Mrt_prepared	12	2,381	1,777	18	4,130	14,9

В Додатку Л, на рис. Л.10 представлено тестове зображення Mrt_prepared, використане в аналізі, з позначеними фрагментами, на які слід звернути увагу при аналізі. Збільшений фрагмент в зеленому кадрі після компресії із забезпеченням

PSNR=38 дБ представлений в Додатку Л, на рис. Л.11, а, де внесені втрати невидимі.

Цей же фрагмент після стиснення із забезпеченням PSNR=36 дБ представлений в Додатку Л, на рис. Л.11, б, де стають помітними внесені втрати. Це підтверджує, що PSNR близько 37 дБ є порогом видимості спотворень.

Таблиця 4.8 – Дані для $MSE_{HVS-M des} = 10$

Зображення	KK_{init}	$MSE_{HVS-M \Sigma}$	$MSE_{HVS-M AGU}$	KK_{fin}	MSE_{prov}	КС
Aerial	19	6,136	4,579	28	14,33	7,72
Airfield	19	7,984	5,958	25	13,79	5,61
Baboon	19	6,626	4,945	27	14,27	5,95
Barbara	19	7,937	5,923	25	10,38	13,2
Diego	19	5,785	4,317	29	18,50	6,34
Frisco	19	8,292	6,188	24	9,132	30,8
Goldhill	19	9,066	6,765	23	11,24	12,1
Lena	19	8,949	6,678	23	8,909	19,2
Mrt_prepared	19	6,126	4,572	28	10,82	21,2

Зверніть увагу, що для розглянутого тестового зображення кодер AGU забезпечує КС, більший, ніж JPEG, приблизно в 1,8 рази. AGU також забезпечує КС більше, ніж JPEG2000, приблизно в 1,2 рази для PSNR від 30 до 40 дБ. Аналогічно, PSNR для AGU є значно більшим, ніж для JPEG2000 або SPIHT [87] для того самого КС. Щоб частково підтвердити це, в Додатку Л, на рис. Л.12, а, представлені залежності розглянутих метрик від КС для тестового зображення Mrt_prepared для кодера AGU. В Додатку Л, на рис. Л.12, б, представлені аналогічні залежності для методу стиснення SPIHT. Як видно, всі показники зменшуються, якщо КС зростає. Для КС приблизно дорівнює 10, характеристики введених помилок починають відрізнятися від АБГІІІ (PSNR і PSNR-HVS починають значно відрізнятися). При КС близько 25 ефект маскування зникає (PSNR і PSNR-HVS-M стають приблизно однаковими).

Порівняння даних в Додатку Л, на рис. Л.12 показує переваги кодера AGU. Значення відповідних показників на 1...2 дБ кращі (більші), ніж для SPIHT за того ж ступеня стиснення (для цього можна порівняти значення PSNR або PSNR-HVS-M для AGU і SPIHT в Додатку Л, на рис. Л.12 відповідно, $КС \approx 20$, що приблизно відповідає випадку, коли спотворення починають бути помітні спостерігачам).

4.2 Інтелектуальне стиснення зображень із втратами з забезпеченням візуально непомітних спотворень

Розглянемо ситуації стиснення зображень, що спотворено АБГШ та просторово-корельованим шумом.

Почнемо з випадку АБГШ. Оскільки дослідження візуальної якості для стиснених з втратами зашумлених зображень не надто поширені, було спочатку вирішено обмежитися аналізом зображень у відтінках сірого без необхідності розглядати особливості HVS, що стосуються сприйняття кольору.

Було вибрано три стандартних тестових зображення розміром 512 x 512 пікселів, а саме Baboon, Barbara та Goldhill, які досить відрізняються за вмістом.

Для попереднього аналізу було вибрано кодери AGU, ADCT, SPIHT. При стисненні зображень з шумом існує позитивний ефект фільтрації, але також існує ефект, коли вносяться деякі спотворення (легке змазування контрастних меж і деталей). Ці спотворення здебільшого викликані квантуванням коефіцієнтів ДКП, і це є негативним ефектом. Спільний внесок цих двох ефектів визначає якість стиснутих зображень і залежить від КК. Однак аналітично отримати параметри та статистику цих спотворень важко.

Ефективність розглянутих кодерів можна порівнювати по-різному. Наприклад, можна отримати залежності розглянутих метрик візуальної якості від БНП. Це хороший варіант для кодера SPIHT, для якого БНП (бітрейт) виступає в якості ПКС. Але це не найкращий варіант для кодерів AGU і ADCT, де КК виконує функцію ПКС. Таким чином, щоб порівняти характеристики кодерів ADCT і AGU, ПКС було встановлено для них як $КК = b\sigma$, де коефіцієнт b був обраний рівним 0,5,

1,5, 2,5, 3,5, 4,5, 5,5 і 6,5, σ – СКВ АБГШ. Використання більшого b призводить до більшого КК і, відповідно, більшого КС. Зверніть увагу, що для того самого КК і стисненого зображення кодер ADCT забезпечує приблизно на 1,03...1,07 більше КС, ніж AGU [272]. Складніше порівнювати ефективність кодера SPIHT з AGU і ADCT. Для цього спочатку було здійснено стиснення даного зашумленого зображення кодером AGU зі значеннями КК, визначеними вище. Потім було визначено значення КС, отримані для AGU, і перераховано до відповідних значень БНП. Після цього було застосовано кодер SPIHT із встановленням обчисленого набору БНП. Отримані набори значень представлені в Додатку Л, в таблиці Л.4. Оскільки більшість типових практичних значень дисперсії адитивного шуму знаходяться в межах від 25 до 200 [122], у експериментах було розглянуто дисперсії $\sigma^2=50$ і $\sigma^2=100$.

Було штучно змодельовано АБГШ із вищезазначеними дисперсіями, який було додано до тестових зображень Baboon, Barbara, Goldhill, що дало загалом шість зображень, спотворених шумом.

Після цього кожне зображення було стиснуте всіма трьома кодерами з кількома ступенями стиснення. Нижче показано залежності PSNR, PSNR-HVS, PSNR-HVS-M і WSNR [273] від b на одному графіку (оскільки всі ці показники надані в дБ) для кодерів AGU і ADCT. Подібним чином представлено й всі ці залежності від БНП для кодера SPIHT. Залежності MS-SSIM від b або БНП будуть представлені в таблицях. Залежності метрик розраховано між істинним зображенням (без шуму) та стисненим зашумленим зображенням. Почнемо з найпростішого випадку, тестового зображення Goldhill. Графіки представлені в Додатку Л, рис. Л.13, Л.14, а також у таблиці Л.5.

По-перше, варто зазначити, що для $\sigma^2=100$ має місце максимальне значення PSNR. Для кодерів ADCT і AGU ці максимуми спостерігаються для b близько 4, тобто для $\text{КК} \approx 40$. Нагадаємо, що в роботі [274] було рекомендовано встановити $\text{КК} = 4,5\sigma$ для кодерів із КК, які використовуються як ПКС. Для SPIHT також спостерігається максимум PSNR. Це відбувається для БНП приблизно 0,44, що відповідає $\text{КК} \approx 45$ для кодера AGU (див. в Додатку Л криву PSNR на рис. Л.14 і дані

в таблиці Л.4). Ці результати підтверджують, що OPT (згідно з критерієм PSNR) існують для різних кодерів. В Додатку К, на рис. К.6, б, також видно, що принаймні один показник візуальної якості, PSNR-HVS, подібно до стандартного PSNR, має максимальне значення для кодера ADCT, хоча цей максимум не є явним. Цей максимум спостерігається для $b \approx 3,5$, тобто для трохи менших значень b , ніж для стандартного PSNR. Таким чином, можна стверджувати, що існують такі ситуації (зображення, дисперсії шуму та, принаймні, одна метрика візуальної якості), що для них спостерігається OPT. Для кодера AGU значення PSNR-HVS залишаються майже постійними при $b < 3,5$, при більших b вони починають зменшуватися.

Для інших показників візуальної якості немає максимумів (OPT) для зображення Goldhill, якщо $\sigma^2 = 100$. Усі вони при збільшенні b мають монотонно спадний характер. Крім того, порівнюючи відповідні залежності для кодерів AGU і ADCT, можна побачити, що всі характеристики ADCT дещо кращі, ніж для AGU для всього розглянутого діапазону КС. Зверніть увагу, що максимум метрики MS-SSIM також існує для SPIHT.

Аналіз залежностей в Додатку Л, на рис. Л.14 показує, що для кодера SPIHT всі метрики візуальної якості не мають максимумів. Для БНП $> 0,44$ вони зменшуються повільно. Тоді, для менших БНП, зменшення стає швидшим.

Розглянемо тепер випадок $\sigma^2 = 50$. Результати для ADCT представлені в Додатку Л, в таблиці Л.5 для зображення Goldhill. Як видно, стандартний PSNR знову має максимум, який спостерігається для $b \approx 3,5$. PSNR-HVS, PSNR-HVS-M і WSNR монотонно зменшуються, якщо b збільшується, як у випадку $\sigma^2 = 100$ (див. графіки в Додатку Л, на рис. Л.13). Метрика MS-SSIM має максимум, який, знову ж таки, спостерігається при $b \approx 3,5$. Таким чином, згідно з метрикою MS-SSIM, OPT також може існувати. Для кодерів AGU і SPIHT всі залежності монотонно спадають.

Розглянемо тепер найскладніше тестове зображення Baboon. Графіки метрик якості для випадку $\sigma^2 = 50$ представлені для кодерів AGU та ADCT в Додатку Л, на рис. Л.15. Як видно для цих дисперсій шуму всі розглянуті метрики зменшуються, якщо b (і, відповідно, КС) збільшується. Зверніть увагу, що швидкість спадання

зростає, якщо b стає більшим. Це можна пояснити наступним чином. Позитивний ефект фільтрації шуму менший, ніж негативний ефект внесених спотворень у цьому випадку, оскільки шум не надто інтенсивний, а вміст зображення складний. Для кодера SPIHT ситуація така ж, тобто за всіма показниками якості декомпресованого зображення стає гіршою, якщо БНП зменшується.

В Додатку Л, в таблиці Л.6 представлені дані для кодера ADCT для зображення Baboon, $\sigma^2=100$. Як видно, усі показники також зменшуються (погіршуються), коли b і КС збільшуються. Такі ж тенденції спостерігаються і для інших кодерів. Це означає, що для таких складних зображень, як Baboon, стиснення з втратами призводить до зниження їх візуальної якості, навіть якщо вони спочатку містять шум. Це спостереження породжує завдання аналізу вмісту зображення перед стисненням і налаштування відповідних параметрів кодера. А тепер проаналізуємо дані, отримані для тестового зображення Barbara. Частина отриманих результатів наведена в Додатку Л, в таблицях Л.7 і Л.8. Згідно з даними в таблиці Л.7 (кодер AGU, $\sigma^2=50$), метрики PSNR і MS-SSIM мають максимуми для $b=4,5$.

Що стосується метрик PSNR-HVS, PSNR-HVS-M і WSNR, то всі вони монотонно спадають, але для b менше 3,5 залежність майже рівна. Подібні результати спостерігаються для кодерів ADCT і SPIHT. Для кодера ADCT максимум PSNR-HVS спостерігається при $b=3,5$.

Завершуючи попередній аналіз, можна констатувати наступне. По-перше, ОПТ можна спостерігати у випадках стиснення із втратами зашумлених зображень не тільки для стандартної метрики PSNR, але й для інших метрик (PSNR-HVS, MS-SSIM), які враховують HVS. По-друге, ці ОПТ можуть мати місце для кодерів, які базуються як на вейвлетах (наприклад, SPIHT), так і на ДКП (наприклад, AGU). По-третє, більша ймовірність мати ОПТ спостерігається у випадках стиснення більш простих за структурою зображень, спотворених шумом з досить великою дисперсією. У протилежних випадках залежності показників візуальної якості є монотонно спадними функціями КС. По-четверте, найчастіше ОПТ спостерігаються для $b=3,5$, якщо для даного кодера ПКС є КК (або для відповідного

БНП). Якщо слідувати цій рекомендації для простого зображення, кодер забезпечує стиснення з втратами в околі OPT (у сенсі візуальної якості стисненого зашумленого зображення). Якщо стиснене зображення має складну структуру, то стиснення з втратами з $KK=3,5\sigma$ призведе до невеликого погіршення візуальної якості стисненого зображення (див., наприклад, дані для зображення Baboon). Останній висновок служить передумовою для автоматичної процедури стиснення з втратами зашумлених зображень.

Хоча в Додатку Л, рис. Л.13 і Л.15 дозволяють провести деякі порівняння ефективності кодерів, вважаємо за краще представити більш ретельний аналіз. В Додатку Л, на рис. Л.16 наведено залежності метрики PSNR-HVS для всіх трьох розглянутих методів стиснення для зображення Goldhill, $\sigma^2=100$. Як видно, ADCT забезпечує, загалом, кращу візуальну якість (згідно з метрикою PSNR-HVS), ніж інші кодери для найбільш цікавої області $b<4,5$.

Зверніть увагу, що в цьому випадку надається дві переваги. По-перше, PSNR-HVS покращується в порівнянні з оригінальним зашумленим зображенням або, що еквівалентно, оригінальним зображенням, стисненим без втрат.

По-друге, для $b=3,5$ кодер AGU забезпечує $KC=9,5$, а кодер ADCT має $KC=10$. Це означає, що KC приблизно в 7,5 разів більше, ніж для стиснення без втрат (RAR дає $KC=1,13$, ZIP дає $KC=1,12$, RKIM дає $KC=1,35$, JPEG-LS забезпечує $KC=1,34$).

Наведемо відповідний приклад. В Додатку Л, на рис. Л.17, а, та Л.17, б, представлено відповідно безшумне та зашумлене зображення Goldhill. У свою чергу, в Додатку Л, рисунки Л.17, в, і Л.17, г, демонструють стиснені зображення для кодерів ADCT і AGU відповідно для оптимального $KK=35$. На зображенні на рис. Л.17, б, візуально помітний шум, який, безумовно, погіршує візуальну якість зображення. Для обох кодерів добре помітний ефект фільтрації зображення, особливо у верхній частині, яка відповідає небу. Трохи краща візуальна якість зображення на рис. Л.17, в, проявляється в чіткіших краях і деталях завдяки використанню схеми розбиття, яка дозволяє адаптуватися до вмісту зображення.

Проаналізуємо ефективність кодера з точки зору іншого показника. В Додатку Л, на рис. Л.18 представлені залежності PSNR-HVS і WSNR від b для

тестового зображення Barbara ($\sigma^2=100$). За метрикою PSNR-HVS (рис. Л.18, а) кодер ADCT знову перевершує інші кодери для розглянутого діапазону b від 2 до 5,5. ADCT трохи кращий і за метрикою WSNR (рис. Л.18, б). Однак ні для одного з розглянутих кодерів максимум не спостерігається. Може здатися, що згідно з метриками PSNR-HVS-M і WSNR, стиснення ні в якому разі не дає покращення (для будь-якої дисперсії зображення, кодера та шуму). Беручи до уваги результати досліджень, наведені вище, було вирішено розглянути більшу дисперсію шуму для тестового зображення Goldhill. Отримані залежності PSNR-HVS-M та WSNR від b для $\sigma^2=200$ представлені на рис. 4.6. Аналіз показує, що PSNR-HVS-M має максимум для кодера ADCT і, знову ж таки, спостерігається при $b=3,5$ (рис. 4.6, а). Таким чином, існують практичні ситуації, коли PSNR-HVS-M теж має ОПТ, як і інші метрики візуальної якості. Навпаки, залежності WSNR від b залишаються монотонно спадаючими (рис. 4.6, б).

Розглянемо найскладніший випадок – текстурне тестове зображення Baboon, спотворене не дуже інтенсивним шумом з дисперсією 50. Через ефект маскування шум видно лише в центральній частині цього зображення (див. рис. 4.7, б).

Якщо застосовано стиснення з втратами, ефект фільтрації спостерігається в цій частині (див. стиснені зображення на рис. 4.7, в, для кодера ADCT, $b=3,5$, і на рис. 4.7, г, для кодера SPIHT, БНП=1,52). В інших частинах стисненого зображення Baboon візуальні спотворення відсутні або спостерігаються дуже незначні артефакти.

У сукупності наведені результати показують, що якщо бажано стискати зображення, спотворені адитивним шумом, з гарною візуальною якістю, доцільно проводити стиснення з установкою $KK=3,5\sigma$ для кодерів, керованих кроком квантування. Це можливо виконати в автоматичному режимі, якщо на першому етапі оцінка стандартного відхилення шуму $\hat{\sigma}$ отримана сліпим (автоматичним) способом з високою точністю.

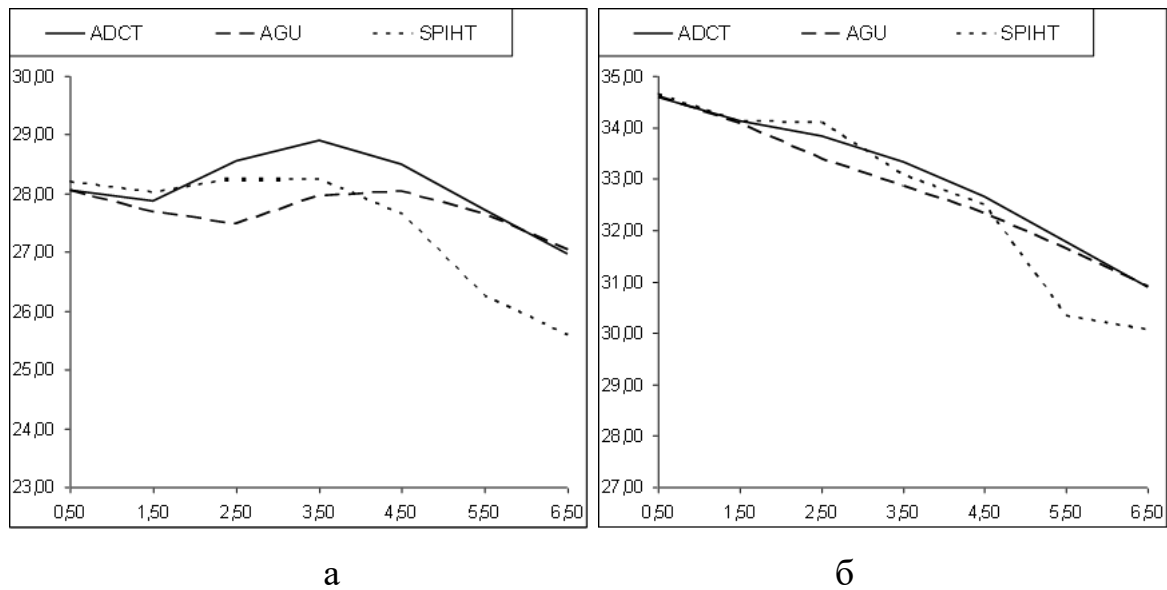


Рисунок 4.6 – Залежності PSNR-HVS-M (а) та WSNR (б) від b для різних кодерів, $\sigma^2=200$, зображення Goldhill

Таким чином, для AGU і ADCT автоматичне стиснення можна виконати досить легко. Якщо кодер керується БНП, як це зазвичай робиться для SPIHT і JPEG2000, ситуація складніша. Однак є декілька виходів. По-перше, вейвлет-кодери, як правило, також дозволяють встановлювати крок квантування, і він повинен бути пропорційним σ (якщо відомий апріорі) або його оцінці $\hat{\sigma}$. Фактор пропорційності залежить від алгоритму, який використовується для нормалізації вейвлет-коефіцієнта в даному кодері. По-друге, можна використовувати AGU з $KK=3,5\sigma$, визначити для нього КС AGU, а потім встановити БНП= $8/КС$ для SPIHT або JPEG2000.

Ця можливість враховує той факт, що кодери AGU і SPIHT або JPEG2000 забезпечують приблизно однаковий КС. Недоліки цього методу полягають у тому, що потрібно мати в розпорядженні AGU і стиснення проводити двічі, перший раз AGU, а потім SPIHT або JPEG2000. Третій спосіб полягає в наступному.

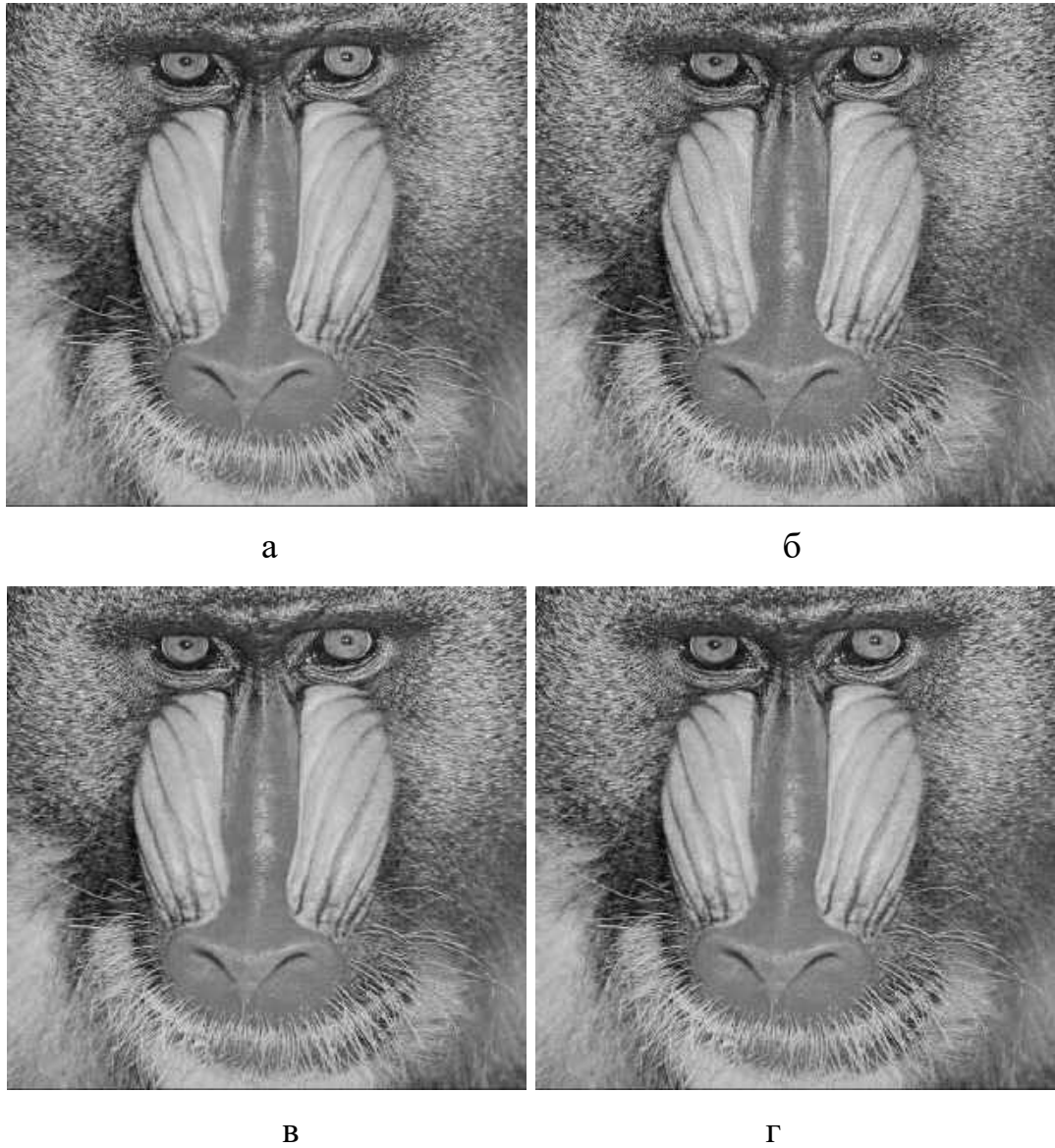


Рисунок 4.7 – Зображення Вабон: безшумне (а); зашумлене, $\sigma^2=50$ (б); стиснуте ADCT, $b=3,5$ (в); стиснуте SPIHT, БНП= 1,52 (г)

У роботі [275] була запропонована ітераційна процедура для надання заданого PSNR або MSE для стиснутого зображення відносно вихідного (за умови стиснення). Пропонується застосувати цю процедуру для забезпечення $MSE \approx 0,8\hat{\sigma}$, де MSE розраховується для стисненого/декомпресованого зображення щодо відповідного зображення з шумом. Зауважте, що ця процедура передбачає виконання багаторазового стиснення/декомпресії, доки не буде досягнуто БНП, що забезпечує $MSE \approx 0,8\hat{\sigma}$. Процедуру слід починати з досить великого БНП₁ (приблизно 6 для 8-бітних зображень) і продовжувати з БНП_i= БНП_{i-1} - ΔБНП, де

$\Delta\text{БНП}$ означає крок зменшення БНП, БНП_i позначає БНП, використаний на i -й ітерації.

Узагальнюючи отримані результати та враховуючи практичні аспекти стиснення з втратами, можна зробити наступні висновки. У сенсі кращої якості зображення та більшого коефіцієнта стиснення кращим є кодер ADCT. Однак AGU і, особливо, ADCT вимагають більше обчислень і, по суті, їх можна розглядати як деякі розширені версії JPEG. Тим часом, JPEG2000 також є стандартом. Таким чином, практичний вибір відповідного кодера є відкритим питанням і залежить від поточної програми та пріоритету вимог. Також варто взяти до уваги останні досягнення. Зокрема, нещодавно [34] (а також вище) було показано, що MSE внесених спотворень для AGU і ADCT можна досить точно і швидко спрогнозувати до стиснення. Крім того, цей MSE приблизно дорівнює $\text{KK}^2/12$ для зображень складної структури (для $\text{KK} \leq 20$) і зображень, спотворених шумом (для $\text{KK} \leq 5\sigma$). Для більшого KK MSE менший, ніж $\text{KK}^2/12$. Тож OPT (якщо вона існує) виникає, якщо MSE внесених спотворень приблизно дорівнює дисперсії шуму σ^2 . Тоді для $\text{KK}_{\text{OPT}}^2/12 \approx \sigma^2$ випливає, що $\text{KK}_{\text{OPT}} \approx 3,5\sigma$ або трохи більше.

Вище були представлені результати для трьох кодерів, розроблених 15...25 років тому. Теорія і практика стиснення зображень неухильно просуваються вперед і з'являються нові, в деякому сенсі кращі методики. Одним із них є better portable graphics (BPG) кодер [276], який є частиною стандарту відеокодування HEVC. Цей метод стиснення та відповідний алгоритм мають низку переваг, а саме: він перевершує багато існуючих методів стиснення, має версії для зображень у відтінках сірого та кольорових, підтримує різні формати тощо.

Залежності PSNR та PSNR-HVS-M від Q (істинне від зашумленого стисненого) наведені в Додатку Л, на рис. Л.19 і Л.20 відповідно. Як бачимо, усі OPT для різних зображень для однієї дисперсії АБГШ відбуваються для одного Q_{OPT} (рис. Л.19). Якщо дисперсія АБГШ збільшується, Q_{OPT} також збільшується. Можна показати, що Q для OPT можна визначити як $Q_{\text{OPT}} \approx [14,9 + 20 \log_{10}(\sigma)]$, де $[]$ означає округлення до найближчого цілого числа.

Як можна побачити, OPT існує для багатьох тестових зображень згідно PSNR і, рідше, згідно PSNR-HVS-M. Дослідження показали, що OPT може існувати для метрики MS-SSIM, і це трапляється рідше, ніж для PSNR, і частіше, ніж для метрики PSNR-HVS-M. Таким чином, можна стверджувати, що OPT також може існувати відповідно до кількох показників візуальної якості для кодера BPG. Як і раніше, це зазвичай відбувається для зображень простої структури, спотворених досить інтенсивним шумом. Відповідний приклад наведено в Додатку Л, на рис. Л.21.

Тепер розглянемо випадок стиснення зображень із втратами, якщо зображення спотворені просторово-корельованим шумом. Додамо до аналізу, крім кодерів AGU, ADCT, SPIHT, кодери AGU-M та ADCT-M.

Однією з проблем при розгляді просторово-корельованого шуму є те, що він може характеризуватися нескінченною кількістю спектрів або автокореляційних функцій шуму з різними параметрами. Ця спроба є однією з перших спроб проаналізувати стиснення із втратами зображення, спотвореного просторово-корельованим шумом, вивчається окремий випадок, коли реалізації просторово-корельованого шуму були отримані шляхом фільтрації двовірного АБГШ лінійно-усереднюючим фільтром з вікном 3×3 пікселі.

Почнемо з аналізу випадку шуму помірної інтенсивності з дисперсією, що дорівнює 100. Визначимо IPSNR, як вказано у формулах (2.17) та (2.18). Нагадаємо, що $IPSNR \approx 3$ дБ спостерігалось для тестового зображення Barbara, АБГШ, кодер AGU [242]. Трохи менший показник має місце для тестового зображення Goldhill, крім цього, OPT не спостерігається для тестового зображення Baboon. Приблизно такі ж результати спостерігаються для кодера ADCT[242].

Розглянемо тепер дані для просторово-корельованого шуму з дисперсією 100. Залежності PSNR(KK) (або PSNR(KM) для AGU-M і ADCT-M) представлені у вигляді PSNR(b) де $KK=b\sigma$ або $KM=b\sigma$. Графіки надані в Додатку Л, на рис. Л.22, Л.23 та Л.24.

В Додатку Л, на рис. Л.22 є 5 графіків для всіх порівнюваних кодерів. Аналіз графіків показує наступне. Кодери AGU та ADCT забезпечують OPT, яка має місце

для $KK \approx 5\sigma$. Ці OPT не є «очевидними», тобто спостережувані IPSNR дорівнюють лише 0,2 і 0,4 дБ відповідно. Кодер SPIHT не забезпечує OPT. Кодери AGU-M і ADCT-M також не забезпечують OPT, і результати стиснення PSNR(b) для них навіть трохи гірші, ніж для SPIHT. Якість стиснутих зображень поступово знижується, коли b (і KC) стає більшим. Для графіків для AGU та ADCTC в Додатку Л, на рис. Л.23 присутній ефект «граничної лінії», коли є максимуми залежності зі значеннями, які практично дорівнюють $PSNR(KK \rightarrow 0)$. Формально ADCT дає OPT, а AGU — ні. Максимум спостерігається при $b \approx 6$, тобто $KK \approx 6\sigma$. Знову ж таки, кодери SPIHT і AGU-M не забезпечують OPT. Крива для ADCT-M має локальний максимум, що спостерігається для $b \approx 1,5$, але формально це не OPT. Загалом ці три кодери мають приблизно однакову ефективність.

Для тестового зображення Baboon всі залежності монотонно спадають (рис. Л.24, Додаток Л). Однак кодери ADCT і AGU дають значно кращі результати, ніж інші три проаналізовані кодери. З представлених результатів можна припустити, що OPT неможливий для кодерів SPIHT, AGU-M і ADCT-M, якщо шум просторово корельований. Але не відповідає дійсності. На рис. 4.8 представлено графіки для тестового зображення Goldhill, спотвореного просторово-корельованим шумом із дисперсією 200. OPT спостерігаються для всіх п'яти методів стиснення, де найбільший IPSNR досягає 1 дБ. Це менше, ніж у відповідному випадку з АБГШ [242]. У цьому випадку OPT досягаються для $b \approx 8$. Результати для AGU-M, ADCT-M і SPIHT трохи гірші, ніж для AGU і ADCT.

Водночас виникає питання, чи можна покращити візуальну якість завдяки стисненню з втратами зображень, спотворених просторово-корельованим шумом?

Нижче представлено графіки (рис. Л.25-Л.27, Додаток Л) для трьох метрик візуальної якості для тестового зображення Goldhill, спотвореного просторово-корельованим шумом з дисперсією, що дорівнює 200. Через брак місця неможливо надати результати для інших зображень і значень дисперсії шуму. OPT спостерігаються для всіх кодерів для тестового зображення Barbara при дисперсії 200. Для кодера ADCT OPT існує для зображень Barbara та Goldhill за умови $\sigma^2 = 100$. OPT не спостерігається для тестового зображення Baboon, навіть якщо $\sigma^2 =$

200. Для метрики PSNR-HVS-M (рис. Л.25, Додаток Л) є OPT для всіх п'яти проаналізованих методів стиснення та вони спостерігаються для $b \approx 7$. PSNR-HVS-M покращується приблизно на 1,2 дБ для кодера ADCT, який працює найкраще.

Таким чином, ймовірність існування OPT зростає для більш простих зображень, більшої дисперсії та кращих кодерів. В Додатку Л, рисунок Л.26 демонструє, що OPT також може існувати для метрики візуальної якості WSNR. Знову це відбувається для $b \approx 7$. І знову ймовірність існування OPT більша для кращих кодерів (наприклад, ADCT), більшої дисперсії та зображення простішої структури. Нарешті, графіки в Додатку Л, на рис. Л.27 показують, що OPT також може існувати для метрики MSSIM. Висновки ті ж самі: OPT має місце для кодера ADCT при $b \approx 7$.

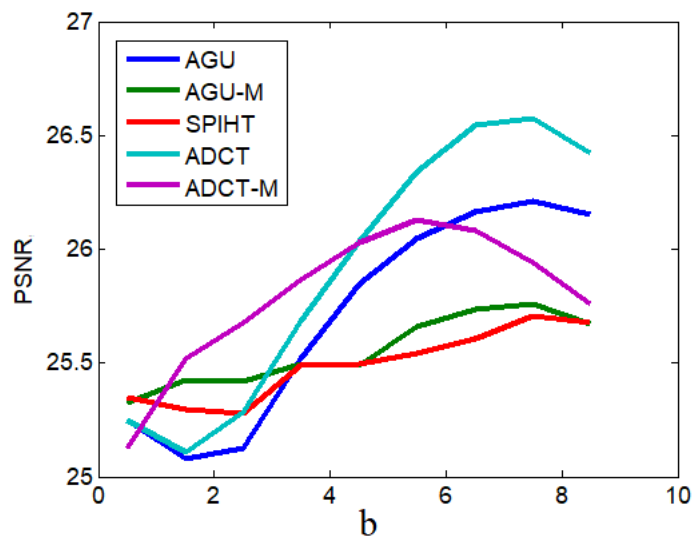


Рисунок 4.8 – Залежності PSNR від b для досліджуваних кодерів, Goldhill, спотворено просторово-корельованим шумом з дисперсією рівною 200

Таким чином, було продемонстровано існування OPT у стисненні з втратами зображень, спотворених просторово-корельованим шумом відповідно до показників візуальної якості. Покращення OPT, якщо воно є, невелике. Кодери, орієнтовані на поліпшення якості зображення, не найкращі для розглянутого застосування.

Висновки до розділу 4

Цей розділ присвячено питанням підвищення ефективності стиснення із втратами зображень дистанційного зондування Землі в умовах присутності складних завад і обмежених обчислювальних ресурсів з урахуванням візуальної якості. Проведено статистичний та просторово-спектральний аналіз спотворень, внесених двома кодерами на основі ДКП. Досліджено випадки візуального стиснення зображень без втрат і малопомітних спотворень. Було продемонстровано, що спотворення можуть бути як просторово некорельованими, так і корельованими, а їх локальна дисперсія залежить від локальної активності зображення. Показано, що дисперсія спотворень значно менша, ніж $KK^2/12$ в локально пасивних областях. Водночас локальна дисперсія введених спотворень становить близько $KK^2/12$ для локально активних областей, для яких $\sigma_{im}/KK \geq 1$. Показано, що досить точне прогнозування можливе не лише для MSE, але й для метрик візуальної якості з еталоном.

Проведений аналіз дозволив запропонувати метод для прогнозування MSE, який був інтенсивно перевірений для зображень різного походження. Показано, що висока точність прогнозування забезпечується для двох кодерів на основі ДКП. Додатковими перевагами цього прогнозування є його висока швидкість обробки. Це дозволяє не тільки прогнозувати MSE, але й обчислити KK, який буде встановлено для даного зображення, щоб забезпечити бажаний MSE.

Проаналізовано особливості стиснення з втратами для медичних та інших типів зображень. Показано, що один і той же KK або той самий КС можуть відповідати суттєво різній візуальній якості стиснутих зображень. Якщо потрібно виконати стиснення зображення з втратами із забезпеченням бажаного рівня візуальної якості, необхідно використовувати метрики візуальної якості та прогнозувати їх значення. Показано, що це можна зробити шляхом простого та швидкого аналізу даних у домені ДКП для обмеженої кількості блоків зображення (наприклад, 500 або 1000), вибраних випадковим чином. Така статистика дозволяє

включати властивості зображення та встановлювати ПКС адаптивно до складності зображення.

Аналіз отриманих результатів показує, що за показниками візуальної якості MS-SSIM, PSNR-HVS та PSNR-HVS-M, максимумами цих показників можуть існувати у випадку стиснення зображень, спотворених адитивним шумом. Якщо ПКС для кодера є крок квантування, тоді рекомендованим вибором є налаштування $KK=3,5\sigma$. Вищезгадані максимумами можуть (але не обов'язково) мати місце для відносно простих зображень, спотворених досить інтенсивним шумом. Для високотекстурних зображень, пошкоджених неінтенсивним шумом, залежності метрик від KK монотонно спадають. Однак навіть у цих випадках вибір $KK=3,5\sigma$ є розумним (для кодера BPG налаштування $Q=28$ створює візуально невидимі або слабо помітні спотворення). Потім ПКС і КС повинні бути адаптовані до статистики шуму. Кодер ADCT забезпечує найкращу візуальну якість стиснених зображень серед розглянутих кодерів. Кодер BPG забезпечує таку ж або, навіть, кращу ефективність. Також бажано адаптувати ПКС до глобальних характеристик (складності, контексту) зображень, що стискаються.

ОРТ може існувати при стисненні з втратами зображень, спотворених просторово-корельованим адитивним шумом. Крім того, ОРТ може існувати відповідно до різних метрик і для різних методів стиснення. З більшою ймовірністю ОРТ існує для кращих методів стиснення, більш простих структурних зображень і більш інтенсивного шуму. Можливо, методи стиснення можна адаптувати до спектральних властивостей шуму, і це покращить стиснення. Слід також звернути увагу на значення КС – вони були приблизно 5...20 для рекомендованих $b \approx 7$.

Основні наукові результати, наведені у четвертому розділі, опубліковано в роботах автора [20], [25], [27], [32], [33], [37], [43], [55], зазначених в анотації.

РОЗДІЛ 5 МЕТОДИ ІНТЕЛЕКТУАЛЬНОГО СТИСНЕННЯ МЕДИЧНИХ ЗОБРАЖЕНЬ

В даному розділі розглянуто питання стиснення медичних зображень, в тому числі спотворених шумами різної природи. Зокрема, проаналізовані особливості стиснення медичних зображень, що сформовані сучасними медичними системами реєстрації зображень. Також виявлені особливості шумів, що присутні на більшості медичних зображень. Розроблено підходи і надано рекомендації щодо стиснення таких зображень із урахування візуальної якості. Основні результати автора по даній тематиці наведені у роботах [23], [24], [26], [39], [53], [56], [61], [62].

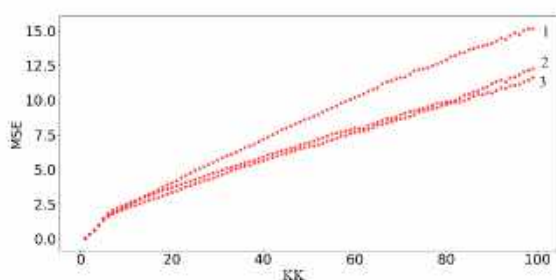
5.1 Використання сучасних кодерів для роботи з медичними зображеннями

5.1.1 Стиснення незашумлених медичних зображень

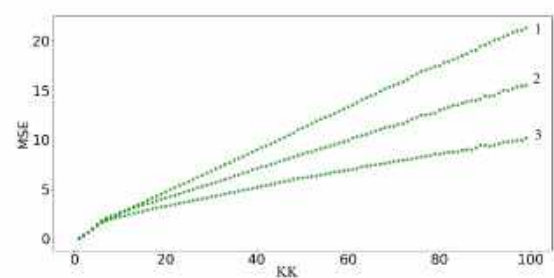
Мета цього підрозділу - розглянути застосовність розроблених підходів до медичних зображень. Зверніть увагу, що використання нестандартного кодера AGU має як переваги, так і недоліки. Одна з переваг полягає в тому, що можна досягти більшого КС у порівнянні з JPEG. Першими розглянемо стиснення кольорових зображень сітківки.

На самому початку було вирішено проаналізувати більш простий випадок застосування AGU окремо до R, G та B компонентів зображень сітківки. Було запропоновано розглянути три зображення сітківки ока дещо різної складності (з різною кількістю деталей, важливих для діагностики). MSE внесених втрат як функції КК для трьох кольорових компонентів трьох тестових зображень представлені на рис. 5.1. Аналіз цих залежностей показує, що вони дещо відрізняються для різних кольорових компонентів і різних кольорових зображень. Залежності також відрізняються для компонентів кольору, де найбільші

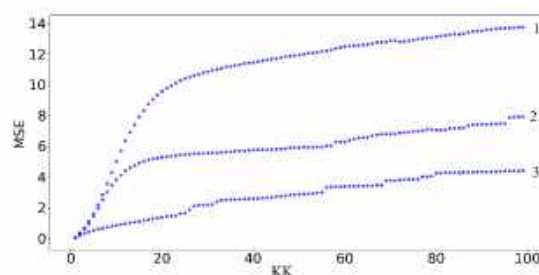
спотворення спостерігаються для синього компонента [277]. Стверджується, що КС може бути до 30 для JPEG і до 50 для вейвлет-кодера [277]. Це дозволяє визначити, що MSE внесених втрат може становити близько 6 для червоних і зелених компонентів і до 10 для «менш важливого» синього компоненту. Згідно з даними на рис. 5.1, це має місце для КК близько 30 для червоної компоненти, для КК близько 30 для зеленої компоненти та близько 20 для синьої складової одного тестового зображення. Подивимося, який КС можна забезпечити в цих випадках.



а



б



в

Рисунок 5.1 – Залежності MSE від КК для трьох різних тестових зображень (1, 2, 3) для R (а), G (б) і B (в) компонентів кольору

В Додатку М, на рис. М.1 представлені залежності КС від КК для червоної та синьої компонент (результати для зеленої компоненти є проміжними). Видно, що червоний компонент для КК близько 30 стискається з КС більше 80 (зелений компонент для КК близько 30 стискається з КС близько 90). Проблема полягає в зображенні синього компонента, яке гірше стискається, і причина в тому, що це зображення компонента достатньо зашумлене (в Додатку М, на рис. М.2 представлено вихідне, стиснене та різницеве зображення для синього компонента).

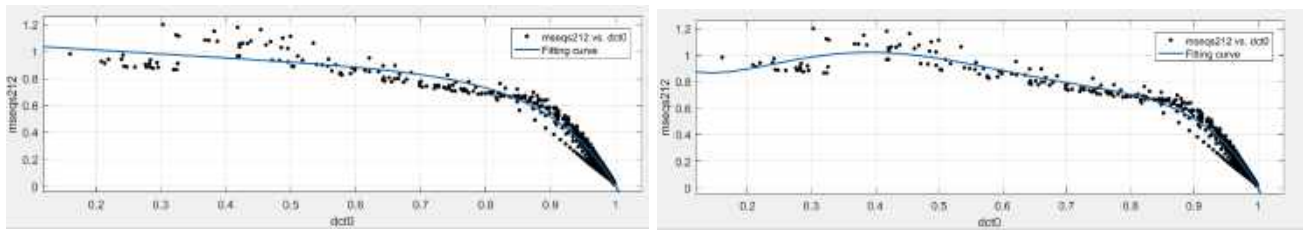
Як видно, шум частково видаляється шляхом стиснення з втратами. Цей ефект був розглянутий в попередніх розділах.

Таким чином, перше завдання полягає в тому, щоб забезпечити бажану MSE при окремому стисненні компонентних зображень. В попередніх розділах було встановлено, що MSE приблизно дорівнює $KK^2/12$ для малих КК. Також у розділі 3.2 було показано, що параметри стиснення (зокрема, КС) сильно залежать від відсотка нулів P_0 для квантованих ДКП коефіцієнтів. По аналогії з методикою, запропонованою у розділі 3 проведемо регресійний аналіз на базі кольорових зображень сітківки з метою визначення КК як функції відсотка нулів коефіцієнтів ДКП.

Користувач вищезазначених інструментів, який контролює та аналізує результати підгонки, має вирішити, які функції використовувати, які обмеження на значення функції та її поведінку. Аналіз даних в Додатку К, на рис. К.8 показує, що функція підгонки має бути досить плавною, вона повинна бути або монотонною, або мати максимум для P_0 близько 0,4, а значення функції повинні бути додатними для всього діапазону можливих значень P_0 .

Враховуючи вищезазначене, проаналізували декілька типів функцій для підгонки кривої до діаграми розсіювання на рис. 5.2, а. Ці функції є сумою двох експонент з чотирма регульованими параметрами ($R^2=0,982$, $RMSE=0,030$), підгонкою ряду Фур'є ($R^2=0,956$, $RMSE=0,047$), поліномом п'ятого порядку ($R^2=0,985$, $RMSE=0,027$, рис. 5.2, б), і степеневою функцією ($R^2=0,980$, $RMSE=0,032$). Сума двох експоненціальних функцій ($f(P_0)=-6,125e^{-05*\exp(9,477*x)} + 1,073*\exp(-0,2914*x)$) і поліном п'ятого порядку є однаково хорошим вибором.

Пригадаємо проблеми з компонентним зображенням В для RGB-представлення даних і покомпонентного стиснення. Нагадаємо також, що JPEG, застосований до кольорових зображень, спочатку перетворює їх на іншу систему кольорів.



а

б

Рисунок 5.2 – Графіки розсіювання та апроксимаційні криві для суми двох експонент (а) і полінома п'ятого порядку (б)

Отже, перетворимо оригінальне зображення в систему кольорів YCrCb (вважається однією з найкращих) і розглянемо стиснення кожного компонента. Результати компресії (залежності MSE від КК і КС від КС) представлені в Додатку М, на рис. М.3 і М.4 відповідно. Як видно, результати по всіх трьох компонентах стали ближчими. Це пов'язано з тим, що шум, присутній у зображенні синього компонента, поширився між усіма трьома компонентами після перетворення. Крім того, шум став практично непомітним. Це добре видно в Додатку М, на рис. М.5, де представлені всі три складові зображення. Крім того, видно, що всі три компоненти містять цінну інформацію, яку необхідно зберегти за допомогою стиснення з втратами.

Аналіз даних в Додатку М, на рис. М.3 показує, що бажана MSE може бути забезпечена для КК близько 15...20. При цьому значення КС для складових зображення становлять від 20 до 70 (див. дані в Додатку М, на рис. М.4), тобто на тому ж рівні, що й у [277]. В Додатку М, на рис. М.6 представлено оригінальне кольорове зображення та стиснене зображення, де для всіх трьох компонентів використовувався КК=15. Отриманий КС дорівнює 114, що практично відповідає рекомендованому в [277] для вейвлет-кодерів. Внесені спотворення практично непомітні. Подібні значення КС від 90 до 120 спостерігалися для інших досліджуваних зображень сітківки. Середній час стиснення зображення сітківки за допомогою комп'ютера з Intel(c) Core (c) i7-4710 HQ 2,5 ГГц займає менше 2 секунд. Декомпресія займає приблизно стільки ж часу.

Вищезазначені експоненціальні та поліноміальні наближення відповідають діаграмі розсіювання з R^2 приблизно 0,976 і середньоквадратичним значенням приблизно 0,040. Таким чином, дійсна та сама процедура визначення КК, як і описана в підрозділі 3.2.

Аналогічно зображенням сітківки перевіримо прогнозування метрик візуальної якості для стоматологічних зображень. Окрім кодера AGU, долучимо до експерименту кодер AGU-M. Великі панорамні зображення зубів було розбито на 20 фрагментів 512x512 пікселів, які було використано в якості тестових. Було стиснуто два фрагменти (в Додатку М, на рис. М.7). Ці фрагменти відрізняються своєю складністю: перший фрагмент (dental1_part23) менш складний (містить менше текстур, країв і деталей), а другий (dental2_part13) має більш контрастні деталі. MSE внесених втрат і КС як функції КК для кодера AGU представлені в Додатку М, на рис. М.8. Аналогічно, залежності для AGU-M від КМ наведені на рис. М.9.

Аналіз графіків показує наступне. Залежності MSE від КК і MSE від КМ для розглянутих фрагментів поведуться специфічним чином. Для малих КК або КМ, MSE більше для більш простого фрагмента структури (див. графіки в Додатку М, на рис. М.8, а, і М.9, а). У свою чергу, для більших КК або КМ, MSE більше для фрагмента в Додатку М, на рис. М.7, а. MSE внесених спотворень для даного фрагмента для AGU та AGU-M майже однакові для $КК=КМ$. Залежності MSE від КК і КМ зростають швидше для малих КК і КМ. КС для даного КК або КМ є меншим для більш складного фрагмента (див. графіки в Додатку М, на рис. М.8, б, і М.9, б), особливо якщо КК або КМ достатньо великі. При цьому вони майже однакові для $КК<10$ і $КМ<10$. Ці результати показують, що якість стисненого зображення, що характеризується MSE, значною мірою залежить від КК або КМ, але також залежить від характеристик зображення.

Проаналізуємо якість стисненого зображення з точки зору показників візуальної якості. Для цієї мети будемо використовувати MSE_{HVS} та MSE_{HVS-M} . Нагадаємо, що MSE і MSE_{HVS-M} приблизно рівні, якщо спотворення мають

властивості, близькі до АБГШ. Графіки MSE_{HVS} і MSE_{HVS-M} від КК для кодера AGU представлені в Додатку М, на рис. М.10.

Аналогічно, графіки MSE_{HVS} і MSE_{HVS-M} від КМ для кодера AGU-M наведені на рис. 5.3. Перше спостереження полягає в тому, що криві для обох тестових фрагментів поводяться майже однаково для $KK < 40$ і $KM < 40$, тобто лише для значень КК і КМ, які представляють інтерес для поточного випадку. По-друге, MSE_{HVS} і MSE_{HVS-M} , здається, приблизно пропорційні КК і КМ для кодів AGU і AGU-M, відповідно. Таким чином, перше завдання полягає в тому, щоб забезпечити бажане значення метрики, яка характеризує якість стиснутого зображення.

Тут варто нагадати, що існують приблизні порогові значення для кількох метрик, які визначають поріг невидимості для зображень, спотворених різним чином [245]. Наприклад, передбачається, що MSE має бути приблизно 20 (PSNR приблизно 35 дБ), тоді як MSE_{HVS-M} має бути приблизно 4, щоб забезпечити PSNR-HVS-M приблизно 42 дБ.

Оскільки MSE_{HVS-M} є більш адекватною мірою візуальної якості, ніж MSE, зацікавлені в аналізі можливості прогнозування MSE_{HVS-M} . Було отримано діаграму розсіювання MSE_{HVS-M} від КК для кодера AGU, представлену на рис. 5.4. Кожна точка цієї діаграми розсіювання відповідає одному тестовому зображенню (загалом 20 тестових зображень зубів), стисненому з одним значенням КК (від 1 до 40, крок 1). Як видно, точки розміщені компактно, їх розбіжність зростає із збільшенням КК. Очевидною є тенденція зростання MSE_{HVS-M} при зростанні КК.

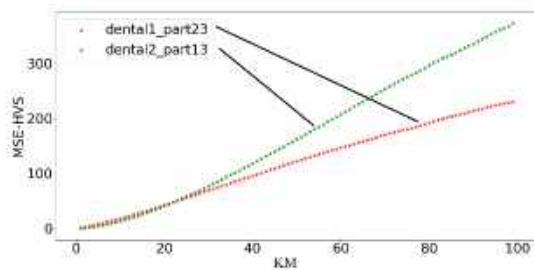
Завдяки компактності діаграми розсіювання легко побудувати криву, що описує залежність MSE_{HVS-M} від КК (приклад наведено на рис. 5.4).

В результаті регресійної підгонки були отримані 3 достатньо ефективні моделі:

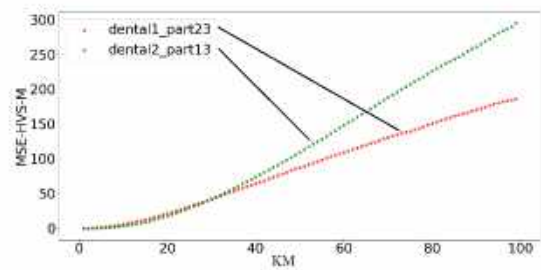
$$MSE_{HVS-M} = 0,02896KK^{1,976}, \quad (5.1)$$

$$MSE_{HVS-M} = 0,02505KK^2 + 0,077KK - 0.7803, \quad (5.2)$$

$$MSE_{HVS-M} = 42,49 \exp(-((KK - 45,79)/21,56)^2). \quad (5.3)$$



а



б

Рисунок 5.3 – Залежності MSE_{HVS} (а) і MSE_{HVS-M} (б) від KM для двох тестових зображень, кодер AGU-M

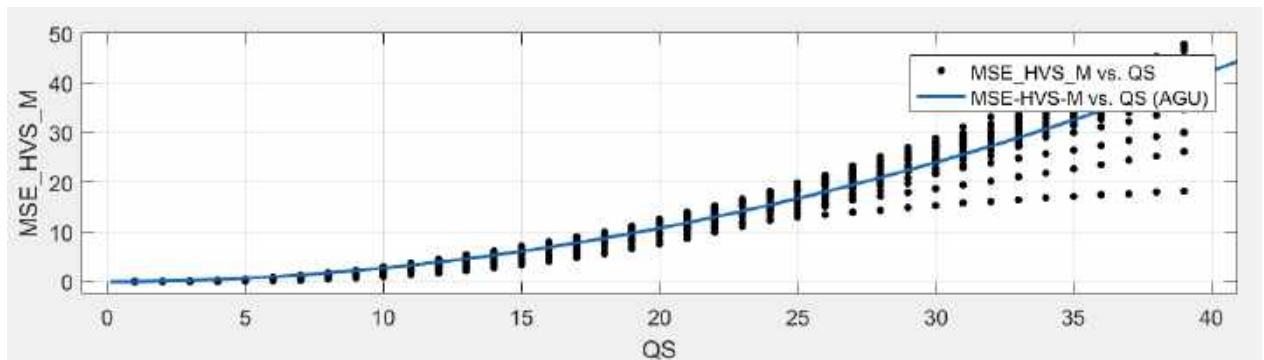


Рисунок 5.4 – Діаграма розсіювання MSE_{HVS-M} від KK для стиснення з втратами зображень у градаціях сірого, кодер AGU

Короткий аналіз цих виразів показує, що MSE_{HVS-M} приблизно пропорційна KK^2 . Для всіх трьох наведених вище виразів значення R^2 практично однакові і дорівнюють 0,95, тобто досить високі. Значення RMSE дорівнюють 2,89, 2,88 і 2,91, тобто всі три апроксимації (5.1) – (5.3) майже однаково добрі. Тому для вибору найкращого наближення необхідно враховувати інші фактори. Апроксимація (5.1) є найкращою, оскільки вона дає MSE_{HVS-M} дуже близько до нуля для $KK=0$ і всі значення є позитивними (на відміну від апроксимації (5.2)).

Розглянемо тепер дані для кодера AGU-M. Діаграма розсіювання та один приклад підігнаної кривої представлені на рис. 5.5. Основні спостереження такі ж, як і в попередньому випадку. Отримані наближення такі:

$$MSE_{HVS-M} = 0,1195KM^{1,976}, \quad (5.4)$$

$$MSE_{HVS-M} = 0,0275KM^2 + 0,5307KM - 2.161, \quad (5.5)$$

$$MSE_{HVS-M} = 61,4\exp(-((KM - 44,92)/22,73)^2). \quad (5.6)$$

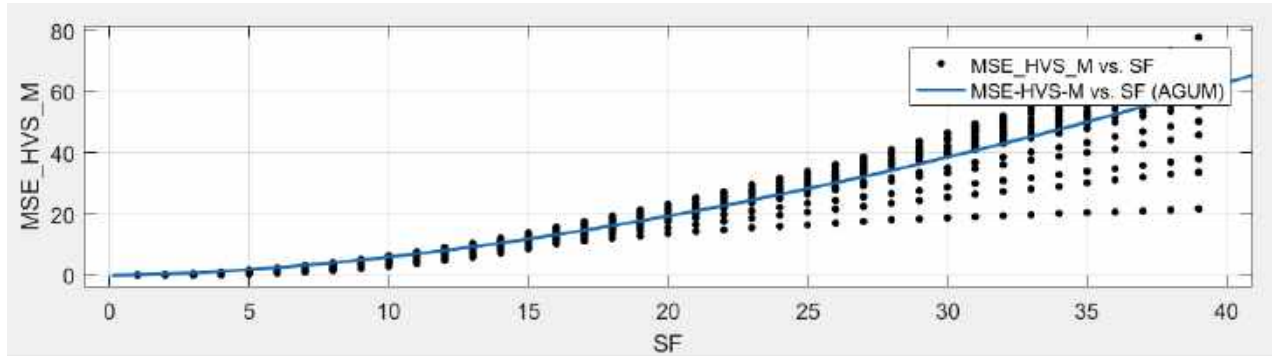


Рисунок 5.5 – Діаграма розсіювання MSE_{HVS-M} від KM для стиснення з втратами зображень у градаціях сірого, кодер AGU-M

Отримані значення R^2 є меншими, ніж у попередньому випадку, і вони становлять близько 0,91, тоді як усі RMSE приблизно дорівнюють 6. Виходячи з аналізу, наближення (5.4) є найкращим. Тоді, щоб забезпечити $MSE_{HVS-M}=4$, KM має бути близько 8. Ці результати добре узгоджуються з усередненими залежностями, представленими на рис. 16 в роботі [168] для $PSNR-HVS-M_{des}$ приблизно 42 дБ, що відповідає $MSE_{HVS-M}=4$ для 8-розрядного представлення оброблених зображень (для стоматологічних медичних зображень). Можна стверджувати, що значення RMSE для отриманих наближень занадто великі, а отже, прогноз недостатньо точний. Однак аналіз результатів підгонки показує, що найбільші похибки мають місце для $KK>30$ і $KM>25$, тобто для таких значень, при яких внесені спотворення є помітними, і таке стиснення з втратами є неприйнятним.

З'ясуємо значення KC , отримані шляхом стиснення з втратами з рекомендованими параметрами. Кодер AGU з $KK=12$ для 20 розглянутих тестових фрагментів дав значення KC від 6,2 до 24,2, де MSE_{HVS-M} варіювалося від 1,78 до 4,52. Таким чином, точність забезпечення бажаного MSE_{HVS-M} близько 4 досить

висока, а отримані значення КС досить великі. Аналогічно для розглянутих тестових фрагментів були проаналізовані значення КС для кодера AGU-M з $KM=8$.

Найбільший $КС=31,3$, тоді як найменший $КС=8,9$. Це означає, що КС значно залежить від вмісту (складності) зображення. MSE_{HVS-M} коливався від 1,53 до 4,12. Стиснені зображення представлені в Додатку М, на рис. М.11, а, та рис. М.11, б. Порівнюючи їх із відповідними зображеннями в Додатку М, на рис. М.7, можна констатувати, що видимих спотворень не вноситься. Крім того, щоб підтвердити збереження діагностично важливої інформації (деталі), на рис. М.11, в, та рис. М.11, г, показано збільшений фрагмент до і після стиснення. Як бачимо, ніяких відмінностей виявити не вдається. Найменший КС спостерігався для високотекстурного тестового фрагмента в Додатку М, на рис. М.12, де показано оригінальні та стиснені зображення. Різниці не спостерігається. Порівнюючи результати для AGU та AGU-M, можемо стверджувати, що AGU-M забезпечує трохи більший КС для заданої якості або трохи кращу якість для даного КС. Середній час стиснення фрагмента 512×512 пікселів за допомогою CPU Intel(c) Core (c) i7-4710 HQ 2,5 ГГц займає менше 1,32 секунди для AGU і менше 1,33 секунди для AGU-M. Декомпресія займає приблизно стільки ж часу.

5.1.2 Стиснення медичних зображень з урахуванням шумів

Стоматологічні зображення, отримані за допомогою сучасних систем візуалізації, мають великий розмір, принаймні 1 мегапіксель [53], [278], [279]. Зокрема, це стосується зображень, отриманих системою панорамного рентгену Morita Veraviewerocs 3D R100 J [280]. Такі зображення мають загалом прямокутну форму, де співвідношення розмірів залежить від системи формування зображення та режиму її роботи. Тим часом може знадобитися стиснути фрагмент зображення в певній області, яка може бути квадратною або прямокутною та мати різний контекст. Приклади таких фрагментів наведено на рис 5.6.

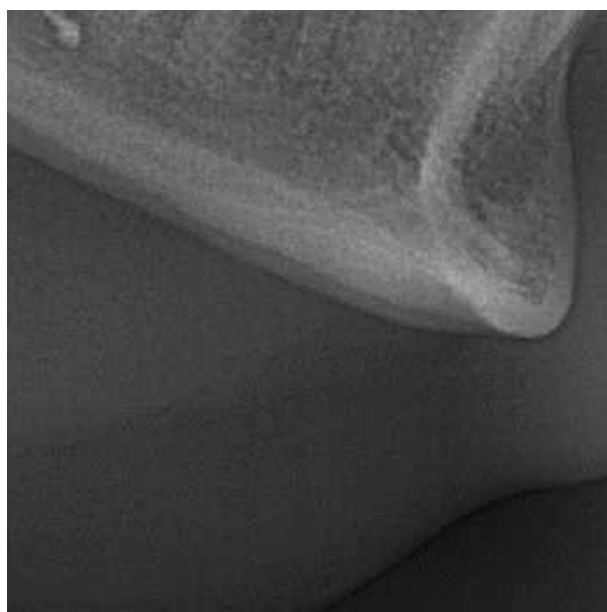
На зображеннях зубів присутній шум, як видно з цих фрагментів, який може частково маскувати важливу діагностичну інформацію. Цей шум виникає через

обмежений час фіксації зображення об'єкта, а властивості шуму є складними [53], [281], [282]. У найпростішому випадку можна змодельовати шум на дентальних зображеннях як такий, що має Пуассонівський закон розподілу ймовірностей [281]. Однак більш детальний аналіз показує, що він, загалом, залежить від сигналу, але має більш складну природу. Він має складову СНШ і складову СЗШ, дисперсія якої може бути пропорційна інтенсивності сигналу, що відповідає моделі (2.19). Наприклад, дисперсія може бути представлена як $\hat{\sigma}^2 = \hat{\sigma}_\mu^2 \cdot I_{ij}^2 + \hat{k} \cdot I_{ij} + \hat{\sigma}_a^2$ (модель 3), де I_{ij} — справжнє значення інтенсивності ij -го зображення пікселя; $\hat{\sigma}_\mu^2$ — оцінка відносної дисперсії мультиплікативного шуму; $\hat{\sigma}_a^2$ — оцінка дисперсії адитивного шуму; $\hat{k}$ — параметр квазіпуассонівського шуму.

Експерименти в роботі [53] показали, що прості моделі типу $\hat{\sigma}^2 = \hat{\sigma}_\mu^2 \cdot I_{ij}^2 + \hat{\sigma}_a^2$ (модель 1) та $\hat{\sigma}^2 = \hat{k} \cdot I_{ij} + \hat{\sigma}_a^2$ (модель 2) часто відповідають реальним даним краще, ніж складна загальна модель, наведена вище. Експерименти, проведені з системою Morita, показали, що еквівалентна дисперсія шуму для найзагальнішої моделі залежить як від властивостей зображення, так і від шуму та приблизно дорівнює:

$$\hat{\sigma}_{eq}^2 = \hat{\sigma}_a^2 + \hat{\sigma}_\mu^2 \sum_{i=1}^N \sum_{j=1}^M I_{ij}^2 / NM + \hat{k} \sum_{i=1}^N \sum_{j=1}^M I_{ij} / NM. \quad (5.7)$$

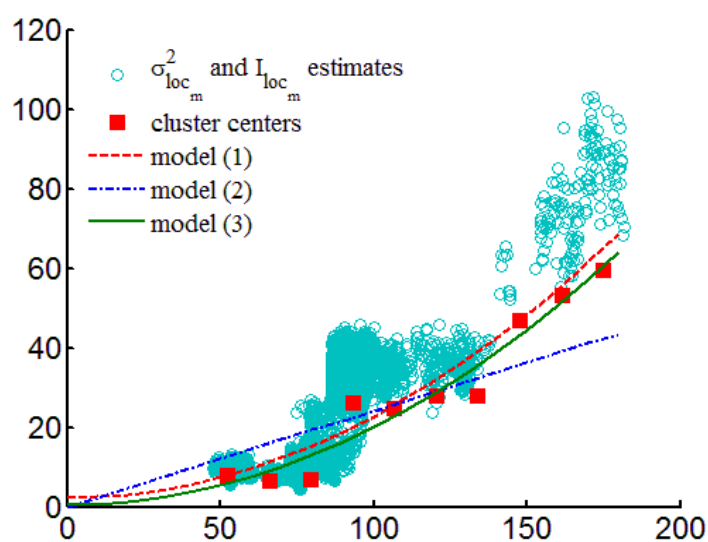
На рис. 5.6, в, представлено діаграму розсіювання оцінок локальної дисперсії локальних середніх оцінок для фрагмента на рис. 5.6, а. Ці оцінки були отримані лише в локально-пасивних (квазіоднорідних) блоках цього фрагмента — відповідна карта розташування блоків наведена на рис. 5.6, д, криві для всіх трьох моделей підігнані (рис. 5.6, в), і видно, що моделі 1 і 3 забезпечують майже ті самі наближення, які є значно кращими, ніж для моделі 2 (це не проблема визначити найкращу відповідність, див. [53] для більш детальної інформації). Визначені параметри для моделі 1 для фрагмента на рис. 5.6, в, такі: $\hat{\sigma}_\mu^2 = 0,0020$, $\hat{\sigma}_a^2 = 2,27$ і еквівалентна дисперсія відповідно дорівнює 15,7.



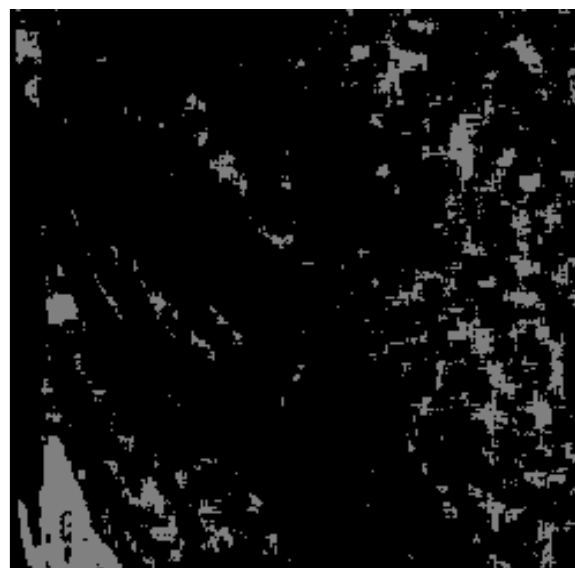
а



б



в



г

Рисунок 5.6 – Фрагменти розміром 512×512 пікселів різного контексту (а, б) і діаграма розсіювання оцінок локальної дисперсії (в) для однорідних областей (г), сірий колір) фрагмента (а)

Подібні операції дають $\hat{\sigma}_{eq}^2 = 80,7$ для фрагмента на рис. 5.6, б ($\hat{\sigma}_{\mu}^2 = 0,0043, \hat{\sigma}_a^2 = 37,49$).

Зверніть увагу, що шумові характеристики знімків зубів системи Morita аналізувалися для двох режимів її роботи. Для одного режиму $\hat{\sigma}_{eq}^2$ змінюється в діапазоні 5–20 для зображень із низькою середньою інтенсивністю (рис. 5.6, а) та в діапазоні 50–70 для зображень із високою середньою інтенсивністю. Для іншого режиму $\hat{\sigma}_{eq}^2$ становить приблизно 30 для зображень із низькою середньою інтенсивністю та знаходиться в діапазоні 60–200 для зображень із низькою та високою середньою інтенсивністю (див. фрагмент рис. 5.6, б). У більшості випадків шум добре видно на візуалізованих даних. Детальний аналіз, проведений у [53], також показав, що шум є просторово корельованим там, де спектральні властивості залежать від режиму формування зображення, і можуть бути артефакти відсікання. Зверніть увагу, що всі вищезазначені властивості та ефекти частково пояснюються специфічним нелінійним перетворенням, яке здійснюється в пристроях для формування зображень з метою кращої візуалізації.

Будемо використовувати кодер ADCT для проведення дослідження. На рис. 5.7 показано схеми розбиття для досліджуваних фрагментів зображень.

Максимальний розмір блоку становить 64×64 пікселя, мінімальний розмір — 8×8 пікселів, усі блоки мають прямокутну форму зі сторонами, що пропорційні ступеню двійки. Більший блок ділиться на два менших на основі ентропійного аналізу (наприклад, коли блок містить різномірні частини).

Залежності PSNR і PSNR-HVS-M від КС для фрагментів на рис. 5.6 представлені в Додатку М, на рис. М.13 для ADCT, JPEG2000 і SPIHT. Аналіз цих кривих показує наступне:

1) ADCT забезпечує значно кращі результати, ніж JPEG2000 і SPIHT: більші значення PSNR і PSNR-HVS-M для КС в діапазоні інтересів для обох тестових зображень;

2) ці залежності можна трактувати по-іншому — для даної якості забезпечується більший КС (фіксовані значення PSNR або PSNR-HVS-M); виграв в КС може сягати 20%.

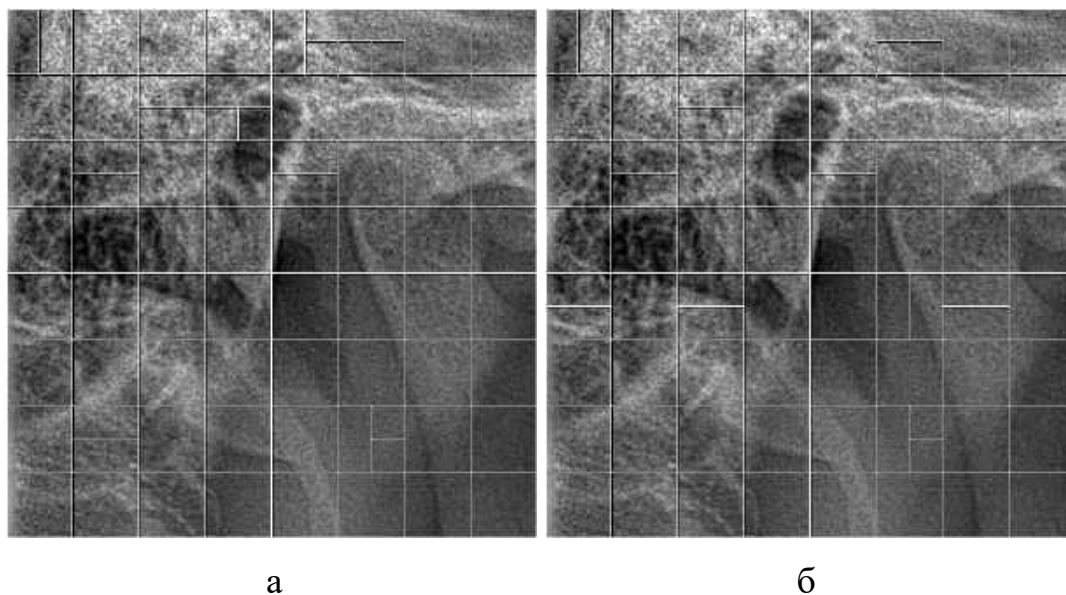


Рисунок 5.7 – Фрагмент розміром 512×512 пікселів із двома схемами розділення отримано для $KK = 3$ (а) та $KK = 13$ (б)

Можна згадати, що ADCT використовує ДКП, і метрика PSNR-HVS-M також використовує ДКП. Через це порівняння ADCT з кодерами JPEG2000 і SPIHT, обидва на основі вейвлетів, з точки зору PSNR-HVS-M може здатися некоректним. Через це також було проведено подібний аналіз, використовуючи індекс схожості ознак FSIM, який є низькорівневим показником оцінки якості зображення на основі ознак. Залежності наведено в Додатку М, на рис. М.14, і їх аналіз підтверджує, що ефективність ADCT також є кращою за метрикою FSIM.

Як було зазначено вище, більші значення KK відповідають більшим ступеням стиснення. Щоб продемонструвати це, наведемо графіки, отримані для цілочисельних значень KK в межах від 1 до 40 для двадцяти фрагментів зображення зубів 512×512 різної складності (рис. 5.8, а). Усі залежності (KC від KK для заданого фрагмента зображення) є монотонними.

При цьому для даного KK значення KC змінюються в широких межах залежно від контексту фрагмента зображення. Наприклад, для $KK = 10$ KC змінюється від 6 до 22, а для $KK = 20$ KC змінюється від 12 до 78. Тобто KC може відрізнитися на один порядок або навіть більше. PSNR монотонно зменшується, якщо KK зростає (рис. 5.8, б).

Якщо $KK \leq 5$, значення PSNR для однакових KK майже однакові для всіх фрагментів. Для більших KK значення PSNR для однакових KK можуть сильно відрізнятися. Наприклад, для $KK = 20$ значення PSNR знаходяться в межах від 34 до 40 дБ (тобто вони змінюються в широких межах). Це означає, що встановлення заданого KK призводить до широкого розкиду KC і PSNR, якщо KK великий. Іншими словами, встановлюючи деякий KK , нелегко забезпечити бажану якість стиснутих зображень, принаймні, відповідно до метрики PSNR і для великих KK . Наприклад, $KK = 20$ призводить до невидимості спотворень для одних зображень і їх видимості для інших зображень (зверніть увагу, що, згідно з [245], спотворення з високою ймовірністю є непомітними, якщо $PSNR > 36$ дБ).

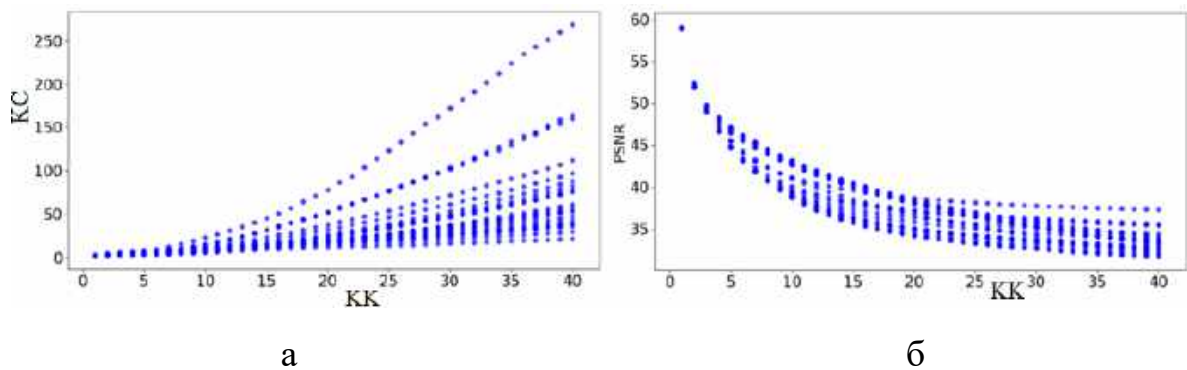


Рисунок 5.8 – Залежності KC (а) та PSNR (б) від KK для всіх фрагментів зображення зубів, кодер ADCT

Що стосується порогів видимості: експериментальні дані, представлені в роботі [245], показують, що спотворення практично непомітні, коли значення PSNR-HVS-M перевищує 41 дБ або $FSIM > 0,99$. Ще однією умовою є те, що спотворення мають бути «рівномірно» розподілені по стисненому зображенню. В попередніх розділах були запропоновані ефективні безітераційні методи для забезпечення бажаної візуальної якості стиснених зображень. Це важливо для ADCT, оскільки він вимагає більше обчислень, ніж JPEG, через оптимізацію схеми розділення та використання ДКП різних розмірів. Зверніть увагу, що під «швидким методом» із застосуванням до ADCT мається на увазі ефективний безітераційний метод стиснення. Розглянемо залежності PSNR-HVS-M та FSIM від KK . Вони

представлені для всіх двадцяти фрагментів зображення на рис. 5.9. Як видно з цього рисунка, точки кривої розміщені компактно. Практично гарантовано, що $\text{PSNR-HVS-M} > 40$ дБ і $\text{FSIM} > 0,99$, якщо $\text{KK} \leq 12$. Винятком є фрагменти зображення з дуже простою структурою. Це можна підтвердити й іншими результатами. Вище було знайдено, що $\text{MSE}_{\text{HVS-M}} = 0,02896\text{KK}^{1,976}$.

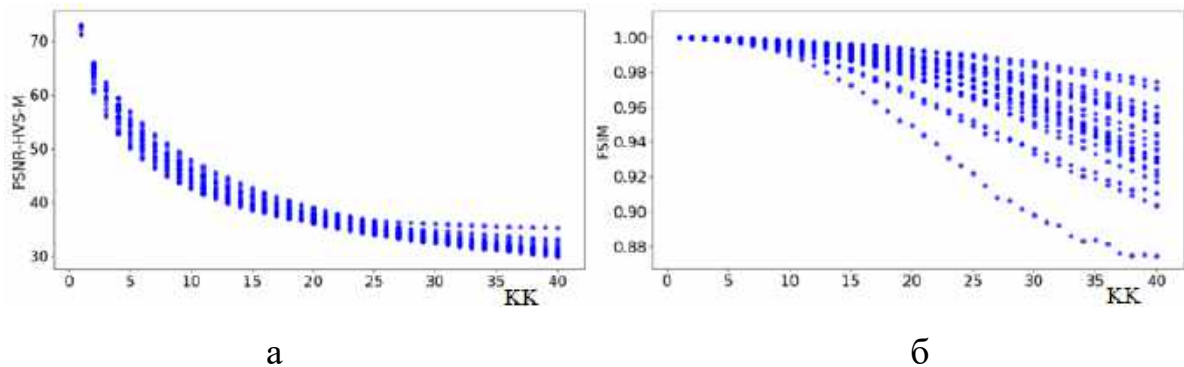


Рисунок 5.9 – Залежності PSNR-HVS-M (а) та FSIM (б) від KK для всіх фрагментів зображення зубів, кодер ADCT

Цей вираз отримано для кодера AGU, але він також приблизно справедливий для ADCT. Тоді маємо $\text{KK}_{\text{des}} \approx (\text{MSE}_{\text{HVS-Mdes}}/0,02896)^{0,5}$, де $\text{MSE}_{\text{HVS-Mdes}}$ є бажаним $\text{MSE}_{\text{HVS-M}}$ (для PSNR-HVS-M близько 41 дБ це відповідає 5) отримуємо $\text{KK}_{\text{des}} \approx 13$.

Вже було обговорено ефект фільтрації шуму, якщо до зашумленого зображення застосовується стиснення з втратами. Останні дослідження показують, що люди практично не виявляють змін рівня шуму, якщо ці зміни на порядок менші за рівень початкового шуму [283] (тобто в цьому випадку, внесені спотворення мають бути на порядок меншими, ніж шум $\hat{\sigma}_{eq}^2$ в оригінальних зображеннях). Тоді для значення $\hat{\sigma}_{eq}^2$ близько 50 потрібно забезпечити PSNR на рівні 40 дБ. Це відбувається, якщо KK становить близько 11 (див. дані на рис. 5.8б). Таким чином, можемо загалом встановити KK на рівні значення 12. Далі перевіримо цей висновок для фрагментів зображення зубів. Але спочатку порівняємо ефективність запропонованого методу з іншими підходами. Ефективність кодера можна

аналізувати та порівнювати різними способами. Метою є те, щоб метод стиснення з втратами гарантував стиснення з втратами з невидимими спотвореннями за одну ітерацію (тобто був швидким) і з високими значеннями КС. ADCT з обраним КК = 12 було застосовано до набору фрагментів зображення зубів. Потім було визначено середнє PSNR-HVS-M ($\text{PSNR-HVS-M}_{\text{mean}}$) і середньоквадратичне його відхилення ($\text{RMSE}_{\text{PSNR-HVS-M}}$). Аналогічно було розраховано середнє КС (КС_{mean}) і середньоквадратичне значення ($\text{RMSE}_{\text{КС}}$): $\text{PSNR-HVS-M}_{\text{mean}} = 42,5$ дБ, $\text{RMSE}_{\text{PSNR-HVS-M}} = 1,42$ (значення PSNR-HVS-M змінювалися від 40,5 до 45,6 дБ (тобто було практично гарантовано, що внесені спотворення невидимі)), $\text{КС}_{\text{mean}} = 12,82$, $\text{RMSE}_{\text{КС}} = 5,57$ (КС змінюється від 7,5 до 20,6 залежно від складності зображення).

Для коректного порівняння було встановлено такий КС для JPEG2000, що $\text{PSNR-HVS-M}_{\text{mean}} = 42,5$ дБ. Цей КС виявився рівним 6,5 (тобто значно меншим, ніж $\text{КС}_{\text{mean}} = 12,82$). Для JPEG2000 $\text{RMSE}_{\text{PSNR-HVS-M}} = 1,76$ (значення PSNR-HVS-M змінювалися від 40,5 до 46,9 дБ), тобто метод на основі JPEG2000 із фіксованим КС або бітрейтом створює менший КС і гіршу точність у забезпеченні бажаної якості зображення, що характеризується PSNR-HVS-M. Зауважте, що точність для JPEG2000 можна ще покращити, якщо застосувати ітераційну процедуру стиснення; однак те саме можна зробити і для ADCT.

Сім спеціалістів-стоматологів (ортодонти та педодонти з досвідом понад 3 роки) пройшли підготовку й анонімно оцінили 140 рентгенівських знімків (95 панорамних та 45 бічних цефалографічних). Оцінка проводилася за такими критеріями якості:

- оптимальні для діагностики: 89 знімків;
- адекватні для діагностики: 38 знімків;
- погані, але піддаються діагностиці: 13 знімків;
- нерозпізнані: 0 знімків.

Для аналізу використовували монітори наступних пристроїв: ноутбук ASUS (15,6", 1920×1080, Full HD, IPS), iPhone XR та iPad. Порівняння оригінальних і стиснених зображень показало ідентичні результати.

Візуально в більшості випадків відмінностей виявлено не було (див. приклад на рис. 5.10). Навіть у тих випадках, коли були виявлені незначні зміни (менше 3% усіх випадків), вони не були пов'язані з погіршенням будь-якої корисної діагностичної інформації або артефактами (помічено деяке придушення шуму). Таким чином, отримані результати не показали відмінностей, які мали б вплинути на встановлення діагнозу, а стиснуті зображення були кваліфіковані так само, як і оригінал відповідно до Таблиці оцінки якості клінічного зображення та оцінки спеціалістів-стоматологів.

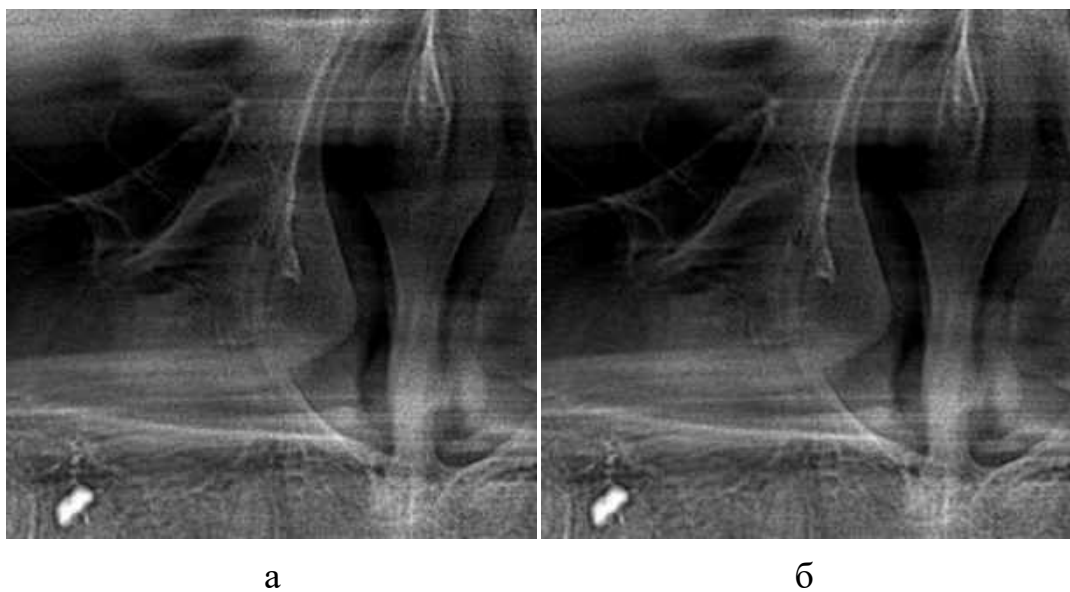


Рисунок 5.10 – Фрагмент зображення зуба до (а) і після (б) компресії з рекомендованим КК, КС = 13,5

Якщо оригінальні зображення мають представлення з більш ніж 8 бітами, рекомендований КК можна обчислити як $\text{ДД}/20$. Зверніть увагу, що отриманий КС був приблизно на один порядок більший, ніж КС для стиснення без втрат (ZIP забезпечує КС, що змінюється в діапазоні від 1,5 до 2,1). Розглянемо випадок інших тестових медичних зображень з шумом та кодери AGU і BPG. При виборі тестових зображень необхідно вибирати зображення різної складності. Зображення простої структури краще стискаються з меншими втратами, вони мають вищу якість для

заданого КС для вибраного методу стиснення та мають краще придушення шумів [128]. Зображення складної структури мають протилежні властивості.

Було вибрано шість зображень (деякі тестові зображення було запозичено з сайту <https://radiopaedia.org/>, решта – отримані в УСЦ ХНМУ) як тестові зображення у форматі PNG, з яких `foot_512.png` і `hand_512.png` можна класифікувати як простий тип зображення, `knee_joint_512.png` – як складні за структурою зображення, а `MRI_Lung_512.png`, `Shpenov_512.png`, `Zuzana_512.png` – зображення середньої складності. Приклади зображень наведено на рис. 5.11.

Будемо використовувати для аналізу метрику PSNR. Було розглянуто випадок АБГШ різних інтенсивностей. Нагадаємо, що важко помітити шум із дисперсією близько 25, якщо зображення є високотекстурним. Цю властивість врахуємо наступним чином. Шум практично непомітний, якщо PSNR більше 36 дБ для зображень простої структури та більше 30...33 дБ для високотекстурних зображень. Це зумовлює діапазон значень дисперсій завад, що аналізуються.

На рис. 5.12 представлені приклади таких випадків. Усі зображення спотворено шумом із $\sigma^2 = 64$. Шум практично непомітний на текстурному зображенні `MRI_Lung_512.png`, тоді як його можна помітити в однорідних областях тестового зображення з шумом `hand_512.png`. Залежності традиційного PSNR, розрахованого між незашумленим і зашумленим стисненим зображенням, від параметра Q для цих тестових зображень наведено на рис. 5.13 (зауважте, що Q є лише цілим числом і його можна змінювати від 1 до 51 для BPG, КК змінювався від 1 до 100 для AGU).

Як видно, найкращий (найбільший) PSNR для $Q > 15$ (що є цікавим для практики) спостерігається для тестового зображення `foot_512.png` із великими однорідними областями. Тестові зображення середньої та високої складності в основній зоні інтересу ($Q > 20$) мають значення PSNR, менші на 3-5 дБ порівняно з PSNR для `foot_512.png` для декодера BPG і 1-10 дБ для декодера AGU. Для $Q > 40$ і $KK > 40$ стиснені зображення переважно характеризуються низькою якістю (PSNR < 40 дБ для обох кодерів), а внесені спотворення є дратівливими. Таким чином, повинні зосередитися на $10 < Q < 40$ і $5 < KK < 40$.



Рисунок 5.11 – Незашумлені тестові зображення, використані в дослідженні: а – hand_512.png, б – knee_joint_512.png, в – Zuzana_512.png, г – MRI_Lung_512.png

Були розраховані відповідні залежності для шести тестових зашумлених зображень з дисперсіями АБГШ 25 і 100 відповідно. PSNR було розраховано між зашумленим істинним зображенням та зашумленим стисненим зображенням. Залежності для кодера BPG наведені на рис. 5.14. У загальному сенсі ці залежності є типовими КШС, де якість відповідно до PSNR погіршується, якщо Q збільшується. Однак варто звернути увагу на два моменти. По-перше, для певного Q (більше певного порогу T_1 і менше певного порогу T_2) криві для всіх тестових

зображень «йдуть разом» з досить високим абсолютним значенням першої похідної. Для кривих на рис. 5.14, а, це має місце при $6 \leq Q \leq 27$, тоді як для другого випадку (рис. 5.14, б) спостерігається більший інтервал такої поведінки ($5 \leq Q \leq 33$).

Якщо дисперсія шуму відома заздалегідь або попередньо оцінена з відповідною точністю, можна спрогнозувати PSNR для певного діапазону значень Q для кодера BPG за допомогою кривих на рис. 5.14. Це також можна розглядати іншим способом. Можна легко встановити Q для забезпечення бажаного PSNR. Наприклад, якщо потрібно забезпечити $\text{PSNR} = 40$ дБ, рекомендується використовувати $Q = 23$. Зауважте, що ця рекомендація дійсна для більшості тестових зображень, які практично вільні від шуму. По-друге, при $Q > T2$ криві починають розходитися (рис. 5.14) і різниця між значеннями PSNR для розглянутих тестових зображень стає більшою при більшому Q .

Ці відмінності можуть досягати 5...7 дБ (див. дані для $Q = 40$ на рис. 5.14, а, PSNR для більш простих структурних зображень є більшим, найменші PSNR спостерігаються для найскладніших структурних зображень).

У [182] визначено приблизний поріг непомітності спотворень для PSNR. Таким чином, якщо PSNR між спотвореним і відповідним еталонним зображенням перевищує 36 дБ, то дуже ймовірно, що спотворення не видно. У всіх випадках, на рис. 5.14, це відбувається при $Q \leq 25$ (і це добре узгоджується з даними [168]). Також метрику PSNR можна розрахувати між істинним та зашумленим стисненим зображенням.

Залежності PSNR (Q) для тих самих тестових зображень і дисперсії шуму, що дорівнює 25 і 100, представлені на рис. 5.15. Їх аналіз показує наступне. По-перше, залежності поводяться дуже по-різному залежно від властивостей тестового зображення та дисперсії шуму.



Рисунок 5.12 – Зашумлені тестові зображення, використані в дослідженні: hand_512.png (а), knee_joint_512.png (б), Zuzana_512.png (в), MRI_Lung_512.png (г), $\sigma^2 = 64$

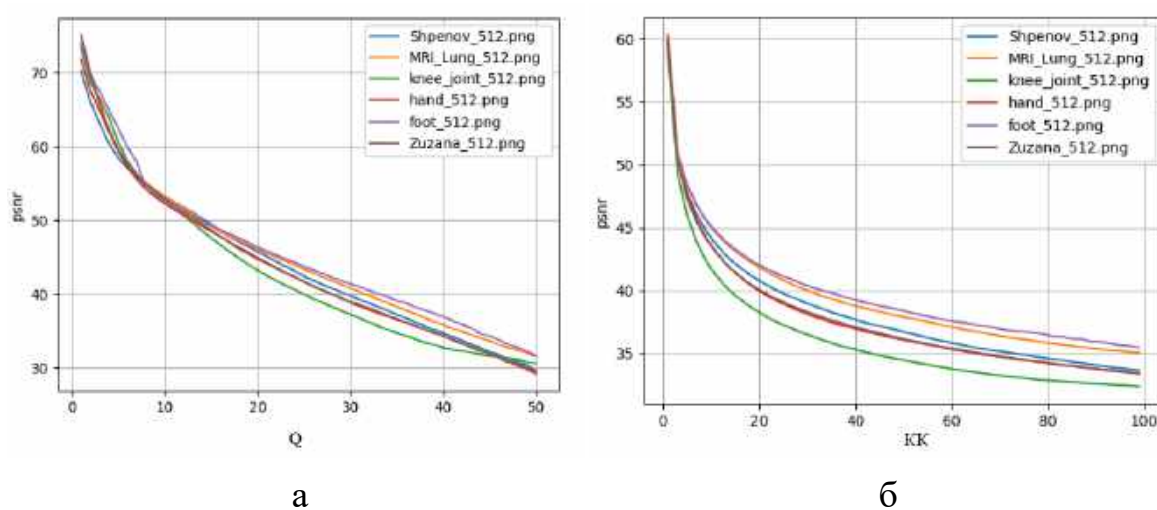


Рисунок 5.13 – КШС для шести тестових незашумлених зображень: кодер VPG (а) та кодер AGU (б)

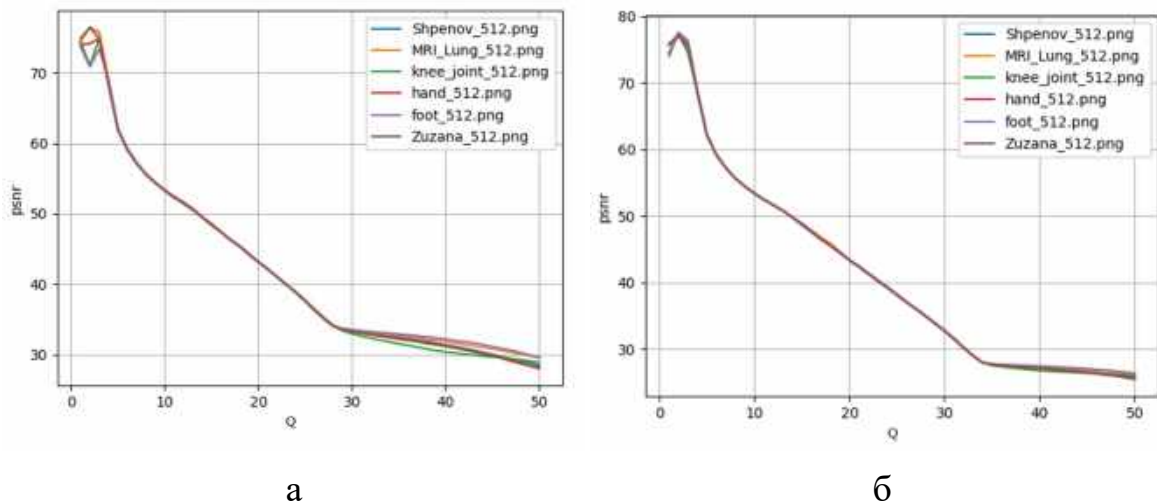


Рисунок 5.14 – КШС для шести тестових зашумлених зображень (PSNR розраховано між зашумленим нестисненим та зашумленим стисненим), кодер BPG: $\sigma^2 = 25$ (а) та $\sigma^2 = 100$ (б)

Єдина спільна річ полягає в тому, що вони починаються з певного значення, «йдуть разом» з невеликим зменшенням PSNR до певного Q_{com} , а потім розходяться. Для $Q > Q_{\text{com}}$ PSNR для зображень простіших структур більше, ніж для зображень складних структур, і ця різниця може досягати 5...8 дБ для великих значень дисперсії шуму. Зауважте, що Q_{com} становить приблизно 16 для $\sigma^2=25$ і 20 для $\sigma^2=100$ для кодера BPG. Для кодера AGU висновки підтверджуються невеликим відхиленням значень KK_{com} . Таким чином, KK_{com} становить приблизно 7 для $\sigma^2=25$ і 17 для $\sigma^2=100$ для кодера AGU.

По-друге, можна припустити, що для стандарту стиснення з втратами BPG, як і для інших стандартів стиснення з втратами, можливі три практичні ситуації: сприятлива, нейтральна та несприятлива. Для сприятливої ситуації існує очевидна ОПТ, тобто існує максимум PSNR(Q): $\text{PSNR}(Q_{\text{OPT}}) > \text{PSNR}_{\text{inp}}$. Тоді є очевидна причина для виконання стиснення з втратами в ОПТ (або в його околах, якщо ОПТ з якоїсь причини не може бути точно встановлено) – одночасно досягаються дві переваги: якість стисненого зображення краща, ніж оригінальне зображення з шумом і досягається досить високий КС.

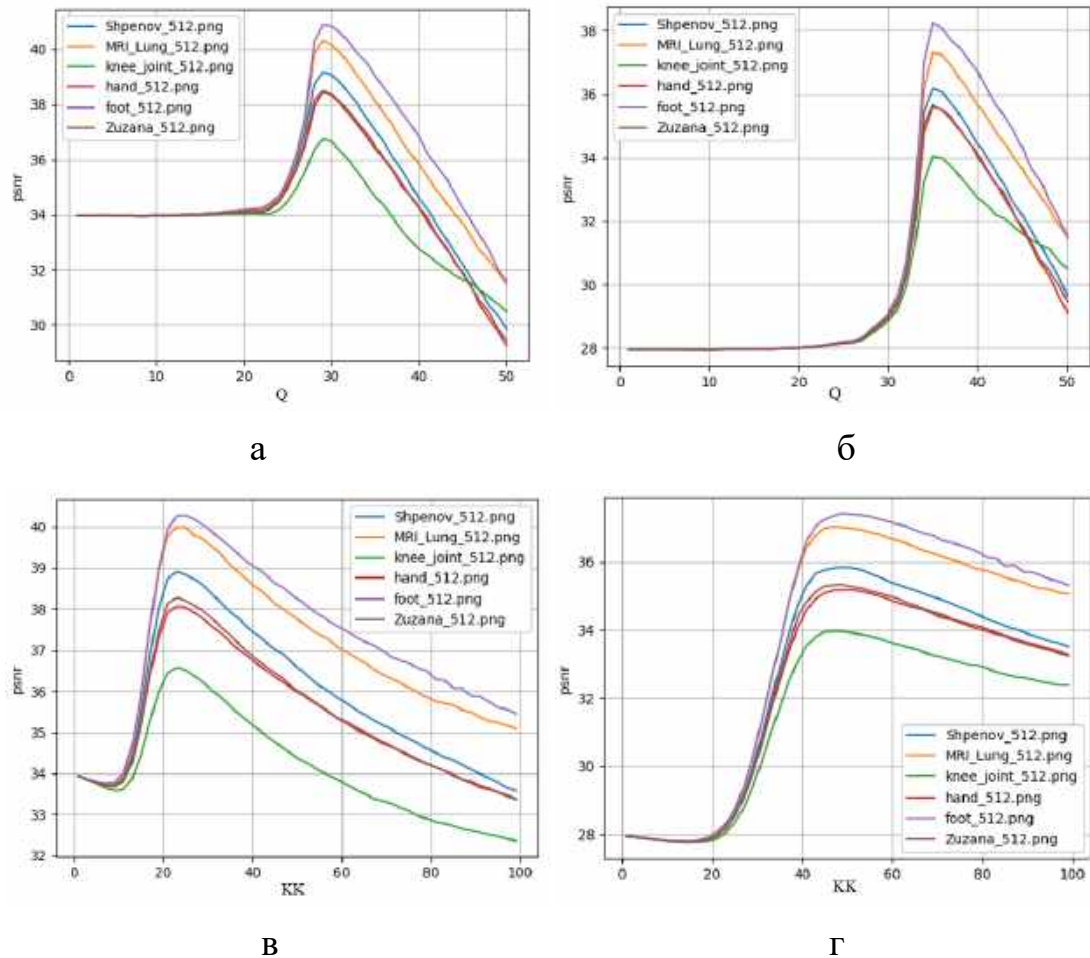


Рисунок 5.15 – КШС для шести тестових зашумлених зображень (PSNR розраховано між істинним та зашумленим стисненням), кодер BPG: $\sigma^2 = 25$ (а) та $\sigma^2 = 100$ (б), кодер AGU: $\sigma^2 = 25$ (в) та $\sigma^2 = 100$ (г)

Нейтральна ситуація – якщо локальний максимум $\text{PSNR}(Q)$ існує, але цей максимум приблизно дорівнює або менший ніж PSNR_{inp} . Тоді виникає питання, яка різниця Δ між цими двома величинами і який пріоритет вимог до стиснення з втратами. Припустимо, що $\text{PSNR}_{\text{inp}} - \text{PSNR}_{\text{max}} < \Delta$, і головна вимога до стиснення з втратами полягає в тому, щоб забезпечити якомога більший КС із прийнятним погіршенням якості зображення, де попереднє налаштування Δ лише контролює це можливе погіршення (наприклад, Δ може дорівнювати 1 або 2 дБ). У протилежному випадку, тобто якщо $\text{PSNR}_{\text{inp}} - \text{PSNR}_{\text{max}} \geq \Delta$ і не дуже потрібно надавати великий КС, необхідно вирішити, що робити. Одним із можливих способів є використання Q , що дорівнює 25, щоб забезпечити невидимість внесених спотворень. Це гарантує, що стиснене зображення буде схоже на вхідне зашумлене зображення.

Нарешті, ситуація під назвою «несприятлива» полягає в наступному. Залежність $PSNR(Q)$ постійно зменшується. Це означає, що стиснене зображення стає все більш і більш відмінним від оригіналу та безшумним, якщо Q збільшується. У цьому випадку також рекомендуємо застосувати Q , що дорівнює 25 або трохи менше, щоб уникнути видимих спотворень. До речі, в таких випадках шум може бути ефективно видалений за рахунок використання фільтрації після декомпресії зображень [284].

Наразі не можна дати уявлення про те, як розрізняти вищезазначені ситуації перед початком стиснення за допомогою BPG. Порівнюючи криві на рис. 5.15 для сприятливих ситуацій, тобто коли існує ОРТ, також було зроблено наступне спостереження. Якщо дисперсія шуму збільшується, ОРТ переміщується в область більшого Q . Для $\sigma^2=25$ ОРТ спостерігається для Q близько 29 (рис. 5.15, а), для $\sigma^2=64$ ОРТ має місце для Q близько 33 (рис. 5.15, б), а для $\sigma^2=100$ ОРТ спостерігається для Q близько 35. Це показує, що, швидше за все, параметри ОРТ можна визначити за умови, що дисперсія шуму апіорі відома або заздалегідь оцінена з відповідною точністю. Ще один приклад наведено на рис. 5.16.



Рисунок 5.16 – Стиснені тестові зображення з шумом $\sigma^2 = 64$: $Q = 28$ (а), $Q = 45$ (б), кодер BPG

На цьому рисунку наведено результати стиснення тестового зображення `hand_512.png`. Використовуються два значення Q : $Q = 28$, що відповідає OPT, і $Q = 45$, що забезпечує більший КС, але вносить більше спотворень. Як видно з рис. 5.16, стиснення в оптимальній робочій точці зберігає деталі зашумленого тестового зображення та усуває шум, тоді як стиснення з $Q=45$ значно спотворює зображення.

5.2 Удосконалення методів стиснення із втратами в умовах шумів різної природи та рівнів

Проаналізуємо більш детально властивості шуму, що спотворює зображення зубів, створені системою Morita. В цьому дослідженні використовували зображення зубів з високою роздільною здатністю, отримані з різними рівнями випромінювання. Розглянемо ті самі 20 фрагментів великого зображення, що було використано вище. Деякі приклади наведено в Додатку М, на рис. М.15.

Перша характерна особливість представлених зображень полягає в тому, що вони містять досить багато однорідних (пласких) фрагментів або ділянок з невеликим градієнтом (деякі з них позначені червоними прямокутниками), де можна побачити шум досить високої інтенсивності. Шум більш помітний у світлих областях і менш помітний у темних областях; це свідчить на користь його сигнально-залежної природи. Крім того, шум має специфічну зернисту структуру, що дозволяє припустити, що він певною мірою просторово корельований.

Друге спостереження полягає в тому, що видимість зернистої структури суттєво відрізняється для різних фрагментів. Це дає підстави підозрювати залежність ступеня просторової кореляції шуму від деяких параметрів отримання (і, можливо, трансформації) зображення.

Третій аспект, на який варто звернути увагу, це наявність кількох дуже світлих (близьких до білого) плоских ділянок у місцях закладення металевих або металоподібних елементів зуба. Це вказує на можливу наявність ефекту відсікання, який з'являється через перетворення зашумлених зображень у формат `bmp` (використовується за замовчуванням у системі Morita). Цей ефект може негативно

вплинути на подальші етапи обробки зображення; отже, це слід враховувати. Було застосовано деякі сліпі (автоматичні) методи оцінки характеристик шуму до аналізу зображень зубів, щоб побачити, чи підтверджуються припущення щодо властивостей шуму на цих зображеннях.

Аналіз просторової кореляції шуму наведено в Додатку В, де доведено, що в медичних системах реєстрації зображень шум є просторово корельованим. Дослідимо поведінку кодерів із втратами для медичних та оптичних зображень в умовах просторово корельованого шуму. Залучимо до аналізу ДКП-кодери AGU, AGU-M, ADCT, ADCT-M.

В Додатку М, на рис. М.16 представлені залежності $PSNR(\beta)$, де $KK = \beta\sigma$ для кодерів AGU і ADCT, та $KM = \beta\sigma$ для кодерів AGU-M і ADCT-M. Ці залежності були отримані для тестового зображення MRT_Prepared, спотвореного АБГШ з $\sigma=10$. Усі чотири залежності мають максимуми, що спостерігаються при $\beta \approx 4$ (KK близько 4σ) для кодерів AGU і ADCT і при $\beta \approx 2$ (тобто KM близько 2σ) для кодерів AGU-M і ADCT-M відповідно. Також зауважимо, що найкращий (максимальний) $PSNR_{tc}$ ОПТ забезпечують кодери ADCT-M і ADCT (близько 33,5 дБ). Тут варто нагадати, що ОПТ існує не завжди. Наприклад, відповідно до $PSNR_{tc}$, ОПТ не існує для жодної з розглянутих методик стиснення для стандартного оптичного зображення Baboon (яке є високотекстурним), спотвореного АБГШ з дисперсією 100 (тестові зображення доступні на <http://ponomarenko.info/agu.htm>). Для інших розглянутих показників (візуальної якості) ОПТ також не спостерігаються.

Розглянемо тепер поведінку $MSE_{cn}(\beta)$, розрахованого між зашумленим нестисненим та зашумленим стисненим. Усі чотири залежності представлені в Додатку М, на рис. М.17. Видно, що всі чотири залежності швидко зростають. Сукупний аналіз залежностей в Додатку М, на рис. М.16 і М.17 показує, що максимуми спостерігаються, коли відповідний $MSE_{cn}(\beta)$ приблизно дорівнює дисперсії адитивного шуму ($\sigma^2=100$ у розглянутому випадку). Нарешті, на рис. М.18 представлені залежності для всіх чотирьох кодерів. Кодер ADCT забезпечує БНП менше, ніж AGU для того самого β і, відповідно, KK . Подібним чином ADCT-M забезпечує БНП, менший, ніж AGU-M, для того самого β і, відповідно, KM . Одна

очевидна проблема в аналізі ефективності обробки зображень для випадку просторово корельованого шуму полягає в тому, що такий шум може мати численні набори своїх параметрів, які характеризують його кореляційну функцію або спектр. Тому, вивчаючи обробку зображень, спотворених просторово корельованим шумом, необхідно враховувати різні набори параметрів шуму.

Мета полягає в тому, щоб показати, що існує декілька відмінностей у залежностях для випадків просторово корельованого шуму та некорельованого шуму. Крім того, параметри просторової кореляції значною мірою впливають на основні тенденції. Вище було розглянуто чотири методи стиснення. Нижче, на додаток до них, також буде проаналізовано кодер SPIHT. Залежності були представлені як функції β . Для кодера SPIHT в якості ПКС використовується БНП.

Щоб надати можливість порівняти результати для SPIHT та інших кодерів, було зроблено наступне. Результати спочатку були отримані для кодера AGU, який забезпечує КШС, подібні до кодера SPIHT. Потім для кожного тестового зображення, кодера та СКВ шуму визначали КС для AGU, перераховували в БНП, і цей БНП використовували в SPIHT для стиснення. Після цього були визначені параметри стиснення кодера SPIHT. Існують різні способи моделювання 2D просторово корельованого шуму. Один із найпростіших надається Matlab (функція 'fspecial' з типом='gaussian', window=7), де потрібно встановити лише параметр χ функції Гауса, яка визначає вагові коефіцієнти 2D фільтра з кінцевою імпульсною характеристикою. Було використано два значення: $\chi=0,7$ (цей випадок умовно назвемо «середнім кореляційним шумом») і $\chi=1,0$ (цей випадок умовно назвемо «сильним кореляційним шумом»). Вертикальні та горизонтальні головні перерізи 2D автокореляційної функції змодельованого корельованого шуму є ідентичними (хоча на практиці вони можуть бути різними і часто відрізняються).

Почнемо з розгляду «середнього корельованого шуму» (СКШ). Оскільки вже є залежності для випадку АБГШ, представлені в Додатку М, на рис. М.16 для тестового зображення MRT_Prepared, наведемо залежності, отримані для того самого тестового зображення та тієї ж дисперсії шуму, але для випадку СКШ. Вони представлені в Додатку М, рис. М.19 для всіх п'яти розглянутих методів стиснення

з втратами. Перше спостереження полягає в тому, що ОПТ існує для всіх п'яти кодерів. Однак перша відмінність полягає в тому, що це відбувається для більших $KK_{\text{ОПТ}}$ і $KM_{\text{ОПТ}}$, ніж для випадку АБГШ (порівняйте з даними в Додатку М, рис. М.16). Оптимальний β для СКШ (в Додатку М, рис. М.19) становить близько 6 для кодерів AGU і ADCT, тоді як оптимальний β становить близько 4 для кодерів AGU-M і ADCT-M.

Друга відмінність полягає в тому, що максимально досяжний PSNR між істинним та зашумленим стисненим зображенням є значно меншим, ніж у випадку незалежного і однаково розподіленого (НОР) шуму. Навіть для найкращого методу стиснення (ADCT) максимальний PSNR досяг 30,6 дБ. Третя відмінність полягає в тому, що ОПТ для розглянутих кодерів спостерігаються для більших значень КС (близько 30...35), ніж для АБГШ.

Між тим, є й подібні спостереження. По-перше, спостерігається швидке зростання $MSE_{\text{сн}}(\beta)$ для всіх досліджуваних методів стиснення (див. залежності в Додатку М, на рис. М.20 для тестового зображення MRT_Prepared). По-друге, важливо, що $MSE_{\text{сн}}(\beta_{\text{ОПТ}})$ практично збігається з дисперсією шуму (можна проаналізувати дані в Додатку М, на рис. М.19 і М.20 разом для вибраного методу стиснення).

В Додатку М, на рис. М.21 представлені залежності БНП(β) для всіх п'яти кодерів. Зрозуміло, що БНП стає меншим (КС стає більшим), якщо β (і, відповідно, КК або КМ) зростає. Подібно до випадку білого шуму, кодер ADCT створює БНП трохи менше, ніж AGU для того самого β . Як і раніше, ADCT-M забезпечує менше БНП, ніж AGU-M для того самого β . Беручи до уваги відмінності в оптимальному β для кодерів AGU та ADCT і оптимальному β для кодерів AGU-M та ADCT-M, отримані значення КС для ОПТ близькі одне до одного. Тоді кодер ADCT забезпечує найкращий компроміс між КС і $PSNR_{\text{тс}}$ в районі ОПТ. SPIHT не є найкращим кодером для розглянутого випадку.

В Додатку М, на рис. М.22, а, показано тестове зображення, спотворене СКШ з дисперсією, рівною 100. Зображення, стиснене кодером ADCT з $KK=65$,

представлено на рис. М.22, б ($KC=42$). Видно, що шуми достатньо придушені, хоча невеликі спотворення інформації також спостерігаються.

Для зображень, спотворених СКІШ, ОПТ може існувати для інших метрик. В Додатку М, залежності на рис. М.23 і М.24 підтверджують існування ОПТ для метрик PSNR-HVS-M і MSSIM. ОПТ спостерігаються практично для тих же значень β , що і для $PSNR_{tc}$.

Щоб довести, що подібні залежності спостерігаються для інших тестових зображень, в Додатку М, на графіках на рис. М.25 і М.26 представлені результати зображення Barbara, спотвореного СКІШ з дисперсією 200.

Зверніть увагу, що максимуми спостерігаються для практично тих же значень β , що і для тестового медичного зображення (в Додатку М можна порівняти графіки на рис. М.25 і М.19). Подібним чином, ОПТ спостерігаються, коли відповідна $MSE_{cn}(\beta)$ досягає дисперсії шуму.

Розглянемо результати для іншого (більшого) ступеня кореляції шуму ($\chi=1,0$), ВКІШ. Деякі дані для тестового зображення MRT_Prepared, спотвореного ВКІШ, представлені на рис. 5.17, а також в Додатку М, на рис. М.27 та М.28.

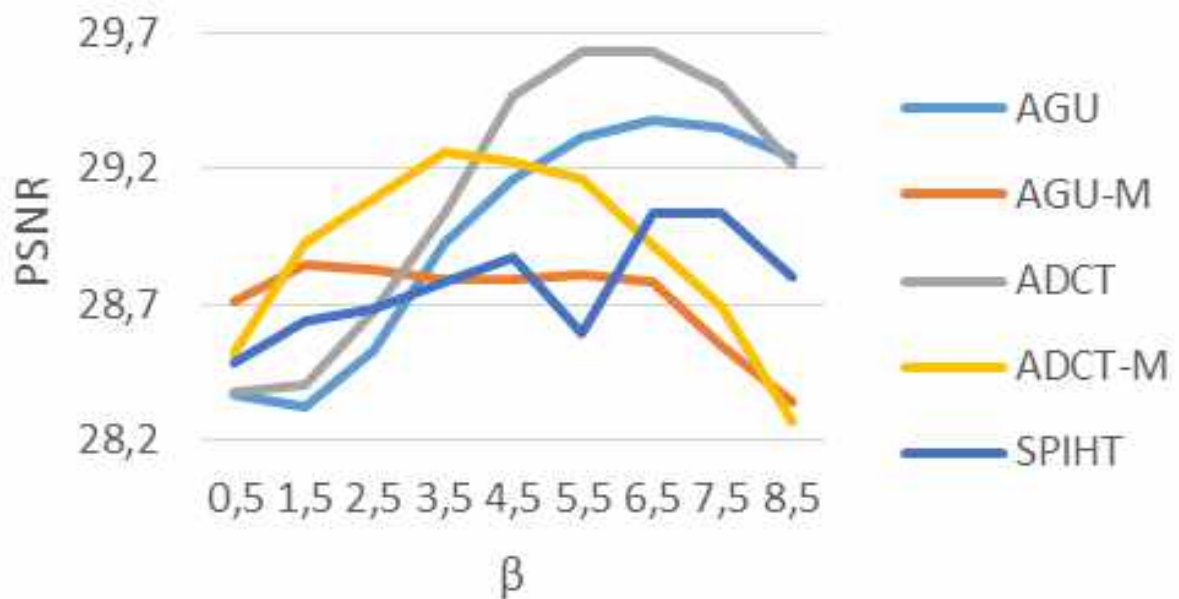


Рисунок 5.17 – Залежності $PSNR_{tc}(\beta)$ для п'яти методів стиснення тестового зображення MRT_Prepared, ВКІШ, $\sigma=10$

Для кодерів AGU, ADCT і SPIHT (див. графіки на рис. 5.17) OPT існують, хоча максимальні $PSNR_{tc}$ менші, ніж для СКШ (можна порівняти з відповідними графіками в Додатку М, на рис. М.16 і М.19). Для інших двох кодерів (рис. 5.17) існують широкі межі варіації β , де значення $PSNR_{tc}$ залишаються майже постійними.

Оптимальні значення β зсуваються до більших значень порівняно з випадком СКШ. Характер залежностей (Додаток М, рис. М.27) залишається незмінним (порівняно з графіками в Додатку М, на рис. М.17), тобто при зростанні β спостерігається швидке зростання MSE (збільшуються КК або КМ).

Спільний аналіз відповідних графіків на рис. 5.17 і М.27 знову показує, що OPT спостерігається для таких β , які дорівнюють дисперсії шуму (100 для цього випадку). Ця властивість може бути використана для досягнення OPT за допомогою ітераційної процедури, розглянутої у попередніх розділах.

Висновки до розділу 5

Цей розділ присвячено питанням підвищення ефективності стиснення із втратами зображень медичної діагностики в умовах присутності складних завад з урахуванням візуальної якості. Розглянуто особливості стиснення зображення сітківки ока з втратами, з метою одночасно забезпечити досить великий КС і внесення практично непомітних спотворень. Показано, що завдяки запропонованій процедурі прогнозування MSE можна забезпечити бажану MSE внесених втрат без ітераційного стиснення та декодування зображень. Ця процедура використовує відсоток нульових коефіцієнтів ДКП після квантування для заданого КК, щоб визначити необхідний КК і виконати стиснення з його використанням. Прогнозування є швидким, оскільки воно використовує 2D ДКП в обмеженій кількості блоків 8x8 пікселів.

Розглянуто особливості стиснення з втратами зображень зубів. Показано, що параметри стиснення (КС і метрики, що характеризують якість зображення) досить сильно залежать від властивостей зображення. Загальний спосіб внесення

спотворень, які є «невидимими», представлено для двох кодерів на основі ДКП – AGU та AGU-M. Надано рекомендації щодо налаштування КК або КМ. Показано, що досягнуті значення КС можуть бути досить великими: вони можуть досягати 30 і навіть більше, що значно перевищує значення КС, які зазвичай рекомендуються для JPEG (близько 10) для стиснення медичних зображень. При цьому значення КС індивідуальні для кожного зображення (або його фрагмента) – від 8,9 до 31,3 для розглянутої сукупності тестових зображень.

Розглянуто проблему ефективного безітераційного стиснення зображень зубів без візуально помітних спотворень. Було показано, що цю проблему можна вирішити шляхом застосування кодеру ADCT із запропонованим налаштуванням КК. Після встановлення КК додаткові ітерації для схеми стиснення не були потрібні, і непомітність внесених спотворень була практично гарантована. Це було підтверджено відповідно до об'єктивних показників візуальної якості, які сильно корелюють з MOS, підтвердженими візуальними прикладами та експериментами, проведеними групою стоматологів. Жодних внесених артефактів та/або втрат діагностично цінної інформації виявлено не було. Водночас були надані високі значення КС (від 7,5 до 20,6 залежно від контексту зображення).

Розглянуто стиснення BPG із втратами одноканальних зображень, спотворених АБГШ. Описано особливості такого стиснення. Показано, що можливі сприятлива, нейтральна і несприятлива ситуації стиснення зашумлених зображень. Аналіз отриманих даних свідчить про можливість прогнозування деяких параметрів, що характеризують ефективність стиснення з втратами. Це дозволяє дати рекомендації щодо вибору параметрів стиснення для кожного конкретного зображення.

Розглянуто шумові властивості знімків зубів, отриманих сучасною цифровою рентгеновською системою Morita. Було показано, що шум на таких зображеннях є просторово корельованим із середнім або високим рівнем кореляції. Було також показано, що шум у стоматологічних зображеннях залежить від сигналу, але модель залежності сигналу є досить складною, нелінійною та, іноді, немонотонною

через наявність ефекту відсікання та методів оптимізації (контрастування) зображення, застосованих внутрішньо у обладнанні.

Досліджено окремий випадок стиснення зображень з втратами, спотворених просторово корельованим шумом. Показано, що просторова кореляція шуму має певний вплив на ОРТ та його властивості. По-перше, ОРТ спостерігається для більшого КК або КМ, ніж для АБГШ шуму, що спотворює те саме зображення та має однакову дисперсію. По-друге, через це значення КС в ОРТ виявляються більшими, ніж для АБГШ шуму. Але ОРТ для просторово корельованого шуму є менш «очевидним», тобто забезпечується менш ефективна фільтрація шуму.

Основні наукові результати, наведені у п'ятому розділі, опубліковано в роботах автора [23], [24], [26], [39], [53], [56], [61], [62], зазначених в анотації.

РОЗДІЛ 6 ПРОГРАМНО-АЛГОРИТМІЧНА РЕАЛІЗАЦІЯ ТА ПІДВИЩЕННЯ ШВИДКОДІЇ ПРИ СТИСНЕННІ ТА ФІЛЬТРАЦІЇ ЗОБРАЖЕНЬ

В даному розділі розглянуто питання програмно-алгоритмічних реалізацій та прискорення при стисненні та фільтрації зображень. Зокрема, розглянуто питання автоматичного стиснення в задачі розпізнавання цілей, а також запропоновано алгоритмічну реалізацію деяких із розроблених методів фільтрації зображень, спотворених просторово корельованим шумом із невідомим спектром і стандартним відхиленням, яке змінюється для різних областей зображення, за допомогою комбінації згорткових нейронних мереж. Основні результати автора представлено у роботах [4], [8], [38], [41], [51], [54], [62], [64], [67].

6.1 Реалізація автоматичного стиснення із втратами при детектуванні цілей

Незважаючи на те, що стиснення із втратами може погіршити якість зображення, у доступній літературі є невелика кількість робіт, присвячених аналізу впливу стиснення на ефективність виявлення цілі людиною. Вплив стиснення на розпізнавання цілі людьми-спостерігачами було проаналізовано в [285] шляхом порівняння результатів для JPEG і зображень, стиснених за допомогою фрактальних методів. Показано, що втрата здатності виявляти цілі на зображеннях, стиснених за допомогою фрактальних методів, значно вища, ніж на зображеннях, стиснених JPEG. Метою статті [286] було проведення добре контрольованого експерименту, за допомогою якого проаналізовано вплив стиснення на роботу людини-спостерігача. Експерименти показали, що для заданого рівня стиснення та завдання ідентифікації стиснення JPEG2000 перевершує традиційне стиснення JPEG. Хасе та ін. [287] показали, що стандарти стиснення з втратами JPEG і H.264/AVC підходять для виявлення пішоходів в інфрачервоних відеопотоках, де H.264/AVC забезпечує прийнятні результати виявлення з вищим ступенем

стиснення, ніж JPEG. Вплив п'яти типів деградації зображення на ефективність класифікації було проаналізовано за допомогою чотирьох моделей глибокої нейронної мережі [288]. Показано, що моделі більш стійкі до впливу стиснення, ніж до впливу шуму та розмиття. У дослідженні [289] було показано, що зображення можна стискати в середньому в 7, 16 і 40 разів за допомогою кодерів JPEG, JPEG2000 і BPG відповідно, зберігаючи ефективність моделі класифікації згорткової нейронної мережі (ЗНМ).

Один цікавий висновок з [290] полягає в тому, що детектор «бачить» так само, як людське око, тобто ефективність детектора значно падає в таких же умовах, коли спостерігач починає бачити (детектувати) візуальну різницю між нестисненим і стисненим зображенням. Концепція ледь помітної різниці (ЛПР) стосується визначення мінімальної різниці (порогу видимості або візуальної помітності) між двома візуальними сигналами, яку помітить середній спостерігач [291]. Застосування цієї концепції дає змогу визначити максимальний ступінь стиснення, який можна застосувати до того, як отримане зображення буде виглядати погіршеним. Моделі прогнозування ЛПР можна застосовувати попіксельно [292], на рівні блоку (поблоково) [293] і на глобальному рівні (на рівні зображень) [294]. Крім того, прогнозування положення першої точки ЛПР стиснених зображень JPEG може бути досягнуто за допомогою прогнозування ФЯ, прогнозування представлення зображення в БНП і прогнозування якості з використанням деяких об'єктивних показників [295].

Надійне прогнозування положення першої точки ЛПР в [296] виконується для забезпечення простоти і високої швидкодії з використанням лише однієї ознаки, отриманої з нестисненого зображення (зображення, що планується тискати), яка є середнім значенням величини градієнта. На його основі визначається бажане значення PSNR зображення, стисненого JPEG (підхід на рівні зображень).

Наявність завад або спотворень на зображенні суттєво впливає на ефективність виявлення цілі в системах візуалізації із залученням людини [297]. Завади стосуються об'єктів або фонових елементів, схожих на ціль, які можуть

зменшити ймовірність виявлення цілі, збільшити ймовірність помилкової тривоги та збільшити час пошуку цілі на зображенні. Тому дуже важливо визначити залежність між рівнем завад на зображенні та виявлення цілі оператором. Метрики завад можна застосовувати на глобальному рівні (на рівні всього зображення і без попереднього знання цілі) і на локальному рівні (поблизу цілі, де необхідно знати додаткову інформацію про ціль). Ці показники є математично визначеними показниками, які вимагають апріорного знання цілі (зображення цілі), і на основі подібності (або відмінності) між ціллю та фоном виконується прогнозування. Цей підрозділ поєднує два підходи, один пов'язаний із прогнозуванням позиції першої точки ЛПР для стиснених зображень JPEG [296], а другий підхід пов'язаний із застосуванням об'єктивних заходів оцінки якості для прогнозування ймовірності виявлення цілі, частоти помилкових тривог і середній час пошуку цілі на зображенні оператором [298].

Загальнодоступна база даних зображень, відома як Search_2 [299], часто використовується для аналізу ефективності виявлення цілей. Однак вплив стиснення на розпізнавання цілі в роботах, у яких використовувалася ця база даних, не аналізувався. Проаналізуємо вплив стиснення JPEG на схожість між ціллю та фоном, як основу для оцінки ефективності виявлення цілі на зображенні оператором. Залежність об'єктивних показників якості від коефіцієнта якості стиснення JPEG проаналізовано в повному динамічному діапазоні ФЯ від 1 до 100. Додатково для чотирьох випадків значення подібності порівнюватимуться з результатами суб'єктивних тестів – без стиснення і три випадки зі стисненням (для двох фіксованих значень ФЯ (мінімального та максимального, що спостерігалися в експериментах) і для значення ФЯ, що відповідає положенню першої точки ЛПР). Перший напрямок цього дослідження полягає в тому, щоб вибрати підходи до оцінки якості зображення, які можна використовувати для надійного прогнозування ефективності виявлення цілі. Оскільки стиснення впливає на якість зображення, другим напрямком є визначення кінцевих меж, до яких можна виконувати стиснення JPEG, не впливаючи на результати пошуку цілі.

База Search_2 має 44 зображення високої роздільної здатності (6144x4096 пікселів), де спостерігачі мали завдання виявити одну з дев'яти можливих військових цілей (танки, БМП та всюдиходи) на цих зображеннях. Зображення отримані в реальних сільських умовах. Для кожного з вихідних зображень доступна інформація про положення цілі, її ширину та висоту, а також бінарні маски цілі, отримані шляхом ручної сегментації. Крім того, дослідникам доступна додаткова інформація про умови, в яких були отримані зображення – відстань, на якій розташована ціль, її ракурс, освітленість сцени, цілі та її оточення. На рис. 6.1 показано оригінальне зображення з бази даних Search_2 з регіоном розташування цілі, бінарним зображенням, отриманим шляхом ручної сегментації (шляхом відділення цілі від фону), а також зображеннями сегментованої цілі та її фону. З наявних 44 зображень дослідники в основному використовують 39 зображень, тобто не використовують зображення під номерами 7, 15, 23 і 26, на яких були виявлені повторювані цілі, і не використовують зображення 39, для якого ймовірність виявлення цілі низька (14,5%) [300].

В експериментах брали участь приблизно однакова кількість жінок і чоловіків, і всі вони мали нормальний або скоригований до нормального зір. Спостерігачі могли вільно обирати стратегію пошуку. Спостерігачі додатково вибирали регіон, у якому виявили ціль. Тривалість одного показу зображення була обмежена 60 с.

Зібраний цільовий час пошуку на зображенні представлений як середнє, середнє геометричне та медіана. Правильність відповідей піддослідних оброблялася через правильні, помилкові та пропущені виявлення. За результатами суб'єктивних тестів можна визначити ймовірність виявлення та частоту помилкової тривоги цілі для кожного вихідного зображення [301]:

$$P_d = \frac{N_{correct}}{N_{correct} + N_{false} + N_{missed}}, \quad (6.1)$$

де $N_{correct}$ — кількість правильних виявлень; N_{false} — кількість помилкових виявлень; N_{missed} — кількість пропущених цілей. Імовірність помилкової тривоги визначається як:

$$FAR = \frac{N_{false}}{N_{correct} + N_{false} + N_{missed}}, \quad (6.2)$$

при цьому загальна ймовірність виявлення становить:

$$P_{d\ total} = P_d + FAR. \quad (6.3)$$

Виявлення цілі залежить від її розмірів і контрасту [302]. Показано, що для великих цілей ефект впливу завад відсутній [303]. Якщо завади не суміжні з ціллю, їх вплив також незначний (навіть для малих цілей). Однак, якщо ціль знаходиться поруч із перешкодами (завадами), ефективність виявлення значно падає, якщо схожість між ціллю та перешкодами висока.

Оцінку якості зображення спочатку було виконано за допомогою показника TSSIM [300], який був створений шляхом адаптації відомого індексу структурної подібності SSIM [95]. Процедура застосування оцінки якості зображення при визначенні ступеня завад на зображенні, яка використовується для прогнозування ймовірності виявлення, частоти помилкових тривог і середнього часу пошуку цілі на зображенні, проілюстрована на рис. 6.2. На цьому зображенні ціль, яка цікавить, позначена червоним прямокутником (подвійна висота та ширина цілі, сегментованої вручну), а області, які використовуються для визначення подібності до зображення цілі, позначено білими прямокутниками.

Подібність між зображенням цілі визначається для кожного з блоків, що не перекриваються, і кінцеве значення отримується як середня (am) або середньоквадратична (rms) подібність, що є двома поширеними підходами до показників завад. Якщо схожість між ціллю і фоном велика, ймовірність виявлення цілі на зображенні менша і час її виявлення збільшується.

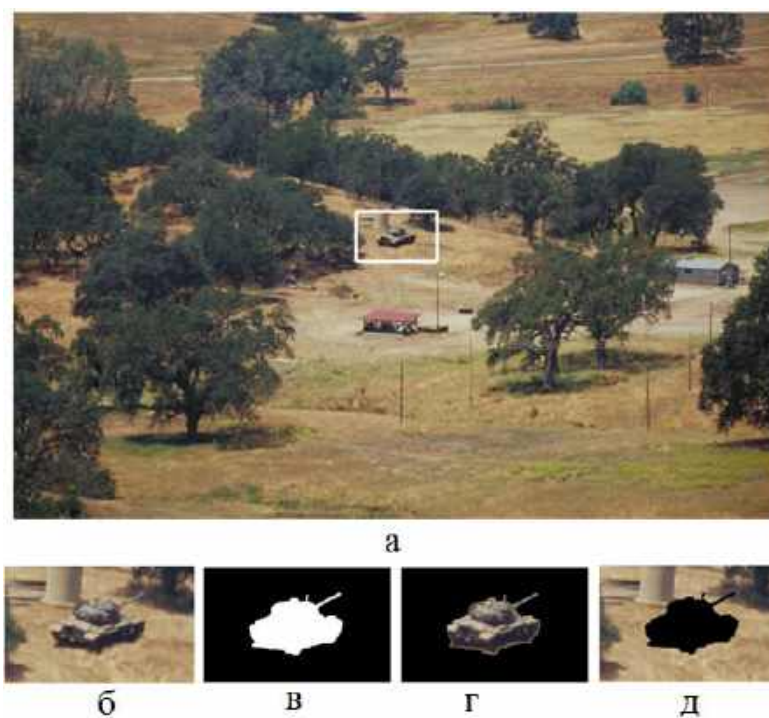


Рисунок 6.1 – Оригінальне зображення (а), зображення цілі (б), бінарне зображення, отримане ручною сегментацією (в), цільове та фонове зображення, отримані за допомогою бінарного зображення (г, д)

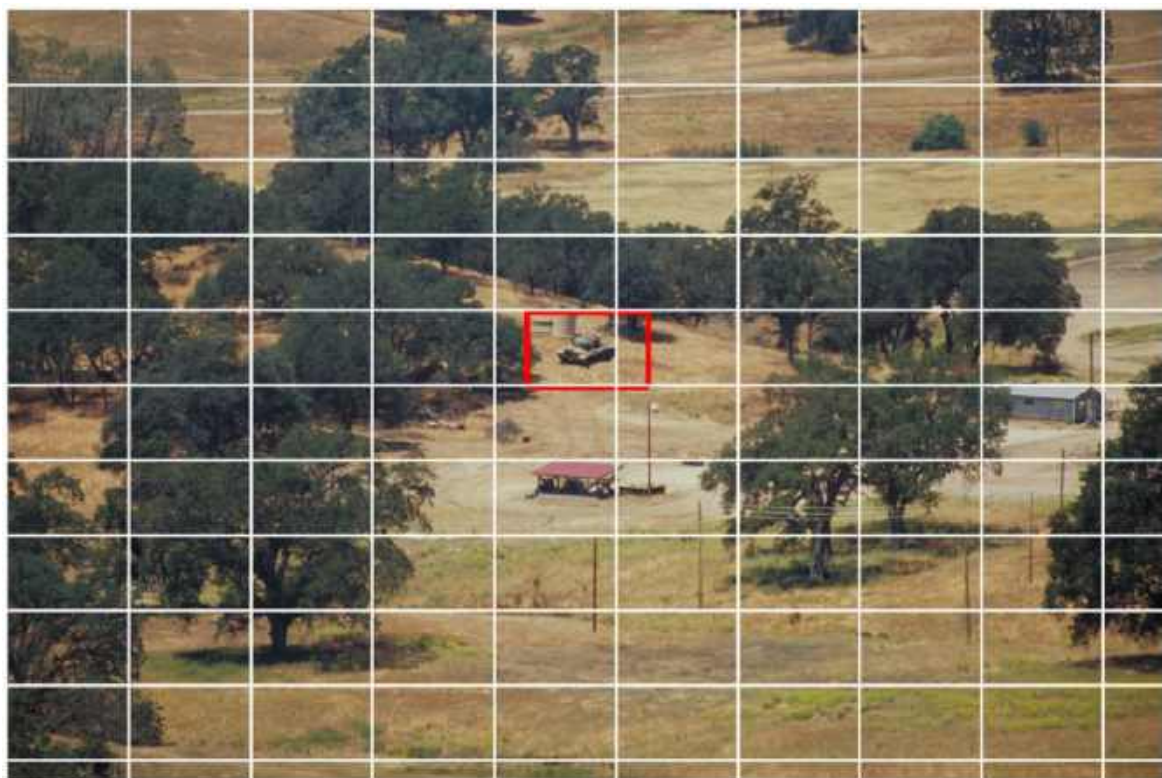


Рисунок 6.2 – Ілюстрація застосування метрик оцінки якості як метрики захищеності

На відміну від вимірювання подібності метрикою TSSIM, де ймовірність виявлення обернено пропорційна значенням TSSIM, у деяких метриках якості їх значення прямо пропорційні ймовірності виявлення. Нещодавній аналіз показав, що надійні прогнози ймовірності виявлення, частоти помилкових тривог і середнього часу пошуку цілі на нерухомих зображеннях можна отримати шляхом оцінки якості зображення [298].

Такі підходи до прогнозування досягли кращих результатів, ніж підходи, що використовують знаходження ознак і обчислення контрасту, визначені на локальному (навколо цілі) або глобальному рівні (рівень цілого зображення).

Зв'язок між показниками завад і ймовірністю виявлення (P_d), частотою помилкових тривог (FAR) і середнім часом пошуку цілі (mST) можна аналізувати за допомогою рекомендованих регресійних моделей [304]:

$$P_{d\ pred} = \frac{(C/C_{50})^E}{1+(C/C_{50})^E}, \quad (6.4)$$

$$FAR_{pred} = P_{d\ total} - \frac{(C/C_{50})^E}{1+(C/C_{50})^E}, \quad (6.5)$$

$$mST_{pred} = uC^v + w, \quad (6.6)$$

де $P_{d\ pred}$, FAR_{pred} , mST_{pred} – прогнозні значення P_d , FAR, mST на основі метрики захищеності C ; $P_{d\ total} = 0,988$ є загальною ймовірністю виявлення цілі; E , C_{50} , u , v , w – параметри регресійних моделей. Лінійний коефіцієнт кореляції Пірсона (LCC) та коефіцієнт рангової кореляції Спірмена (SROCC) використовувалися як кількісні міри ступеня узгодженості між прогнозами та істинними даними. Надійний підхід, описаний у [296] використовувався для визначення положення першої точки ЛПР зображень, стиснених JPEG. За допомогою цього підходу прогнозується значення PSNR першої точки ЛПР, після чого метрика якості визначається ітераційним процесом для досягнення прогнозованого значення. Для прогнозування використовується середня величина градієнта (MGM) оригінального нестиснутого зображення в градаціях сірого.

У цьому підході вихідне кольорове (RGB) зображення спочатку перетворюється на зображення у градаціях сірого:

$$F(n, m) = 0,299R(n, m) + 0,587G(n, m) + 0,114B(n, m), \quad (6.7)$$

після чого зображення зберігалися в 64-бітному форматі, з динамічним діапазоном від 0 до 1. Використовуючи фільтр Собеля, отримано градієнти (відгуки) по горизонтальній і вертикальній осях:

$$g_x(n, m) = F(n + 1, m + 1) + 2F(n + 1, m) + F(n + 1, m - 1) - F(n - 1, m + 1) - 2F(n - 1, m) - F(n - 1, m - 1), \quad (6.8)$$

$$g_y(n, m) = F(n - 1, m + 1) + 2F(n, m + 1) + F(n + 1, m + 1) - F(n - 1, m - 1) - 2F(n, m - 1) - F(n + 1, m - 1), \quad (6.9)$$

і на їх основі отримують MGM як:

$$MGM = \frac{1}{NM} \sum_{n,m} \frac{1}{g_{max}} \sqrt{g_x^2(n, m) + g_y^2(n, m)}, \quad (6.10)$$

де g_{max} позначає експериментально визначену максимальну величину градієнта, отриману шляхом застосування фільтра Собеля ($g_{max} = 4,472$).

Прогнозований PSNR можна представити за допомогою регресії другого порядку [296]:

$$PSNR = \begin{cases} 2115,5MGM^2 - 377MGM + 46,4, & MGM \leq 0,0896 \\ 29,58, & MGM > 0,0896 \end{cases}, \quad (6.11)$$

і, як бачимо, це значення не може бути нижче 29,58 дБ. Регресійна залежність отримана на основі результатів суб'єктивних тестів ЛПР. Нижчі значення MGM відповідають вищим значенням прогнозів PSNR, тобто більш високі значення

MGM відповідають нижчим значенням PSNR. Це означає, що на зображеннях із неоднорідним вмістом можна допускати вищий ступінь погіршення, ніж на зображеннях з однорідним вмістом.

Використовуючи підхід [296] для оригінального зображення, показаного на рис. 6.1, а, положення першої точки ЛПР спостерігається для $PSNR = 40,53$ дБ, і воно досягається з $\Phi Я = 35$. Візуальні відмінності між оригіналом зображення та його стисненою версією JPEG помітні лише тоді, коли вони значно збільшені. На рис. 6.3 показано фрагменти оригінального зображення (рис. 6.1, а), що стиснене і відповідає положенню першої точки ЛПР, де спостерігаються відмінності, викликані квантуванням коефіцієнтів ДКП під час стиснення JPEG, і з характерною блоковою структурою, яка є властива для стиснення JPEG. Блокова структура зображення негативно впливає на якість зображення та ефективність отримання цільових даних. На рис. 6.4 показано залежності PSNR від якості стиснення JPEG для всіх 44 зображень бази Search_2. Позиції перших точок ЛПР, отримані шляхом прогнозування значень PSNR за допомогою підходу, представленого вище, представлені на цьому рис. (вони позначені чорними символами x). На відміну від інших баз даних зображень, на яких проводилися тести ЛПР (MCL-JCI, JND-Рапо, JND VVC) [305], тут можна помітити відсутність суттєвого розкиду кривих PSNR, що, ймовірно, є наслідком схожого вмісту оригінальних зображень, які були отримані в сільському середовищі. Мінімальний і максимальний прогнози значень PSNR перших точок ЛПР становлять 39,54 дБ і 42,29 дБ, а відповідні значення $\Phi Я$ – 29 і 57.

Досліджено, як стиснення JPEG впливає на подібність (або відмінність) між об'єктом і фоном. Подібності розраховуються для $\Phi Я$ JPEG від 1 до 100 і для випадку без стиснення, використовуючи добре відомі метрик якості зображення. У таблиці 6.1 наведено ступінь узгодженості між середньоквадратичними показниками подібності, отриманими за допомогою метрик якості, і суб'єктивними результатами тесту (випадок без стиснення).



Рисунок 6.3 – Збільшені: а - фрагмент оригінального зображення з рис. 6.1а та б - фрагмент цього ж, але стисненого, зображення кодером JPEG, що відповідає положенню першої точки ЛПР (ФЯ = 35)

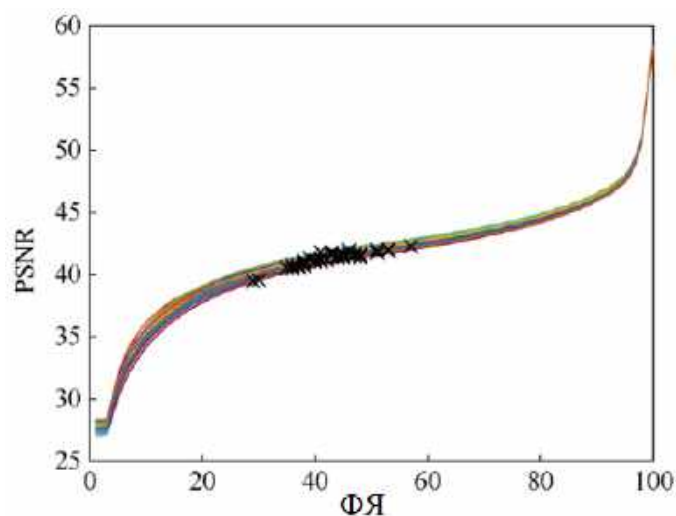


Рисунок 6.4 – КШС PSNR для зображень бази даних Search_2 і прогнозів JPEG PSNR (положення перших точок ЛПР позначені чорними символами x)

Метрика HaarPSI використовує вейвлет-розклад сигналів Хаара. Отримані коефіцієнти використовуються для визначення локальної подібності між порівнюваними сигналами, а також відносної важливості областей зображення. Кінцеві оцінки якості знаходяться в діапазоні $[0, 1]$, де вищі значення відповідають більшій схожості сигналів, що порівнюються. Незважаючи на те, що було обрано добре відомі метрики якості, які мають високий ступінь узгодження з результатами суб'єктивних тестів на стандартних базах даних зображень із суб'єктною оцінкою, з таблиці 6.1 можна зробити висновок, що їх ефективність у базі даних зображень Search_2 гірша. Припускаємо, що ефективність визначення цілі за допомогою

вимірювань метрик якості є низькою через змінні розміри цілі, малі розміри цілі, поганий контраст і текстуру цілі. Це додатково вказує на потребу в додатковій адаптації метрик якості для застосування в ефективному виявленні цілей.

Таблиця 6.1 – Ступінь узгодженості (LCC і SROCC) між метриками якості і суб'єктивними результатами тестування (випадок без компресії)

	P _d		FAR		mST		Середнє
	LCC	SROCC	LCC	SROCC	LCC	SROCC	
PSNR-HVS _{rms}	0,6169	0,6803	0,6984	0,7232	0,5979	0,6386	0,6592
PSNR-HVS-M _{rms}	0,6163	0,6842	0,6977	0,7271	0,5970	0,6412	0,6606
FSIM _{rms}	0,6773	0,5433	0,7160	0,5578	0,6537	0,5484	0,6161
FSIMC _{rms}	0,6348	0,5395	0,6927	0,5606	0,6285	0,5685	0,6041
HaarPSI _{rms}	0,8738	0,7441	0,8623	0,7596	0,8792	0,7631	0,8137

Виходячи із залежності метрик подібності від стиснення JPEG, можна зробити висновок, що подібність залежить від ФЯ, але значення подібності дещо змінюються, починаючи від випадку без стиснення, тобто стиснення з коефіцієнтом якості 100, до першої точки ЛПР. Оскільки перша точка ЛПР представляє межу між стисненням без візуальних втрат і стисненням із візуальними втратами, слід очікувати, що зі стисненими зображеннями JPEG у першій точці ЛПР ймовірність виявлення та час пошуку цілі на зображенні залишаються на рівні відповідних значень, отриманих у суб'єктивних тестах, проведених на зображеннях без стиснення.

Розмір кожного оригінального нестиснутого кольорового зображення становить 73 МБ, тоді як розмір зображень, отриманих стисненням на основі положення першої точки ЛПР, становить приблизно 1 МБ. Таким чином досягається величезна економія пам'яті та комунікаційних ресурсів, але без впливу на кінцеву ефективність виявлення. Крім того, з усіма об'єктивними метриками подібності можна спостерігати, що, починаючи з зображення без стиснення та зі зменшенням ФЯ до перших точок ЛПР, динамічний діапазон подібності

залишається приблизно однаковим. При подальшому зниженні якості динамічний діапазон істотно змінюється.

За допомогою вимірювання NaarPSI зафіксовано існування екстремумів (мінімуми) подібності для низьких значень ФЯ (близько 5). Через природу стиснення JPEG, де для низьких ФЯ отримують майже гладкі області (фон і ціль), для цієї метрики подібність між ціллю і фоном з мінімальної позиції починає збільшуватися зі зменшенням ФЯ. Звичайно, такий низький ФЯ не буде використовуватися в реальних ситуаціях. Результати суб'єктивних тестів порівнювали з об'єктивними оцінками, розрахованими для чотирьох випадків: (1) без стиснення (БС), (2) з використанням стиснення JPEG з фіксованим ФЯ 57 (максимальне значення, отримане за допомогою прогнозування, рис. 6.4), (3) використання стиснення JPEG із фіксованим ФЯ 29 (мінімальне значення, отримане за допомогою прогнозування) та (4) застосування стиснення JPEG із коефіцієнтом якості, що відповідає першій точці ЛПР, спрогнозованою за допомогою [296]. Ступінь узгодженості між об'єктивними балами та результатами суб'єктивних тестів було визначено за допомогою кількісних показників LCC та SROCC. Для кожного об'єктивного показника є дві частини таблиці: одна, якщо кінцева оцінка отримана як середнє значення (am), а інша, де кінцеву оцінку отримано як корінь із середньоквадратичних значень (rms) оцінок подібності.

Аналізуючи результати з таблиці 6.2, можна зробити висновок, що ступінь узгодженості між показниками подібності та результатами суб'єктивних тестів є найнижчим для стиснутих зображень JPEG для ФЯ = 29. Базуючись на цих результатах, краще розрахувати об'єктивні значення на стиснених зображеннях JPEG для коефіцієнта якості, що відповідає першим точкам ЛПР. У будь-якому випадку, визначення подібності для перших стиснених зображень ЛПР JPEG не суттєво погіршує ефективність прогнозування виявлення цілі на зображенні. Крім того, для метрик завад, наведених у таблиці 6.2, можна зробити висновок, що немає суттєвої різниці в ефективності, якщо кінцева оцінка визначається як середнє значення або середньоквадратичне значення локальних оцінок.

Найвищого ступеня збігу було досягнуто між результатами суб'єктивних тестів і результатами об'єктивної метрики HaarPSI. Крім того, можна зробити висновок, що ефективність, досягнута шляхом застосування HaarPSI, без будь-яких адаптацій до оцінки завад, відповідає рівню ефективності найсучасніших заходів, спеціально розроблених для аналізу завад. Більш детально результати цього дослідження наведено у роботі [67].

Таблиця 6.2 – Ступінь узгодженості між суб'єктивними результатами тесту та метрикою HaarPSI для бази даних Search_2

		Нестиснене	ФЯ=57	ФЯ ЛПП#1	ФЯ=29
HaarPSI _{am}					
P _d	LCC	0,8751	0,8795	0,8752	0,8601
	SROCC	0,7460	0,7460	0,7416	0,7235
FAR	LCC	0,8626	0,8615	0,8637	0,8423
	SROCC	0,7619	0,7620	0,7602	0,7401
mST	LCC	0,8806	0,8852	0,8831	0,8671
	SROCC	0,7640	0,7644	0,7702	0,7564
HaarPSI _{rms}					
P _d	LCC	0,8738	0,8781	0,8738	0,8584
	SROCC	0,7441	0,7441	0,7438	0,7233
FAR	LCC	0,8623	0,8611	0,8634	0,8419
	SROCC	0,7596	0,7597	0,7620	0,7401
mST	LCC	0,8792	0,8838	0,8815	0,8652
	SROCC	0,7631	0,7627	0,7697	0,7562

Рисунок 6.5, як наочна ілюстрація, показує зв'язок між HaarPSI_{rms} (для стиснених зображень JPEG у перших точках ЛПП) і суб'єктивними результатами тесту. На цих діаграмах розсіювання кожна точка представляє одне оригінальне тестове зображення (серед 39 зображень Search_2), а горизонтальна та вертикальна осі представляють значення метрики HaarPSI_{rms} та суб'єктивні дані (P_d, FAR та mST), відповідно.

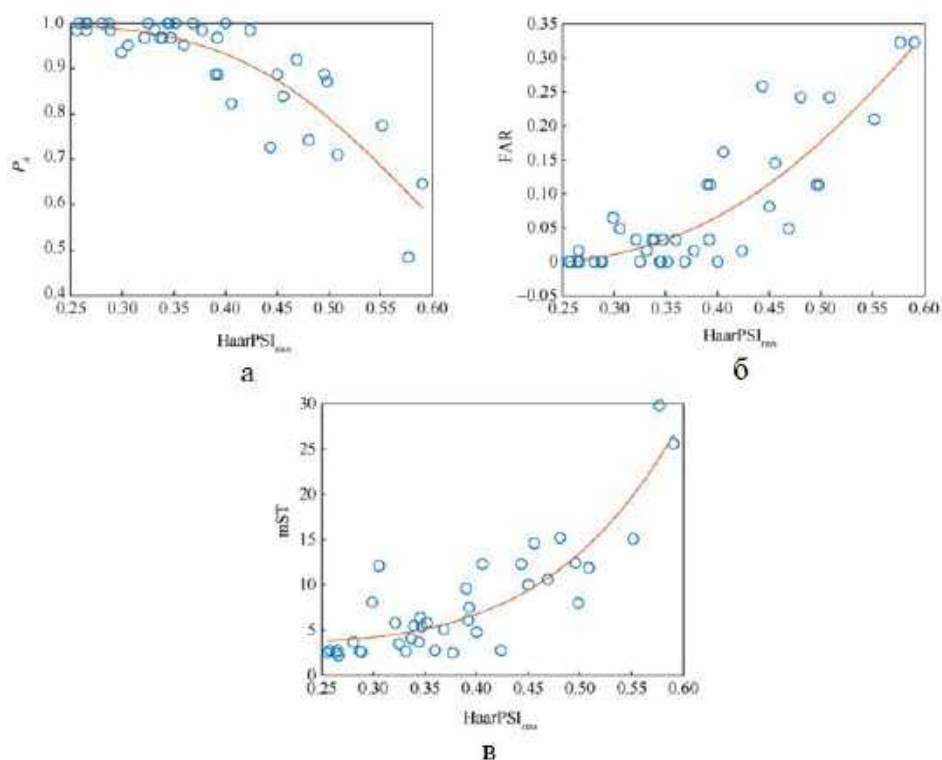


Рисунок 6.5 – Залежність метрики HaarPSI_{rms} від P_d , FAR та mST

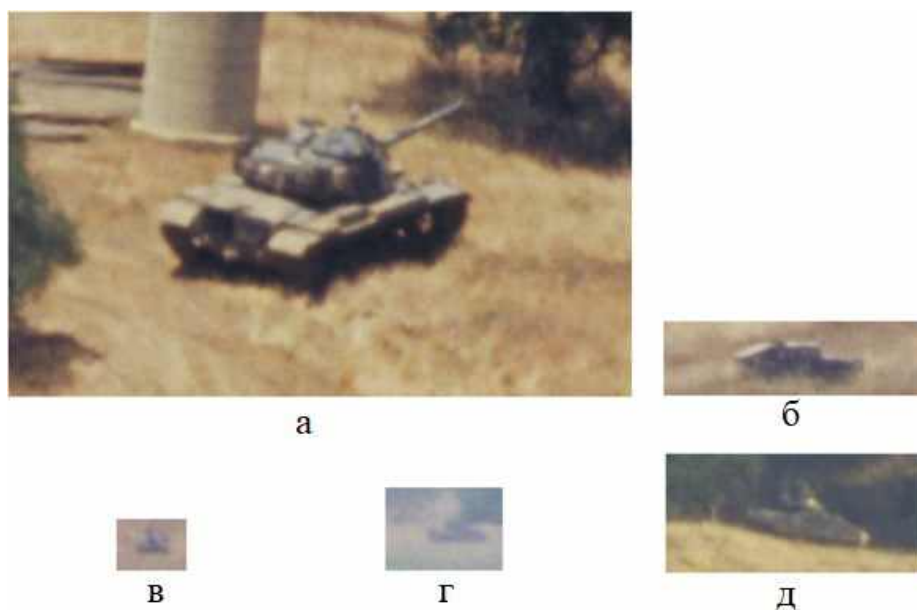


Рисунок 6.6 – Деякі зображення цілей у Search_2, які розглянуто: ціль # 5, 322x199 пікселів, $P_d = 1$, mST = 2,8 с, HaarPSI_{rms} = 0,258 (а); ціль #17, 139x39 пікселів, $P_d = 1$, mST = 4,8 с, HaarPSI_{rms} = 0,4 (б); ціль #4, 38x28 пікселів, $P_d = 0,484$, FAR = 0,323, mST = 29,8 с, HaarPSI_{rms} = 0,577 (в); ціль #22, 59x42 пікселів, $P_d = 0,645$, FAR = 0,323, mST = 25,6 с, HaarPSI_{rms} = 0,591 (г); ціль #38, 139x64 пікселів, $P_d = 0,95$, FAR = 0,048, mST = 12,1 с, HaarPSI_{rms} = 0,305 (д)

Оптимальні криві регресії (6.4)-(6.6) також показані на цих діаграмах розсіювання. З рис. 6.5 можна зробити висновок, що якщо подібність між цільовим зображенням і фоном вища (вищі значення $\text{NaarPSI}_{\text{rms}}$), ймовірність виявлення зменшується, частота помилкових тривог зростає, а середній час пошуку також збільшується. Існує 11 зображень, де ймовірність виявлення дорівнює одиниці, і для яких значення $\text{NaarPSI}_{\text{rms}}$ становлять від 0,258 до 0,4, рис. 6.6, а. У цьому випадку мінімальне значення $\text{NaarPSI}_{\text{rms}}$ було отримано для цілі, що є високотекстурною та з великими розмірами (рис. 6.6, а). Тоді як максимальне значення $\text{NaarPSI}_{\text{rms}}$ було отримано для цілі, що добре контрастує з навколишнім середовищем і є малотекстурною (рис. 6.6, б). Крім того, з рис. 6.5, а, можна зробити висновок, що два найвищі значення $\text{NaarPSI}_{\text{rms}}$ (0,577 і 0,591) були отримані для випадків, коли ймовірності виявлення цілі є найнижчими ($P_d = 0,484$ і $P_d = 0,645$), і вони відповідають максимальній частоті помилкової тривоги, рис. 6.5, б ($\text{FAR} = 0,323$), і максимального часу пошуку цілі 29,8 с і 25,6 с, рис. 6.5, в. Ці два випадки мали місце з цілями малих розмірів, зі слабким контрастом до фону та з складною текстурою, рис. 6.6, в-г. З діаграм розсіювання на рис. 6.5, а, б, видно, що розкид точок навколо кривих регресії збільшується зі збільшенням метрик якості. З рис. 6.5, в, видно, що одна точка відхиляється від основного тренду, для якого $mST = 12,1$ і $\text{NaarPSI}_{\text{rms}} = 0,305$. Це відповідає цілі, яка закамуйфльована фоном, рис. 6.6д, і це призвело до нижчого значення подібності NaarPSI з фоном, тоді як ймовірність виявлення висока $P_d = 0,95$.

Репозиторій https://github.com/krivenkoz/grayscale_compression містить проект, розроблений на кафедрі інформаційних та комунікаційних технологій Національного аерокосмічного університету «Харківський авіаційний інститут». Цей проект присвячений дослідженню та реалізації методів автоматизації стиснення зображень у градаціях сірого, зокрема, з використанням ДКП та прогнозування спотворень у алгоритмах стиснення із втратами (рис. 6.7).

Проект реалізовано у вигляді Jupyter Notebook під назвою `k504_grayscale_compression.ipynb`, який призначений для виконання в середовищі Google Colab. Для початку роботи користувачеві необхідно створити папку XAI у

кореневому каталозі свого Google Drive та скопіювати туди вміст репозиторію. Після цього можна запускати ноутбук у Colab для проведення експериментів зі стиснення зображень (рис. 6.8).

Основна мета проекту — розробка швидкого та точного методу прогнозування спотворень, що виникають під час стиснення зображень із втратами. Це дозволяє оцінити якість стисненого зображення без необхідності його декомпресії, що значно економить обчислювальні ресурси та час.

Репозиторій містить декілька каталогів та файлів, зокрема:

- Dental: містить підкаталоги з медичними зображеннями для тестування алгоритмів стиснення;
- Images: містить тестові зображення у градаціях сірого для проведення експериментів;
- Results: каталог для збереження результатів роботи алгоритмів стиснення та оцінки якості;
- Supplementary materials A Fast and Accurate Prediction of Distortions in DCT-based Lossy Image Compression: додаткові матеріали до наукової статті, що описує розроблений метод передбачення спотворень;
- k504_grayscale_compression.ipynb: основний Jupyter Notebook з реалізацією алгоритмів та експериментів;
- readme.md: файл з описом проекту та інструкціями щодо використання.

Розроблений метод базується на аналізі коефіцієнтів ДКП та їх змін після квантування під час стиснення. Основні етапи методу включають:

- 1) розбиття зображення на блоки: вихідне зображення розбивається на невеликі блоки (зазвичай 8x8 пікселів), до кожного з яких застосовується ДКП;
- 2) обчислення коефіцієнтів ДКП: для кожного блоку обчислюються коефіцієнти ДКП, які характеризують частотні компоненти зображення;
- 3) квантування коефіцієнтів: отримані коефіцієнти піддаються квантуванню з використанням матриці квантування, що призводить до втрат інформації та спотворень;

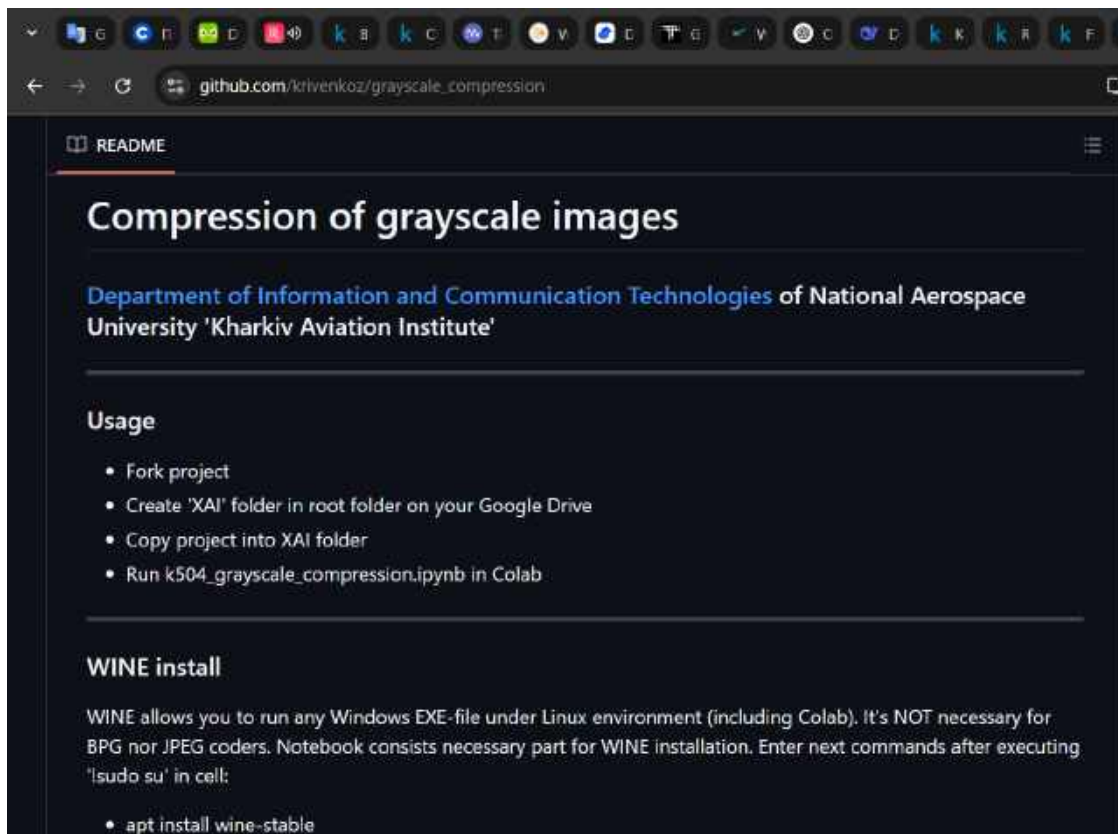


Рисунок 6.7 – Загальний вигляд першої сторінки проекту



Рисунок 6.8 – Приклад роботи проекту

1) аналіз спотворень: порівняння оригінальних та квантованих коефіцієнтів дозволяє оцінити рівень спотворень у кожному блоці;

2) прогнозування спотворень: на основі аналізу квантованих коефіцієнтів розроблено модель, яка прогнозує загальний рівень спотворень у всьому зображенні без необхідності його декомпресії.

Такий підхід дозволяє швидко та точно оцінювати якість стиснених зображень, що є важливим у реальних застосуваннях, де необхідно швидко приймати рішення щодо прийнятного рівня.

6.2 Покращення швидкодії фільтрації зображень

Існують практичні ситуації, коли ефективність обчислень становить головний інтерес, тоді як інші характеристики ефективності фільтрації також важливі. Операції, які використовуються для придушення шумів, мають бути стандартними, простими, швидкими та/або добре узгоджуватися з іншими операціями, що використовуються на початкових етапах обробки зображення, як оцінка параметрів шуму та/або стиснення. Це значно зменшує кількість класів застосовних методів придушення шумів, що робить необхідним пошук компромісних рішень для фільтрів на базі ковзних вікон [77], [121] і методів усунення шумів на основі простих і ефективних ортогональних перетворень [129], [130]. Придушення шумів зображення на основі ДКП є хорошим кандидатом.

Хоча ДКП у блоках 8×8 можна вважати стандартною і швидкою операцією, ДКП у блоках, що повністю перекриваються, іноді може бути недостатньо швидким. Є варіанти прискорення придушення шумів завдяки використанню обробки даних у блоках, які частково перекриваються. Це призводить до можливого значного прискорення обробки, але погіршується ефективність придушення шуму [130]. Таким чином, потрібно знайти належний компроміс між втратою ефективності фільтрації за кількісними показниками та перевагами в обчислювальній ефективності.

Питання втрати ефективності для фільтрів на основі ДКП з частковим перекриттям блоків порівняно з випадком повного перекриття було коротко розглянуто в [130]. Було показано, що ця втрата може досягати 3...4 дБ з точки зору пікового відношення сигнал/шум (PSNR) для граничного випадку блоків, що не перекриваються (для яких можуть також виникати блокові артефакти, подібні до тих, що характерні для стиснення JPEG). Між тим, випадок часткового перекриття блоків ретельно не вивчений. Таким чином, зосередимось на аналізі якості зображення на виході, яка може бути досягнута у випадку придушення шуму на основі ДКП із використанням часткового перекриття блоків. Зокрема, досліджуються можливості підвищення ефективності фільтрації за рахунок застосування так званої комбінованої порогової обробки та локальної адаптації. Крім того, аналіз виконується з використанням як стандартних метрик якості (PSNR), так і метрик візуальної якості. Хоча в основному вивчається обробка зображень у відтінках сірого, спотворених АБГШ, також наводиться приклад того, що запропоновані методи можуть бути застосовані, якщо на зображенні присутній сигнально-залежний шум.

Нижче припускається, що дисперсія АБГШ відома заздалегідь або попередньо оцінена з відповідною точністю. Тоді можливі декілька способів встановлення порогу: м'який, жорсткий, а також адаптивний або комбінований. М'яке порогове визначення є найменш ефективним для усунення шумів на основі ДКП [129]. Тому наведемо вирази для жорсткого та комбінованого порогів для випадку АБГШ. Для жорсткого порогу:

$$D_{thr}(n, m, k, l) = \begin{cases} 0, & \text{якщо } |D(n, m, k, l)| < \beta(n, m)\sigma \\ D(n, m, k, l), & \text{якщо } |D(n, m, k, l)| \geq \beta(n, m)\sigma \end{cases}, \quad (6.12)$$

де $\beta(n, m)$ – параметр, який, в загальному випадку, може бути різним для різних блоків.

Комбінований поріг визначається як:

$$D_{thr}(n, m, k, l) = \begin{cases} D^3(n, m, k, l) / \beta^2 \sigma^2(n, m, k, l), & \text{якщо } |D(n, m, k, l)| < \beta(n, m) \sigma \\ D(n, m, k, l), & \text{якщо } |D(n, m, k, l)| \geq \beta(n, m) \sigma \end{cases}, \quad (6.13)$$

де амплітуди менших ДКП-коефіцієнтів зменшуються більше, ніж амплітуди відносно великих ДКП-коефіцієнтів. Можна використовувати фіксований $\beta(n, m)$, тоді рекомендовані значення дорівнюють 2,6 для (6.12) і 4,0 для (6.13). Однак існує можливість налаштувати $\beta(n, m)$ локально адаптивним способом. Як буде показано нижче, це може надати деякі переваги. Питання полягає в тому, як це зробити без значної втрати ефективності обчислень.

Дослідження було проведено для п'яти стандартних тестових зображень у градаціях сірого: Lena, Goldhill, Baboon, Barbara та Peppers. Розглянуто три значення дисперсії шуму: 25, 49 і 100. Перше значення відповідає випадку, коли АБГШ ледь візуально помітний на зображеннях. Проаналізовані метрики: вихідна MSE позначається як MSE_{out} , а вихідні HVS-метрики PSNR-HVS-M та MSSIM.

Наведемо деякі дані для придушення шумів на основі ДКП з повним перекриттям блоків. Дані для дисперсій 49 та 100 представлені в таблиці 6.3 для жорсткого та комбінованого порогів ($\beta(n, m)$ є фіксованим і дорівнює 2,6 та 4,0 відповідно). Як видно, придушення шуму з жорстким і комбінованим порогом забезпечують практично однакову вихідну MSE. Зменшення MSE стає більшим для більшої дисперсії шуму та є великим для зображень простої структури, таких як Lenna та Peppers. Найменше зниження спостерігається для найскладнішого за структурою (текстурного) тестового зображення Baboon.

Тепер розглянемо візуальну якість відфільтрованих зображень. Отримані результати представлені в таблиці 6.4. Їх аналіз демонструє, що існує дуже невелика різниця у візуальній якості вихідних зображень для фільтрів на основі ДКП, які використовують жорстке та комбіноване порогове значення (різницю у візуальній якості двох зображень можна помітити, якщо відповідні значення PSNR-HVS-M відрізняються на 0,5 дБ або більше та/або якщо залишкові спотворення мають різні типи, наприклад, є певні артефакти для одного вихідного зображення,

а ці артефакти відсутні в іншому зображенні). Якщо дисперсія шуму дорівнює 49, можна помітити залишкові спотворення, але вони не дратують.

Таблиця 6.3 – MSE_{out} після фільтрації, повне перекриття

Зображення	σ^2	Жорсткий поріг	Комбінований поріг
Lenna	49	12,99	13,32
	100	18,96	20,39
Baboon	49	34,29	34,14
	100	59,37	58,81
Barbara	49	14,96	15,38
	100	24,01	25,77
Peppers	49	16,27	15,99
	100	22,32	22,90
Goldhill	49	20,22	19,95
	100	30,56	30,74

Для більшої дисперсії шуму залишковий шум і спотворення вихідних зображень чітко видно. Зверніть увагу, що для заданої дисперсії шуму найменші PSNR-HVS-M і MSSIM спостерігаються для тестових зображень найскладнішої структури Baboon і Goldhill. Це тому, що, як і будь-які інші фільтри, усунення шумів на основі DCT вносить певні спотворення разом із позитивним ефектом придушення шуму.

Проаналізуємо ефективність придушення шумів на основі ДКП для блоків, що не перекриваються. Результати для значення дисперсії 100 представлені в таблиці 6.5, значення $\beta(n, m)$ фіксовані і дорівнюють 2,6 і 4,0 відповідно. Перше спостереження полягає в тому, що значення MSE_{out} для жорсткого порогового значення приблизно вдвічі більші, ніж у випадку повного перекриття (див. дані в таблиці 6.3). Крім того, MSE_{out} для тестового зображення Baboon більше, ніж σ^2 , тобто зменшення шуму вносить більше спотворень, ніж придушує шум.

Для цього є дві основні причини. Якщо блоки не перекриваються (усереднення не використовується), діє тільки один механізм придушення шумів.

Крім того, фіксований $\beta(n, m) = 2,6$ не є оптимальним (і навіть квазіоптимальним) для різних блоків. Зокрема, існує досить багато блоків, для яких оптимальне (що дає мінімальну MSE_{out} для даного блоку) β значно відрізняється від 2,6 і може бути навіть близьким до 0 (тобто фільтрація для таких блоків взагалі не потрібна). На рис. 6.9 показано гістограму оптимальних локальних значень β для блоків тестового зображення Baboon, де є очевидний максимум для дуже малих значень β . Таким чином, можна очікувати, що деяка локальна адаптація (локальне налаштування $\beta(n, m)$ в залежності від деяких статистичних даних у кожному блоці) може призвести до покращення ефективності.

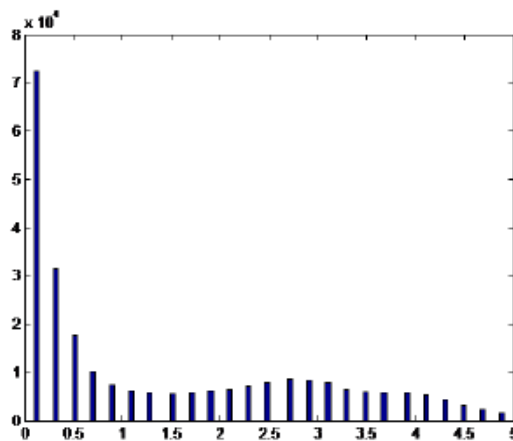
Таблиця 6.4 – PSNR-HVS-M та MSSIM після фільтрації, повне перекриття

Зображення	σ^2	Жорсткий поріг		Комбінований поріг	
		PSNR-HVS-M, дБ	MSSIM	PSNR-HVS-M, дБ	MSSIM
Lenna	49	38,49	0,9869	38,49	0,9869
	100	35,71	0,9811	35,54	0,9804
Baboon	49	37,85	0,9881	37,74	0,9881
	100	34,10	0,9797	34,02	0,9800
Barbara	49	38,58	0,9897	38,59	0,9896
	100	35,56	0,9842	35,53	0,9835
Peppers	49	38,22	0,9848	38,31	0,9852
	100	35,76	0,9795	35,74	0,9793
Goldhill	49	36,96	0,9842	37,17	0,9850
	100	33,86	0,9747	33,98	0,9757

Тим часом аналіз даних у таблиці 6.5 для комбінованого порогового визначення показує, що його ефективність значно краща, ніж для жорсткого порогового визначення. Покращення порівняно з жорстким пороговим значенням може бути до 1,5 разів (приблизно 1,6 дБ з точки зору вихідного PSNR) без будь-якої локальної адаптації. Звичайно, отримані результати для комбінованого порогового значення є значно гіршими, ніж для випадку використання блоків, що повністю перекриваються, за всіма проаналізованими показниками.

Таблиця 6.5 – PSNR-HVS-M та MSSIM після фільтрації, $\sigma^2=100$

Зображення	Жорсткий поріг	Комбінований поріг		
	MSE	MSE	PSNR-HVS-M, dB	MSSIM
Lenna	45,89	31,83	32,44	0,9711
Baboon	109,0	83,53	31,14	0,9691
Barbara	58,12	44,15	32,05	0,9738
Peppers	51,32	34,10	32,50	0,9704
Goldhill	64,51	46,51	31,21	0,9648

Рисунок 6.9 – Гістограма оптимальних значень β в блоках для фільтрації без перекриття блоків для зображення Baboon

Однак значення MSE_{out} переважно кращі, ніж для інших методів швидкого придушення шумів, наприклад, сигма-фільтр з ковзним вікном 7×7 [75] (41,41 для зображення Lenna, 50,36 для Goldhill, 73,93 для Baboon, 53,33 для Barbara та 39,51 для Peppers). Як було зазначено вище, одним із недоліків придушення шумів на основі ДКП є те, що у вихідних зображеннях спостерігаються блочні артефакти. Існування цього ефекту значно обмежує практичне застосування фільтрації в блоках, що не перекриваються, хоча цей підхід є дуже швидким (приблизно в 64 рази швидше, ніж обробка зображень у блоках, що повністю перекриваються). Крім того, чітко простежується різниця в якості зображення (близько 2...3 дБ в перерахунку на метрику PSNR-HVS-M), яку варто зменшити.

На основі попереднього аналізу одразу виникають такі ідеї:

- треба хоча б певною мірою «ввімкнути» механізм усереднення;
- варто спробувати застосувати локальну адаптацію в блоках.

Механізм усереднення може дати певні переваги, якщо використовується часткове перекриття блоків. Ступінь перекриття може бути різним, і в рамках даної роботи важко розглянути всі можливі варіанти. Тому зосередимося на випадку $r=4$ (нагадаємо, що r – це кількість пікселів перекриття в блоці). У таблиці 6.10 представлені характеристики ефективності усунення шумів на основі ДКП з неадаптивним жорстким і комбінованим порогом (значення $\beta(n, m)$ фіксовані і дорівнюють 2,6 і 4,0 відповідно). Перше спостереження для жорсткого порогового значення полягає в тому, що MSE_{out} для нього приблизно на 20% більше (вихідний PSNR приблизно на 0,7 дБ менший), ніж у випадку повного перекриття блоків (див. дані в таблиці 6.3). Для комбінованого порогового значення результати ще кращі, і значення MSE_{out} відрізняються приблизно на 10...12% від випадку використання блоків, що повністю перекриваються.

Таким чином, важко помітити різницю між відповідними результатами. Це означає, що доцільно застосовувати комбіноване порогове значення у випадку фільтрації зображення на основі ДКП з частковим перекриттям блоків. Можна також помітити, що для всіх зображень, спотворених АБГШ з дисперсією 25, MSSIM для вихідних зображень, перевищує 0,99, тобто залишковий шум і спотворення практично непомітні (можна ледь помітити).

Було згадано про локальну адаптацію вище. Існує досить багато можливих підходів, у тому числі й досить складних [103], [122]. Тут зосереджуємося на дуже швидкій обробці, тому алгоритм локальної адаптації, якщо його застосувати, має бути дуже простим і швидким.

У роботі [231] було відмічено, що середня кількість ДКП-коефіцієнтів у блоках, амплітуди яких перевищують заданий поріг, добре корелює з ефективністю фільтрації та глобально оптимальним β для фільтрації на основі ДКП із повним перекриттям. Оскільки для кожного блоку розраховуємо ДКП-коефіцієнти, легко і швидко порівняти їх амплітуди з заданим порогом і визначити кількість ДКП-

коефіцієнтів $N_{DCT}(n, m)$ для кожного nm -го блоку, який використовується в обробці.

Таблиця 6.6 – MSE_{out} та $MSSIM$ після фільтрації, часткове перекриття, $r=4$

Зображення	σ^2	Жорсткий поріг		Комбінований поріг	
		MSE_{out}	$MSSIM$	MSE_{out}	$MSSIM$
Lenna	25	11,51	0,9889	10,28	0,9901
	49	16,78	0,9841	15,31	0,9850
	100	25,01	0,9765	23,70	0,9769
Baboon	25	24,40	0,9914	22,53	0,9920
	49	42,49	0,9851	38,91	0,9865
	100	73,32	0,9743	67,30	0,9763
Barbara	25	12,51	0,9916	11,49	0,9924
	49	19,94	0,9871	18,47	0,9881
	100	32,87	0,9796	30,81	0,9806
Peppers	25	14,39	0,9870	12,69	0,9885
	49	20,41	0,9816	17,98	0,9837
	100	28,80	0,9746	26,19	0,9764
Goldhill	25	16,45	0,9887	14,75	0,9897
	49	25,32	0,9815	22,69	0,9833
	100	38,64	0,9700	34,67	0,9732

Тоді ідея локальної адаптації полягає в тому, щоб розрахувати $\beta(n, m)$ для кожного блоку окремо залежно від $N_{DCT}(n, m)$. Очевидно, що ця залежність має бути спадаючою, і є багато можливих варіантів. Ефективність залежить від попередньо встановленого порогу амплітуди коефіцієнта ДКП $T = b\sigma$, де b є параметром. Було перевірено декілька безперервних монотонно спадаючих функцій $\beta(n, m)$ від $N_{DCT}(n, m)$ і змінювались параметри цих функцій, а також змінювались b , щоб знайти оптимуми для метрик MSE_{out} і $MSSIM$. Проблема полягає в тому, що для різних тестових зображень і дисперсій шуму положення оптимумів не збігаються. Нарешті, було знайдено дуже просте рішення

використовувати $b=2,6$ і наступне жорстке перемикання для локально адаптивного налаштування $\beta(n, m)$:

$$\beta(n, m) = \begin{cases} 2,7, \text{ якщо } N_{DCT}(n, m) \leq 5 \\ 2,3, \text{ якщо } N_{DCT}(n, m) > 5 \end{cases}, \quad (6.14)$$

для випадку жорсткого порогу та

$$\beta(n, m) = \begin{cases} 4,4, \text{ якщо } N_{DCT}(n, m) \leq 6 \\ 3,5, \text{ якщо } N_{DCT}(n, m) > 6 \end{cases}, \quad (6.15)$$

для випадку комбінованого порогу.

Отримані результати для цих локально адаптивних методів представлені в таблиці 6.7. Їх порівняння з відповідними даними в таблиці 6.6 показує, що для всіх тестових зображень і значень дисперсії шуму розглянуті показники дещо покращилися. Однак ці покращення невеликі, і зміни якості зображення практично не помітні. Таким чином, користувач вирішує застосовувати або не застосовувати локальні алгоритми адаптації (6.14) і (6.15) на практиці.

При цьому забезпечена ефективність фільтрації за розглянутими метриками якості дуже близька до аналогічної для випадку блоків, що повністю перекриваються (див. дані в таблицях 6.3 і 6.4). Це означає, що зображення можна відфільтрувати приблизно в 16 разів швидше та практично без зниження ефективності порівняно з повним перекриттям.

Поліпшення ефективності фільтрації завдяки локальній адаптації, досягнутій для часткового перекриття блоків, може наштовхнути на ідею, що те ж саме може бути дійсним для придушення шумів на основі ДКП із повним перекриттям блоків (хоча це не прискорює фільтрацію та, крім того, робить його навіть трохи повільнішою).

Таблиця 6.7 – MSE_{out} та $MSSIM$ після адаптивної фільтрації, часткове перекриття, $r=4$

Зображення	σ^2	Жорсткий поріг		Комбінований поріг	
		MSE_{out}	$MSSIM$	MSE_{out}	$MSSIM$
Lenna	25	11,26	0,9892	10,10	0,9902
	49	16,33	0,9845	14,93	0,9854
	100	24,34	0,9770	22,73	0,9776
Baboon	25	22,75	0,9922	21,27	0,9928
	49	39,83	0,9863	37,15	0,9871
	100	69,44	0,9761	64,31	0,9778
Barbara	25	12,35	0,9917	11,26	0,9925
	49	19,81	0,9870	18,12	0,9882
	100	32,52	0,9799	30,11	0,9810
Peppers	25	13,96	0,9873	12,43	0,9886
	49	20,07	0,9822	17,73	0,9836
	100	28,30	0,9752	25,44	0,9765
Goldhill	25	15,83	0,9895	14,34	0,9903
	49	24,68	0,9827	22,30	0,9841
	100	37,64	0,9715	34,24	0,9740

Перевірено застосовність алгоритмів (6.14) і (6.15) для фільтрації з повним перекриттям. Дуже невелике покращення (зменшення вихідного MSE приблизно на 1% порівняно з неадаптивним випадком, див. дані в таблиці 6.3) було досягнуто для тих самих параметрів, що й у (6.14) і (6.15), відповідно. Таким чином, локальна адаптація дає певні переваги, але не дуже суттєві.

Можна стверджувати, що проведений вище аналіз був виконаний для АБГШ, тоді як шум на практиці може бути, наприклад, залежним від сигналу. У цьому випадку придушення шуму на основі ДКП має бути адаптовано до залежності локального стандартного відхилення $\sigma(n, m)$ від локального середнього $I_{mean}(n, m)$, яке вважається апіорно відомим. Потім у всіх алгоритмах, описаних вище, потрібно замінити σ на $\hat{\sigma}(n, m)$, який обчислюється на основі визначеного

$I_{mean}(n, m)$ для даного nm -го блоку. Наприклад, якщо шум пуассонівський, то $\hat{\sigma}(n, m) = \sqrt{I_{mean}(n, m)}$.

Випадок пуассонівського шуму був розглянутий для перевірки ефективності розглянутого підходу до фільтрації на основі ДКП, включаючи прискорені версії. Тестування було проведено для декількох тестових зображень. Як окремий випадок наведемо дані для тестового зображення Goldhill. Фільтрація на основі ДКП із повним перекриттям блоків забезпечує $MSE_{out}=30,51$ і $MSSIM=0,9759$ для жорсткого порогу, тоді як для комбінованого порогу отримані $MSE_{out}=31,69$ і $MSSIM=0,9754$. Як видно, значення метрик для цих версій придушення шуму на основі ДКП знову досить близькі.

Якщо блоки не перекриваються, метрики для вихідних зображень є такими: $MSE_{out}=63,79$, $MSSIM=0,9544$ для жорсткого порогу та $MSE_{out}=46,05$, $MSSIM=0,9661$ для комбінованого порогу. Таким чином, якість значно гірша, ніж у випадку використання блоків, що повністю перекриваються. Знову ж таки, використання комбінованого порогу дає кращі результати. Для часткового перекриття блоків із $r=4$ отримано наступні результати: $MSE_{out}=39,16$, $MSSIM=0,9703$ для жорсткого порогу та $MSE_{out}=35,36$, $MSSIM=0,9730$ для комбінованого порогу. Як видно, значення метрик значно кращі, ніж у випадку блоків, що не перекриваються, і гірші, ніж у випадку блоків, що повністю перекриваються. Використання локальної адаптації дозволяє підвищити ефективність фільтру. На основі алгоритму адаптації (6.14) і (6.15) було отримано наступні значення метрик: $MSE_{out}=38,72$, $MSSIM=0,9705$ для жорсткого порогу та $MSE_{out}=34,88$, $MSSIM=0,9733$ для комбінованого порогу. Вони ближчі до відповідних значень для блоків, що повністю перекриваються. Різниця настільки мала, що вихідні зображення мають приблизно однакову візуальну якість. Щоб підтвердити це, наведемо декілька прикладів. Фрагмент тестового зображення Goldhill, спотвореного пуассонівським шумом, представлено на рис. 6.10, а. Шум чітко видно, особливо в однорідних областях зображення з відносно великим локальним середнім.

Фрагмент цього зображення після застосування фільтра на основі ДКП з блоками, що не перекриваються, з жорстким порогом показано на рис. 6.10, б. Як видно, шуми частково придушені, але є неприємні блокові артефакти. Такі ж фрагменти на виході фільтра на основі ДКП з $r=4$ і комбінованим порогом показано на рис. 6.10, в. Блокові артефакти відсутні і спостерігається ефективно придушення шумів, а також прийнятне збереження деталей, меж і текстур. Нарешті, на рис. 6.9, г, представлено фрагмент на виході фільтра на основі ДКП із жорстким порогом і повним перекриттям блоків. Він демонструє високу якість зображення на виході, але розрізнити відмінності порівняно з фрагментом на рис. 6.9, в, нелегко навіть для збільшених фрагментів. Таким чином, можливе і доцільне застосування придушення шумів на основі ДКП з високою обчислювальною ефективністю (з $r=4$ і комбінованим порогом) замість фільтрації на основі ДКП з повним перекриттям, яка має значно більше обчислювальне навантаження. Прискорений варіант, за умови відповідної адаптації, може бути використаний для придушення не тільки АБГШ, а й сигнально-залежних шумів.

На практиці не тільки СКВ шуму, але і його спектр часто є невідомим і нестационарним. Це є проблемою для методів придушення шуму, оскільки вони зазвичай спираються на конкретні припущення щодо характеристик шуму. Наприклад, як показано на рис. 6.11, а, шум у зображенні, отриманому за допомогою фазового зображення [306], є просторово корельованим у центрі зображення, але є майже білим біля кутів зображення.

СКВ шуму є низьким у центрі та зростає до кутів, а розподіл майже гаусів у центрі, але стає негаусовим (з великими хвостами) біля кутів. Крім того, шум на кутах можна вважати імпульсним. У таких випадках ні ЗНМ, навчені шуму з певним розподілом і спектром, ні методи оцінки карти СКВ шуму не є ефективними.



а

б



в

г

Рисунок 6.10 – Фрагмент зображення Goldhill: зашумлене (а) і виходи фільтрів на основі ДКП з блоками, що не перекриваються (б), частково перекриваються (в) і повністю перекриваються (г)

Інший приклад показано на рис. 6.11, б, де показано зображення, отримане за допомогою ланцюга обробки зображень у цифровій камері Canon EOS 250D. Це зображення автоматично об'єднується з декількох зашумлених зображень за допомогою процедури зіставлення блоків, що призводить до видимого залишкового просторово корельованого шуму з невідомим спектром.



а

б

Рисунок 6.11 – Приклади зображень шумом, характеристики якого не визначені: фазове (а) та отримане за допомогою Canon EOS 250D (б)

СКВ шуму відрізняється в різних областях зображення, і шум частково згладжується через стиснення JPEG із втратами. У цьому дослідженні спеціально розглядаємо завдання зменшення шуму на рентгенівських зображеннях зубів системи Morita. Шум на цих зображеннях, як показано в [53], є просторово корельованим, а СКВ шуму змінюється по різних регіонах.

Запропонований підхід базується на ідеї отримання карт рівнів СКВ (КСКВ) для трьох різних масштабів зображення (рис. 6.12). Це забезпечує 3-параметрову оцінку спектру шуму для кожного пікселя, яка, хоч і є грубою, може бути достатньою для ефективного придушення шуму. Пропонований метод базується на двох ЗНМ (рис. 6.13), які навчаються та застосовуються послідовно. Однак ЗНМ можна застосовувати окремо для використання в інших завданнях обробки зображень. Архітектура обох ЗНМ схожа на архітектуру DRUNet [148].

Перша ЗНМ, NSPNet, у парі прогнозує три двовимірні КСКВ шуму: для повнорозмірного зображення («КСКВ1»), для зображення, зменшеного в два рази («КСКВ2»), і для зображення зменшеного масштабу в чотири рази («КСКВ3»). Разом ці три карти шуму забезпечують оцінку спектру шуму та СКВ шуму для кожного пікселя зображення.

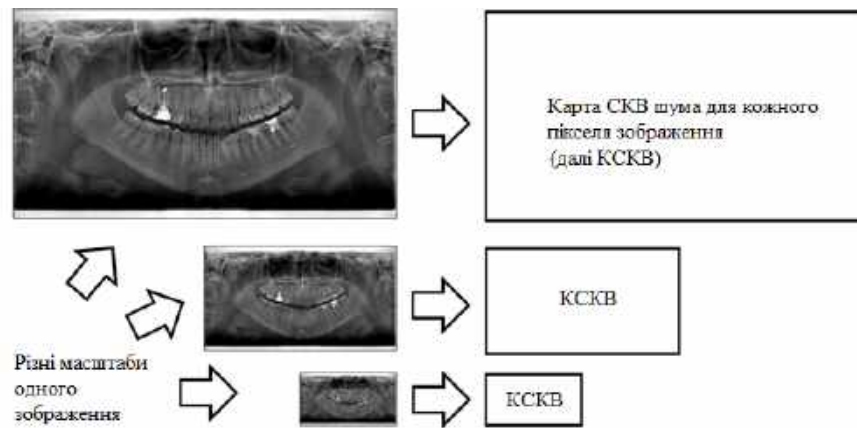


Рисунок 6.12 – Процедура оцінки спектру та рівнів шуму р трьох карт значень СКВ для трьох різних масштабів зображення

Друга ЗНМ, GDNet, придушує шум, приймаючи на вхід зашумлені зображення та КСКВ1, КСКВ2, КСКВ3. Карти КСКВ2 та КСКВ3 також підключені до внутрішніх шарів мережі, які обробляють відповідні масштаби зображення.

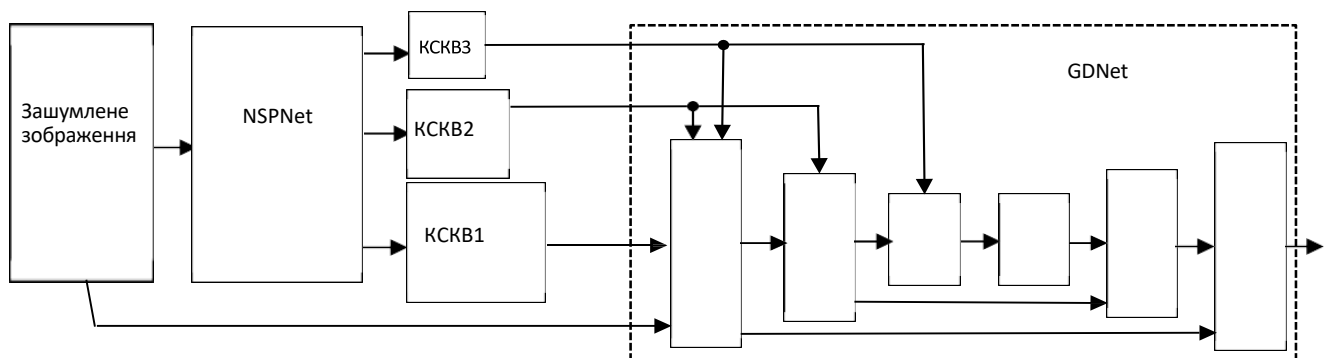


Рисунок 6.13 – Структурна схема запропонованого сліпого придушення шуму

Для навчання GDNet було використано карти СКВ, оцінені вже попередньо навченими NSPNet. Це забезпечує додаткову надійність навчній GDNet.

Для навчання NSPNet було використано безшумні зображення зі штучним додаванням шуму. Таким чином, відомі базові карти істинності «КСКВ1», «КСКВ2» і «КСКВ3». Щоб навчити NSPNet, було вибрано 4238 зображень із різних наборів даних, як описано нижче. Було використано 1000 зображень із бази даних зображень Tampere21 [307], яка складається з 1000 майже вільних від шумів кольорових зображень. Другий набір із 1000 зображень було вибрано з чотирьох

наборів даних: KonIQ10k [308], FLIVE [309], NRTID [310] і SPAQ [311]. Об'єднані середні оцінки думок експертів (MOS) цих чотирьох баз даних були використані для вибору високоякісних зображень із найкращими значеннями MOS. Решта зображень були зібрані з різних джерел, в тому числі 1795 зображень з різних фотохостингів. Метрика KonCepT512, попередньо навчена на шести базах даних (KonIQ10k, Live-in-the-Wild, FLIVE, NRTID, HTID і SPAQ), використовувалася для збору високоякісних зображень із згаданих вище джерел зображень.

Для простоти моделювання шуму в навчальних наборах просторово корельований шум був змодельований як $n \otimes k$, де n представляє шум, а k представляє ротаційно-симетричний фільтр низьких частот Гауса зі стандартним відхиленням σ_k . Згідно з цією моделлю рівень шуму характеризується значенням σ , тоді як ступінь кореляції між значеннями шуму в сусідніх пікселях зображення, що називається «шумовим зерном», характеризується значенням σ_k . У даних навчання 20% шуму становили АБГШ, 60% – просторово корельований шум із різними рівнями кореляції, а 20% – високочастотний шум. 7% шуму були адитивним шумом із постійним СКВ, 7% були мультиплікативним шумом із постійним СКВ і 7% були пуассонівським шумом з постійним СКВ. Решта 79% були адитивним шумом із СКВ-картою випадково вибраної форми. Цей тренінг надав NSPNet можливість ефективно оцінювати спектр і рівень шуму всліпу. GDNet було навчено на тих самих патчах із вхідними картами СКВ, оціненими попередньо навченою NSPNet. Ця схема навчання забезпечила GDNet часткову стійкість до помилок оцінки, зроблених NSPNet. Навчання проводилося в середовищі Matlab 2022 з використанням спеціального циклу навчання, що призвело до 100 000 ітерацій із розміром міні-пакета 32. MSE було використано як функцію втрат, щоб мінімізувати кількість викидів в оцінках карт СКВ.

Щоб оцінити якість навчання NSPNet, було застосовано мережу для оцінки стандартного відхилення АБГШ із постійним СКВ шуму. У таблиці 6.8 наведено результати порівняння NSPNet із найкращими методами для цього завдання. Результати показують, що NSPNet демонструє точність оцінки, яка дуже близька до найсучаснішої мережі SDNet [307], яка була спеціально розроблена для АБГШ.

Рисунок 6.13 ілюструє приклад оцінки спектра та рівня шуму NSPNet для зображення з просторово корельованим шумом із фіксованим СКВ шуму. Медіанні значення СКВ для істинних карт СКВ і розрахованих карт СКВ (σ_1 , σ_2 і σ_3) можна порівняти. Для цього зображення NSPNet оцінює спектр шуму з високою точністю. Візуальний аналіз розрахованих карт підтверджує, що дрібні деталі та текстури зображення вносять мінімальні помилки в розраховані карти СКВ.

Таблиця 6.8 – Відносна похибка СКВ оцінки АБГШ з постійним СКВ

σ	IEDD[270]	PCA[312]	WTP[313]	VDNet[149]	SDNet[307]	NSPNet
3	0,407	0,573	0,507	1,782	0,150	0,192
5	0,204	0,274	0,269	0,905	0,094	0,091
7	0,114	0,166	0,174	0,559	0,062	0,050
10	0,073	0,097	0,122	0,323	0,034	0,028
15	0,071	0,053	0,075	0,168	0,013	0,021
20	0,077	0,038	0,057	0,105	0,016	0,020
30	0,114	0,036	0,055	0,066	0,011	0,012
50	0,157	0,069	0,103	0,089	0,009	0,009
75	0,207	0,172	0,176	0,150	0,007	0,009

Було змодельовано шум, характерний для рентгенівських зображень, отриманих за допомогою стоматологічної системи Morita [280]. Шум складається із суміші просторово корельованого адитивного гаусового шуму з $\sigma=2$ і просторово-корельованого пуссонівського шуму, а також геометричних спотворень. Ця модель була використана для створення 20 наборів із трьох зашумлених зображень (з трьома різними шумовими спектрами) для 20 відібраних вручну зображень без шуму. На рис. 6.14, 6.15 показаний приклад змодельованих безшумних і зашумлених зображень із шумом, подібним до рентгенівських знімків зубів із $\sigma_k=0,9$.

Згенерований тестовий набір використовувався для порівняння запропонованого методу (нейромережі) NSPNet+GDNet з кількома найсучаснішими методами придушення шумів. Результати представлені в таблиці

6.9. Параметр σ_k у цьому контексті пов'язаний зі ступенем кореляції або «зерном» просторово-корельованого шуму. Слід зазначити, що як сліпа, так і фіксована сигма версії найсучаснішого фільтра Restormer [150] не працюють добре в цьому випадку, оскільки вони були навчені для АБГШ.



Рисунок 6.14 – Приклад зображення з просторово корельованим шумом і спектром шуму (карти СКВ для трьох масштабів), оцінені NSPNet. Справжній спектр: $\sigma_1=10$, $\sigma_2=8,5$, $\sigma_3=5,7$. Розрахунковий спектр: $\sigma_1=9,7$, $\sigma_2=8,3$, $\sigma_3=5,7$

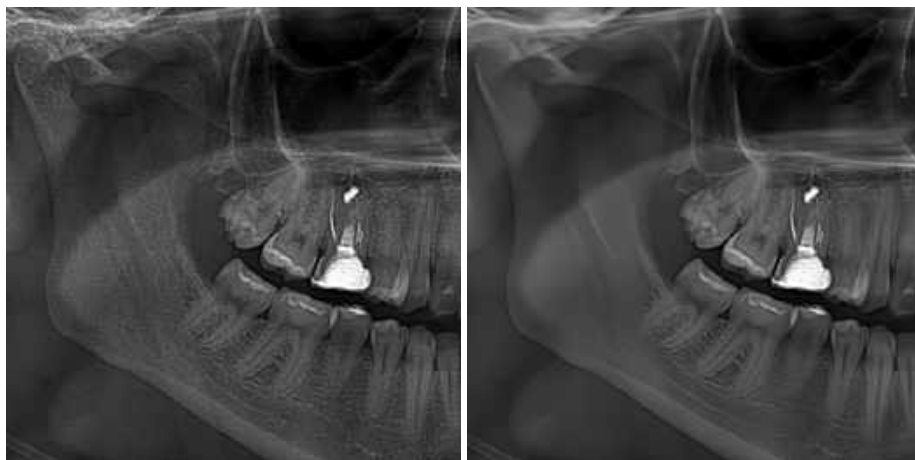


Рисунок 6.15 – Змодельоване безшумне зображення та змодельоване зашумлене зображення з шумом, характерним для рентгенівського зображення зубів

Рис. 6.16 ілюструє результат зменшення шуму фрагмента реального рентгенівського зображення зубів, отриманого за допомогою системи Morita.

Таблиця 6.9 – Придушення шуму, характерного для рентгенівських знімків зубів, для 3 різних рівнів кореляції шуму, PSNR, дБ

Фільтр	$\sigma_k=0,7$	$\sigma_k=0,8$	$\sigma_k=0,9$
Зашумлене зображення	29,2	29,2	29,2
NSPNet+GDNet	36,5	36,5	36,7
DRUNet з постійним СКВ	35,2	34,6	34,0
Нелокалізоване середнє з фіксованим порогом	33,9	33,4	32,9
Двосторонній фільтр Гауса зі збереженням меж	33,5	33,4	33,2
Фільтр ДКП з фіксованим порогом	33,3	32,8	32,3
Вейвлет-фільтр із фіксованим порогом	31,8	31,5	31,2
DRUNET+NLNET+M2	30,2	29,9	29,7
Restormer з сігмою 5	30,2	29,9	29,8
Restormer, сліпий	29,2	29,2	29,2



а

б

Рисунок 6.16 – Приклад придушення шумів реального зображення зубів системи Morita: оригінал (а) та після придушення шумів (б)

Навпаки, поєднання NSPNet + GDNet є ефективним і забезпечує PSNR краще, ніж у DRUNet, на 2,7 дБ для більшої кореляції шуму. На зображенні з шумом немає видимих залишкових шумів.

Висновки до розділу 6

В цьому розділі розглянуто питання підвищення ефективності інтелектуальної фільтрації та стиснення із втратами зображень дистанційного зондування Землі та медичної діагностики в умовах присутності складних завад і обмежених обчислювальних ресурсів з урахуванням візуальної якості. Представлено огляд досліджень, які розглядають вплив стиснення на ефективність виявлення цілі, з деякими новими ідеями. А саме, розглядається вплив стиснення в околі точки ледь помітної різниці, де ціль та фон схожі. Показано, що стиснення в цьому випадку не впливає суттєво на подібність між ціллю та фоном, тому слід очікувати, що воно також не вплине на ефективність виявлення цілі людиною. Використовувалися чотири метрики складності фону, які досягли високого ступеня збігу з результатами суб'єктивних тестів. Ці показники вперше використовуються як метрики завад, і серед них метрика HaarPSI, яка моделює увагу людини до фрагментів зображення, досягає найкращих результатів із ефективністю, яка знаходиться в діапазоні об'єктивних показників, спеціально розроблених для ефективності виявлення цілі. Метрики, засновані на оцінці якості зображення, досягають певної згоди з результатами суб'єктивних тестів, виміряних за допомогою лінійного коефіцієнта кореляції 90%.

Запропоновано та досліджено шляхи прискорення придушення шумів на основі ДКП за допомогою обробки зображень у блоках, що частково перекриваються. Показано, що комбіноване порогове визначення забезпечує кращу ефективність, а локальна адаптація здатна додатково покращити її. Запропоновані методи здатні придушувати як АБГШ, так і сигнально-залежний шум із прийнятним збереженням корисної інформації в зображеннях.

Запропоновано новий швидкий підхід для сліпого придушення шумів складного виду на рентгенівських знімках зубів, які спотворені просторово корельованим шумом із невідомим спектром і стандартним відхиленням, яке змінюється для різних областей зображення. Запропоновано дві глибокі згорточні нейронні мережі, NSPNet і GDNet, які дозволяють забезпечити ефективне вирішення цієї проблеми. Результати показали, що запропонований підхід забезпечує PSNR, який є кращим, ніж інші сучасні методи придушення шуму, більш ніж на 2,7 дБ. Запропонований підхід є високоефективним і може бути використаний для придушення будь-якого типу просторово-корельованого шуму з невідомим спектром і стандартним відхиленням, яке змінюється для різних областей зображення.

Основні наукові результати, наведені в шостому розділі, опубліковано в роботах автора [4], [8], [38], [41], [51], [54], [62], [64], [67], зазначених в анотації.

ВИСНОВКИ

Згідно з поставленою метою в дисертаційній роботі вирішено важливу науково-практичну проблему ефективного інтелектуального оброблення (фільтрації та стиснення із втратами) зображень дистанційного зондування Землі та медичної діагностики в умовах присутності складних завад і обмежених обчислювальних ресурсів з урахуванням візуальної якості. Отримано такі наукові та практичні результати:

1) методи детектування текстур на зашумлених зображеннях. Було показано, що оптимальним є спільне використання локальних параметрів (відносної локальної дисперсії, квазірозмаху та параметра на основі косинусного дискретного перетворення) з обчисленням у вікні розміром 5×5 . Удосконалення алгоритмів навчання класифікаторів, зокрема через агрегування параметрів у ковзному вікні або використання потужних класифікаторів, таких як машина опорних векторів, дозволяє підвищити якість класифікації. Крім того, використання нейронних мереж на основі декількох локальних параметрів забезпечує високу ефективність із меншими обчислювальними витратами.

2) фільтрація спекл-шуму у зображеннях дистанційного зондування. Для боротьби зі спекл-шумом запропоновано застосування методу Level Set для локалізації неоднорідностей зображень. Цей підхід дозволяє адаптивно застосовувати частотно-залежну фільтрацію до однорідних ділянок, забезпечуючи збереження важливих деталей. Використання вейвлет-фільтра Хаара та інтегрально-диференціального методу показало гарні результати як за кількісними метриками (MSE, MSSIM), так і за якістю збереження меж. Запропоновано способи суттєвого підвищення швидкодії обробки.

3) прогнозування ефективності фільтрації. Розроблено метод прогнозування ефективності придушення сигнально-залежного шуму. Це дозволяє оцінювати покращення традиційних та візуальних метрик якості, таких як PSNR і PSNR-HVS-M, до фільтрації, що скорочує час обробки та знижує обчислювальні витрати. Метод прогнозування продемонстрував високу точність для шумів різної природи

(адитивного, мультиплікативного) і забезпечує адаптацію до особливостей зображень.

4) стиснення зображень із втратами. Запропоновано двоетапну процедуру стиснення, яка дозволяє забезпечити бажаний рівень якості зображень згідно з метриками PSNR-HVS-M і PSNR з похибкою менше 1 дБ. Ефективність процедури підтверджено для різних типів зображень, включаючи медичні дані, і показано, що вона є значно швидшою за ітераційні підходи. У випадках, коли потрібна модифікація для забезпечення «недратівливих» спотворень, метод залишається універсальним і може бути адаптований для інших показників візуальної якості.

5) аналіз характеристик шуму та адаптація алгоритмів. Проведено просторово-спектральний аналіз шумів у зображеннях, отриманих сучасними рентгенівськими системами. Показано, що шум є просторово корельованим і залежить від сигналу, що вимагає розробки адаптивних методів фільтрації. Запропоновано нові алгоритми, такі як NSPNet і GDNet, для сліпого придушення шуму, які забезпечують покращення PSNR на 2,7 дБ у порівнянні з існуючими методами.

6) стиснення для задач медичної діагностики. Визначено, що стиснення зображень зубів із використанням кодера ADCT із правильно підібраними параметрами дозволяє досягти високих значень коефіцієнта стиснення без втрат діагностично важливої інформації. Застосування запропонованих підходів до медичних зображень гарантує непомітність спотворень, що підтверджено експериментами з групами стоматологів.

7) контроль якості стиснення для мультиспектральних даних. Запропоновано методику поканального аналізу мультиспектральних даних, що дозволяє визначити оптимальні параметри стиснення для кожного спектрального каналу. Показано, що для зображень із простою структурою стиснення може навіть покращити результати класифікації завдяки придушенню шумів і високочастотних варіацій.

Результати дослідження дозволяють розробляти практичні рекомендації щодо вибору параметрів кодерів залежно від властивостей даних. Це сприяє оптимізації процесів обробки зображень у реальних системах, таких як медична

візуалізація, дистанційне зондування, робототехніка тощо. Дослідження відкривають перспективи для створення більш точних і адаптивних методів обробки, які можуть бути застосовані для 3D-даних і відеопослідовностей.

Дисертаційна робота вирішує актуальну проблему прогнозування та обробки зображень, яка є критично важливою для сучасних систем дистанційного зондування та медичної діагностики. Отримані результати забезпечують наукову та практичну базу для подальшого розвитку високотехнологічних рішень у цих галузях, а також створюють умови для підвищення ефективності реальних систем обробки даних.

СПИСОК ВИКОРИСТАНИХ ДЖЕРЕЛ

1. С. К. Абрамов, Р. А. Кожемякин, С. С. Кривенко, Н. Н. Пономаренко, and В. В. Лукин, “Эффективность фильтрации одновзглядовых РСА-изображений при пространственно-коррелированных помехах,” *Радіоелектронні і комп’ютерні системи*, no. 3, pp. 18–25, 2012.
2. В. В. Лукин, А. В. Наumenко, and С. С. Кривенко, “Адаптивная фильтрация РСА-изображений с использованием нейросетевых детекторов границ,” *Радіоелектронні і комп’ютерні системи*, no. 4, pp. 121–129, 2012.
3. С. С. Кривенко, Е. О. Колганова, and В. В. Лукин, “Фильтрация изображений с использованием преобразования Хаара в блоках,” *Системи обробки інформації*, no. 8, pp. 63–69, 2013.
4. V. V. Lukin, S. S. Krivenko, E. O. Kolganova, and M. K. Chobanu, “Acceleration and efficiency analysis of dct-based filtering,” *Telecom Rad Eng*, vol. 72, no. 7, pp. 613–626, 2013, doi: 10.1615/TelecomRadEng.v72.i7.50.
5. S. Abramov, S. Krivenko, A. Roenko, V. Lukin, I. Djurovic, and M. Chobanu, “Prediction of filtering efficiency for DCT-based image denoising,” in *2013 2nd Mediterranean Conference on Embedded Computing (MECO)*, Budva, Montenegro: IEEE, Jun. 2013, pp. 97–100. doi: 10.1109/MECO.2013.6601327.
6. D. V. Fevrlev, S. S. Krivenko, V. V. Lukin, R. Marques, and F. D. Medeiros, “Combining level sets and orthogonal transform for despeckling sar images,” *Авиационно-космическая техника и технология*, no. 2, pp. 103–112, 2013.
7. R. A. Kozhemiakin, S. S. Krivenko, V. V. Lukin, R. C. P. Marques, F. N. S. Medeiros, and B. Vozel, “Efficiency analysis of combined despeckling of single-look sar images,” *AKTT*, no. 5, pp. 102–111, 2013.
8. V. V. Lukin *et al.*, “Approaches to automatic data processing in hyperspectral remote sensing,” *Telecom Rad Eng*, vol. 73, no. 13, pp. 1125–1139, 2014, doi: 10.1615/TelecomRadEng.v73.i13.10.
9. S. Krivenko, V. Lukin, B. Vozel, and K. Chehdi, “Filtering efficiency prediction for images corrupted by signal-dependent noise,” in *12th International Conference*

Modern Problems Of Radio Engineering, Telecommunications, and Computer Science, 2014, pp. 615–617. Accessed: Feb. 08, 2024. [Online]. Available: <https://hal.science/hal-01111060/>

10. S. Krivenko, V. Lukin, B. Vozel, and K. Chehdi, “Prediction of DCT-based denoising efficiency for images corrupted by signal-dependent noise,” in *2014 IEEE 34th International Scientific Conference on Electronics and Nanotechnology (ELNANO)*, Kyiv, Ukraine: IEEE, Apr. 2014, pp. 254–258. doi: 10.1109/ELNANO.2014.6873920.

11. V. V. Lukin *et al.*, “An approach to prediction of signal-dependent noise removal efficiency by dct-based filter,” *Telecom Rad Eng*, vol. 73, no. 18, pp. 1645–1659, 2014, doi: 10.1615/TelecomRadEng.v73.i18.40.

12. O. S. Rubel, R. O. Kozhemiakin, S. S. Krivenko, V. V. Lukin, B. Vozel, and K. Chehdi, “A method for predicting denoising efficiency for color images,” in *2015 IEEE 35th International Conference on Electronics and Nanotechnology (ELNANO)*, Kyiv, Ukraine: IEEE, Apr. 2015, pp. 304–309. doi: 10.1109/ELNANO.2015.7146897.

13. A. Naumenko, S. Krivenko, V. Lukin, and K. Egiazarian, “Texture region detection by trained neural network,” in *2016 9th International Kharkiv Symposium on Physics and Engineering of Microwaves, Millimeter and Submillimeter Waves (MSMW)*, Kharkiv, Ukraine: IEEE, Jun. 2016, pp. 1–4. doi: 10.1109/MSMW.2016.7538174.

14. A. Naumenko, S. Krivenko, N. Ponomarenko, A. Zelensky, and V. Lukin, “Texture detection in noisy images by combining several local parameters,” in *2015 Second International Scientific-Practical Conference Problems of Infocommunications Science and Technology (PIC S&T)*, Kharkiv, Ukraine: IEEE, Oct. 2015, pp. 230–233. doi: 10.1109/INFOCOMMST.2015.7357321.

15. С. С. Кривенко, А. В. Наumenко, and В. В. Лукин, “Обнаружение текстурных участков SVM-классификатором на изображениях при наличии помех,” *Радіоелектронні і комп’ютерні системи*, no. 2, pp. 50–57, 2015.

16. A. V. Naumenko, V. V. Lukin, M. S. Zriakhov, and S. S. Krivenko, “Feature aggregation for noisy image to improve ‘texture/non-texture’ classification,” in *2017 IEEE 37th International Conference on Electronics and Nanotechnology (ELNANO)*, Kyiv, Ukraine: IEEE, Apr. 2017, pp. 461–464. doi: 10.1109/ELNANO.2017.7939808.

17. S. K. Abramov, A. Naumenko, V. V. Lukin, S. S. Krivenko, and I. Kaluzhynov, "Neural network based detection of heterogeneities in noisy images," *Telecom Rad Eng*, vol. 79, no. 19, pp. 1691–1705, 2020, doi: 10.1615/TelecomRadEng.v79.i19.20.

18. B. Vozel *et al.*, "Processing of multichannel Remote-sensing images with prediction of performance parameters," in *Environmental Applications of Remote Sensing*, IntechOpen, 2016. Accessed: Feb. 08, 2024. [Online]. Available: <https://www.google.com/books?hl=uk&lr=&id=Wm-QDwAAQBAJ&oi=fnd&pg=PA373&dq=Processing+of+Multichannel+Remote-Sensing+Images+with+Prediction+of+Performance+Parameters&ots=ClwhdRf6aV&sig=1jtYRgPSY3irBSaTTqyoJQ-L3n4>

19. V. V. Lukin *et al.*, "Application of filtering efficiency prediction to hyperspectral data pre-processing," *Ukrainian journal of remote sensing*, no. 7, pp. 4–11, 2015.

20. S. Krivenko, O. Kolganova, and V. Lukin, "Lossy compression of images corrupted by spatially correlated noise," in *2016 13th International Conference on Modern Problems of Radio Engineering, Telecommunications and Computer Science (TCSET)*, Lviv, Ukraine: IEEE, Feb. 2016, pp. 698–702. doi: 10.1109/TCSET.2016.7452157.

21. M. Zriakhov, S. Kryvenko, and B. Лукін, "Peculiarities of multispectral image lossy compression," *Вісник Хмельницького національного університету. Серія: Технічні науки*, vol. 331, no. 1, pp. 370–376, Feb. 2024, doi: 10.31891/2307-5732-2024-331-56.

22. V. Lukin, B. Kovalenko, S. Kryvenko, V. Naumenko, and B. Vozel, "Prediction of Optimal Operation Point Existence and Its Parameters in BPG-Based Automatic Lossy Compression of Noisy Images," in *Current Overview on Science and Technology Research Vol. 9*, Dr. G. Bucci, Ed., B P International (a part of SCIENCEDOMAIN International), 2022, pp. 1–36. doi: 10.9734/bpi/costr/v9/4316A.

23. S. S. Krivenko, V. V. Lukin, and E. Kolganova, "Peculiarities of lossy compression of optical and medical images corrupted by noise," in *2016 IEEE 36th*

International Conference on Electronics and Nanotechnology (ELNANO), Kyiv, Ukraine: IEEE, Apr. 2016, pp. 251–256. doi: 10.1109/ELNANO.2016.7493060.

24. S. Krivenko, V. Lukin, O. Krylova, and V. Shutko, “Visually Lossless Compression of Retina Images,” in *2018 IEEE 38th International Conference on Electronics and Nanotechnology (ELNANO)*, Kiev: IEEE, Apr. 2018, pp. 255–260. doi: 10.1109/ELNANO.2018.8477459.

25. S. Krivenko, M. Zriakhov, V. Lukin, and B. Vozel, “MSE prediction in DCT-based lossy compression of noise-free and noisy remote sensing,” in *2018 14th International Conference on Advanced Trends in Radioelectronics, Telecommunications and Computer Engineering (TCSET)*, Lviv-Slavske, Ukraine: IEEE, Feb. 2018, pp. 883–888. doi: 10.1109/TCSET.2018.8336338.

26. V. Naumenko, V. Lukin, and S. Kryvenko, “Comparison of Noisy Medical Image Lossy Compression by BPG and AGU,” in *Integrated Computer Technologies in Mechanical Engineering - 2022*, vol. 657, M. Nechyporuk, V. Pavlikov, and D. Kritskiy, Eds., in *Lecture Notes in Networks and Systems*, vol. 657, Cham: Springer Nature Switzerland, 2023, pp. 636–648. doi: 10.1007/978-3-031-36201-9_53.

27. V. Lukin, B. Kovalenko, S. Kryvenko, N. Ponomarenko, K. Egiazarian, and J. Astola, “Compression of Noisy Images Taking into Account Visual Quality: A Comprehensive Study,” *Research Developments in Science and Technology Vol. 8*, pp. 89–109, 2022.

28. B. Kovalenko, V. Lukin, S. Kryvenko, V. Naumenko, and B. Vozel, “Prediction of Parameters in Optimal Operation Point for BPG-based Lossy Compression of Noisy Images,” *UJRS*, vol. 9, no. 2, pp. 4–12, Jun. 2022, doi: 10.36023/ujrs.2022.9.2.212.

29. B. Kovalenko, V. Lukin, S. Kryvenko, V. Naumenko, and B. Vozel, “BPG-Based Automatic Lossy Compression of Noisy Images with the Prediction of an Optimal Operation Existence and Its Parameters,” *Applied Sciences*, vol. 12, no. 15, p. 7555, Jul. 2022, doi: 10.3390/app12157555.

30. B. Kovalenko, V. Lukin, V. Naumenko, and S. Krivenko, “Analysis of noisy image lossy compression by BPG using visual quality metrics,” in *2021 IEEE 3rd*

International Conference on Advanced Trends in Information Theory (ATIT), Kyiv, Ukraine: IEEE, Dec. 2021, pp. 20–25. doi: 10.1109/ATIT54053.2021.9678575.

31. С. К. Абрамов, Н. В. Бурцев, С. С. Кривенко, А. Н. Земляченко, and В. В. Лукин, “Автоматическое сжатие в окрестности оптимальной рабочей точки изображений с шумом кодерами типа SPIHT и JPEG2000,” *Авиационно-космическая техника и технология*, no. 2, pp. 99–109, 2016.

32. S. Krivenko, V. Lukin, and B. Vozel, “Prediction of Introduced Distortions Parameters in Lossy Image Compression,” in *2018 International Scientific-Practical Conference Problems of Infocommunications. Science and Technology (PIC S&T)*, Kharkiv, Ukraine: IEEE, Oct. 2018, pp. 447–451. doi: 10.1109/INFOCOMMST.2018.8632130.

33. S. Krivenko, M. Zriakhov, V. Lukin, and B. Vozel, “MSE and PSNR prediction for ADCT coder applied to lossy image compression,” in *2018 IEEE 9th International Conference on Dependable Systems, Services and Technologies (DESSERT)*, Kyiv, Ukraine: IEEE, May 2018, pp. 613–618. doi: 10.1109/DESSERT.2018.8409205.

34. S. S. Krivenko, O. Krylova, E. Bataeva, and V. V. Lukin, “Smart lossy compression of images based on distortion prediction,” *Telecom Rad Eng*, vol. 77, no. 17, pp. 1535–1554, 2018, doi: 10.1615/TelecomRadEng.v77.i17.40.

35. V. Lukin, A. Zemliachenko, S. Krivenko, B. Vozel, and K. Chehdi, “Lossy compression of remote sensing images with controllable distortions,” *Satellite Information Classification and Interpretation*, 2018, Accessed: Feb. 11, 2024. [Online]. Available:

https://www.google.com/books?hl=uk&lr=&id=lzj8DwAAQBAJ&oi=fnd&pg=PA27&dq=Lossy+compression+of+remote+sensing+images+with+controllable+distortions&ots=irka3mNKud&sig=ljJ_EiGSZkYlW4tNWjpzfsvU3dg

36. S. Krivenko, V. Lukin, and O. Krylova, “Visually Lossless Compression of Dental Images,” in *2019 IEEE 39th International Conference on Electronics and Nanotechnology (ELNANO)*, Kyiv, Ukraine: IEEE, Apr. 2019, pp. 394–399. doi: 10.1109/ELNANO.2019.8783218.

37. S. Krivenko, M. Zriakhov, N. Kussul, and V. Lukin, “Prediction of Visual Quality for Lossy Compressed Images,” in *2019 IEEE 15th International Conference on the Experience of Designing and Application of CAD Systems (CADSM)*, Polyana, Ukraine: IEEE, Feb. 2019, pp. 1–5. doi: 10.1109/CADSM.2019.8779266.

38. O. Krylova, L. Kryvenko, S. Krivenko, and V. Lukin, “A Fast Noniterative Visually Lossless Compression of Dental Images Using AGU-M Coder,” in *2021 IEEE 16th International Conference on the Experience of Designing and Application of CAD Systems (CADSM)*, Lviv, Ukraine: IEEE, Feb. 2021, pp. 6–10. doi: 10.1109/CADSM52681.2021.9385257.

39. S. Krivenko, V. Lukin, O. Krylova, L. Kryvenko, and K. Egiazarian, “A Fast Method of Visually Lossless Compression of Dental Images,” *Applied Sciences*, vol. 11, no. 1, Art. no. 1, Jan. 2021, doi: 10.3390/app11010135.

40. S. S. Krivenko, S. K. Abramov, V. V. Lukin, B. Vozel, and K. Chehdi, “Lossy DCT-based compression of remote sensing images with providing a desired visual quality,” in *Image and Signal Processing for Remote Sensing XXV*, L. Bruzzone, F. Bovolo, and J. A. Benediktsson, Eds., Strasbourg, France: SPIE, Oct. 2019, p. 36. doi: 10.1117/12.2532726.

41. F. Li, S. Krivenko, and V. Lukin, “A Modified Algorithm of Visual Quality Providing in Lossy Compression of Remote Sensing Images,” in *2021 IEEE International Conference on Information and Telecommunication Technologies and Radio Electronics (UkrMiCo)*, Odesa, Ukraine: IEEE, Nov. 2021, pp. 16–21. doi: 10.1109/UkrMiCo52950.2021.9716684.

42. S. Krivenko, D. Demchenko, I. Dyogtev, and V. Lukin, “A Two-Step Approach to Providing a Desired Quality of Lossy Compressed Images,” in *Integrated Computer Technologies in Mechanical Engineering*, vol. 1113, M. Nechyporuk, V. Pavlikov, and D. Kritskiy, Eds., in *Advances in Intelligent Systems and Computing*, vol. 1113, Cham: Springer International Publishing, 2020, pp. 482–491. doi: 10.1007/978-3-030-37618-5_41.

43. S. Krivenko, F. Li, V. Lukin, B. Vozel, and O. Krylova, “Prediction of Visual Quality Metrics in Lossy Image Compression,” in *2020 IEEE 40th International*

Conference on Electronics and Nanotechnology (ELNANO), Kyiv, Ukraine: IEEE, Apr. 2020, pp. 478–483. doi: 10.1109/ELNANO50318.2020.9088819.

44. F. Li, S. Krivenko, and V. Lukin, “A Two-step Procedure for Image Lossy Compression by ADCTC With a Desired Quality,” in *2020 IEEE 11th International Conference on Dependable Systems, Services and Technologies (DESSERT)*, Kyiv, Ukraine: IEEE, May 2020, pp. 307–312. doi: 10.1109/DESSERT50317.2020.9125000.

45. F. Li, S. S. Krivenko, and V. V. Lukin, “Analysis of two-step approach for compressing texture images with desired quality,” *AKTT*, no. 1, pp. 50–58, Jan. 2020, doi: 10.32620/aktt.2020.1.08.

46. V. Lukin *et al.*, “Lossy Compression of Multichannel Remote Sensing Images with Quality Control,” *Remote Sensing*, vol. 12, no. 22, p. 3840, Nov. 2020, doi: 10.3390/rs12223840.

47. F. Li, S. Krivenko, and V. Lukin, “Adaptive two-step procedure of providing desired visual quality of compressed image,” in *Proceedings of the 2020 4th International Conference on Electronic Information Technology and Computer Engineering*, Xiamen China: ACM, Nov. 2020, pp. 407–414. doi: 10.1145/3443467.3443791.

48. F. Li, S. Krivenko, and V. Lukin, “An Approach to Better Portable Graphics (BPG) Compression with Providing a Desired Quality,” in *2020 IEEE 2nd International Conference on Advanced Trends in Information Theory (ATIT)*, Kyiv, Ukraine: IEEE, Nov. 2020, pp. 13–17. doi: 10.1109/ATIT50783.2020.9349289.

49. F. Li, S. Krivenko, and V. Lukin, “A Fast Method for Visual Quality Prediction and Providing in Image Lossy Compression by SPIHT,” in *Integrated Computer Technologies in Mechanical Engineering - 2020*, vol. 188, M. Nechyporuk, V. Pavlikov, and D. Kritskiy, Eds., in *Lecture Notes in Networks and Systems*, vol. 188. , Cham: Springer International Publishing, 2021, pp. 17–29. doi: 10.1007/978-3-030-66717-7_2.

50. F. Li, S. Krivenko, and V. Lukin, “Two-step providing of desired quality in lossy image compression by SPIHT,” *Радіоелектронні і комп’ютерні системи*, no. 2, pp. 22–32, Apr. 2020, doi: 10.32620/reks.2020.2.02.

51. F. Li, S. Krivenko, and V. Lukin, “An Automatic Optimization Method for BPG Compression Based on Visual Perception,” in *Mathematical Modeling and*

Simulation of Systems, vol. 344, S. Shkarlet, A. Morozov, A. Palagin, D. Vinnikov, N. Stoianov, M. Zhelezniak, and V. Kazymyr, Eds., in *Lecture Notes in Networks and Systems*, vol. 344. , Cham: Springer International Publishing, 2022, pp. 213–225. doi: 10.1007/978-3-030-89902-8_17.

52. A. Kharkov, V. Oliynyk, V. V. Lukin, and S. S. Krivenko, “Blind estimation of noise variance for 1d signal denoising,” *Telecom Rad Eng*, vol. 79, no. 7, pp. 567–581, 2020, doi: 10.1615/TelecomRadEng.v79.i7.30.

53. V. Abramova, S. Krivenko, V. Lukin, and O. Krylova, “Analysis of Noise Properties in Dental Images,” in *2020 IEEE 40th International Conference on Electronics and Nanotechnology (ELNANO)*, Kyiv, Ukraine: IEEE, Apr. 2020, pp. 511–515. doi: 10.1109/ELNANO50318.2020.9088768.

54. M. Ponomarenko, O. Miroshnichenko, V. Lukin, S. Krivenko, and K. Egiazarian, “Blind denoising of dental X-ray images,” *ei*, vol. 35, no. 9, pp. 299-1-299–6, Jan. 2023, doi: 10.2352/EI.2023.35.9.IPAS-299.

55. V. Abramova, V. Lukin, S. Abramov, S. Kryvenko, P. Lech, and K. Okarma, “A Fast and Accurate Prediction of Distortions in DCT-Based Lossy Image Compression,” *Electronics*, vol. 12, no. 11, p. 2347, May 2023, doi: 10.3390/electronics12112347.

56. V. Lukin, S. Krivenko, I. Kaluzhinov, O. Krylova, and L. Kryvenko, “Lossy Compression of Remote Sensing and Dental Images Corrupted by Spatially Correlated Noise,” in *Integrated Computer Technologies in Mechanical Engineering - 2021*, vol. 367, M. Nechyporuk, V. Pavlikov, and D. Kritskiy, Eds., in *Lecture Notes in Networks and Systems*, vol. 367. , Cham: Springer International Publishing, 2022, pp. 1003–1014. doi: 10.1007/978-3-030-94259-5_77.

57. G. Proskura, O. Rubel, S. Kryvenko, and V. Lukin, “On classifier performance for remote sensing images compressed by different coders,” *aktt*, no. 1, pp. 67–77, Feb. 2023, doi: 10.32620/aktt.2023.1.07.

58. V. Lukin, S. Abramov, S. Krivenko, A. Kurekin, and O. Pogrebnyak, “Analysis of classification accuracy for pre-filtered multichannel remote sensing data,” *Expert*

Systems with Applications, vol. 40, no. 16, pp. 6400–6411, Nov. 2013, doi: 10.1016/j.eswa.2013.05.061.

59. V. Lukin, F. Li, G. Proskura, S. Kryvenko, and B. Vozel, “Peculiarities of Classification of Lossy Compressed Multichannel Remote Sensing Images Using Trained Neural Networks,” in *Digital Ecosystems: Interconnecting Advanced Networks with AI Applications*, vol. 1198, A. Luntovskyy, M. Klymash, I. Melnyk, M. Beshley, and A. Schill, Eds., in *Lecture Notes in Electrical Engineering*, vol. 1198. , Cham: Springer Nature Switzerland, 2024, pp. 127–143. doi: 10.1007/978-3-031-61221-3_7.

60. V. Lukin, F. Li, J. Zhu, and S. Kryvenko, “Peculiarities of SVM-Based Classification of BPG Compressed Three-Channel Images,” in *Proceedings of the 8th International Symposium of Space Optical Instruments and Applications*, vol. 191, H. P. Urbach, D. Li, and D. Yu, Eds., in *Springer Proceedings in Physics*, vol. 191. , Singapore: Springer Nature Singapore, 2024, pp. 352–365. doi: 10.1007/978-981-97-6718-2_31.

61. L. Kryvenko, O. Krylova, V. Lukin, and S. Kryvenko, “Intelligent visually lossless compression of dental images,” *Adv. Opt. Technol.*, vol. 13, p. 1306142, Feb. 2024, doi: 10.3389/aot.2024.1306142.

62. L. Kryvenko, O. Krylova, S. Krivenko, and V. Lukin, “A Fast Noniterative Visually Lossless Compression of Dental Images Using BPG Coder,” in *2022 11th Mediterranean Conference on Embedded Computing (MECO)*, Budva, Montenegro: IEEE, Jun. 2022, pp. 1–6. doi: 10.1109/MECO55406.2022.9797114.

63. B. Bondžulić, V. Lukin, D. Bujaković, F. Li, S. Kryvenko, and B. Pavlović, “On Visually Lossless JPEG Image Compression,” in *2023 Zooming Innovation in Consumer Technologies Conference (ZINC)*, Novi Sad, Serbia: IEEE, May 2023, pp. 113–118. doi: 10.1109/ZINC58345.2023.10174090.

64. V. Abramova, S. Kryvenko, S. Abramov, V. Makarichev, and V. Lukin, “A Fast Procedure for Image Lossy Compression by ADCTC Using Prediction of Distortions’ MSE,” in *Integrated Computer Technologies in Mechanical Engineering - 2022*, vol. 657, M. Nechyporuk, V. Pavlikov, and D. Kritskiy, Eds., in *Lecture Notes in Networks and Systems*, vol. 657. , Cham: Springer Nature Switzerland, 2023, pp. 625–635. doi: 10.1007/978-3-031-36201-9_52.

65. F. Li, S. Kryvenko, and V. Lukin, “Lossy Compression of Three-Channel Remote Sensing Images Based on Controlled Visual Quality,” in *Integrated Computer Technologies in Mechanical Engineering - 2022*, vol. 657, M. Nechyporuk, V. Pavlikov, and D. Kritskiy, Eds., in Lecture Notes in Networks and Systems, vol. 657. , Cham: Springer Nature Switzerland, 2023, pp. 649–661. doi: 10.1007/978-3-031-36201-9_54.

66. V. Lukin, S. Krivenko, F. Li, S. Abramov, and V. Makarichev, “On Image Complexity in Viewpoint of Image Processing Performance,” in *2022 3rd International Workshop on Intelligent Information Technologies and Systems of Information Security*, 2022, pp. 458–473. Accessed: Feb. 12, 2024. [Online]. Available: <https://ceur-ws.org/Vol-3156/paper35.pdf>

67. B. Bondžulić, N. Stojanović, V. Lukin, S. A. Stankevich, D. Bujaković, and S. Kryvenko, “Target acquisition performance in the presence of JPEG image compression,” *Defence Technology*, p. S2214914723003306, Dec. 2023, doi: 10.1016/j.dt.2023.12.006.

68. F. Li, S. Kryvenko, and V. Lukin, “Remote Sensing Image Lossy Compression Based on JPEG with Controlled Visual Quality,” in *Proceedings of the 7th International Symposium of Space Optical Instruments and Applications*, vol. 295, H. P. Urbach and H. Jiang, Eds., in Springer Proceedings in Physics, vol. 295. , Singapore: Springer Nature Singapore, 2023, pp. 8–19. doi: 10.1007/978-981-99-4098-1_2.

69. S. Kryvenko and V. Lukin, “Peculiarities of noisy image lossy compression,” *Вісник Хмельницького національного університету. Серія: Технічні науки*, vol. 333, no. 2, pp. 278–283, Apr. 2024, doi: 10.31891/2307-5732-2024-333-2-44.

70. B. Bondžulić, N. Stojanović, V. Lukin, and S. Kryvenko, “JPEG and BPG visually lossless image compression via KonJND-1k database,” *Vojnotehnički glasnik*, vol. 72, no. 3, pp. 1214–1241, 2024, doi: 10.5937/vojtehg72-50300.

71. F. Li, V. Lukin, S. Kryvenko, B. Bondzulich, D. Bujakovic, and B. Pavlovic, “Strange Images in Remote Sensing and Their Properties,” *UJRS*, vol. 10, no. 2, pp. 12–18, Jun. 2023, doi: 10.36023/ujrs.2023.10.2.240.

72. M. Partio, E. Guldogan, O. Guldogan, and M. Gabbouj, “Applying texture and color features to natural image retrieval,” in *2003 Finnish Signal Processing Symposium*,

2003, pp. 199–203. Accessed: Apr. 04, 2024. [Online]. Available: <https://www.academia.edu/download/32042510/applying-texture.pdf>

73. C. Oliver and S. Quegan, *Understanding synthetic aperture radar images*. SciTech Publishing, 2004. Accessed: Apr. 11, 2024. [Online]. Available: <https://www.google.com/books?hl=uk&lr=&id=IeGKe40S77AC&oi=fnd&pg=PR17&dq=Understanding+Synthetic+Aperture+Radar+Images&ots=3fCZo39u8K&sig=iBrD8p mXuM-UGT-rHU7E8mnFGsA>

74. B. Goyal, A. Dogra, S. Agrawal, and B. S. Sohi, “Noise Issues Prevailing in Various Types of Medical Images,” *Biomed. Pharmacol. J.*, vol. 11, no. 3, pp. 1227–1237, Sep. 2018, doi: 10.13005/bpj/1484.

75. J.-S. Lee, “Digital image smoothing and the sigma filter,” *Computer Vision, Graphics, and Image Processing*, vol. 24, no. 2, pp. 255–269, Nov. 1983, doi: 10.1016/0734-189X(83)90047-6.

76. Z. Sun, Z. Zhang, Y. Chen, S. Liu, and Y. Song, “Frost Filtering Algorithm of SAR Images With Adaptive Windowing and Adaptive Tuning Factor,” *IEEE Geosci. Remote Sensing Lett.*, vol. 17, no. 6, pp. 1097–1101, Jun. 2020, doi: 10.1109/LGRS.2019.2939208.

77. J. Astola and P. Kuosmanen, *Fundamentals of nonlinear digital filtering*, 1st ed. Boca Raton: CRC Press, 2020.

78. K. Dabov, A. Foi, V. Katkovnik, and K. Egiazarian, “Image Denoising by Sparse 3-D Transform-Domain Collaborative Filtering,” *IEEE Trans. on Image Process.*, vol. 16, no. 8, pp. 2080–2095, Aug. 2007, doi: 10.1109/TIP.2007.901238.

79. A. Rubel, V. Lukin, and O. Pogrebnyak, “Efficiency of DCT-Based Denoising Techniques Applied to Texture Images,” in *Pattern Recognition*, vol. 8495, J. F. Martínez-Trinidad, J. A. Carrasco-Ochoa, J. A. Olvera-Lopez, J. Salas-Rodríguez, and C. Y. Suen, Eds., in Lecture Notes in Computer Science, vol. 8495. , Cham: Springer International Publishing, 2014, pp. 261–270. doi: 10.1007/978-3-319-07491-7_27.

80. L. Fan, F. Zhang, H. Fan, and C. Zhang, “Brief review of image denoising techniques,” *Vis. Comput. Ind. Biomed. Art*, vol. 2, no. 1, p. 7, Dec. 2019, doi: 10.1186/s42492-019-0016-7.

81. S. Doss, S. Pal, D. Akila, S. Jeyalakshmi, T. N. Jabeen, and G. Suseendran, "Satellite image remote sensing for identifying aircraft using SPIHT and NSCT," *Journal of Critical Reviews*, vol. 7, no. 5, pp. 631–634, 2020.
82. J. Wei, L. Mi, Y. Hu, J. Ling, Y. Li, and Z. Chen, "Effects of Lossy Compression on Remote Sensing Image Classification Based on Convolutional Sparse Coding," *IEEE Geosci. Remote Sensing Lett.*, vol. 19, pp. 1–5, 2022, doi: 10.1109/LGRS.2020.3047789.
83. P. Muruganantham, S. Wibowo, S. Grandhi, N. H. Samrat, and N. Islam, "A Systematic Literature Review on Crop Yield Prediction with Deep Learning and Remote Sensing," *Remote Sensing*, vol. 14, no. 9, p. 1990, Apr. 2022, doi: 10.3390/rs14091990.
84. X. Wang *et al.*, "Multi-level uncertainty aware learning for semi-supervised dental panoramic caries segmentation," *Neurocomputing*, vol. 540, p. 126208, Jul. 2023, doi: 10.1016/j.neucom.2023.03.069.
85. N. Yokoya and A. Iwasaki, "Airborne hyperspectral data over Chikusei," *Space Appl. Lab., Univ. Tokyo, Tokyo, Japan, Tech. Rep. SAL-2016-05-27*, vol. 5, no. 5, p. 5, 2016.
86. E. Magli, "The JPEG family of coding standards," *Document and image compression*, pp. 87–112, 2006.
87. D. S. Taubman and M. W. Marcellin, *JPEG2000: image compression fundamentals, standards, and practice*. 2013.
88. G. Ginesu, M. Pintus, and D. D. Giusto, "Objective assessment of the WebP image coding algorithm," *Signal Processing: Image Communication*, vol. 27, no. 8, pp. 867–874, Sep. 2012, doi: 10.1016/j.image.2012.01.011.
89. A. Indradjad, A. S. Nasution, H. Gunawan, and A. Widipaminto, "A comparison of Satellite Image Compression methods in the Wavelet Domain," *IOP Conf. Ser.: Earth Environ. Sci.*, vol. 280, no. 1, p. 012031, Jun. 2019, doi: 10.1088/1755-1315/280/1/012031.
90. A. Altamimi and B. Ben Youssef, "Lossless and Near-Lossless Compression Algorithms for Remotely Sensed Hyperspectral Images," *Entropy*, vol. 26, no. 4, p. 316, Apr. 2024, doi: 10.3390/e26040316.

91. T. Fawcett, "An introduction to ROC analysis," *Pattern Recognition Letters*, vol. 27, no. 8, pp. 861–874, Jun. 2006, doi: 10.1016/j.patrec.2005.10.010.
92. A. Colin Cameron and F. A. G. Windmeijer, "An R-squared measure of goodness of fit for some common nonlinear regression models," *Journal of Econometrics*, vol. 77, no. 2, pp. 329–342, Apr. 1997, doi: 10.1016/S0304-4076(96)01818-0.
93. K. Egiazarian, J. Astola, V. Lukin, F. Battisti, and M. Carli, "A new full-reference quality metrics based on hvs," Jan. 2006.
94. N. Ponomarenko, F. Silvestri, K. Egiazarian, M. Carli, J. Astola, and V. Lukin, "On between-coefficient contrast masking of dct basis functions," p. 4.
95. Z. Wang, A. C. Bovik, H. R. Sheikh, and E. P. Simoncelli, "Image Quality Assessment: From Error Visibility to Structural Similarity," *IEEE Trans. on Image Process.*, vol. 13, no. 4, pp. 600–612, Apr. 2004, doi: 10.1109/TIP.2003.819861.
96. Z. Wang, E. P. Simoncelli, and A. C. Bovik, "Multiscale structural similarity for image quality assessment," in *The Thrity-Seventh Asilomar Conference on Signals, Systems & Computers, 2003*, Pacific Grove, CA, USA: IEEE, 2003, pp. 1398–1402. doi: 10.1109/ACSSC.2003.1292216.
97. Z. Wang, L. Lu, and A. C. Bovik, "Video quality assessment based on structural distortion measurement," *Signal Processing: Image Communication*, vol. 19, no. 2, pp. 121–132, Feb. 2004, doi: 10.1016/S0923-5965(03)00076-6.
98. Lin Zhang, Lei Zhang, Xuanqin Mou, and D. Zhang, "FSIM: A Feature Similarity Index for Image Quality Assessment," *IEEE Trans. on Image Process.*, vol. 20, no. 8, pp. 2378–2386, Aug. 2011, doi: 10.1109/TIP.2011.2109730.
99. H. Ziaei Nafchi, A. Shahkolaei, R. Hedjam, and M. Cheriet, "Mean Deviation Similarity Index: Efficient and Reliable Full-Reference Image Quality Evaluator," *IEEE Access*, vol. 4, pp. 5579–5590, 2016, doi: 10.1109/ACCESS.2016.2604042.
100. O. Ieremeiev, V. Lukin, K. Okarma, K. Egiazarian, and B. Vozel, "On properties of visual quality metrics in remote sensing applications," *ei*, vol. 34, no. 10, pp. 354-1-354–6, Jan. 2022, doi: 10.2352/EI.2022.34.10.IPAS-354.

101. N. Ponomarenko *et al.*, “Image database TID2013: Peculiarities, results and perspectives,” *Signal Processing: Image Communication*, vol. 30, pp. 57–77, Jan. 2015, doi: 10.1016/j.image.2014.10.009.
102. O. Ieremeiev, V. Lukin, K. Okarma, and K. Egiazarian, “Full-Reference Quality Metric Based on Neural Network to Assess the Visual Quality of Remote Sensing Images,” *Remote Sensing*, vol. 12, no. 15, p. 2349, Jul. 2020, doi: 10.3390/rs12152349.
103. A. Buades, B. Coll, and J. M. Morel, “A Review of Image Denoising Algorithms, with a New One,” *Multiscale Model. Simul.*, vol. 4, no. 2, pp. 490–530, Jan. 2005, doi: 10.1137/040616024.
104. A. K. Boyat and B. K. Joshi, “A Review Paper: Noise Models in Digital Image Processing,” 2015, doi: 10.48550/ARXIV.1505.03489.
105. P. Jain and V. Tyagi, “A survey of edge-preserving image denoising methods,” *Inf Syst Front*, vol. 18, no. 1, pp. 159–170, Feb. 2016, doi: 10.1007/s10796-014-9527-0.
106. M. Diwakar and M. Kumar, “A review on CT image noise and its denoising,” *Biomedical Signal Processing and Control*, vol. 42, pp. 73–88, Apr. 2018, doi: 10.1016/j.bspc.2018.01.010.
107. R. O. Green *et al.*, “Imaging Spectroscopy and the Airborne Visible/Infrared Imaging Spectrometer (AVIRIS),” *Remote Sensing of Environment*, vol. 65, no. 3, pp. 227–248, Sep. 1998, doi: 10.1016/S0034-4257(98)00064-9.
108. C. Elachi and J. Van Zyl, *Introduction to the Physics and Techniques of Remote Sensing*, 3rd ed. in Wiley Series in Remote Sensing and Image Processing Ser. Hoboken: Wiley, 2021. doi: 10.1002/9781119523048.
109. B. R. Corner, R. M. Narayanan, and S. E. Reichenbach, “Noise estimation in remote sensing imagery using data masking,” *International Journal of Remote Sensing*, vol. 24, no. 4, pp. 689–702, Jan. 2003, doi: 10.1080/01431160210164271.
110. P. Singh, M. Diwakar, A. Shankar, R. Shree, and M. Kumar, “A Review on SAR Image and its Despeckling,” *Arch Computat Methods Eng*, vol. 28, no. 7, pp. 4633–4653, Dec. 2021, doi: 10.1007/s11831-021-09548-z.

111. R. Werninghaus and S. Buckreuss, "The TerraSAR-X Mission and System Design," *IEEE Trans. Geosci. Remote Sensing*, vol. 48, no. 2, pp. 606–614, Feb. 2010, doi: 10.1109/TGRS.2009.2031062.
112. C. C. Freitas, A. C. Frery, and A. H. Correia, "The polarimetric $\mathcal{G}$ distribution for SAR data analysis," *Environmetrics*, vol. 16, no. 1, pp. 13–31, Feb. 2005, doi: 10.1002/env.658.
113. N. N. Ponomarenko, V. V. Lukin, K. O. Egiazarian, and J. T. Astola, "A method for blind estimation of spatially correlated noise characteristics," presented at the IS&T/SPIE Electronic Imaging, J. T. Astola and K. O. Egiazarian, Eds., San Jose, California, Feb. 2010, p. 753208. doi: 10.1117/12.847986.
114. S. Aja-Fernández, G. Vegas-Sánchez-Ferrero, M. Martín-Fernández, and C. Alberola-López, "Automatic noise estimation in images using local statistics. Additive and multiplicative cases," *Image and Vision Computing*, vol. 27, no. 6, pp. 756–770, May 2009, doi: 10.1016/j.imavis.2008.08.002.
115. S. Abramov, M. Uss, V. Lukin, B. Vozel, K. Chehdi, and K. Egiazarian, "Enhancement of Component Images of Multispectral Data by Denoising with Reference," *Remote Sensing*, vol. 11, no. 6, p. 611, Mar. 2019, doi: 10.3390/rs11060611.
116. O. Rubel, V. Lukin, A. Rubel, and K. Egiazarian, "Selection of Lee Filter Window Size Based on Despeckling Efficiency Prediction for Sentinel SAR Images," *Remote Sensing*, vol. 13, no. 10, p. 1887, May 2021, doi: 10.3390/rs13101887.
117. G. Fracastoro, E. Magli, G. Poggi, G. Scarpa, D. Valsesia, and L. Verdoliva, "Deep Learning Methods For Synthetic Aperture Radar Image Despeckling: An Overview Of Trends And Perspectives," *IEEE Geosci. Remote Sens. Mag.*, vol. 9, no. 2, pp. 29–51, Jun. 2021, doi: 10.1109/MGRS.2021.3070956.
118. R. Lukac and K. N. Plataniotis, Eds., *Color image processing: methods and applications*. in Image processing series. Boca Raton, FL: CRC/Taylor & Francis, 2007.
119. W. K. Pratt, *Digital Image Processing: PIKS Scientific Inside*, 1st ed. Wiley, 2007. doi: 10.1002/0470097434.

120. A. C. Bovik, *Handbook of Image and Video Processing*, 2nd ed. in Communications, Networking and Multimedia Ser. Burlington: Elsevier Science & Technology, 2010.
121. S. Jia, Z. Ji, Y. Qian, and L. Shen, "Unsupervised Band Selection for Hyperspectral Imagery Classification Without Manual Band Removal," *IEEE J. Sel. Top. Appl. Earth Observations Remote Sensing*, vol. 5, no. 2, pp. 531–543, Apr. 2012, doi: 10.1109/JSTARS.2012.2187434.
122. A. Foi, *Pointwise shape-adaptive DCT image filtering and signal-dependent noise estimation*. Tampere University of Technology, 2007. Accessed: Sep. 03, 2024. [Online]. Available: <https://trepo.tuni.fi/handle/10024/113882>
123. D. V. Fevraleov, N. N. Ponomarenko, V. V. Lukin, S. K. Abramov, K. O. Egiazarian, and J. T. Astola, "Efficiency analysis of color image filtering," *EURASIP J. Adv. Signal Process.*, vol. 2011, no. 1, p. 41, Dec. 2011, doi: 10.1186/1687-6180-2011-41.
124. P. Zhong and R. Wang, "Multiple-Spectral-Band CRFs for Denoising Junk Bands of Hyperspectral Imagery," *IEEE Trans. Geosci. Remote Sensing*, vol. 51, no. 4, pp. 2260–2275, Apr. 2013, doi: 10.1109/TGRS.2012.2209656.
125. V. Zlokolica, A. Pizurica, and W. Philips, "Wavelet-Domain Video Denoising Based on Reliability Measures," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 16, no. 8, pp. 993–1007, Aug. 2006, doi: 10.1109/TCSVT.2006.879994.
126. D. Petrescu and J. Pincenti, "Quality and noise measurements in mobile phone video capture," presented at the IS&T/SPIE Electronic Imaging, D. Akopian, R. Creutzburg, C. G. M. Snoek, N. Sebe, and L. Kennedy, Eds., San Francisco Airport, California, USA, Feb. 2011, p. 788105. doi: 10.1117/12.876745.
127. M. Elad, *Sparse and redundant representations: from theory to applications in signal and image processing*. New York: Springer, 2010.
128. P. Chatterjee and P. Milanfar, "Is denoising dead?," *IEEE Trans Image Process*, vol. 19, no. 4, pp. 895–911, Apr. 2010, doi: 10.1109/TIP.2009.2037087.

129. L. Sendur and I. W. Selesnick, "Bivariate shrinkage with local variance estimation," *IEEE Signal Process. Lett.*, vol. 9, no. 12, pp. 438–441, Dec. 2002, doi: 10.1109/LSP.2002.806054.
130. V. V. Lukin, R. Oktem, N. N. Ponomarenko, and K. O. Egiazarian, "Image Filtering Based on Discrete Cosine Transform," *Telecom Rad Eng*, vol. 66, no. 18, pp. 1685–1701, 2007, doi: 10.1615/TelecomRadEng.v66.i18.70.
131. U. Erkan, D. N. H. Thanh, L. M. Hieu, and S. Enginoglu, "An Iterative Mean Filter for Image Denoising," *IEEE Access*, vol. 7, pp. 167847–167859, 2019, doi: 10.1109/ACCESS.2019.2953924.
132. Tao Chen, Kai-Kuang Ma, and Li-Hui Chen, "Tri-state median filter for image denoising," *IEEE Trans. on Image Process.*, vol. 8, no. 12, pp. 1834–1838, Dec. 1999, doi: 10.1109/83.806630.
133. H. Zhang, A. Nosratinia, and R. O. Wells, "Image denoising via wavelet-domain spatially adaptive FIR Wiener filtering," in *2000 IEEE International Conference on Acoustics, Speech, and Signal Processing. Proceedings (Cat. No.00CH37100)*, Istanbul, Turkey: IEEE, 2000, pp. 2179–2182. doi: 10.1109/ICASSP.2000.859269.
134. S. Banerjee, S. S. Chaudhuri, R. Mehra, and A. Misra, "A Comprehensive Survey on Frost Filter and its Proposed Variants," in *2020 5th International Conference on Communication and Electronics Systems (ICCES)*, Coimbatore, India: IEEE, Jun. 2020, pp. 109–114. doi: 10.1109/ICCES48766.2020.9137869.
135. B. Rasti, Y. Chang, E. Dalsasso, L. Denis, and P. Ghamisi, "Image Restoration for Remote Sensing: Overview and toolbox," *IEEE Geosci. Remote Sens. Mag.*, vol. 10, no. 2, pp. 201–230, Jun. 2022, doi: 10.1109/MGRS.2021.3121761.
136. J.-S. Lee, "Speckle analysis and smoothing of synthetic aperture radar images," *Computer Graphics and Image Processing*, vol. 17, no. 1, pp. 24–32, Sep. 1981, doi: 10.1016/S0146-664X(81)80005-6.
137. V. S. Frost, J. A. Stiles, K. S. Shanmugan, and J. C. Holtzman, "A Model for Radar Images and Its Application to Adaptive Digital Filtering of Multiplicative Noise," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. PAMI-4, no. 2, pp. 157–166, Mar. 1982, doi: 10.1109/TPAMI.1982.4767223.

138. D. Kuan, A. Sawchuk, T. Strand, and P. Chavel, "Adaptive restoration of images with speckle," *IEEE Trans. Acoust., Speech, Signal Process.*, vol. 35, no. 3, pp. 373–383, Mar. 1987, doi: 10.1109/TASSP.1987.1165131.

139. V. P. Melnik, "Local activity indicators for hard-switching adaptive filtering of images with mixed noise," *Opt. Eng.*, vol. 40, no. 8, p. 1441, Aug. 2001, doi: 10.1117/1.1385815.

140. A. Achim, P. Tsakalides, and A. Bezerianos, "SAR image denoising via Bayesian wavelet shrinkage based on heavy-tailed modeling," *IEEE Trans. Geosci. Remote Sensing*, vol. 41, no. 8, pp. 1773–1784, Aug. 2003, doi: 10.1109/TGRS.2003.813488.

141. R. Öktem, K. Egiazarian, V. V. Lukin, N. N. Ponomarenko, and O. V. Tsymbal, "Locally Adaptive DCT Filtering for Signal-Dependent Noise Removal," *EURASIP J. Adv. Signal Process.*, vol. 2007, no. 1, p. 042472, Dec. 2007, doi: 10.1155/2007/42472.

142. R. C. P. Marques, F. N. S. De Medeiros, and D. M. Ushizima, "Target Detection in SAR Images Based on a Level Set Approach," *IEEE Trans. Syst., Man, Cybern. C*, vol. 39, no. 2, pp. 214–222, Mar. 2009, doi: 10.1109/TSMCC.2008.2006685.

143. V. Lukin, S. Abramov, N. Ponomarenko, K. Egiazarian, and J. Astola, "Image filtering: Potential efficiency and current problems," in *2011 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, Prague, Czech Republic: IEEE, May 2011, pp. 1433–1436. doi: 10.1109/ICASSP.2011.5946683.

144. M. L. Uss, B. Vozel, V. V. Lukin, and K. Chehdi, "Local Signal-Dependent Noise Variance Estimation From Hyperspectral Textural Images," *IEEE J. Sel. Top. Signal Process.*, vol. 5, no. 3, pp. 469–486, Jun. 2011, doi: 10.1109/JSTSP.2010.2104312.

145. N. Acito, M. Diani, and G. Corsini, "Signal-Dependent Noise Modeling and Model Parameter Estimation in Hyperspectral Images," *IEEE Trans. Geosci. Remote Sensing*, vol. 49, no. 8, pp. 2957–2971, Aug. 2011, doi: 10.1109/TGRS.2011.2110657.

146. S. Abramov *et al.*, “Methods for Blind Estimation of Speckle Variance in SAR Images: Simulation Results and Verification for Real-Life Data,” in *Computational and Numerical Simulations*, J. Awrejcewicz, Ed., InTech, 2014. doi: 10.5772/57040.
147. M. Colom, A. Buades, and J.-M. Morel, “Nonparametric noise estimation method for raw images,” *J. Opt. Soc. Am. A*, vol. 31, no. 4, p. 863, Apr. 2014, doi: 10.1364/JOSAA.31.000863.
148. K. Zhang, Y. Li, W. Zuo, L. Zhang, L. Van Gool, and R. Timofte, “Plug-and-Play Image Restoration With Deep Denoiser Prior,” *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 44, no. 10, pp. 6360–6376, Oct. 2022, doi: 10.1109/TPAMI.2021.3088914.
149. Z. Yue, H. Yong, Q. Zhao, D. Meng, and L. Zhang, “Variational Denoising Network: Toward Blind Noise Modeling and Removal,” in *Advances in Neural Information Processing Systems*, H. Wallach, H. Larochelle, A. Beygelzimer, F. d’Alché-Buc, E. Fox, and R. Garnett, Eds., Curran Associates, Inc., 2019. [Online]. Available: https://proceedings.neurips.cc/paper_files/paper/2019/file/6395ebd0f4b478145ecfbaf939454fa4-Paper.pdf
150. S. W. Zamir, A. Arora, S. Khan, M. Hayat, F. S. Khan, and M.-H. Yang, “Restormer: Efficient Transformer for High-Resolution Image Restoration,” in *2022 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, New Orleans, LA, USA: IEEE, Jun. 2022, pp. 5718–5729. doi: 10.1109/CVPR52688.2022.00564.
151. V. V. Lukin, “Methods and automatic procedures for processing images based on blind evaluation of noise type and characteristics,” *J. Appl. Remote Sens.*, vol. 5, no. 1, p. 053502, Jan. 2011, doi: 10.1117/1.3539768.
152. J. Hegde, “Efficient Denoising Method to Improve The Resolution of Satellite Images,” 2024, *arXiv*. doi: 10.48550/ARXIV.2411.10476.
153. I. Blanes, E. Magli, and J. Serra-Sagrista, “A Tutorial on Image Compression for Optical Space Imaging Systems,” *IEEE Geosci. Remote Sens. Mag.*, vol. 2, no. 3, pp. 8–26, Sep. 2014, doi: 10.1109/MGRS.2014.2352465.

154. G. Yu, T. Vladimirova, and M. N. Sweeting, "Image compression systems on board satellites," *Acta Astronautica*, vol. 64, no. 9–10, pp. 988–1005, May 2009, doi: 10.1016/j.actaastro.2008.12.006.
155. B. Aiazzi, L. Alparone, S. Baronti, C. Lastri, and M. Selva, "Spectral Distortion in Lossy Compression of Hyperspectral Data," *Journal of Electrical and Computer Engineering*, vol. 2012, no. 1, p. 850637, Jan. 2012, doi: 10.1155/2012/850637.
156. S. K. Abramov, M. L. Uss, V. V. Abramova, V. V. Lukin, B. Vozel, and K. Chehdi, "On noise properties in hyperspectral images," in *2015 IEEE International Geoscience and Remote Sensing Symposium (IGARSS)*, Milan, Italy: IEEE, Jul. 2015, pp. 3501–3504. doi: 10.1109/IGARSS.2015.7326575.
157. A. Mahmood and M. Sears, "Estimation of correlated signal-dependent noise statistics in hyperspectral images," *Remote Sensing Letters*, vol. 12, no. 10, pp. 961–969, Oct. 2021, doi: 10.1080/2150704X.2021.1946616.
158. M. L. Uss, B. Vozel, V. V. Lukin, and K. Chehdi, "Image informative maps for component-wise estimating parameters of signal-dependent noise," *J. Electron. Imaging*, vol. 22, no. 1, p. 013019, Feb. 2013, doi: 10.1117/1.JEI.22.1.013019.
159. V. V. Lukin, S. K. Abramov, R. A. Kozhemiakin, B. Vozel, B. Djurovic, and I. Djurovic, "Optimal operation point in 3d dct- based lossy compression of color and multichannel remote sensing images," *Telecom Rad Eng*, vol. 74, no. 20, pp. 1803–1821, 2015, doi: 10.1615/TelecomRadEng.v74.i20.40.
160. V. Lukin and E. Bataeva, "Challenges in pre-processing multichannel remote sensing terrain images," *Importance of GEO Initiatives and Montenegrin Capacities in this Area. The Section for Natural Sciences Book*, no. 16, pp. 63–76, 2012.
161. M. A. Popov, S. A. Stankevich, L. P. Lischenko, V. V. Lukin, and N. N. Ponomarenko, "Processing of Hyperspectral Imagery for Contamination Detection in Urban Areas," in *Environmental Security and Ecoterrorism*, H. Alpas, S. M. Berkowicz, and I. Ermakova, Eds., in NATO Science for Peace and Security Series C: Environmental Security. , Dordrecht: Springer Netherlands, 2011, pp. 147–156. doi: 10.1007/978-94-007-1235-5_12.

162. C. Fu and B. Du, "Remote Sensing Image Compression Based on the Multiple Prior Information," *Remote Sensing*, vol. 15, no. 8, p. 2211, Apr. 2023, doi: 10.3390/rs15082211.
163. V. Alves De Oliveira *et al.*, "Satellite Image Compression and Denoising With Neural Networks," *IEEE Geosci. Remote Sensing Lett.*, vol. 19, pp. 1–5, 2022, doi: 10.1109/LGRS.2022.3145992.
164. E. Christophe, "Hyperspectral Data Compression Tradeoff," in *Optical Remote Sensing*, S. Prasad, L. M. Bruce, and J. Chanussot, Eds., Berlin, Heidelberg: Springer Berlin Heidelberg, 2011, pp. 9–29. doi: 10.1007/978-3-642-14212-3_2.
165. A. N. Zemliachenko *et al.*, "Lossy compression of hyperspectral images based on noise parameters estimation and variance stabilizing transform," *J. Appl. Remote Sens.*, vol. 8, no. 1, p. 083571, Aug. 2014, doi: 10.1117/1.JRS.8.083571.
166. A. N. Zemliachenko *et al.*, "Prediction of Compression Ratio for DCT-Based Coders With Application to Remote Sensing Images," *IEEE J. Sel. Top. Appl. Earth Observations Remote Sensing*, vol. 11, no. 1, pp. 257–270, Jan. 2018, doi: 10.1109/JSTARS.2017.2781906.
167. F. Khelifi, A. Bouridane, and F. Kurugollu, "Joined Spectral Trees for Scalable SPIHT-Based Multispectral Image Compression," *IEEE Trans. Multimedia*, vol. 10, no. 3, pp. 316–329, Apr. 2008, doi: 10.1109/TMM.2008.917357.
168. A. Zemliachenko, V. Lukin, N. Ponomarenko, K. Egiazarian, and J. Astola, "Still image/video frame lossy compression providing a desired visual quality," *Multidim Syst Sign Process*, vol. 27, no. 3, pp. 697–718, Jul. 2016, doi: 10.1007/s11045-015-0333-8.
169. C. Shi, L. Wang, J. Zhang, F. Miao, and P. He, "Remote Sensing Image Compression Based on Direction Lifting-Based Block Transform with Content-Driven Quadtree Coding Adaptively," *Remote Sensing*, vol. 10, no. 7, p. 999, Jun. 2018, doi: 10.3390/rs10070999.
170. Z. Chen, Y. Hu, and Y. Zhang, "Effects of Compression on Remote Sensing Image Classification Based on Fractal Analysis," *IEEE Trans. Geosci. Remote Sensing*, vol. 57, no. 7, pp. 4577–4590, Jul. 2019, doi: 10.1109/TGRS.2019.2891679.

171. J. Garcia-Sobrino, V. Laparra, J. Serra-Sagrista, X. Calbet, and G. Camps-Valls, “Improved Statistically Based Retrievals via Spatial-Spectral Data Compression for IASI Data,” *IEEE Trans. Geosci. Remote Sensing*, vol. 57, no. 8, pp. 5651–5668, Aug. 2019, doi: 10.1109/TGRS.2019.2901396.

172. V. Makarichev, I. Vasilyeva, V. Lukin, B. Vozel, A. Shelestov, and N. Kussul, “Discrete Atomic Transform-Based Lossy Compression of Three-Channel Remote Sensing Images with Quality Control,” *Remote Sensing*, vol. 14, no. 1, p. 125, Dec. 2021, doi: 10.3390/rs14010125.

173. G. K. Wallace, “The JPEG still picture compression standard,” *IEEE Trans. Consumer Electron.*, vol. 38, no. 1, pp. xviii–xxxiv, Feb. 1992, doi: 10.1109/30.125072.

174. Beong-Jo Kim and W. A. Pearlman, “An embedded wavelet video coder using three-dimensional set partitioning in hierarchical trees (SPIHT),” in *Proceedings DCC '97. Data Compression Conference*, Snowbird, UT, USA: IEEE Comput. Soc. Press, 1997, pp. 251–260. doi: 10.1109/DCC.1997.582048.

175. N. Ponomarenko, V. Lukin, K. Egiazarian, and J. Astola, “DCT Based High Quality Image Compression,” in *Image Analysis*, vol. 3540, H. Kalviainen, J. Parkkinen, and A. Kaarna, Eds., in *Lecture Notes in Computer Science*, vol. 3540. , Berlin, Heidelberg: Springer Berlin Heidelberg, 2005, pp. 1177–1185. doi: 10.1007/11499145_119.

176. N. Ponomarenko, V. Lukin, K. Egiazarian, and J. Astola, “ADCTC: advanced DCT-based coder,” in *Proceedings of the 2008 International Workshop on Local and Non-Local Approximation in Image Processing, LNLA2008*, Lausanne, Switzerland, 23-24 August 2008, 2008, p. 6. Accessed: Aug. 02, 2024. [Online]. Available: <https://researchportal.tuni.fi/en/publications/adctc-advanced-dct-based-coder>

177. “BPG Image format.” Accessed: May 05, 2022. [Online]. Available: <https://bellard.org/bpg/>

178. U. Albalawi, S. P. Mohanty, and E. Kougianos, “Energy-Efficient Design of the Secure Better Portable Graphics Compression Architecture for Trusted Image Communication in the IoT,” in *2016 IEEE Computer Society Annual Symposium on VLSI*

(ISVLSI), Pittsburgh, PA, USA: IEEE, Jul. 2016, pp. 302–307. doi: 10.1109/ISVLSI.2016.21.

179. G. Pastuszak and A. Abramowski, “Algorithm and Architecture Design of the H.265/HEVC Intra Encoder,” *IEEE Trans. Circuits Syst. Video Technol.*, vol. 26, no. 1, pp. 210–222, Jan. 2016, doi: 10.1109/TCSVT.2015.2428571.

180. J. Lainema, M. M. Hannuksela, V. K. M. Vadakital, and E. B. Aksu, “HEVC still image coding and high efficiency image file format,” in *2016 IEEE International Conference on Image Processing (ICIP)*, Phoenix, AZ, USA: IEEE, Sep. 2016, pp. 71–75. doi: 10.1109/ICIP.2016.7532321.

181. N. Barman and M. G. Martini, “An Evaluation of the Next-Generation Image Coding Standard AVIF,” in *2020 Twelfth International Conference on Quality of Multimedia Experience (QoMEX)*, Athlone, Ireland: IEEE, May 2020, pp. 1–4. doi: 10.1109/QoMEX48832.2020.9123131.

182. V. V. Lukin, M. S. Zriakhov, N. N. Ponomarenko, S. S. Krivenko, and M. Zhenjiang, “Lossy compression of images without visible distortions and its application,” in *IEEE 10th INTERNATIONAL CONFERENCE ON SIGNAL PROCESSING PROCEEDINGS*, Beijing, China: IEEE, Oct. 2010, pp. 698–701. doi: 10.1109/ICOSP.2010.5655751.

183. N. N. Ponomarenko, K. O. Egiazarian, V. V. Lukin, and J. T. Astola, “High-Quality DCT-Based Image Compression Using Partition Schemes,” *IEEE Signal Process. Lett.*, vol. 14, no. 2, pp. 105–108, Feb. 2007, doi: 10.1109/LSP.2006.879861.

184. N. De Lange, “Remote Sensing and Digital Image Processing,” in *Geoinformatics in Theory and Practice*, in Springer Textbooks in Earth Sciences, Geography and Environment. , Berlin, Heidelberg: Springer Berlin Heidelberg, 2023, pp. 435–510. doi: 10.1007/978-3-662-65758-4_10.

185. E. Christophe, D. Leger, and C. Mailhes, “Quality criteria benchmark for hyperspectral imagery,” *IEEE Trans. Geosci. Remote Sensing*, vol. 43, no. 9, pp. 2103–2114, Sep. 2005, doi: 10.1109/TGRS.2005.853931.

186. A. N. Zemliachenko, S. K. Abramov, V. V. Lukin, B. Vozel, and K. Chehdi, “Lossy compression of noisy remote sensing images with prediction of optimal operation

point existence and parameters,” *J. Appl. Remote Sens.*, vol. 9, no. 1, p. 095066, Aug. 2015, doi: 10.1117/1.JRS.9.095066.

187. E. Christophe, C. Mailhes, and P. Duhamel, “Hyperspectral Image Compression: Adapting SPIHT and EZW to Anisotropic 3-D Wavelet Coding,” *IEEE Trans. on Image Process.*, vol. 17, no. 12, pp. 2334–2346, Dec. 2008, doi: 10.1109/TIP.2008.2005824.

188. S. Thayammal and D. Selvathi, “Multispectral band image compression using adaptive wavelet transform - Tetrolet transform,” in *2014 International Conference on Electronics and Communication Systems (ICECS)*, Coimbatore: IEEE, Feb. 2014, pp. 1–5. doi: 10.1109/ECS.2014.6892610.

189. N. Ponomarenko, V. Lukin, M. Zriakhov, and A. Kaarna, “Improved Grouping and Noise Cancellation for Automatic Lossy Compression of AVIRIS Images,” in *Advanced Concepts for Intelligent Vision Systems*, vol. 6475, J. Blanc-Talon, D. Bone, W. Philips, D. Popescu, and P. Scheunders, Eds., in *Lecture Notes in Computer Science*, vol. 6475, Berlin, Heidelberg: Springer Berlin Heidelberg, 2010, pp. 261–271. doi: 10.1007/978-3-642-17691-3_24.

190. B. Vozel, K. Chehdi, L. Klaine, V. V. Lukin, and S. K. Abramov, “Noise Identification and Estimation of its Statistical Parameters by Using Unsupervised Variational Classification,” in *2006 IEEE International Conference on Acoustics Speech and Signal Processing Proceedings*, Toulouse, France: IEEE, 2006, p. II-841-II-844. doi: 10.1109/ICASSP.2006.1660474.

191. R. Kozhemiakin, S. Abramov, V. Lukin, I. Djurovic, and B. Vozel, “Peculiarities of 3D compression of noisy multichannel images,” in *2015 4th Mediterranean Conference on Embedded Computing (MECO)*, Budva, Montenegro: IEEE, Jun. 2015, pp. 331–334. doi: 10.1109/MECO.2015.7181936.

192. B. Kovalenko, V. Lukin, and B. Vozel, “BPG-Based Lossy Compression of Three-Channel Noisy Images with Prediction of Optimal Operation Existence and Its Parameters,” *Remote Sensing*, vol. 15, no. 6, p. 1669, Mar. 2023, doi: 10.3390/rs15061669.

193. H. Jiang, K. Yang, T. Liu, and Y. Zhang, "Quality Prediction of DWT-Based Compression for Remote Sensing Image Using Multiscale and Multilevel Differences Assessment Metric," *Mathematical Problems in Engineering*, vol. 2014, pp. 1–15, 2014, doi: 10.1155/2014/593213.
194. J. Minguillo'n, "JPEG standard uniform quantization error modeling with applications to sequential and progressive operation modes," *J. Electron. Imaging*, vol. 10, no. 2, p. 475, Apr. 2001, doi: 10.1117/1.1344592.
195. A. N. Zemliachenko, S. K. Abramov, V. V. Lukin, B. Vozel, and K. Chehdi, "Improved compression ratio prediction in DCT-based lossy compression of remote sensing images," in *2016 IEEE International Geoscience and Remote Sensing Symposium (IGARSS)*, Beijing, China: IEEE, Jul. 2016, pp. 6966–6969. doi: 10.1109/IGARSS.2016.7730817.
196. J. Malik, S. Belongie, T. Leung, and J. Shi, "Contour and Texture Analysis for Image Segmentation," *International Journal of Computer Vision*, vol. 43, no. 1, pp. 7–27, 2001, doi: 10.1023/A:1011174803800.
197. T. Leung and J. Malik, "Representing and Recognizing the Visual Appearance of Materials using Three-dimensional Textons," *International Journal of Computer Vision*, vol. 43, no. 1, pp. 29–44, 2001, doi: 10.1023/A:1011126920638.
198. M. Datcu, D. Luca, and K. Seidel, "Multiresolution analysis of SAR images," in *Proc. EUSAR*, 1996, pp. 375–378. Accessed: Apr. 04, 2024. [Online]. Available: https://www.researchgate.net/profile/Klaus-Seidel/publication/224803420_Multiresolution_Analysis_of_SAR_Images/links/0deec533453ed28d55000000/Multiresolution-Analysis-of-SAR-Images.pdf
199. R. M. Haralick, K. Shanmugam, and I. Dinstein, "Textural Features for Image Classification," *IEEE Trans. Syst., Man, Cybern.*, vol. SMC-3, no. 6, pp. 610–621, Nov. 1973, doi: 10.1109/TSMC.1973.4309314.
200. B. Mičušík and A. Hanbury, "Supervised Texture Detection in Images," in *Computer Analysis of Images and Patterns*, vol. 3691, A. Gagalowicz and W. Philips, Eds., in *Lecture Notes in Computer Science*, vol. 3691. , Berlin, Heidelberg: Springer Berlin Heidelberg, 2005, pp. 441–448. doi: 10.1007/11556121_54.

201. J. C. Nunes, O. Niang, Y. Bouaoune, E. Delechelle, and Ph. Bunel, "Texture analysis based on the bidimensional empirical mode decomposition with gray-level co-occurrence models," in *Seventh International Symposium on Signal Processing and Its Applications, 2003. Proceedings.*, Paris, France: IEEE, 2003, pp. 633–635 vol.2. doi: 10.1109/ISSPA.2003.1224962.
202. V. V. Lukin, O. V. Tsymbal, N. N. Ponomarenko, K. O. Egiazarian, and J. T. Astola, "Image processing with texture feature preservation by three-state locally adaptive filter," presented at the Remote Sensing, L. Bruzzone, Ed., Barcelona, Spain, Feb. 2004, p. 120. doi: 10.1117/12.511518.
203. Y. Yoshida and Y. Wu, "Classification of rotated and scaled textured images using invariants based on spectral moments," *IEICE transactions on fundamentals of electronics, communications and computer sciences*, vol. 81, no. 8, pp. 1661–1666, 1998.
204. O. V. Tsymbal, V. V. Lukin, N. N. Ponomarenko, A. A. Zelensky, K. O. Egiazarian, and J. T. Astola, "Three-State Locally Adaptive Texture Preserving Filter for Radar and Optical Image Processing," *EURASIP J. Adv. Signal Process.*, vol. 2005, no. 8, p. 851609, Dec. 2005, doi: 10.1155/ASP.2005.1185.
205. B. Aiazzi, L. Alparone, S. Baronti, and R. Carla, "Adaptive texture-preserving filtering of multitemporal ERS-1 SAR images," in *IGARSS'97. 1997 IEEE International Geoscience and Remote Sensing Symposium Proceedings. Remote Sensing - A Scientific Vision for Sustainable Development*, Singapore: IEEE, 1997, pp. 2066–2068. doi: 10.1109/IGARSS.1997.609224.
206. K.-H. Yap, *Adaptive Image Processing: A Computational Intelligence Perspective*, 2nd ed. CRC Press, 2009. doi: 10.1201/9781315218625.
207. S. Larnier and J. Fehrenbach, "Edge detection and image restoration with anisotropic topological gradient," in *2010 IEEE International Conference on Acoustics, Speech and Signal Processing*, Dallas, TX: IEEE, Mar. 2010, pp. 1362–1365. doi: 10.1109/ICASSP.2010.5495448.
208. G. Gilboa, N. Sochen, and Y. Y. Zeevi, "Variational denoising of partly textured images by spatially varying constraints," *IEEE Trans. on Image Process.*, vol. 15, no. 8, pp. 2281–2289, Aug. 2006, doi: 10.1109/TIP.2006.875247.

209. W. Zuo, L. Zhang, C. Song, and D. Zhang, "Texture enhanced image denoising via gradient histogram preservation," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 2013, pp. 1203–1210. Accessed: Apr. 04, 2024. [Online]. Available: https://www.cv-foundation.org/openaccess/content_cvpr_2013/html/Zuo_Texture_Enhanced_Image_2013_CVPR_paper.html

210. B. Schölkopf and A. J. Smola, *Learning with kernels: support vector machines, regularization, optimization, and beyond*. MIT press, 2002. Accessed: Apr. 04, 2024. [Online]. Available: <https://www.google.com/books?hl=uk&lr=&id=y8ORL3DWt4sC&oi=fnd&pg=PR13&dq=Learning+with+Kernels+&ots=bNtS7tS5AF&sig=mpyaK50oMBSkjR31ZORvfEKLkl8>

211. X. Wu *et al.*, "Top 10 algorithms in data mining," *Knowl Inf Syst*, vol. 14, no. 1, pp. 1–37, Jan. 2008, doi: 10.1007/s10115-007-0114-2.

212. K. Karu, A. K. Jain, and R. M. Bolle, "Is there any texture in the image?," *Pattern Recognition*, vol. 29, no. 9, pp. 1437–1446, Sep. 1996, doi: 10.1016/0031-3203(96)00004-0.

213. Xin Kang, Chongzhao Han, Yi Yang, and Tangfei Tao, "SAR Image Edge Detection by Ratio-based Harris Method," in *2006 IEEE International Conference on Acoustics Speech and Signal Processing Proceedings*, Toulouse, France: IEEE, 2006, p. II-837-II–840. doi: 10.1109/ICASSP.2006.1660473.

214. J.-S. Lee, "Digital Image Enhancement and Noise Filtering by Use of Local Statistics," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. PAMI-2, no. 2, pp. 165–168, Mar. 1980, doi: 10.1109/TPAMI.1980.4766994.

215. Sangwoon Yun and Hyenkyun Woo, "A New Multiplicative Denoising Variational Model Based on m th Root Transformation," *IEEE Trans. on Image Process.*, vol. 21, no. 5, pp. 2523–2533, May 2012, doi: 10.1109/TIP.2012.2185942.

216. M. Makitalo, A. Foi, D. Fevrale, and V. Lukin, "Denoising of single-look SAR images based on variance stabilization and nonlocal filters," in *2010 International*

Conference on Mathematical Methods in Electromagnetic Theory, Kyiv, Ukraine: IEEE, Sep. 2010, pp. 1–4. doi: 10.1109/MMET.2010.5611418.

217. N. N. Ponomarenko, V. V. Lukin, and K. O. Egiazarian, “Visually lossless compression of synthetic aperture radar images,” in *2011 VIII International Conference on Antenna Theory and Techniques*, Kyiv, Ukraine: IEEE, Sep. 2011, pp. 263–265. doi: 10.1109/ICATT.2011.6170755.

218. O. Dogan and M. Kartal, “Time Domain SAR Raw Data Simulation of Distributed Targets,” *EURASIP J. Adv. Signal Process.*, vol. 2010, no. 1, p. 784815, Dec. 2010, doi: 10.1155/2010/784815.

219. A. Rubel, O. Rubel, V. Abramova, G. Proskura, and V. Lukin, “Improved Noisy Image Quality Assessment Using Multilayer Neural Networks,” in *2019 IEEE 2nd Ukraine Conference on Electrical and Computer Engineering (UKRCON)*, Lviv, Ukraine: IEEE, Jul. 2019, pp. 1046–1051. doi: 10.1109/UKRCON.2019.8879950.

220. O. Rubel, V. Lukin, and K. Egiazarian, “Additive spatially correlated noise suppression by robust block matching and adaptive 3D filtering (JIST-first),” *Electronic Imaging*, vol. 2019, no. 11, pp. 60401-1-60401–11, Jan. 2019, doi: 10.2352/J.ImagingSci.Technol.2018.62.6.060401.

221. O. Rubel, A. Rubel, V. Lukin, and K. Egiazarian, “Blind DCT-based prediction of image denoising efficiency using neural networks,” in *2018 7th European Workshop on Visual Information Processing (EUVIP)*, Tampere: IEEE, Nov. 2018, pp. 1–6. doi: 10.1109/EUVIP.2018.8611710.

222. V. V. Lukin *et al.*, “Performance evaluation for blind methods of noise characteristic estimation for TerraSAR-X images,” presented at the SPIE Remote Sensing, L. Bruzzone, Ed., Prague, Czech Republic, Oct. 2011, p. 81800X. doi: 10.1117/12.897730.

223. R. I. Saye and J. A. Sethian, “A review of level set methods to model interfaces moving under complex physics: Recent challenges and advances,” in *Handbook of Numerical Analysis*, vol. 21, Elsevier, 2020, pp. 509–554. doi: 10.1016/bs.hna.2019.07.003.

224. J.-M. Park, W. J. Song, and W. A. Pearlman, "Speckle filtering of SAR images based on adaptive windowing," *IEE Proc., Vis. Image Process.*, vol. 146, no. 4, p. 191, 1999, doi: 10.1049/ip-vis:19990550.
225. J. A. Sethian, *Level set methods and fast marching methods*, vol. 98. Cambridge Cambridge UP, 1999. Accessed: Apr. 18, 2024. [Online]. Available: <https://hrcak.srce.hr/file/69388?origin=publicationDetail>.
226. V. Lukin, N. Ponomarenko, D. Fevraleov, B. Vozel, K. Chehdi, and A. Kurekin, "Classification of pre-filtered multichannel remote sensing images," *Remote Sensing—Advanced Techniques and Platforms. Austria: InTech*, pp. 75–98, 2012.
227. N. Ponomarenko, V. Lukin, K. Egiazarian, J. Astola, M. Carli, and F. Battisti, "Color image database for evaluation of image quality metrics," in *2008 IEEE 10th Workshop on Multimedia Signal Processing*, Cairns, Qld: IEEE, Oct. 2008, pp. 403–408. doi: 10.1109/MMSP.2008.4665112.
228. L. Klaine, B. Vozel, and K. Chehdi, "An Integro-Differential Method for Adaptive Filtering of Additive or Multiplicative Noise," in *Proceedings. (ICASSP '05). IEEE International Conference on Acoustics, Speech, and Signal Processing, 2005.*, Philadelphia, Pennsylvania, USA: IEEE, 2005, pp. 1001–1004. doi: 10.1109/ICASSP.2005.1415576.
229. C.-A. Deledalle, F. Tupin, and L. Denis, "Patch similarity under non Gaussian noise," in *2011 18th IEEE International Conference on Image Processing*, Brussels, Belgium: IEEE, Sep. 2011, pp. 1845–1848. doi: 10.1109/ICIP.2011.6115825.
230. N. Ponomarenko, V. Lukin, and K. Egiazarian, "HVS-metric-based performance analysis of image denoising algorithms," in *3rd European Workshop on Visual Information Processing*, Paris, France: IEEE, Jul. 2011, pp. 156–161. doi: 10.1109/EuVIP.2011.6045554.
231. O. Pogrebnyak and V. V. Lukin, "Wiener discrete cosine transform-based image filtering," *J. Electron. Imaging*, vol. 21, no. 4, p. 043020, Dec. 2012, doi: 10.1117/1.JEI.21.4.043020.

232. R. Touzi, "A review of speckle filtering in the context of estimation theory," *IEEE Trans. Geosci. Remote Sensing*, vol. 40, no. 11, pp. 2392–2404, Nov. 2002, doi: 10.1109/TGRS.2002.803727.
233. C. M. Bishop, "Pattern recognition and machine learning," *Springer google schola*, vol. 2, pp. 645–678, 2006.
234. O. S. Rubel and V. V. Lukin, "An improved prediction of dct-based image filters efficiency using regression analysis," *Information and Telecommunication Sciences*, vol. 0, no. 1, pp. 30–41, Oct. 2014, doi: 10.20535/2411-2976.12014.30-41.
235. S. Rahardja and B. J. Falkowski, "Comparative study of discrete orthogonal transforms in adaptive signal processing," *IEICE transactions on fundamentals of electronics, communications and computer sciences*, vol. 82, no. 8, pp. 1386–1390, 1999.
236. D. V. Fevrale, N. N. Ponomarenko, V. V. Lukin, S. K. Abramov, K. O. Egiazarian, and J. T. Astola, "Efficiency analysis of DCT-based filters for color image database," presented at the IS&T/SPIE Electronic Imaging, J. T. Astola and K. O. Egiazarian, Eds., San Francisco Airport, California, USA, Feb. 2011, p. 78700R. doi: 10.1117/12.871944.
237. E. T. Jaynes and G. L. Bretthorst, *Probability theory: the logic of science*. Cambridge, UK ; New York, NY: Cambridge University Press, 2003.
238. P. Symes and P. Symes, *Digital video compression*. New York: McGraw-Hill, 2004.
239. M. Carli, P. Campisi, and A. Neri, "Perceptual aspects in data hiding," *Telecommun Syst*, vol. 33, no. 1–3, pp. 117–129, Nov. 2006, doi: 10.1007/s11235-006-9009-x.
240. Zhenzhong Chen and C. Guillemot, "Perceptually-Friendly H.264/AVC Video Coding Based on Foveated Just-Noticeable-Distortion Model," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 20, no. 6, pp. 806–819, Jun. 2010, doi: 10.1109/TCSVT.2010.2045912.
241. H. R. Wu, W. Lin, and L. J. Karam, "An Overview of Perceptual Processing for Digital Pictures," in *2012 IEEE International Conference on Multimedia and Expo*

Workshops, Melbourne, Australia: IEEE, Jul. 2012, pp. 113–120. doi: 10.1109/ICMEW.2012.27.

242. N. Ponomarenko, S. Krivenko, V. Lukin, K. Egiazarian, and J. T. Astola, “Lossy Compression of Noisy Images Based on Visual Quality: A Comprehensive Study,” *EURASIP J. Adv. Signal Process.*, vol. 2010, no. 1, p. 976436, Dec. 2010, doi: 10.1155/2010/976436.

243. A. Mittal, A. K. Moorthy, and A. C. Bovik, “Visually Lossless H.264 Compression of Natural Videos,” *The Computer Journal*, vol. 56, no. 5, pp. 617–627, May 2013, doi: 10.1093/comjnl/bxs105.

244. N. Ponomarenko, A. Zemlyachenko, V. Lukin, K. Egiazarian, and J. Astola, “Performance analysis of visually lossless image compression,” *Proceedings of VPQM, Scottsdale, USA*, pp. 1–6, 2012.

245. N. Ponomarenko, V. Lukin, J. Astola, and K. Egiazarian, “Analysis of HVS-Metrics’ Properties Using Color Image Database TID2013,” in *Advanced Concepts for Intelligent Vision Systems*, vol. 9386, S. Battiato, J. Blanc-Talon, G. Gallo, W. Philips, D. Popescu, and P. Scheunders, Eds., in *Lecture Notes in Computer Science*, vol. 9386, Cham: Springer International Publishing, 2015, pp. 613–624. doi: 10.1007/978-3-319-25903-1_53.

246. M. Alhihi, A. Zemliachenko, S. K. Abramov, B. Vozel, K. O. Egiazarian, and V. V. Lukin, “Pre-requisites for smart lossy compression of noisy remote sensing images,” *Telecom Rad Eng*, vol. 77, no. 3, pp. 225–241, 2018, doi: 10.1615/TelecomRadEng.v77.i3.40.

247. V. Lukin, A. Zemliachenko, S. Abramov, B. Vozel, and K. Chehdi, “Automatic lossy compression of noisy images by spiht or jpeg2000 in optimal operation point neighborhood,” in *2016 6th European Workshop on Visual Information Processing (EUVIP)*, Marseille, France: IEEE, Oct. 2016, pp. 1–6. doi: 10.1109/EUVIP.2016.7764581.

248. B. Vozel, R. A. Kozhemiakin, S. K. Abramov, V. V. Lukin, and K. Chehdi, “Output MSE and PSNR prediction in DCT-based lossy compression of remote sensing images,” in *Image and Signal Processing for Remote Sensing XXIII*, L. Bruzzone, F.

Bovolo, and J. A. Benediktsson, Eds., Warsaw, Poland: SPIE, Oct. 2017, p. 84. doi: 10.1117/12.2278002.

249. R. Kozhemiakin, V. Lukin, and B. Vozel, “Image quality prediction for DCT-based compression,” in *2017 14th International Conference The Experience of Designing and Application of CAD Systems in Microelectronics (CADSM)*, Lviv - Polyana, Ukraine: IEEE, 2017, pp. 225–228. doi: 10.1109/CADSM.2017.7916121.

250. P. Kitsos, N. S. Voros, T. Dagiuklas, and A. N. Skodras, “A high speed FPGA implementation of the 2D DCT for Ultra High Definition video coding,” in *2013 18th International Conference on Digital Signal Processing (DSP)*, Fira, Santorini, Greece: IEEE, Jul. 2013, pp. 1–5. doi: 10.1109/ICDSP.2013.6622742.

251. A. Zemliachenko, R. Kozhemiakin, B. Vozel, and V. Lukin, “Prediction of compression ratio in lossy compression of noisy images,” in *2016 13th International Conference on Modern Problems of Radio Engineering, Telecommunications and Computer Science (TCSET)*, Lviv, Ukraine: IEEE, Feb. 2016, pp. 693–697. doi: 10.1109/TCSET.2016.7452156.

252. N. Ozah and A. Kolokolova, “Compression Improves Image Classification Accuracy,” in *Advances in Artificial Intelligence*, vol. 11489, M.-J. Meurs and F. Rudzicz, Eds., in *Lecture Notes in Computer Science*, vol. 11489, Cham: Springer International Publishing, 2019, pp. 525–530. doi: 10.1007/978-3-030-18305-9_55.

253. A. Zabala, X. Pons, R. Diaz-Delgado, F. Garcia, F. Auli-Llinas, and J. Serra-Sagrasta, “Effects of JPEG and JPEG2000 Lossy Compression on Remote Sensing Image Classification for Mapping Crops and Forest Areas,” in *2006 IEEE International Symposium on Geoscience and Remote Sensing*, Denver, CO, USA: IEEE, Jul. 2006, pp. 790–793. doi: 10.1109/IGARSS.2006.203.

254. V. Lukin, “Processing of Multichannel Remote Sensing Data for Environment Monitoring,” in *GeoSpatial Visual Analytics*, R. D. Amicis, R. Stojanovic, and G. Conti, Eds., in *NATO Science for Peace and Security Series C: Environmental Security*, Dordrecht: Springer Netherlands, 2009, pp. 129–138. doi: 10.1007/978-90-481-2899-0_10.

255. P. S. Sisodia, V. Tiwari, and A. Kumar, “Analysis of Supervised Maximum Likelihood Classification for remote sensing image,” in *International Conference on Recent Advances and Innovations in Engineering (ICRAIE-2014)*, Jaipur, India: IEEE, May 2014, pp. 1–4. doi: 10.1109/ICRAIE.2014.6909319.

256. J. Sun, J. Yang, C. Zhang, W. Yun, and J. Qu, “Automatic remotely sensed image classification in a grid environment based on the maximum likelihood method,” *Mathematical and Computer Modelling*, vol. 58, no. 3–4, pp. 573–581, Aug. 2013, doi: 10.1016/j.mcm.2011.10.063.

257. G. Proskura, I. Vasilyeva, L. Fangfang, and V. Lukin, “Classification of Compressed Multichannel Images and Its Improvement,” in *2020 30th International Conference Radioelektronika (RADIOELEKTRONIKA)*, Bratislava, Slovakia: IEEE, Apr. 2020, pp. 1–6. doi: 10.1109/RADIOELEKTRONIKA49387.2020.9092371.

258. B. R. Parresol, *Recovering parameters of Johnson’s SB distribution*, vol. 31. US Department of Agriculture, Forest Service, Southern Research Station, 2003. Accessed: Aug. 17, 2024. [Online]. Available: https://www.google.com/books?hl=uk&lr=&id=zzjxAAAAMAAJ&oi=fnd&pg=PT4&dq=Recovering+Parameters+of+Johnson%27s+SB+Distribution&ots=CRwdYkSKRN&sig=0D38YPh0r9ortsnettgf_l8_YrA

259. A. J. Lemonte and J. L. Bazán, “New class of Johnson distributions and its associated regression model for rates and proportions,” *Biometrical J*, vol. 58, no. 4, pp. 727–746, Jul. 2016, doi: 10.1002/bimj.201500030.

260. J. M. Bioucas-Dias *et al.*, “Hyperspectral Unmixing Overview: Geometrical, Statistical, and Sparse Regression-Based Approaches,” *IEEE J. Sel. Top. Appl. Earth Observations Remote Sensing*, vol. 5, no. 2, pp. 354–379, Apr. 2012, doi: 10.1109/JSTARS.2012.2194696.

261. O. K. Al-Shaykh and R. M. Mersereau, “Lossy compression of noisy images,” *IEEE Trans. on Image Process.*, vol. 7, no. 12, pp. 1641–1652, Dec. 1998, doi: 10.1109/83.730376.

262. B. Bondzulich, D. Bujakovic, F. Li, and V. Lukin, “On strange images with application to lossy image compression,” *Radioelectronic and Computer Systems*, no. 4, pp. 143–152, 2022.

263. F. Li, V. Lukin, and X. Liu, “Strange Images with Non-monotonous Rate-Distortion Curves in Lossy Image Compression,” in *2022 IEEE 5th International Conference on Information Systems and Computer Aided Education (ICISCAE)*, Sep. 2022, pp. 11–15. doi: 10.1109/ICISCAE55891.2022.9927685.

264. S. Nan, X. Feng, Y. Wu, and H. Zhang, “Remote sensing image compression and encryption based on block compressive sensing and 2D-LCCCM,” *Nonlinear Dyn*, vol. 108, no. 3, pp. 2705–2729, May 2022, doi: 10.1007/s11071-022-07335-4.

265. S. K. Abramov, V. V. Lukin, N. N. Ponomarenko, and O. B. Pogrebnyak, “Entropy-Like Measure of Background Content for Image Retrieval and Sorting in Large Databases,” *Telecom Rad Eng*, vol. 68, no. 8, pp. 667–675, 2009, doi: 10.1615/TelecomRadEng.v68.i8.20.

266. F. Li, S. Krivenko, and V. Lukin, “A Two-step Approach to Providing a Desired Visual Quality in Image Lossy Compression,” in *2020 IEEE 15th International Conference on Advanced Trends in Radioelectronics, Telecommunications and Computer Engineering (TCSET)*, Lviv-Slavske, Ukraine: IEEE, Feb. 2020, pp. 502–506. doi: 10.1109/TCSET49122.2020.235483.

267. V. Abramova, V. Lukin, S. Abramov, K. Abramov, and E. Bataeva, “Analysis of Statistical and Spatial Spectral Characteristics of Distortions in Lossy Image Compression,” in *2022 IEEE 2nd Ukrainian Microwave Week (UkrMW)*, Ukraine: IEEE, Nov. 2022, pp. 644–649. doi: 10.1109/UkrMW58013.2022.10036949.

268. V. V. Abramova, S. K. Abramov, and V. V. Lukin, “Iterative method for blind evaluation of mixed noise characteristics on images,” *Information and Telecommunication Sciences*, vol. 0, no. 1, pp. 8–14, Jul. 2015, doi: 10.20535/2411-2976.12015.8-14.

269. V. V. Abramova, S. K. Abramov, V. V. Lukin, A. A. Roenko, and B. Vozel, “Automatic estimation of spatially correlated noise variance in spectral domain for

images,” *Telecom Rad Eng*, vol. 73, no. 6, pp. 511–527, 2014, doi: 10.1615/TelecomRadEng.v73.i6.40.

270. M. Ponomarenko, N. Gapon, V. Voronin, and K. Egiazarian, “Blind estimation of white Gaussian noise variance in highly textured images,” *ei*, vol. 30, no. 13, pp. 382-1-382–5, Jan. 2018, doi: 10.2352/ISSN.2470-1173.2018.13.IPAS-382.

271. F. Li, V. V. Lukin, G. Proskura, I. V. Vasilyeva, and G. Chernova, “Image Classification Accuracy Analysis for Three-channel Remote Sensing Data.,” in *IntelITSIS*, 2022, pp. 505–519. Accessed: Sep. 01, 2024. [Online]. Available: https://www.researchgate.net/profile/Vladimir-Lukin-4/publication/361385547_Image_Classification_Accuracy_Analysis_for_Three-channel_Remote_Sensing_Data/links/63145a395eed5e4bd1462d4c/Image-Classification-Accuracy-Analysis-for-Three-channel-Remote-Sensing-Data.pdf

272. V. V. Lukin, N. N. Ponomarenko, A. A. Zelensky, A. A. Kurekin, and K. Lever, “Compression and classification of noisy multichannel remote sensing images,” presented at the SPIE Remote Sensing, L. Bruzzone, C. Notarnicola, and F. Posa, Eds., Cardiff, Wales, United Kingdom, Oct. 2008, p. 71090W. doi: 10.1117/12.799497.

273. N. Damera-Venkata, T. D. Kite, W. S. Geisler, B. L. Evans, and A. C. Bovik, “Image quality assessment based on a degradation model,” *IEEE Trans. on Image Process.*, vol. 9, no. 4, pp. 636–650, Apr. 2000, doi: 10.1109/83.841940.

274. N. Ponomarenko, M. Zriakhov, V. V. Lukin, J. T. Astola, and K. O. Egiazarian, “Estimation of accessible quality in noisy image compression,” in *2006 14th European Signal Processing Conference*, Sep. 2006, pp. 1–4.

275. N. Ponomarenko, V. Lukin, M. Zriakhov, K. Egiazarian, and J. Astola, “Lossy Compression of Images with Additive Noise,” in *Advanced Concepts for Intelligent Vision Systems*, vol. 3708, J. Blanc-Talon, W. Philips, D. Popescu, and P. Scheunders, Eds., in *Lecture Notes in Computer Science*, vol. 3708. , Berlin, Heidelberg: Springer Berlin Heidelberg, 2005, pp. 381–386. doi: 10.1007/11558484_48.

276. D. Yee, S. Soltaninejad, D. Hazarika, G. Mbuyi, R. Barnwal, and A. Basu, “Medical image compression based on region of interest using better portable graphics

(BPG),” in *2017 IEEE International Conference on Systems, Man, and Cybernetics (SMC)*, Banff, AB: IEEE, Oct. 2017, pp. 216–221. doi: 10.1109/SMC.2017.8122605.

277. R. H. Eikelboom *et al.*, “Methods and limits of digital image compression of retinal images for telemedicine,” *Investigative ophthalmology & visual science*, vol. 41, no. 7, pp. 1916–1924, 2000.

278. Anthony Seibert J., “Archiving, Chapter 2: Medical Image Data Characteristics,” Society for Imaging Informatics in Medicine. [Online]. Available: https://siim.org/page/archiving_chapter2

279. “Schick Digital X-ray Image Size,” Patterson Companies. Accessed: Oct. 19, 2020. [Online]. Available: https://pattersonsupport.custhelp.com/app/answers/detail/a_id/1484/~schick-digital-x-ray-image-size

280. “Diagnostic and Imaging Equipment | MORITA.” Accessed: Oct. 19, 2020. [Online]. Available: <https://www.jmoritaeurope.de/en/products/diagnostic-and-imaging-equipment-overview/>

281. M. J. Flynn, S. M. Hames, S. J. Wilderman, and J. J. Ciarelli, “Quantum noise in digital X-ray image detectors with optically coupled scintillators,” *IEEE Trans. Nucl. Sci.*, vol. 43, no. 4, pp. 2320–2325, Aug. 1996, doi: 10.1109/23.531897.

282. W. Huda and R. B. Abrahams, “Radiographic Techniques, Contrast, and Noise in X-Ray Imaging,” *American Journal of Roentgenology*, vol. 204, no. 2, pp. W126–W131, Feb. 2015, doi: 10.2214/AJR.14.13116.

283. M. Ponomarenko, O. Ieremeiev, V. Lukin, and K. Egiazarian, “An expandable image database for evaluation of full-reference image visual quality metrics,” *Electronic Imaging*, vol. 2020, no. 10, pp. 137-1-137–6, Jan. 2020, doi: 10.2352/ISSN.2470-1173.2020.10.IPAS-137.

284. V. Rebrov, B. Vozel, and V. Lukin, “Post-filtering of lossy compressed noisy images and its efficiency prediction,” *A.I.S.*, vol. 8, no. 3, pp. 39–45, Sep. 2024, doi: 10.20998/2522-9052.2024.3.04.

285. G. J. Ewing, "Comparison of JPEG and fractal-based image compression on target acquisition by human observers," *Opt. Eng.*, vol. 35, no. 1, p. 284, Jan. 1996, doi: 10.1117/1.600932.

286. P. D. O'Shea, E. L. Jacobs, and R. L. Espinola, "Effects of image compression on sensor performance," *Opt. Eng.*, vol. 47, no. 1, p. 013202, 2008, doi: 10.1117/1.2829770.

287. T. Hase, W. Hintermaier, A. Frey, T. Strobel, U. Baumgarten, and E. Steinbach, "Influence of Image/Video Compression on Night Vision Based Pedestrian Detection in an Automotive Application," in *2011 IEEE 73rd Vehicular Technology Conference (VTC Spring)*, Budapest, Hungary: IEEE, May 2011, pp. 1–5. doi: 10.1109/VETECS.2011.5956649.

288. S. Dodge and L. Karam, "Understanding how image quality affects deep neural networks," in *2016 Eighth International Conference on Quality of Multimedia Experience (QoMEX)*, Lisbon, Portugal: IEEE, Jun. 2016, pp. 1–6. doi: 10.1109/QoMEX.2016.7498955.

289. M. Dejean-Servières, K. Desnos, K. Abdelouahab, W. Hamidouche, L. Morin, and M. Pelcat, "Study of the impact of standard image compression techniques on performance of image classification with a convolutional neural network," PhD Thesis, INSA Rennes; Univ Rennes; IETR; Institut Pascal, 2017. Accessed: Oct. 01, 2024. [Online]. Available: <https://hal.science/hal-01725126/document>

290. K. Kajak, "Impact of video compression on the performance of object detection algorithms in automotive applications." 2020. Accessed: Oct. 01, 2024. [Online]. Available: <https://www.diva-portal.org/smash/record.jsf?pid=diva2:1417099>

291. J. Wu, G. Shi, and W. Lin, "Survey of visual just noticeable difference estimation," *Front. Comput. Sci.*, vol. 13, no. 1, pp. 4–15, Feb. 2019, doi: 10.1007/s11704-016-6213-z.

292. Z. Wang, T.-H. Tran, P. K. Muthappa, and S. Simon, "A JND-Based Pixel-Domain Algorithm and Hardware Architecture for Perceptual Image Coding," *J. Imaging*, vol. 5, no. 5, p. 50, Apr. 2019, doi: 10.3390/jimaging5050050.

293. S. Bosse, S. Becker, K.-R. Müller, W. Samek, and T. Wiegand, "Estimation of distortion sensitivity for visual quality prediction using a convolutional neural network," *Digital Signal Processing*, vol. 91, pp. 54–65, Aug. 2019, doi: 10.1016/j.dsp.2018.12.005.
294. B. Bondžulić, B. Pavlović, N. Stojanović, and V. Petrović, "Picture-wise just noticeable difference prediction model for JPEG image quality assessment," *Vojnotehnički glasnik*, vol. 70, no. 1, pp. 62–86, 2022, doi: 10.5937/vojtehg70-34739.
295. H. Liu *et al.*, "Deep Learning-Based Picture-Wise Just Noticeable Distortion Prediction Model for Image Compression," *IEEE Trans. on Image Process.*, vol. 29, pp. 641–656, 2020, doi: 10.1109/TIP.2019.2933743.
296. B. Bondžulić, N. Stojanović, V. Petrović, B. Pavlović, and Z. Miličević, "Efficient prediction of the first just noticeable difference point for JPEG compressed images," *Acta Polytechnica Hungarica*, vol. 18, no. 8, pp. 201–220, 2021.
297. Q. Li, C. Yang, and J.-Q. Zhang, "Target acquisition performance in a cluttered environment," *Appl. Opt.*, vol. 51, no. 31, p. 7668, Nov. 2012, doi: 10.1364/AO.51.007668.
298. B. Bondžulić, D. Bujaković, and J. Mihajlović, "Image clutter metrics and target acquisition performance," *Vojnotehnički glasnik*, vol. 71, no. 3, pp. 588–615, 2023, doi: 10.5937/vojtehg71-44117.
299. A. Toet, "The Search_2 dataset." figshare, p. 2298042065 Bytes, 2014. doi: 10.6084/M9.FIGSHARE.1041463.V6.
300. H. Chang, "Evaluation of human detection performance using target structure similarity clutter metrics," *Opt. Eng.*, vol. 45, no. 9, p. 096404, Sep. 2006, doi: 10.1117/1.2353848.
301. H. Chang, J. Zhang, X. Liu, C. Yang, and Q. Li, "Color Image Clutter Metrics for Predicting Human Target Acquisition Performance," in *2010 International Conference on Computational Intelligence and Software Engineering*, Chengdu City, China: IEEE, Sep. 2010, pp. 1–4. doi: 10.1109/WICOM.2010.5600622.

302. Z. Qiu, Y. Ma, F. Fan, J. Huang, M. Wu, and X. Mei, “A pixel-level local contrast measure for infrared small target detection,” *Defence Technology*, vol. 18, no. 9, pp. 1589–1601, Sep. 2022, doi: 10.1016/j.dt.2021.07.002.

303. A. Novak, N. Armstrong, T. Caelli, and I. Blair, “Bayesian Contrast Measures and Clutter Distribution Determinants of Human Target Detection,” *IEEE Trans. on Image Process.*, vol. 26, no. 3, pp. 1115–1126, Mar. 2017, doi: 10.1109/TIP.2016.2644269.

304. J. B. Culpepper, “Texture metric that predicts target detection performance,” *Opt. Eng.*, vol. 54, no. 12, p. 123101, Dec. 2015, doi: 10.1117/1.OE.54.12.123101.

305. B. Pavlović, B. Bondžulić, N. Stojanović, V. Petrović, and D. Bujaković, “Prediction of the First Just Noticeable Difference Point based on Simple Image Features,” in *2023 Zooming Innovation in Consumer Technologies Conference (ZINC)*, Novi Sad, Serbia: IEEE, May 2023, pp. 125–130. doi: 10.1109/ZINC58345.2023.10173865.

306. P. Kocsis, I. Shevkunov, V. Katkovnik, H. Rekola, and K. Egiazarian, “Single-shot pixel super-resolution phase imaging by wavefront separation approach,” *Opt. Express*, vol. 29, no. 26, p. 43662, Dec. 2021, doi: 10.1364/OE.445218.

307. S. G. Bahnemiri, M. Ponomarenko, and K. Egiazarian, “Learning-Based Noise Component Map Estimation for Image Denoising,” *IEEE Signal Process. Lett.*, vol. 29, pp. 1407–1411, 2022, doi: 10.1109/LSP.2022.3169706.

308. V. Hosu, H. Lin, T. Sziranyi, and D. Saupe, “KonIQ-10k: An Ecologically Valid Database for Deep Learning of Blind Image Quality Assessment,” *IEEE Trans. on Image Process.*, vol. 29, pp. 4041–4056, 2020, doi: 10.1109/TIP.2020.2967829.

309. Z. Ying, H. Niu, P. Gupta, D. Mahajan, D. Ghadiyaram, and A. Bovik, “From Patches to Pictures (PaQ-2-PiQ): Mapping the Perceptual Space of Picture Quality,” in *2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, Seattle, WA, USA: IEEE, Jun. 2020, pp. 3572–3582. doi: 10.1109/CVPR42600.2020.00363.

310. N. Ponomarenko, O. Ereemeev, V. Lukin, and K. Egiazarian, “Statistical evaluation of no-reference image visual quality metrics,” in *2010 2nd European*

Workshop on Visual Information Processing (EUVIP), Paris, France: IEEE, Jul. 2010, pp. 50–54. doi: 10.1109/EUVIP.2010.5699121.

311. Y. Fang, H. Zhu, Y. Zeng, K. Ma, and Z. Wang, “Perceptual Quality Assessment of Smartphone Photography,” in *2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, Seattle, WA, USA: IEEE, Jun. 2020, pp. 3674–3683. doi: 10.1109/CVPR42600.2020.00373.

312. S. Pyatykh, J. Hesser, and Lei Zheng, “Image Noise Level Estimation by Principal Component Analysis,” *IEEE Trans. on Image Process.*, vol. 22, no. 2, pp. 687–699, Feb. 2013, doi: 10.1109/TIP.2012.2221728.

313. X. Liu, M. Tanaka, and M. Okutomi, “Noise level estimation using weak textured patches of a single noisy image,” in *2012 19th IEEE International Conference on Image Processing*, Orlando, FL, USA: IEEE, Sep. 2012, pp. 665–668. doi: 10.1109/ICIP.2012.6466947.

314. L. Armi and S. Fekri-Ershad, “Texture image analysis and texture classification methods - A review,” 2019, doi: 10.48550/ARXIV.1904.06554.

315. “SIPI Image Database.” Accessed: Aug. 05, 2024. [Online]. Available: <https://sipi.usc.edu/database/>

316. V. V. Abramova, “A blind method for additive noise variance evaluation based on homogeneous region detection using the fourth central moment analysis,” *Telecom Rad Eng*, vol. 74, no. 18, pp. 1651–1669, 2015, doi: 10.1615/TelecomRadEng.v74.i18.50.

317. R. Tsekhmystro, O. Rubel, V. Abramova, M. Zriakhov, and M. Uss, “Mobile Deployment of NoiseNet: Noise Characteristics Assessment in Real-World Images,” in *2019 IEEE 2nd Ukraine Conference on Electrical and Computer Engineering (UKRCON)*, Lviv, Ukraine: IEEE, Jul. 2019, pp. 1112–1117. doi: 10.1109/UKRCON.2019.8879797.

ДОДАТОК А

ДОСЛІДЖЕННЯ ДВОЕТАПНИХ ПРОЦЕДУР ЗАБЕЗПЕЧЕННЯ НЕОБХІДНОЇ ЯКОСТІ ДЛЯ КОДЕРА ADCT

Розглянемо кодер ADCT. Згідно з двома критеріями якості стисненого зображення – КС (БНП) і PSNR – експериментальні результати доводять, що ADCT кращий за JPEG2000 і JPEG, особливо для зображень із різномірним вмістом [168]. Наведемо один приклад роботи ADCT. В якості тестового було обрано синтетичне дистанційне зображення (рис. 2.13), оскільки воно містить однорідні ділянки, протяжні об'єкти, що імітують дороги або паркани, і насичена текстурна інформація (у нижньому лівому куті). Характеристики таких об'єктів є важливими для завдань класифікації ДЗЗ, і вони повинні досить добре зберігатися на стадії стиснення. PSNR і PSNR-HVS-M дорівнюють 30 дБ для цих тестів, щоб порівняти ефективність кодування ADCT і AGU. Результати експерименту наведені в таблиці А.1. Видно, що КС для ADCT на 7...12% вищий, ніж для AGU. Це очевидна перевага ADCT. Тим часом потрібен деякий час, щоб оптимізувати схему розбиття для ADCT. Таким чином, у порівнянні з іншими методами, компресія та декомпресія за допомогою ADCT відбуваються повільніше, і це головний недолік ADCT, негативний вплив якого необхідно мінімізувати. Тут мається на увазі, що надзвичайно бажано уникнути багатьох кроків ітерації для ADCT.

Таблиця А.1 – Порівняння ефективності кодерів AGU та ADCT для зображення на рис. 2.13 (PSNR-HVS-M_{prov}≈30 дБ)

Кодер	Метрика	Забезпечене значення	КС
ADCT	PSNR	30,002	89,950
ADCT	PSNR-HVS-M	29,999	77,718
AGU	PSNR	29,987	83,300
AGU	PSNR-HVS-M	30,012	69,664

У цьому сенсі метод двоетапного стиснення, запропонований вище для кодера AGU, здається, може досягти хороших результатів як щодо точності, так і ефективності часу в розглянутому випадку. Важливою властивістю є те, що обидва кодери ADCT і AGU базуються на ДКП і обидва використовують крок квантування як ПКС.

Знову будемо використовувати МЯЕ PSNR і PSNR-HVS-M. Отримані КШС показані на рис. А.1; в якості тестових зображень, як і раніше, обрано дев'ять зображень. Серед них є чотири стандартні зображення, чотири зображення дистанційного зондування та одне медичне зображення, які представляють зображення з різними властивостями. Зокрема, є високотекстурні зображення Baboon і Diego, тоді як тестові зображення Frisco і MRT_prepared досить прості.

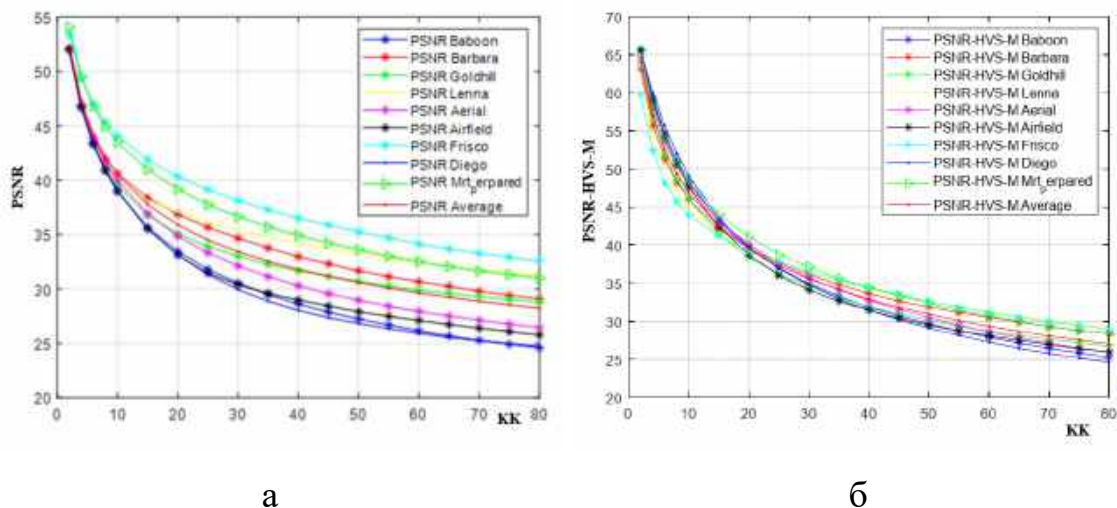


Рисунок А.1 – Залежність PSNR(а) та PSNR-HVS-M(б) від КК для ADCT

На рис. А.1, а, показано, що залежності PSNR від КК є монотонними. Дані для кожного зображення різні, але вони демонструють однакову тенденцію. На основі цих даних можна отримати середню КШС. Можна побачити, що коли значення КК знаходиться в межах 20 ... 80, середня крива та КШС для конкретних зображень приблизно паралельні. На рис. А.1, б, наведено залежності для PSNR-HVS-M. Зі збільшенням КК метрика зменшується. Отримано середню криву. Порівняно з відмінностями кривих на рис. А.1, а, можна виявити, що криві

спотворення PSNR-HVS-M «більш компактні» і майже «паралельні» одна одній у всьому діапазоні.

Згідно з попередніми дослідженнями, для середньої кривої були отримані значення похідної. Нижче показано, як ці дані використовуються для прогнозування та забезпечення бажаної якості.

Таблиця А.2 містить частину статистичних даних для залежності PSNR-HVS-M від КК. Усереднення було виконано для кожного КК для набору тестових зображень, потім були визначені значення похідної (усі вони негативні, і їхні абсолютні значення зменшуються, якщо КК збільшується). На першому етапі стиснення пропонується встановити ПКС відповідно до середньої КШС. Наведемо один приклад. Припустимо, потрібно забезпечити стиснення з бажаним PSNR-HVS-M, рівним 40 дБ, для тестового зображення Baboon. Відповідно до статистики середньої КШС (див. рядок «Середнє» в таблиці А.2), значення КК має бути між 15 і 20. Потім, використовуючи лінійну інтерполяцію, можемо обчислити початкове значення КК за формулою (3.15), де замість Q буде використовуватись КК.

Таблиця А.2 – Залежність PSNR-HVS-M від КК для ADCT

КК	2	15	20
Goldhill	63,94	41,41	38,36
Baboon	65,62	43,036	39,575
Barbara	63,091	42,137	39,594
Lenna	62,433	41,3	39,183
Aerial	64,569	43,257	40,049
Airfield	65,492	42,387	38,756
Frisco	59,728	41,163	39,204
Diego	65,71	43,474	39,711
Mrt_prepared	65,57	43,98	41,06
Середнє	64,01	42,46	39,50
Похідна	-3,471	-0,592	-0,442

Використовуючи дані таблиці А.2, можна розрахувати, що початкове значення КК дорівнює 19,155. Після отримання початкового значення КК зображення Baboon стискається на першому кроці, а потім розпаковується, щоб отримати значення метрики для даного зображення. Це значення $PSNR-HVS-M_{init}$. Для зображення Baboon вона дорівнює 40,08 дБ. Як бачимо, це значення вже дуже близько до бажаного значення метрики візуальної якості (40 дБ). На другому кроці процедури стиснення потрібно обчислити виправлене значення КК для цього зображення, використовуючи значення $PSNR-HVS-M_{init}$, отримане на першому кроці, щоб підвищити точність. Це можна зробити за допомогою формули (3.16), де замість Q буде використовуватись КК.

Виправлене (бажане) значення КК використовується на другому кроці стиснення. Наданий $PSNR-HVS-M_{prov}$ дорівнює 39,993 дБ і є ближчим до бажаного значення, ніж $PSNR-HVS-M_{init}$ (40,08 дБ). Тут показано конкретну реалізацію двоетапного методу. Видно, що після другого етапу стиснення надане значення $PSNR-HVS-M_{prov}$ ближче до бажаного значення $PSNR-HVS-M_{des}$, ніж початкове значення $PSNR-HVS-M_{init}$. Але виникає питання, чи завжди це відбувається?

Перші дев'ять тестових зображень (зображень у бібліотеці, які використовуються для отримання КШС) будуть використані для перевірки застосовності двоетапного методу до ADCTC. Крім того, щоб перевірити його практичну універсальність, були додані деякі текстурні зображення поза основною бібліотекою [314]. Давайте спочатку перевіримо ефективність двоетапного методу на цих дев'яти тестових зображеннях. Взявши за приклад значення візуальної якості $PSNR-HVS-M$ 40 дБ, статистичні дані наведено в таблиці А.3. Треба зауважити, що $\Delta KK = \frac{PSNR-HVS-M_{des} - PSNR-HVS-M_{init}}{M'}$.

Аналіз даних у таблиці А.3 показує, що всі значення метрики візуальної якості після двоетапного стиснення покращилися порівняно з відповідними значеннями візуальної якості, отриманими після першого кроку стиснення. Були розраховані значення дисперсії для першого кроку та для двох кроків.

Таблиця А.3 – Статистичні дані та параметри для PSNR-HVS-M_{des}=40 дБ

Test image	KK _{init}	PSNR-HVS-M _{init}	ΔKK	KK _{des}	PSNR-HVS-M _{prov}
Goldhill	19,16	38,802	-2,02	17,13	40,001
Baboon	19,16	40,080	0,135	19,29	39,993
Barbara	19,16	40,000	0,000	19,15	40,000
Lenna	19,16	39,490	-0,86	18,29	39,895
Aerial	19,16	40,547	0,923	20,08	40,026
Airfield	19,16	39,272	-1,23	17,93	40,116
Frisco	19,16	39,516	-0,82	18,34	39,804
Diego	19,16	40,290	0,490	19,65	39,919
Mrt_prepared	19,16	41,516	2,56	21,71	40,187
Дисперсія	0,6401		0,0131		

Значення дисперсії для двох кроків доводить, що двоетапне стиснення є значно точнішим. Вище для кодера AGU було показано, що двоетапний метод добре працює для забезпечення бажаних значень метрики PSNR-HVS-M, але він стикається з проблемами, якщо бажаний PSNR менше 35 дБ. Експеримент для таких значень PSNR також проводився для ADCT. Подібно до [45] було запропоновано наступну модифікацію. Якщо абсолютне значення ΔKK більше, ніж $KK_{init} / 2$, а ΔKK від'ємне, тоді встановлюємо

$$KK_{rec} = KK_{init} / 2, \quad (A.1)$$

та продовжуємо двоетапну компресію. Завдяки цій модифікації гарантується точність забезпечення потрібної якості. Експериментальні дані, що доводять це, наведені в таблиці А.4. Причини необхідності вдосконалення методу можна пояснити наступним аналізом. Для зображення Baboon в таблиці А.4 ΔKK, розраховане після першого кроку стиснення/декомпресії, дорівнює -38,21, що перевищує половину початкового значення та є від'ємним. Така ж ситуація має місце для іншого високотекстурного зображення Diego, де ΔKK дорівнює -41,62.

Повернемося до КШС PSNR (рис. А.1, а). Можна виявити, що криві Baboon і Diego є найдальшими (нижніми) від КШС, а найдальшими «зверху» є зображення Frisco. Через це значення ΔKK для цих трьох кривих є найбільшими, де значення для Baboon і Diego негативні (позначені червоним), а значення Frisco позитивні (позначені зеленим).

На рис. А.1, а, можна побачити, що коли KK є великим, усі криві спотворення змінюються повільно, а похідні змінюються незначно; коли KK невелика, похідні кривої викривлення сильно змінюються.

Таблиця А.4 – Статистичні дані та параметри для PSNR-HVS- $M_{des}=40$ дБ

Test image	KK_{init}	PSNR-HVS- M_{init}	ΔKK	KK_{des}	PSNR-HVS- M_{prov}
Goldhill	56,05	30,17	1,988	58,033	30,038
Baboon	56,05	26,56	-38,21	17,835	31,018
Barbara	56,05	31,04	11,65	67,691	29,981
Lenna	56,05	32,73	30,33	86,372	30,883
Aerial	56,05	28,29	-18,91	37,134	30,764
Airfield	56,05	27,41	-28,78	27,259	30,775
Frisco	56,05	34,59	50,99	107,03	31,099
Diego	56,05	26,25	-41,62	14,418	30,449
Mrt_prepared	56,05	32,90	32,218	88,263	30,423
Дисперсія	9,1895		0,1643		

Таким чином, коли ΔKK є від’ємним і його абсолютне значення перевищує $KK_{init}/2$ (негативний напрям), прогнозування KK за похідною, що відповідає KK_{init} , більше не є точним, і необхідно прийняти вдосконалені рішення, наприклад для високотекстурних тестових зображень Baboon та Diego. Якщо ΔKK більше, ніж $KK_{init}/2$ (позитивний напрям), це мало впливає на точність, як, наприклад, для зображення Frisco. Шість наборів статистичних даних для метрик PSNR і PSNR-HVS- M представлені в таблиці А.5. Тут M_{des} — бажане значення метрики візуальної якості, що розглядається, VAR_1 — дисперсія метрики візуальної якості

для дев'яти тестових зображень, отриманих після першого етапу. стиснення, VAR_2 — дисперсія показника візуальної якості, отримана після другого (коригувального) кроку стиснення, $MAX_{\Delta_{fin}}$ — максимальна помилка між M_{des} і наданим значенням розглянутого показника візуальної якості.

З аналізу цих результатів можна побачити, що для метрики PSNR-HVS-M двоетапний метод працює досить добре, а дисперсія після двох кроків значно, приблизно на один порядок, менша, ніж дисперсія на першому кроці.

Таблиця А.5 – Статистичні дані, отримані для дев'яти тестових зображень

Метрика	M_{des}	VAR_1	VAR_2	$MAX_{\Delta_{final}}$
PSNR	40	4,064	0,268	1,231
PSNR	35	7,776	0,788	2,368
PSNR	30	9,189	0,164	1,099
PSNR-HVS-M	40	0,640	0,013	0,187
PSNR-HVS-M	35	1,168	0,012	0,336
PSNR-HVS-M	30	2,217	0,058	0,843

При цьому похибка також знаходиться в розумних межах, вона не перевищує 0,85 дБ. Прийнявши модифікований двоетапний метод для цього набору даних із $PSNR_{des}=30$ дБ, дисперсія після двох кроків також приблизно на порядок менша, ніж дисперсія після першого кроку. Максимальна похибка не перевищує 2,4 дБ. Що стосується універсальності цього методу, для перевірки потрібно використовувати зображення поза первинною бібліотекою зображень. Для цього двадцять високотекстурних зображень із різними типами та властивостями текстур (див. рис. А.2) були відібрані як перевірочний тестовий набір зображень [315]. За допомогою цих 20 зображень можна перевірити застосовність двоетапного методу до зображень з високою текстурою, а також можна оцінити діапазон помилок двоетапного методу в практичних застосуваннях. В експерименті три репрезентативні значення візуальної якості 30, 35 і 40 дБ використовувалися як

бажана якість, і було отримано 6 наборів даних метрики візуальної якості PSNR і PSNR-HVS-M.

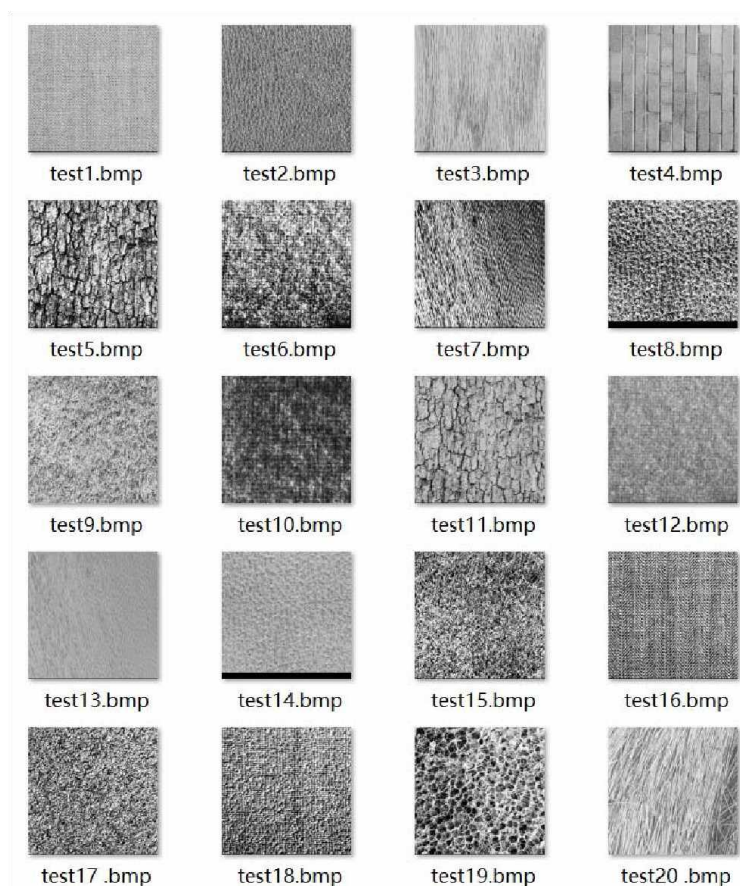


Рисунок А.2 – Тестовий набір з високотекстурними зображеннями

Результати показані в таблицях А.6 та А.7.

Таблиця А.6 – Статистичні дані, отримані для двадцяти високотекстурних зображень

Метрика	M_{des}	VAR_1	VAR_2	$MAX\Delta_{final}$
PSNR	40	0,1677	0,1653	1,314
PSNR	35	1,8509	0,808	2,729
PSNR	30	5,092	0,1121	1,424
PSNR-HVS-M	40	7,3716	0,2819	1,502
PSNR-HVS-M	35	6,3311	0,4281	1,370
PSNR-HVS-M	30	2,9301	0,738	1,463

Таблиця А.7 – Статистичні дані та параметри для $PSNR_{des}=30$ дБ, високотекстурні зображення

Зображення	KK_{init}	$PSNR-HVS-M_{init}$	ΔKK	KK_{des}	$PSNR-HVS-M_{prov}$
Test1	56,045	26,328	-40,800	28,023	30,455
Test2	56,045	25,144	-53,951	28,023	30,0986
Test3	56,045	30, 863	9,590	65,635	30,092
Test4	56,045	27,247	-30,585	28,023	30,624
Test5	56,045	24,959	-56,003	28,023	30,2574
Test6	56,045	24,189	-64,557	28,023	30,2024
Test7	56,045	24,928	-56,347	28,023	29,9809
Tes8	56,045	24,539	-60,671	28,023	30,3694
Test9	56,045	26,728	-36,347	28,023	31,424
Test10	56,045	25,562	-49,311	28,023	30,354
Test11	56,045	26,481	-39,101	28,023	30,794
Test12	56,045	26,158	-42,689	28,023	30,152
Test13	56,045	32,706	30,066	86,111	30,186
Test14	56,045	27,803	-24,411	31,634	30,756
Test15	56,045	25,049	-55,002	28,023	30,427
Test16	56,045	24,473	-61,401	28,023	30,119
Test17	56,045	24,329	-63,006	28,023	30,163
Test18	56,045	25,372	-51,412	28,023	30,502
Test19	56,045	25,425	-50,825	28,023	30,565
Test20	56,045	29,140	-9,559	46,486	30,187
Дисперсія	5,092		0,1121		

У цьому наборі статистичних даних для двох наборів даних було використано модифіковану двоетапну процедуру. Це випадки для $PSNR_{des}$, що дорівнює 30 дБ і 35 дБ. В обох наборах даних були випадки, коли ΔKK був негативним, а його абсолютне значення було більше $KK_{init} / 2$.

З аналізу статистичних даних у таблиці А.6 можна зробити висновок, що для ADCT двоетапний метод є універсальним. Для трьох розглянутих значень PSNR-HVS-M процедура працює добре.

Дисперсія залишкових помилок після двоетапного стиснення зменшується приблизно на порядок. Для трьох розглянутих значень PSNR також доведена ефективність двоетапного методу.

Якщо бажаний PSNR дорівнює 35 дБ або 30 дБ, спостерігається значне покращення, особливо для $\text{PSNR}_{\text{des}}=30$ дБ. Для спрощення детального аналізу дані для цієї групи тестових зображень наведено в таблиці А.7.

Як видно з аналізу даних у таблиці А.7, точність була значно покращена – дисперсія похибок після двоетапного стиснення приблизно в 50 разів менша, ніж після першого кроку. Майже для всіх тестових зображень (крім тестового зображення №3 і тестового зображення №13) $\text{PSNR}_{\text{init}}$ менший за PSNR_{des} . Більшість ΔKK є від’ємними, а їхні абсолютні значення перевищують $\text{KK}_{\text{init}}/2$ (є навіть випадки, коли абсолютні значення ΔKK перевищують KK_{init} і тоді KK_{rec} (відповідно до (3.16)) виявляється від’ємним, що є непрактичним. Застосовано формулу (А.1), що забезпечує $\text{PSNR}_{\text{prov}}$ дуже близько до бажаного значення PSNR (максимальна похибка для $\text{PSNR}_{\text{des}}=30$ дБ дорівнює 1,424 дБ, і це можна вважати нормальним на практиці).

ДОДАТОК Б

РЕЗУЛЬТАТИ КЛАСИФІКАЦІЇ ЗОБРАЖЕНЬ ДЛЯ НМ- КЛАСИФІКАТОРА

Беручи до уваги результати для класифікатора ММП (підрозділ 3.3), також спочатку було розглянуто випадок, коли НМ-класифікатор був навчений для вихідного зображення, а потім застосований до стиснених зображень. Дані, отримані для випадків PSNR-HVS-M, що дорівнюють 42 дБ, 36 дБ і 30 дБ, наведені відповідно в таблицях Б.1–Б.3. Аналіз даних в таблиці Б.1 свідчить про наступне. Як і очікувалося, всі ймовірності змінилися. Розглянемо діагональні елементи, що відповідають ймовірностям правильної класифікації окремих класів. У цьому сенсі є цікаві спостереження (порівняйте відповідні дані в таблицях 3.27 і Б.1): деякі ймовірності значно зменшилися (див. P_{11}), деякі дещо зменшилися (P_{33} , P_{44} , P_{66} , P_{77}). Водночас деякі ймовірності дещо зросли (P_{22} і P_{55}).

Таблиця Б.1 – Матриця плутанини для стисненого зображення, що класифіковано НМ, яка навчена для вихідного зображення та застосована для стиснутого зображення (PSNR-HVS-M = 42 дБ)

Клас	Ймовірність рішення						
	Дороги	Поле-У	Поле-Р	Вода	Трава	Дерева	Текстура
Дороги	0,299	0,699	0	0	0	0	0
Поле-У	0,089	0,627	0	0	0	0	0,276
Поле-Р	0	0,018	0,982	0	0	0	0
Вода	0	0	0	0,980	0	0	0,020
Трава	0	0	0	0	0,753	0,219	0,028
Дерева	0	0	0	0	0,084	0,911	0,006
Текстура	0	0,03	0	0,26	0,116	0,153	0,441

Аналіз даних у таблицях Б.2 і Б.3 показує, що стиснення з втратами негативно впливає на P_{11} , тобто ймовірність правильної класифікації для класу «Дорога» суттєво зменшується. Нагадаємо, що такий же ефект спостерігався для ММП. З

більшим КС (або меншим PSNR-HVS-M) ймовірності P_{33} і P_{77} продовжують зменшуватися. P_{22} повільно зростає. P_{44} залишається практично таким же і дуже високим. P_{66} також залишається практично таким же. Нарешті, P_{55} зменшується. Таким чином, можна стверджувати, що під час навчання для вихідного зображення класифікатор на базі НМ продовжує добре працювати для PSNR-HVS-M близько 40 дБ, а потім його ефективність починає погіршуватися швидше (з подальшим збільшенням КС). На рис. Б.1 показано результати класифікації для стиснутого тестового зображення, що забезпечує три різні значення PSNR-HVS-M.

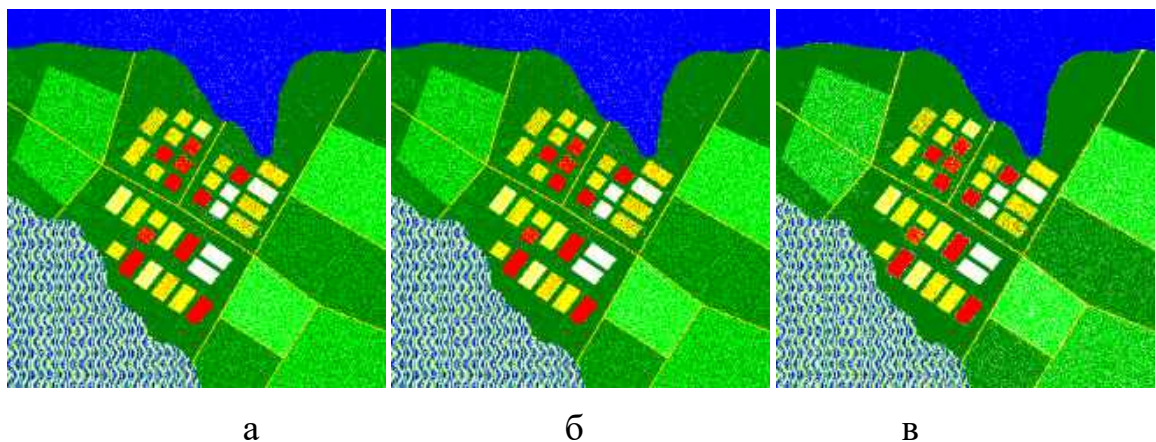


Рисунок Б.1 – Результати класифікації: для стисненого зображення (PSNR-HVS-M = 42 дБ) (а), для стисненого зображення (PSNR-HVS-M = 36 дБ) (б), для стисненого (PSNR-HVS-M = 30 дБ) (в)

Результати досить схожі. Однак можна відзначити деякі відмінності:

- менше неправильних класифікацій спостерігається для PSNR-HVS-M = 30 дБ для класу «Вода»;
- навколо класу «Вода» існує досить багато неправильних класифікацій;
- більше неправильних класифікацій з'являється для класу «Трава», коли КС збільшується;
- багато пікселів і навіть деякі об'єкти, які належать до класу 2 (Поле-У), класифікуються як «Текстура» (білий колір). У цьому прикладі ще раз показано проблеми класифікації зображень, якщо зображення, яке потрібно класифікувати,

містить такий клас, як «Текстура». Проблеми виникають для самого класу «Текстура» і класів, для яких є суттєвий перетин ознак з ознаками класу «Текстура».

Таблиця Б.2 – Матриця плутанини для стисненого зображення, що класифіковано НМ, навченою на вихідному зображенні та застосованою до стиснутого зображення (PSNR-HVS-M = 36 дБ)

Клас	Ймовірність рішення						
	Дороги	Поле-У	Поле-R	Вода	Трава	Дерева	Текстура
Дороги	0,271	0,728	0,001	0	0	0	0
Поле-У	0,078	0,631	0,012	0	0,001	0	0,279
Поле-R	0	0,027	0,973	0	0	0	0
Вода	0	0	0	0,979	0	0	0,021
Трава	0	0	0	0	0,721	0,22	0,059
Дерева	0	0	0	0	0,078	0,915	0,007
Текстура	0	0,049	0	0,253	0,123	0,148	0,426

Давайте тепер перевіримо, що станеться, якщо для навчання використовувати стиснуті зображення. Таблиця Б.4 представляє матрицю плутанини, отриману для випадку навчання НМ із використанням стиснутого зображення, що забезпечує PSNR-HVS-M = 42 дБ. Ці дані можна порівняти з відповідними даними в таблиці Б.1. Якщо для навчання використовується стиснене зображення, ймовірності P_{11} , P_{55} і P_{77} є значно кращими, ймовірність P_{22} значно гіршою, інші ймовірності трохи гіршими. Таким чином, якість класифікації, в цілому, дещо покращується. Карта класифікації представлена на рис. Б.2, б.

Наступним розглянутим випадком є класифікація стисненого зображення із PSNR-HVS-M = 36 дБ. Отримані ймовірності наведені в таблиці Б.5, карта класифікації представлена на рис. Б.2, в. Ймовірності правильної класифікації в таблиці Б.5 майже такі ж, як і відповідні значення в таблиці Б.4.

Таблиця Б.3 – Матриця плутанини для стисненого зображення, класифікованого НМ, навченою на вихідному зображенні та застосованою до стиснутого зображення (PSNR-HVS-M=30 дБ)

Клас	Ймовірність рішення						
	Дороги	Поле-Y	Поле-R	Вода	Трава	Дерева	Текстура
Дороги	0,217	0,766	0	0	0	0	0,017
Поле-Y	0,067	0,646	0,007	0	0,005	0	0,275
Поле-R	0	0,05	0,946	0,004	0	0	0
Вода	0	0	0	0,992	0	0	0,008
Трава	0	0	0	0	0,658	0,21	0,131
Дерева	0	0	0	0	0,051	0,925	0,025
Текстура	0	0,067	0	0,243	0,141	0,131	0,418

Таблиця Б.4 – Матриця плутанини для стисненого зображення, класифікованого НМ, навченою на цьому зображенні (PSNR-HVS-M=42 дБ).

Клас	Ймовірність рішення						
	Дороги	Поле-Y	Поле-R	Вода	Трава	Дерева	Текстура
Дороги	0,455	0,545	0,001	0	0	0	0
Поле-Y	0,137	0,561	0,014	0	0	0	0,289
Поле-R	0	0,026	0,974	0	0	0	0
Вода	0	0	0	0,961	0	0	0,039
Трава	0	0	0	0	0,820	0,165	0,015
Дерева	0	0	0	0	0,068	0,929	0,004
Текстура	0	0,029	0	0,234	0,124	0,153	0,461

Порівнюючи дані в таблицях Б.2 і Б.5, можна стверджувати, що для навчання варто використовувати стиснене зображення (найбільша ймовірність правильної класифікації для всіх класів в цьому випадку). Карту класифікації наведено на рис. Б.2, в, і вона дуже схожа на карту Б.2, а,б. Це означає, що немає значного погіршення точності класифікації порівняно з двома попередніми випадками.

Нарешті, останнім випадком є класифікація стисненого зображення із забезпеченням PSNR-HVS-M = 30 дБ, тобто зі значними спотвореннями. Матриця

плутанини наведена в таблиці Б.6, карта класифікації представлена на рис. Б.2, г. Порівняно з відповідними даними в таблиці Б.3 спостерігається достатнє зростання P_{11} , P_{55} , P_{66} , P_{77} , P_{22} , P_{33} та P_{44} також трохи зросли. Це означає, що варто використовувати стиснене зображення для навчання, особливо якщо стиснення здійснюється з великим КС.

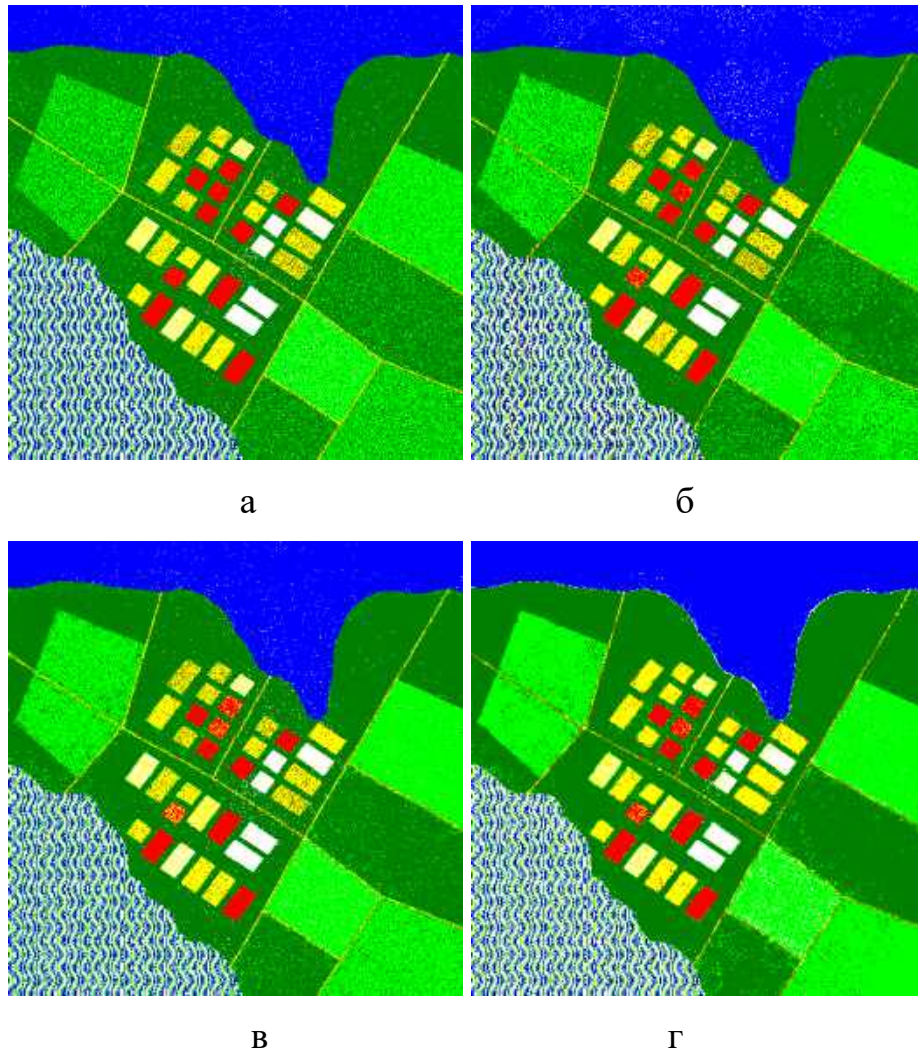


Рисунок Б.2 – Результати класифікації зображень НМ, навченою для того самого зображення: оригінал (а), стиснене із забезпеченням PSNR-HVS-M= 42 дБ (б), стиснене із забезпеченням PSNR-HVS-M = 36 дБ (в), стиснене із забезпеченням PSNR-HVS -M = 30 дБ (г)

Уважний аналіз карти класифікації на рис. Б.2, г, дозволяє помітити досить багато неправильних класифікацій поблизу межі між класами «Вода» та «Дерева»,

де ці неправильно класифіковані пікселі пов'язані з «Текстурою». Пов'язуємо цей ефект із розмиттям меж та іншими ефектами, які виникають поблизу висококонтрастних меж, якщо ступінь стиснення досить високий.

Таблиця Б.5 – Матриця плутанини для стисненого зображення, класифікованого НМ, навченою на цьому зображенні (PSNR-HVS-M=36 дБ).

Клас	Ймовірність рішення						
	Дороги	Поле-У	Поле-R	Вода	Трава	Дерева	Текстура
Дороги	0,507	0,492	0,002	0	0	0	0
Поле-У	0,127	0,562	0,005	0	0,001	0	0,305
Поле-R	0	0,069	0,931	0	0	0	0
Вода	0	0	0	0,989	0	0	0,01
Трава	0	0	0	0	0,901	0,087	0,013
Дерева	0	0	0	0	0,039	0,956	0,005
Текстура	0	0,029	0	0,236	0,142	0,141	0,452

Таблиця Б.6 – Матриця плутанини для стисненого зображення, класифікованого НМ, навченою на цьому зображенні (PSNR-HVS-M=30 дБ).

Клас	Ймовірність рішення						
	Дороги	Поле-У	Поле-R	Вода	Трава	Дерева	Текстура
Дороги	0,692	0,304	0	0	0,002	0,001	0,01
Поле-У	0,055	0,654	0,009	0	0,019	0	0,263
Поле-R	0	0,033	0,967	0,004	0	0	0
Вода	0	0	0	0,994	0	0	0,006
Трава	0	0	0	0	0,939	0,035	0,026
Дерева	0	0	0	0	0,023	0,972	0,005
Текстура	0	0,049	0	0,221	0,113	0,131	0,486

ДОДАТОК В

АНАЛІЗ ПРОСТОРОВОЇ КОРЕЛЯЦІЇ ШУМУ В МЕДИЧНИХ СИСТЕМАХ РЕЄСТРАЦІЇ ЗОБРАЖЕНЬ

Для дослідження просторової кореляції шуму використано метод [269], який дозволяє швидко та ефективно оцінити рівень просторової кореляції шумів, присутніх на зображенні. Перевагою цього методу є його незалежність від типу шуму; отже, його можна застосовувати до зображень, які спотворені незалежним від сигналу, залежним від сигналу та змішаним шумом. Відповідно до [269], якщо значення M_h менше 3,75, шум можна вважати просторово некорельованим; якщо M_h більше 3,75 і менше 5,25, шум має середній рівень кореляції; і, нарешті, якщо значення M_h перевищує 5,25, шум в аналізованому зображенні характеризується високим рівнем просторової кореляції.

Результати оцінки M_h для всіх фрагментів зображення зубів представлені на рис. В.1. По горизонтальній осі наведено індекси фрагментів. Червоні пунктирні лінії відповідають пороговим значенням для середнього та високого рівнів кореляції.

Як видно, шум просторово корельований на всіх зображеннях, крім фрагмента №12, де значення M_h виявилося трохи меншим за нижню межу, і шум формально вважався просторово некорельованим. Вочевидь, існують дві групи фрагментів зображення, які мають суттєво різний рівень кореляції шумів, середній для фрагментів №№ 1 – 12 і високий для фрагментів №№ 13 – 20. Ці фрагменти належать до зображень, отриманих з різними параметрами обладнання (різні поля зору, дози радіації тощо). Це повністю підтверджує припущення про залежність рівня кореляції шуму від параметрів отримання зображення та налаштувань обладнання.

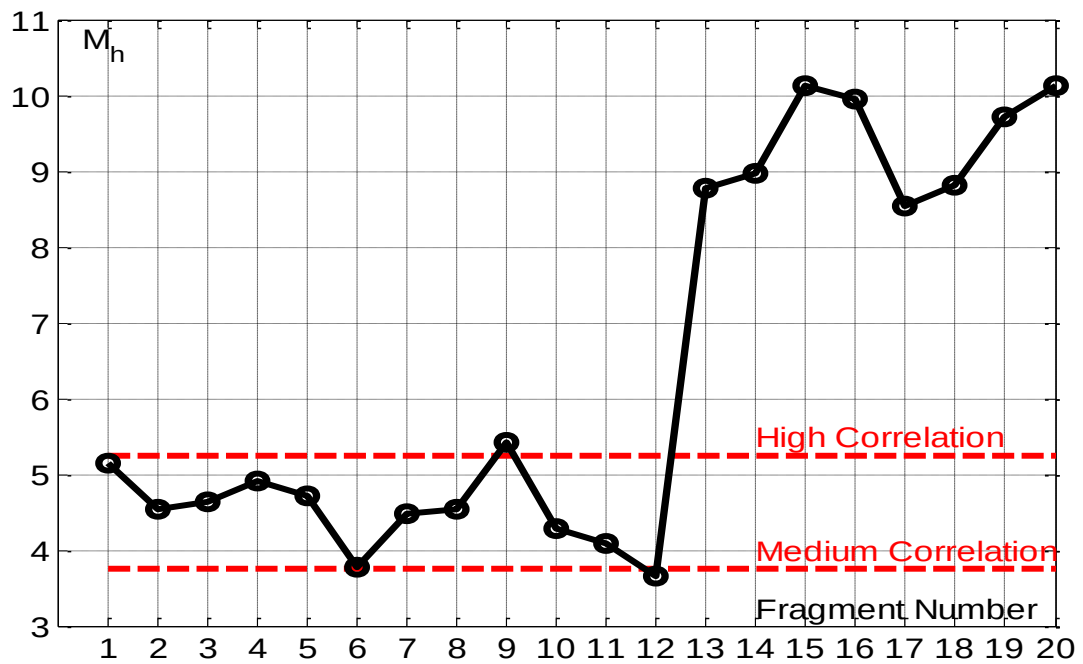


Рисунок В.1 – Результати оцінки рівня просторової кореляції шуму для зображень зубів

Більшість сучасних сліпих методів оцінювання статистичних параметрів шуму або працюють у спектральній області [268], [312], [316], або базуються на використанні нейронних мереж [317]. Спільним недоліком обох груп методів є їхня чутливість до рівня кореляції шуму. Щоб мати можливість застосовувати такі методи для зображень, спотворених просторово корельованим шумом, необхідно усунути просторову кореляцію. Це можна зробити шляхом зменшення дискретизації зображення. Відповідно до [269], для перетворення шуму із середньою кореляцією в некорельований шум необхідне зниження дискретизації в 2 рази, тоді як для висококорельованого шуму необхідне зниження дискретизації в 3 рази. Оскільки ймовірно існування сигнально-залежної складової, але не відомий точний характер цієї залежності, розглянемо ті ж самі моделі шумів (модель 1, модель 2, модель 3), що було використано у розділі 5.1.2.

Для оцінки параметрів цих моделей логічно застосувати метод, заснований на діаграмі розсіювання. Одним із таких методів запропонований у роботі [268]. Ідея методу полягає в тому, щоб отримати діаграму розсіювання локальної дисперсії та локальних оцінок середнього, отриманих для квазіоднорідних блоків

зображень, потім оцінити центри кластерів діаграми розсіювання та використовувати ці опорні точки для підгонки поліноміальної кривої і прийняти її параметри як оцінки характеристик шуму. Особливостями методу [268] є використання детектора на основі аналізу четвертого центрального моменту [316] для виділення блоків, що належать до квазіоднорідних ділянок зображення та застосування методу [316] для оцінки дисперсії шуму в кластери (області зображення з однаковими рівнями інтенсивності). Підгонка поліноміальної кривої забезпечується за допомогою алгоритму подвійного зваженого найменшого середнього квадрата з обмеженнями (DWLMSC) [268]. Приклади карт квазіоднорідних областей, отриманих за допомогою методу [268] для деяких фрагментів зображень зубів, наведених в Додатку М, на рис. М.15, наведені на рис. В.2. Блоки, що належать до виявлених квазіоднорідних областей, позначені сірим кольором.

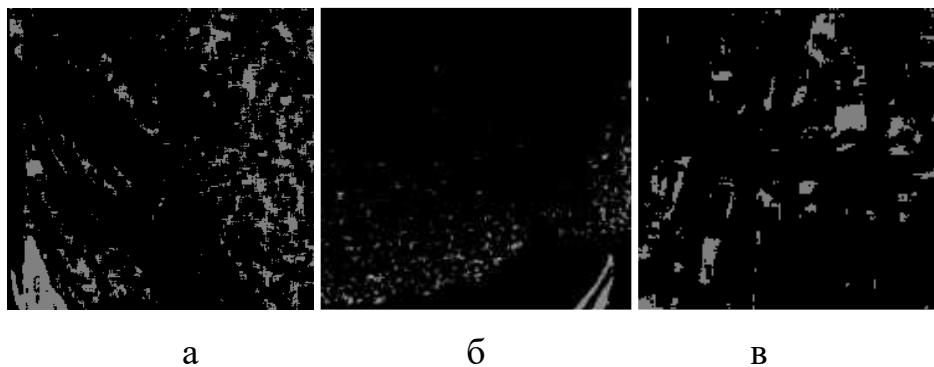


Рисунок В.2 – Приклади карт квазіоднорідних областей, отриманих методом [268] для фрагментів № 2 (а), 12 (б) та 19 (в)

Незважаючи на те, що під час візуального аналізу здавалося, що деякі зображення містили досить багато плоских (квазіоднорідних) ділянок з майже незмінною інтенсивністю, більшість із цих ділянок виявилися градієнтами або текстурою низької інтенсивності, і вони були відкинуті детектором. Відсоток блоків, класифікованих як квазіоднорідні, змінювався залежно від вмісту фрагментів і становив (у середньому) близько 10% від загальної кількості блоків зображення, що перекриваються. На рис. В.3 показані діаграми розсіювання

локальних середніх і оцінок дисперсії для фрагментів зображень зубів, представлених в Додатку М, на рис. М.15, із позначеними центрами кластерів (червоні квадрати) та підігнаними кривими регресії для моделей 1, 2 і 3.

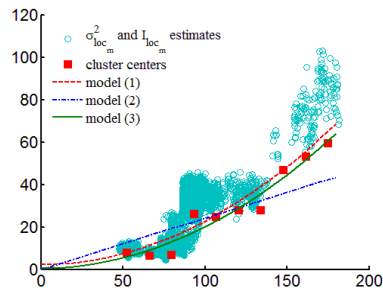
Як видно, дисперсія шуму змінюється в залежності від середнього рівня; отже, шум, вочевидь, є сигнально-залежним. Ще одне спостереження з представлених діаграм розсіювання полягає в тому, що для деяких фрагментів присутні кластери з дисперсіями шуму, близькими до нуля, і такі кластери здебільшого з'являються на середніх рівнях, близьких до 255. Це підтверджує попереднє припущення про наявність ефектів відсікання. Що стосується підігнаних кривих регресії, то ситуація істотно змінюється для різних фрагментів. Вочевидь, що в більшості ситуацій модель 2, що відповідає суміші квазіпуассонівського та адитивного шуму, не дуже узгоджується з отриманими експериментальними даними, хоча згідно з теорією рентгенівського зображення квазіпуассоновський компонент повинен переважати. З двома іншими моделями ситуація також досить неоднозначна. Для деяких фрагментів (наприклад, № 13) модель 3 здається ідеальною, тоді як для інших випадків (фрагменти № 12, 19, 20) вона абсолютно неадекватна. Для ряду ситуацій модель 1 досить добре апроксимує залежність дисперсії від середнього значення. Цікаво, що для деяких фрагментів (наприклад, №2 і 9) моделі 1 і 2 дають дуже схожі результати підгонки; це свідчить про те, що складова, залежна від шуму, переважає.

Деяка оптимізація зображення (контрастування) забезпечується в системі Morita; тому, можливо, ця вбудована обробка призводить до появи деяких додаткових сигнально-залежних компонентів шуму та ускладнює модель шуму.

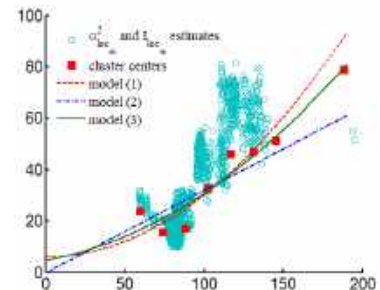
На рис. В.4 наведено результати оцінки параметрів шуму, отримані за методом [268] в припущенні адитивно-мультиплікативної моделі 1. Оцінки адитивної дисперсії шуму наведено на рис. В.4, а, параметру мультиплікативної складової шуму – на рис. В.4, б, по горизонтальній осі позначено індекси фрагментів.

Як видно, оцінки дисперсії адитивного шуму суттєво відрізняються для різних фрагментів; для деяких фрагментів оцінки дисперсії адитивного шуму

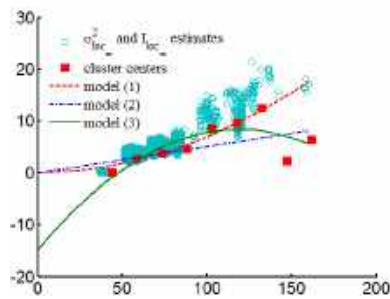
дорівнюють нулю, а в інших ситуаціях вплив цієї складової істотний. Що стосується мультиплікативної компоненти, можна зробити декілька спостережень відповідно до представлених даних.



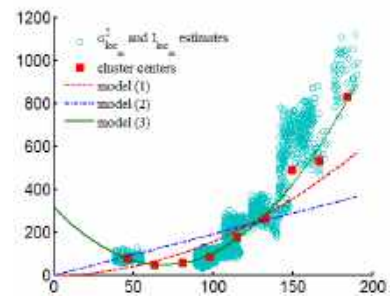
а



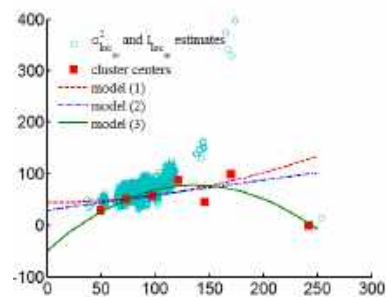
б



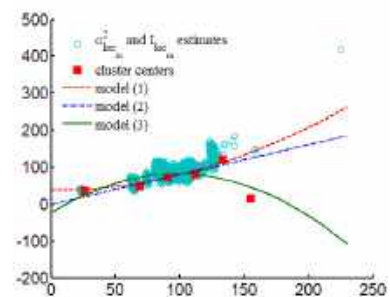
в



г



д

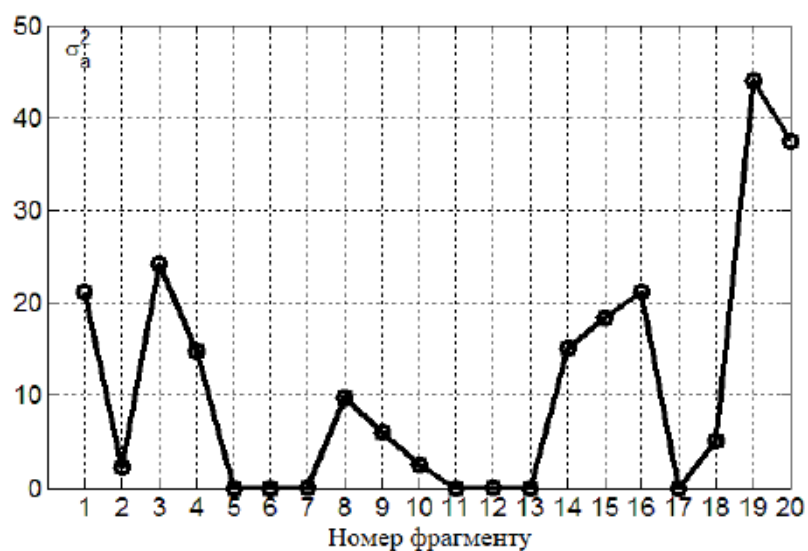


е

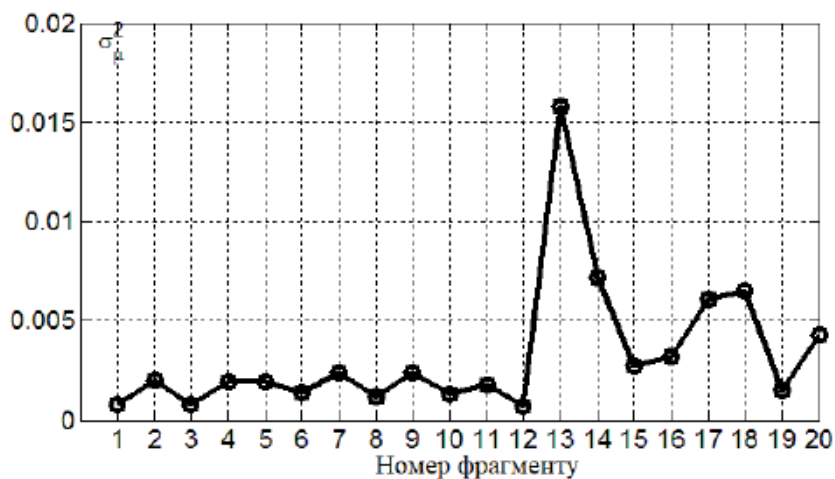
Рисунок В.3 – Діаграми розсіювання локальних оцінок із позначеними центрами кластерів (червоні квадрати) та кривими, підігнаними за припущеннями різних шумових моделей для фрагментів №2 (а), 9 (б), 12 (в), 13 (г), 19 (д), 20 (е)

По-перше, для всіх фрагментів дисперсія мультиплікативного шуму відмінна від нуля. По-друге, для фрагментів №1–12 дисперсії мультиплікативного шуму знаходяться майже на тому самому рівні приблизно 0,0015, тоді як для решти

фрагментів дисперсії, як правило, вищі та змінюються в досить широкому діапазоні. На рис. В.5, а, і В.5, б, показано два зображення зубів з високою роздільною здатністю, отримані в різних умовах. На рис. В.5, в, та В.5, г, наведено карти квазіоднорідних областей для зображень, представлених на рис. В.5, а, та рис. В.5, б.



а



б

Рисунок В.4 – Оцінки параметрів адитивної (а) та мультиплікативної (б) складових дисперсії шуму для фрагментів зображень зубів

Нарешті, рис. В.5, д, та рис. В.5, е, показують діаграми розсіювання локальної дисперсії та середніх оцінок із позначеними центрами кластерів та кривими

регресії, підігнаними відповідно до моделей 1 – 3. Діаграми розсіювання представлених зображень демонструють майже ті ж ефекти, що спостерігалися для фрагментів. Шум є СЗШ і ця залежність досить складна і може істотно відрізнятися від квазіпуассонівського шуму. У фрагментах з низьким середнім рівнем шум можна розглядати як пуассонівський, але для фрагментів із вищими середніми шумова модель трохи відрізняється від нього. Іноді може спостерігатися ефект відсікання, що також слід враховувати.

Щоб забезпечити краще порівняння розглянутих моделей шуму, було розраховано декілька кількісних критеріїв, які традиційно використовуються для оцінки підгонки кривої. Їх значення наведено в таблиці В.1. Як видно, модель 1 демонструє найкращу узгодженість з експериментальними даними за всіма трьома критеріями для зображення на рис. В.5а, тоді як для зображення на рис. В.5, б, модель 3 виявляється найкращою за R^2 та скоригованими R^2 , але програє моделям 1 і 2 за значенням RMSE.

Значення оцінки еквівалентної дисперсії шуму дорівнюють 14,5 для зображення на рис. В.5, а, та 37,76 для зображення на рис. В.5, б. Це відповідає еквівалентному PSNR приблизно 36,52 і 32,36 дБ відповідно.

Таблиця В.1 – Оцінка якості підгонки кривих

Зображення	Модель шуму	Критерії якості підгонки		
		R^2	Скореговане R^2	RMSE
Рисунок В.5, а	Модель 1	0,93	0,92	2,28
	Модель 2	0,84	0,82	3,46
	Модель 3	0,62	0,46	17,59
Рисунок В.5, б	Модель 1	0,82	0,78	3,47
	Модель 2	0,54	0,45	5,44
	Модель 3	0,94	0,92	9,34

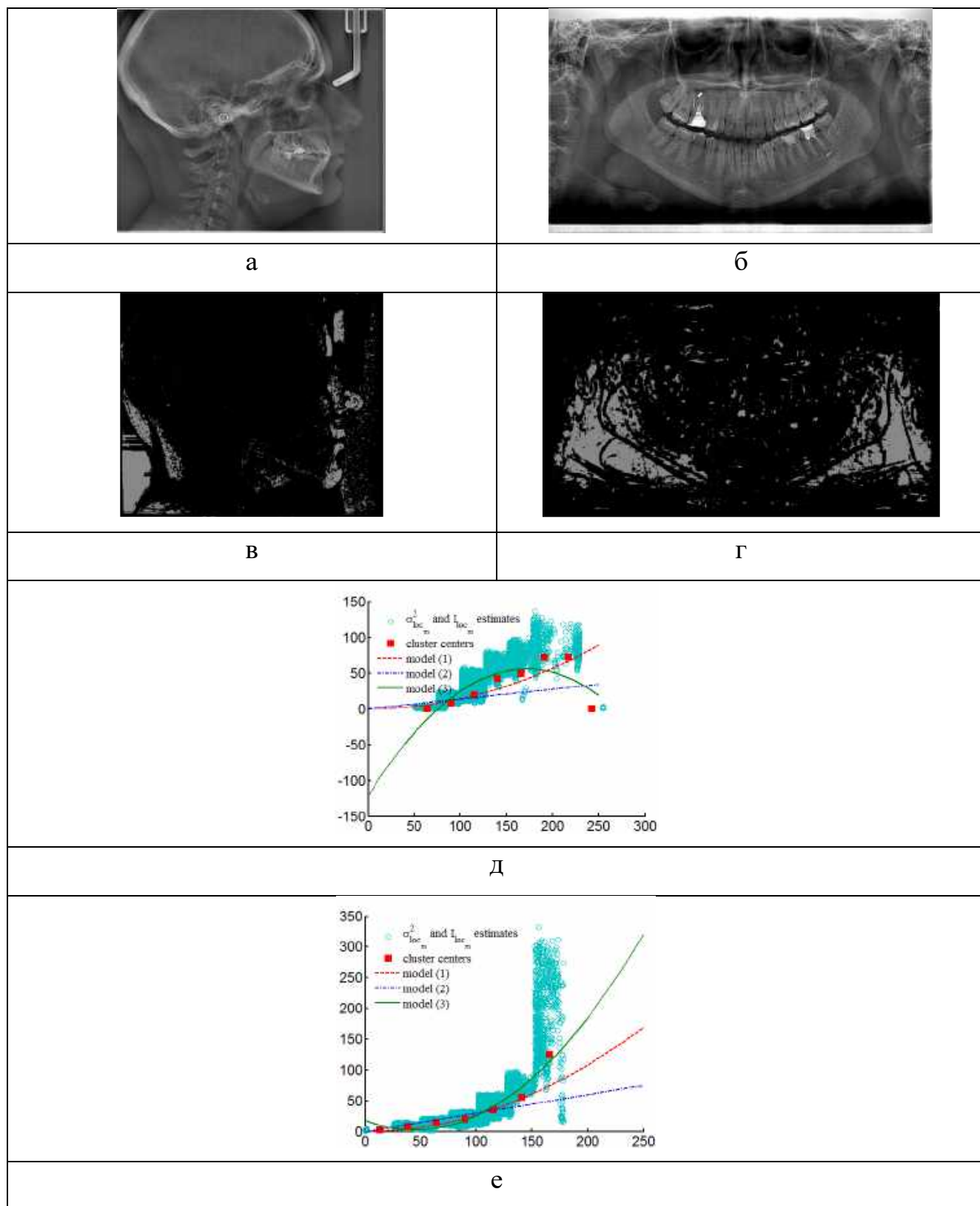


Рисунок В.5 – Стоматологічні зображення високої роздільної здатності, отримані за допомогою системи Morita (а, б), відповідні карти квазіоднорідних областей (в, г), а також локальні середні значення та оцінки дисперсії з позначеними центрами кластерів і лініями регресії, підігнаними відповідно до моделей 1 – 3 (д, е)

Ці значення PSNR ще раз показують, що шум є видимим і, якщо можливо, його потрібно враховувати. Щодо візуального стиснення без втрат, спотворення практично непомітні, якщо внесені спотворення приблизно в 10 разів менші за еквівалентну дисперсію шуму. Це відкриває можливості для автоматичного візуального стиснення без втрат на основі оцінки параметра шуму для даного зображення.

ДОДАТОК Д

СПИСОК ПУБЛІКАЦІЙ ЗДОБУВАЧА ЗА ТЕМОЮ ДИСЕРТАЦІЇ

Праці, в яких опубліковано основні наукові результати дисертації:

1. С. К. Абрамов, Р. А. Кожемякин, С. С. Кривенко, Н. Н. Пономаренко, and В. В. Лукин, “Эффективность фильтрации однозгладовых РСА-изображений при пространственно-коррелированных помехах,” *Радіoeлектронні і комп’ютерні системи*, no. 3, pp. 18–25, 2012.
2. В. В. Лукин, А. В. Наumenко, and С. С. Кривенко, “Адаптивная фильтрация РСА-изображений с использованием нейросетевых детекторов границ,” *Радіoeлектронні і комп’ютерні системи*, no. 4, pp. 121–129, 2012.
3. С. С. Кривенко, Е. О. Колганова, and В. В. Лукин, “Фильтрация изображений с использованием преобразования Хаара в блоках,” *Системи обробки інформації*, no. 8, pp. 63–69, 2013.
4. V. V. Lukin, S. S. Krivenko, E. O. Kolganova, and M. K. Chobanu, “Acceleration and efficiency analysis of dct-based filtering,” *Telecom Rad Eng*, vol. 72, no. 7, pp. 613–626, 2013, doi: 10.1615/TelecomRadEng.v72.i7.50.
5. S. Abramov, S. Krivenko, A. Roenko, V. Lukin, I. Djurovic, and M. Chobanu, “Prediction of filtering efficiency for DCT-based image denoising,” in *2013 2nd Mediterranean Conference on Embedded Computing (MECO)*, Budva, Montenegro: IEEE, Jun. 2013, pp. 97–100. doi: 10.1109/MECO.2013.6601327.
6. D. V. Fevraleev, S. S. Krivenko, V. V. Lukin, R. Marques, and F. D. Medeiros, “Combining level sets and orthogonal transform for despeckling sar images,” *Авиационно-космическая техника и технология*, no. 2, pp. 103–112, 2013.
7. R. A. Kozhemiakin, S. S. Krivenko, V. V. Lukin, R. C. P. Marques, F. N. S. Medeiros, and B. Vozel, “Efficiency analysis of combined despeckling of single-look sar images,” *AKTT*, no. 5, pp. 102–111, 2013.

8. V. V. Lukin et al., “Approaches to automatic data processing in hyperspectral remote sensing,” *Telecom Rad Eng*, vol. 73, no. 13, pp. 1125–1139, 2014, doi: 10.1615/TelecomRadEng.v73.i13.10.

9. S. Krivenko, V. Lukin, B. Vozel, and K. Chehdi, “Filtering efficiency prediction for images corrupted by signal-dependent noise,” in *12th International Conference Modern Problems Of Radio Engineering, Telecommunications, and Computer Science*, 2014, pp. 615–617. Accessed: Feb. 08, 2024. [Online]. Available: <https://hal.science/hal-01111060/>

10. S. Krivenko, V. Lukin, B. Vozel, and K. Chehdi, “Prediction of DCT-based denoising efficiency for images corrupted by signal-dependent noise,” in *2014 IEEE 34th International Scientific Conference on Electronics and Nanotechnology (ELNANO)*, Kyiv, Ukraine: IEEE, Apr. 2014, pp. 254–258. doi: 10.1109/ELNANO.2014.6873920.

11. V. V. Lukin et al., “An approach to prediction of signal-dependent noise removal efficiency by dct-based filter,” *Telecom Rad Eng*, vol. 73, no. 18, pp. 1645–1659, 2014, doi: 10.1615/TelecomRadEng.v73.i18.40.

12. O. S. Rubel, R. O. Kozhemiakin, S. S. Krivenko, V. V. Lukin, B. Vozel, and K. Chehdi, “A method for predicting denoising efficiency for color images,” in *2015 IEEE 35th International Conference on Electronics and Nanotechnology (ELNANO)*, Kyiv, Ukraine: IEEE, Apr. 2015, pp. 304–309. doi: 10.1109/ELNANO.2015.7146897.

13. A. Naumenko, S. Krivenko, V. Lukin, and K. Egiazarian, “Texture region detection by trained neural network,” in *2016 9th International Kharkiv Symposium on Physics and Engineering of Microwaves, Millimeter and Submillimeter Waves (MSMW)*, Kharkiv, Ukraine: IEEE, Jun. 2016, pp. 1–4. doi: 10.1109/MSMW.2016.7538174.

14. A. Naumenko, S. Krivenko, N. Ponomarenko, A. Zelensky, and V. Lukin, “Texture detection in noisy images by combining several local parameters,” in *2015 Second International Scientific-Practical Conference Problems of Infocommunications Science and Technology (PIC S&T)*, Kharkiv, Ukraine: IEEE, Oct. 2015, pp. 230–233. doi: 10.1109/INFOCOMMST.2015.7357321.

15. С. С. Кривенко, А. В. Наumenко, and В. В. Лукин, “Обнаружение текстурных участков SVM-классификатором на изображениях при наличии помех,” *Радіoeлектронні і комп’ютерні системи*, no. 2, pp. 50–57, 2015.

16. A. V. Naumenko, V. V. Lukin, M. S. Zriakhov, and S. S. Krivenko, “Feature aggregation for noisy image to improve ‘texture/non-texture’ classification,” in *2017 IEEE 37th International Conference on Electronics and Nanotechnology (ELNANO)*, Kyiv, Ukraine: IEEE, Apr. 2017, pp. 461–464. doi: 10.1109/ELNANO.2017.7939808.

17. S. K. Abramov, A. Naumenko, V. V. Lukin, S. S. Krivenko, and I. Kaluzhynov, “Neural network based detection of heterogeneities in noisy images,” *Telecom Rad Eng*, vol. 79, no. 19, pp. 1691–1705, 2020, doi: 10.1615/TelecomRadEng.v79.i19.20.

18. B. Vozel *et al.*, “Processing of multichannel Remote-sensing images with prediction of performance parameters,” in *Environmental Applications of Remote Sensing*, IntechOpen, 2016. Accessed: Feb. 08, 2024. [Online]. Available: <https://www.google.com/books?hl=uk&lr=&id=Wm-QDwAAQBAJ&oi=fnd&pg=PA373&dq=Processing+of+Multichannel+Remote-Sensing+Images+with+Prediction+of+Performance+Parameters&ots=ClwhdRf6aV&sig=1jtYRgPSY3irBSaTTqyoJQ-L3n4>

19. V. V. Lukin *et al.*, “Application of filtering efficiency prediction to hyperspectral data pre-processing,” *Ukrainian journal of remote sensing*, no. 7, pp. 4–11, 2015.

20. S. Krivenko, O. Kolganova, and V. Lukin, “Lossy compression of images corrupted by spatially correlated noise,” in *2016 13th International Conference on Modern Problems of Radio Engineering, Telecommunications and Computer Science (TCSET)*, Lviv, Ukraine: IEEE, Feb. 2016, pp. 698–702. doi: 10.1109/TCSET.2016.7452157.

21. M. Zriakhov, S. Kryvenko, and В. Лукін, “Peculiarities of multispectral image lossy compression,” *Вісник Хмельницького національного університету. Серія: Технічні науки*, vol. 331, no. 1, pp. 370–376, Feb. 2024, doi: 10.31891/2307-5732-2024-331-56.

22. V. Lukin, B. Kovalenko, S. Kryvenko, V. Naumenko, and B. Vozel, “Prediction of Optimal Operation Point Existence and Its Parameters in BPG-Based Automatic Lossy Compression of Noisy Images,” in *Current Overview on Science and Technology Research Vol. 9*, Dr. G. Bucci, Ed., B P International (a part of SCIENCEDOMAIN International), 2022, pp. 1–36. doi: 10.9734/bpi/costr/v9/4316A.

23. S. S. Krivenko, V. V. Lukin, and E. Kolganova, “Peculiarities of lossy compression of optical and medical images corrupted by noise,” in *2016 IEEE 36th International Conference on Electronics and Nanotechnology (ELNANO)*, Kyiv, Ukraine: IEEE, Apr. 2016, pp. 251–256. doi: 10.1109/ELNANO.2016.7493060.

24. S. Krivenko, V. Lukin, O. Krylova, and V. Shutko, “Visually Lossless Compression of Retina Images,” in *2018 IEEE 38th International Conference on Electronics and Nanotechnology (ELNANO)*, Kiev: IEEE, Apr. 2018, pp. 255–260. doi: 10.1109/ELNANO.2018.8477459.

25. S. Krivenko, M. Zriakhov, V. Lukin, and B. Vozel, “MSE prediction in DCT-based lossy compression of noise-free and noisy remote sensing,” in *2018 14th International Conference on Advanced Trends in Radioelectronics, Telecommunications and Computer Engineering (TCSET)*, Lviv-Slavske, Ukraine: IEEE, Feb. 2018, pp. 883–888. doi: 10.1109/TCSET.2018.8336338.

26. V. Naumenko, V. Lukin, and S. Kryvenko, “Comparison of Noisy Medical Image Lossy Compression by BPG and AGU,” in *Integrated Computer Technologies in Mechanical Engineering - 2022*, vol. 657, M. Nechyporuk, V. Pavlikov, and D. Kritskiy, Eds., in *Lecture Notes in Networks and Systems*, vol. 657, Cham: Springer Nature Switzerland, 2023, pp. 636–648. doi: 10.1007/978-3-031-36201-9_53.

27. V. Lukin, B. Kovalenko, S. Kryvenko, N. Ponomarenko, K. Egiazarian, and J. Astola, “Compression of Noisy Images Taking into Account Visual Quality: A Comprehensive Study,” *Research Developments in Science and Technology Vol. 8*, pp. 89–109, 2022.

28. B. Kovalenko, V. Lukin, S. Kryvenko, V. Naumenko, and B. Vozel, “Prediction of Parameters in Optimal Operation Point for BPG-based Lossy

Compression of Noisy Images,” *UJRS*, vol. 9, no. 2, pp. 4–12, Jun. 2022, doi: 10.36023/ujrs.2022.9.2.212.

29. B. Kovalenko, V. Lukin, S. Kryvenko, V. Naumenko, and B. Vozel, “BPG-Based Automatic Lossy Compression of Noisy Images with the Prediction of an Optimal Operation Existence and Its Parameters,” *Applied Sciences*, vol. 12, no. 15, p. 7555, Jul. 2022, doi: 10.3390/app12157555.

30. B. Kovalenko, V. Lukin, V. Naumenko, and S. Krivenko, “Analysis of noisy image lossy compression by BPG using visual quality metrics,” in *2021 IEEE 3rd International Conference on Advanced Trends in Information Theory (ATIT)*, Kyiv, Ukraine: IEEE, Dec. 2021, pp. 20–25. doi: 10.1109/ATIT54053.2021.9678575.

31. С. К. Абрамов, Н. В. Бурцев, С. С. Кривенко, А. Н. Земляченко, and В. В. Лукин, “Автоматическое сжатие в окрестности оптимальной рабочей точки изображений с шумом кодерами типа SPIHT и JPEG2000,” *Авиационно-космическая техника и технология*, no. 2, pp. 99–109, 2016.

32. S. Krivenko, V. Lukin, and B. Vozel, “Prediction of Introduced Distortions Parameters in Lossy Image Compression,” in *2018 International Scientific-Practical Conference Problems of Infocommunications. Science and Technology (PIC S&T)*, Kharkiv, Ukraine: IEEE, Oct. 2018, pp. 447–451. doi: 10.1109/INFOCOMMST.2018.8632130.

33. S. Krivenko, M. Zriakhov, V. Lukin, and B. Vozel, “MSE and PSNR prediction for ADCT coder applied to lossy image compression,” in *2018 IEEE 9th International Conference on Dependable Systems, Services and Technologies (DESSERT)*, Kyiv, Ukraine: IEEE, May 2018, pp. 613–618. doi: 10.1109/DESSERT.2018.8409205.

34. S. S. Krivenko, O. Krylova, E. Bataeva, and V. V. Lukin, “Smart lossy compression of images based on distortion prediction,” *Telecom Rad Eng*, vol. 77, no. 17, pp. 1535–1554, 2018, doi: 10.1615/TelecomRadEng.v77.i17.40.

35. V. Lukin, A. Zemliachenko, S. Krivenko, B. Vozel, and K. Chehdi, “Lossy compression of remote sensing images with controllable distortions,” *Satellite Information Classification and Interpretation*, 2018, Accessed: Feb. 11, 2024. [Online]. Available:

https://www.google.com/books?hl=uk&lr=&id=lzj8DwAAQBAJ&oi=fnd&pg=PA27&dq=Lossy+compression+of+remote+sensing+images+with+controllable+distortions&ots=irka3mNKud&sig=ljJ_EiGSZkYlW4tNWjpzfsvU3dg

36. S. Krivenko, V. Lukin, and O. Krylova, “Visually Lossless Compression of Dental Images,” in *2019 IEEE 39th International Conference on Electronics and Nanotechnology (ELNANO)*, Kyiv, Ukraine: IEEE, Apr. 2019, pp. 394–399. doi: 10.1109/ELNANO.2019.8783218.

37. S. Krivenko, M. Zriakhov, N. Kussul, and V. Lukin, “Prediction of Visual Quality for Lossy Compressed Images,” in *2019 IEEE 15th International Conference on the Experience of Designing and Application of CAD Systems (CADSM)*, Polyana, Ukraine: IEEE, Feb. 2019, pp. 1–5. doi: 10.1109/CADSM.2019.8779266.

38. O. Krylova, L. Kryvenko, S. Krivenko, and V. Lukin, “A Fast Noniterative Visually Lossless Compression of Dental Images Using AGU-M Coder,” in *2021 IEEE 16th International Conference on the Experience of Designing and Application of CAD Systems (CADSM)*, Lviv, Ukraine: IEEE, Feb. 2021, pp. 6–10. doi: 10.1109/CADSM52681.2021.9385257.

39. S. Krivenko, V. Lukin, O. Krylova, L. Kryvenko, and K. Egiazarian, “A Fast Method of Visually Lossless Compression of Dental Images,” *Applied Sciences*, vol. 11, no. 1, Art. no. 1, Jan. 2021, doi: 10.3390/app11010135.

40. S. S. Krivenko, S. K. Abramov, V. V. Lukin, B. Vozel, and K. Chehdi, “Lossy DCT-based compression of remote sensing images with providing a desired visual quality,” in *Image and Signal Processing for Remote Sensing XXV*, L. Bruzzone, F. Bovolo, and J. A. Benediktsson, Eds., Strasbourg, France: SPIE, Oct. 2019, p. 36. doi: 10.1117/12.2532726.

41. F. Li, S. Krivenko, and V. Lukin, “A Modified Algorithm of Visual Quality Providing in Lossy Compression of Remote Sensing Images,” in *2021 IEEE International Conference on Information and Telecommunication Technologies and Radio Electronics (UkrMiCo)*, Odesa, Ukraine: IEEE, Nov. 2021, pp. 16–21. doi: 10.1109/UkrMiCo52950.2021.9716684.

42. S. Krivenko, D. Demchenko, I. Dyogtev, and V. Lukin, “A Two-Step Approach to Providing a Desired Quality of Lossy Compressed Images,” in *Integrated Computer Technologies in Mechanical Engineering*, vol. 1113, M. Nechyporuk, V. Pavlikov, and D. Kritskiy, Eds., in *Advances in Intelligent Systems and Computing*, vol. 1113. , Cham: Springer International Publishing, 2020, pp. 482–491. doi: 10.1007/978-3-030-37618-5_41.

43. S. Krivenko, F. Li, V. Lukin, B. Vozel, and O. Krylova, “Prediction of Visual Quality Metrics in Lossy Image Compression,” in *2020 IEEE 40th International Conference on Electronics and Nanotechnology (ELNANO)*, Kyiv, Ukraine: IEEE, Apr. 2020, pp. 478–483. doi: 10.1109/ELNANO50318.2020.9088819.

44. F. Li, S. Krivenko, and V. Lukin, “A Two-step Procedure for Image Lossy Compression by ADCTC With a Desired Quality,” in *2020 IEEE 11th International Conference on Dependable Systems, Services and Technologies (DESSERT)*, Kyiv, Ukraine: IEEE, May 2020, pp. 307–312. doi: 10.1109/DESSERT50317.2020.9125000.

45. F. Li, S. S. Krivenko, and V. V. Lukin, “Analysis of two-step approach for compressing texture images with desired quality,” AKTT, no. 1, pp. 50–58, Jan. 2020, doi: 10.32620/aktt.2020.1.08.

46. V. Lukin et al., “Lossy Compression of Multichannel Remote Sensing Images with Quality Control,” Remote Sensing, vol. 12, no. 22, p. 3840, Nov. 2020, doi: 10.3390/rs12223840.

Праці, що підтверджують апробацію матеріалів дисертації:

47. F. Li, S. Krivenko, and V. Lukin, “Adaptive two-step procedure of providing desired visual quality of compressed image,” in *Proceedings of the 2020 4th International Conference on Electronic Information Technology and Computer Engineering*, Xiamen China: ACM, Nov. 2020, pp. 407–414. doi: 10.1145/3443467.3443791.

48. F. Li, S. Krivenko, and V. Lukin, “An Approach to Better Portable Graphics (BPG) Compression with Providing a Desired Quality,” in *2020 IEEE 2nd International Conference on Advanced Trends in Information Theory (ATIT)*, Kyiv, Ukraine: IEEE, Nov. 2020, pp. 13–17. doi: 10.1109/ATIT50783.2020.9349289.

49. F. Li, S. Krivenko, and V. Lukin, “A Fast Method for Visual Quality Prediction and Providing in Image Lossy Compression by SPIHT,” in *Integrated Computer Technologies in Mechanical Engineering - 2020*, vol. 188, M. Nechyporuk, V. Pavlikov, and D. Kritskiy, Eds., in Lecture Notes in Networks and Systems, vol. 188. , Cham: Springer International Publishing, 2021, pp. 17–29. doi: 10.1007/978-3-030-66717-7_2.

50. F. Li, S. Krivenko, and V. Lukin, “Two-step providing of desired quality in lossy image compression by SPIHT,” *Радіоелектронні і комп’ютерні системи*, no. 2, pp. 22–32, Apr. 2020, doi: 10.32620/reks.2020.2.02.

51. F. Li, S. Krivenko, and V. Lukin, “An Automatic Optimization Method for BPG Compression Based on Visual Perception,” in *Mathematical Modeling and Simulation of Systems*, vol. 344, S. Shkarlet, A. Morozov, A. Palagin, D. Vinnikov, N. Stoianov, M. Zhelezniak, and V. Kazymyr, Eds., in Lecture Notes in Networks and Systems, vol. 344. , Cham: Springer International Publishing, 2022, pp. 213–225. doi: 10.1007/978-3-030-89902-8_17.

52. A. Kharkov, V. Oliinyk, V. V. Lukin, and S. S. Krivenko, “Blind estimation of noise variance for 1d signal denoising,” *Telecom Rad Eng*, vol. 79, no. 7, pp. 567–581, 2020, doi: 10.1615/TelecomRadEng.v79.i7.30.

53. V. Abramova, S. Krivenko, V. Lukin, and O. Krylova, “Analysis of Noise Properties in Dental Images,” in *2020 IEEE 40th International Conference on Electronics and Nanotechnology (ELNANO)*, Kyiv, Ukraine: IEEE, Apr. 2020, pp. 511–515. doi: 10.1109/ELNANO50318.2020.9088768.

54. M. Ponomarenko, O. Miroshnichenko, V. Lukin, S. Krivenko, and K. Egiazarian, “Blind denoising of dental X-ray images,” *ei*, vol. 35, no. 9, pp. 299-1-299–6, Jan. 2023, doi: 10.2352/EI.2023.35.9.IPAS-299.

55. V. Abramova, V. Lukin, S. Abramov, S. Kryvenko, P. Lech, and K. Okarma, “A Fast and Accurate Prediction of Distortions in DCT-Based Lossy Image Compression,” *Electronics*, vol. 12, no. 11, p. 2347, May 2023, doi: 10.3390/electronics12112347.

56. V. Lukin, S. Krivenko, I. Kaluzhinov, O. Krylova, and L. Kryvenko, “Lossy Compression of Remote Sensing and Dental Images Corrupted by Spatially Correlated

Noise,” in *Integrated Computer Technologies in Mechanical Engineering - 2021*, vol. 367, M. Nechyporuk, V. Pavlikov, and D. Kritskiy, Eds., in *Lecture Notes in Networks and Systems*, vol. 367. , Cham: Springer International Publishing, 2022, pp. 1003–1014. doi: 10.1007/978-3-030-94259-5_77.

57. G. Proskura, O. Rubel, S. Kryvenko, and V. Lukin, “On classifier performance for remote sensing images compressed by different coders,” *aktt*, no. 1, pp. 67–77, Feb. 2023, doi: 10.32620/aktt.2023.1.07.

58. V. Lukin, S. Abramov, S. Krivenko, A. Kurekin, and O. Pogrebnyak, “Analysis of classification accuracy for pre-filtered multichannel remote sensing data,” *Expert Systems with Applications*, vol. 40, no. 16, pp. 6400–6411, Nov. 2013, doi: 10.1016/j.eswa.2013.05.061.

59. V. Lukin, F. Li, G. Proskura, S. Kryvenko, and B. Vozel, “Peculiarities of Classification of Lossy Compressed Multichannel Remote Sensing Images Using Trained Neural Networks,” in *Digital Ecosystems: Interconnecting Advanced Networks with AI Applications*, vol. 1198, A. Luntovskyy, M. Klymash, I. Melnyk, M. Beshley, and A. Schill, Eds., in *Lecture Notes in Electrical Engineering*, vol. 1198. , Cham: Springer Nature Switzerland, 2024, pp. 127–143. doi: 10.1007/978-3-031-61221-3_7.

60. V. Lukin, F. Li, J. Zhu, and S. Kryvenko, “Peculiarities of SVM-Based Classification of BPG Compressed Three-Channel Images,” in *Proceedings of the 8th International Symposium of Space Optical Instruments and Applications*, vol. 191, H. P. Urbach, D. Li, and D. Yu, Eds., in *Springer Proceedings in Physics*, vol. 191. , Singapore: Springer Nature Singapore, 2024, pp. 352–365. doi: 10.1007/978-981-97-6718-2_31.

Праці, які додатково відображають наукові результати дисертації:

61. L. Kryvenko, O. Krylova, V. Lukin, and S. Kryvenko, “Intelligent visually lossless compression of dental images,” *Adv. Opt. Technol.*, vol. 13, p. 1306142, Feb. 2024, doi: 10.3389/aot.2024.1306142.

62. L. Kryvenko, O. Krylova, S. Krivenko, and V. Lukin, “A Fast Noniterative Visually Lossless Compression of Dental Images Using BPG Coder,” in *2022 11th*

Mediterranean Conference on Embedded Computing (MECO), Budva, Montenegro: IEEE, Jun. 2022, pp. 1–6. doi: 10.1109/MECO55406.2022.9797114.

63. B. Bondžulić, V. Lukin, D. Bujaković, F. Li, S. Kryvenko, and B. Pavlović, “On Visually Lossless JPEG Image Compression,” in *2023 Zooming Innovation in Consumer Technologies Conference (ZINC)*, Novi Sad, Serbia: IEEE, May 2023, pp. 113–118. doi: 10.1109/ZINC58345.2023.10174090.

64. V. Abramova, S. Kryvenko, S. Abramov, V. Makarichev, and V. Lukin, “A Fast Procedure for Image Lossy Compression by ADCTC Using Prediction of Distortions’ MSE,” in *Integrated Computer Technologies in Mechanical Engineering - 2022*, vol. 657, M. Nechyporuk, V. Pavlikov, and D. Kritskiy, Eds., in *Lecture Notes in Networks and Systems*, vol. 657. , Cham: Springer Nature Switzerland, 2023, pp. 625–635. doi: 10.1007/978-3-031-36201-9_52.

65. F. Li, S. Kryvenko, and V. Lukin, “Lossy Compression of Three-Channel Remote Sensing Images Based on Controlled Visual Quality,” in *Integrated Computer Technologies in Mechanical Engineering - 2022*, vol. 657, M. Nechyporuk, V. Pavlikov, and D. Kritskiy, Eds., in *Lecture Notes in Networks and Systems*, vol. 657. , Cham: Springer Nature Switzerland, 2023, pp. 649–661. doi: 10.1007/978-3-031-36201-9_54.

66. V. Lukin, S. Krivenko, F. Li, S. Abramov, and V. Makarichev, “On Image Complexity in Viewpoint of Image Processing Performance,” in *2022 3rd International Workshop on Intelligent Information Technologies and Systems of Information Security*, 2022, pp. 458–473. Accessed: Feb. 12, 2024. [Online]. Available: <https://ceur-ws.org/Vol-3156/paper35.pdf>

67. B. Bondžulić, N. Stojanović, V. Lukin, S. A. Stankevich, D. Bujaković, and S. Kryvenko, “Target acquisition performance in the presence of JPEG image compression,” *Defence Technology*, p. S2214914723003306, Dec. 2023, doi: 10.1016/j.dt.2023.12.006.

68. F. Li, S. Kryvenko, and V. Lukin, “Remote Sensing Image Lossy Compression Based on JPEG with Controlled Visual Quality,” in *Proceedings of the 7th International Symposium of Space Optical Instruments and Applications*, vol. 295, H. P. Urbach and

H. Jiang, Eds., in Springer Proceedings in Physics, vol. 295. , Singapore: Springer Nature Singapore, 2023, pp. 8–19. doi: 10.1007/978-981-99-4098-1_2.

69. S. Kryvenko and V. Lukin, “Peculiarities of noisy image lossy compression,” Вісник Хмельницького національного університету. Серія: Технічні науки, vol. 333, no. 2, pp. 278–283, Apr. 2024, doi: 10.31891/2307-5732-2024-333-2-44.

70. B. Bondžulić, N. Stojanović, V. Lukin, and S. Kryvenko, “JPEG and BPG visually lossless image compression via KonJND-1k database,” Vojnotehnički glasnik, vol. 72, no. 3, pp. 1214–1241, 2024, doi: 10.5937/vojtehg72-50300.

71. F. Li, V. Lukin, S. Kryvenko, B. Bondzulich, D. Bujakovic, and B. Pavlovic, “Strange Images in Remote Sensing and Their Properties,” UJRS, vol. 10, no. 2, pp. 12–18, Jun. 2023, doi: 10.36023/ujrs.2023.10.2.240.

ДОДАТОК Е
ДОДАТКОВІ МАТЕРІАЛИ ДО РОЗДІЛУ 1

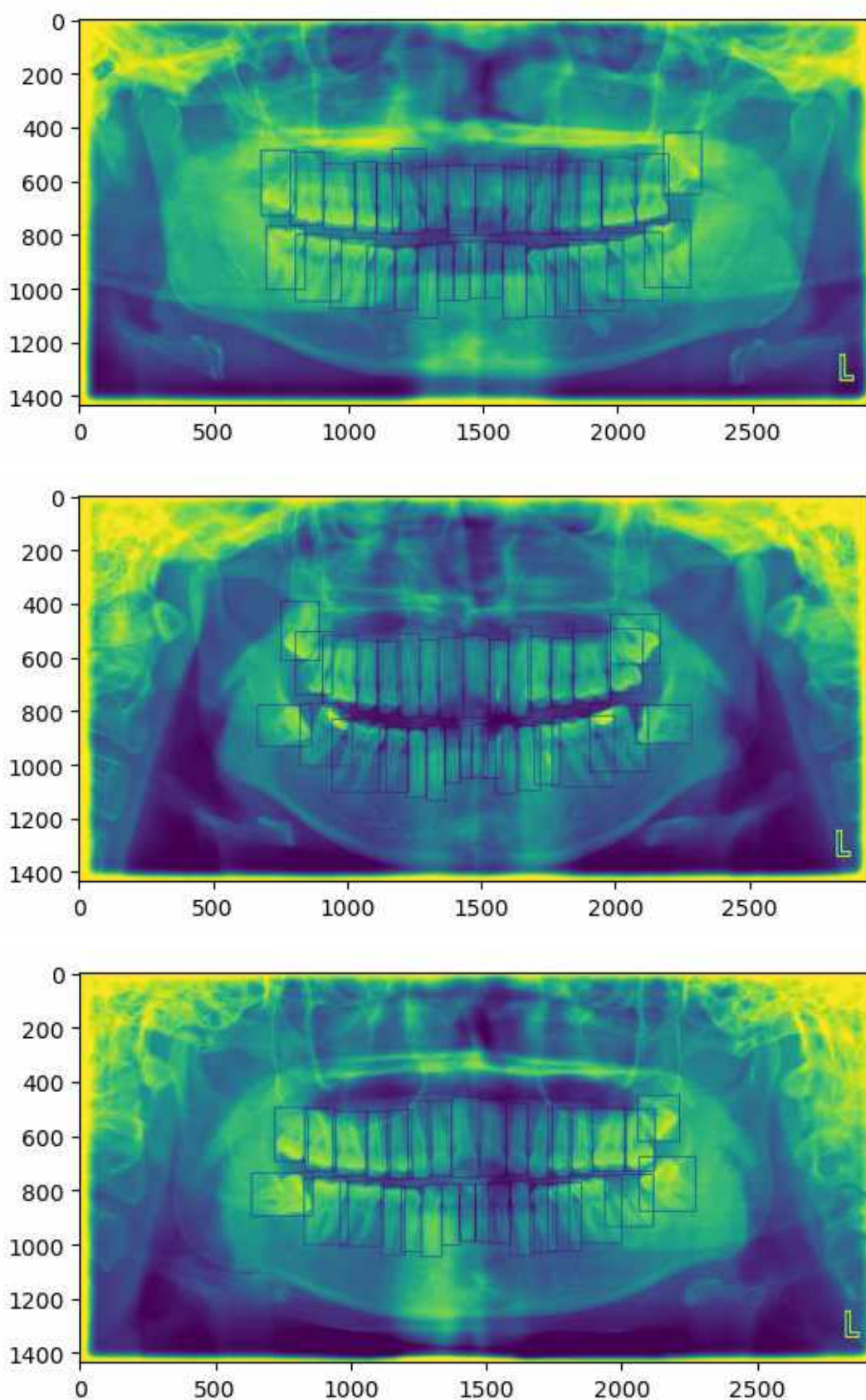


Рисунок Е.1 – Панорамні сегментовані зображення зубів



Рисунок Е.2 – Приклад зображення з датасету Chikusei

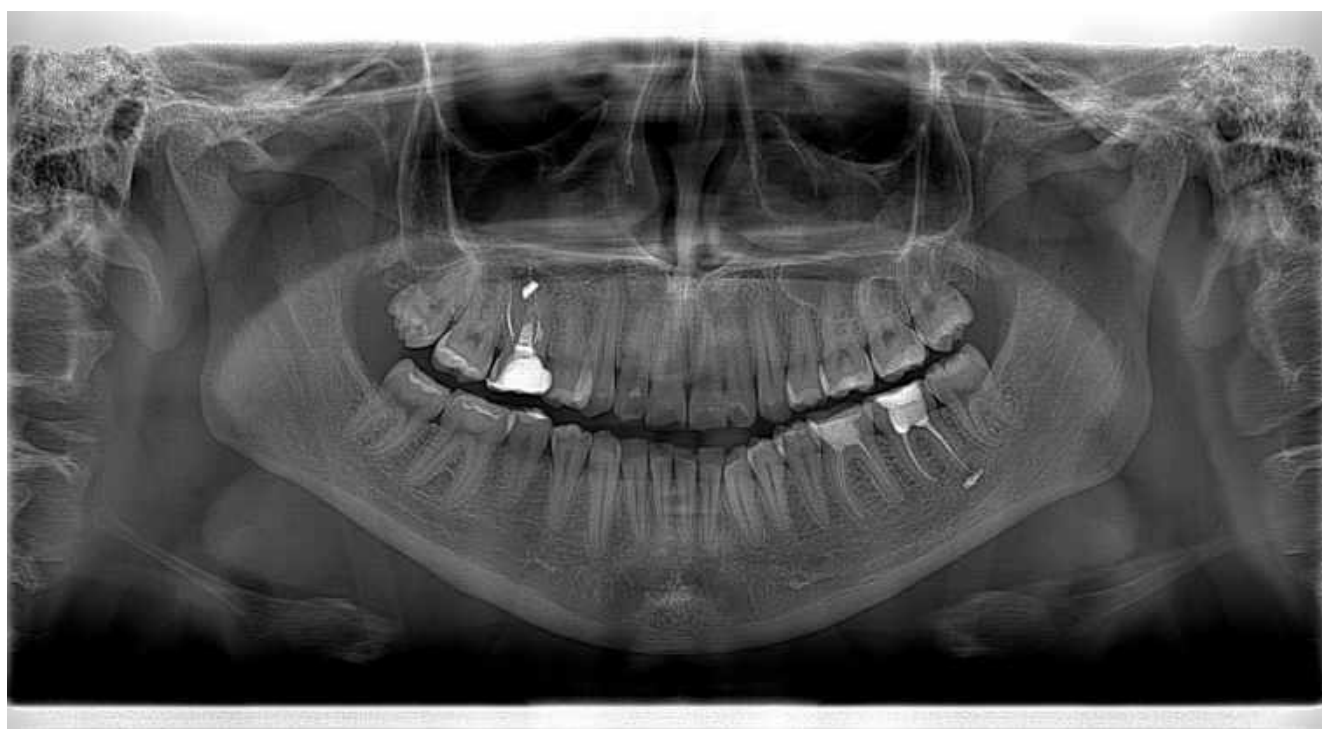
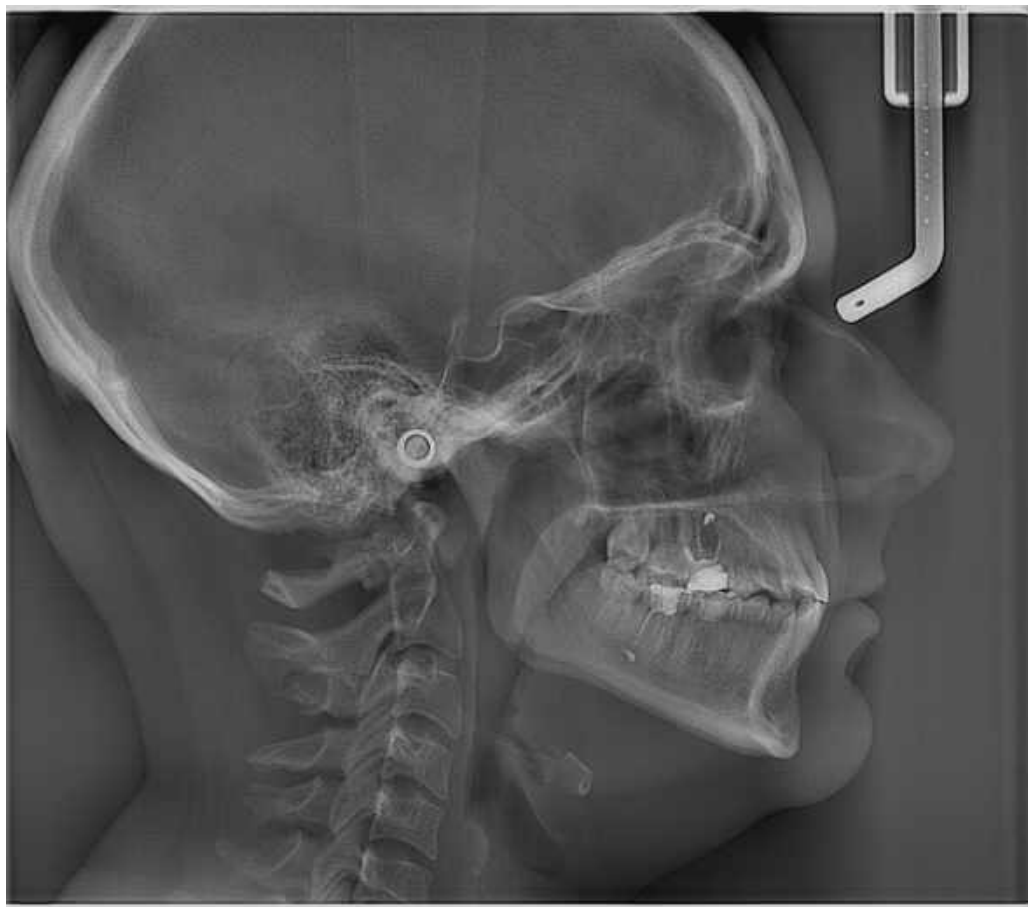


Рисунок Е.3 – Тестові зображення Dental1 та Dental2

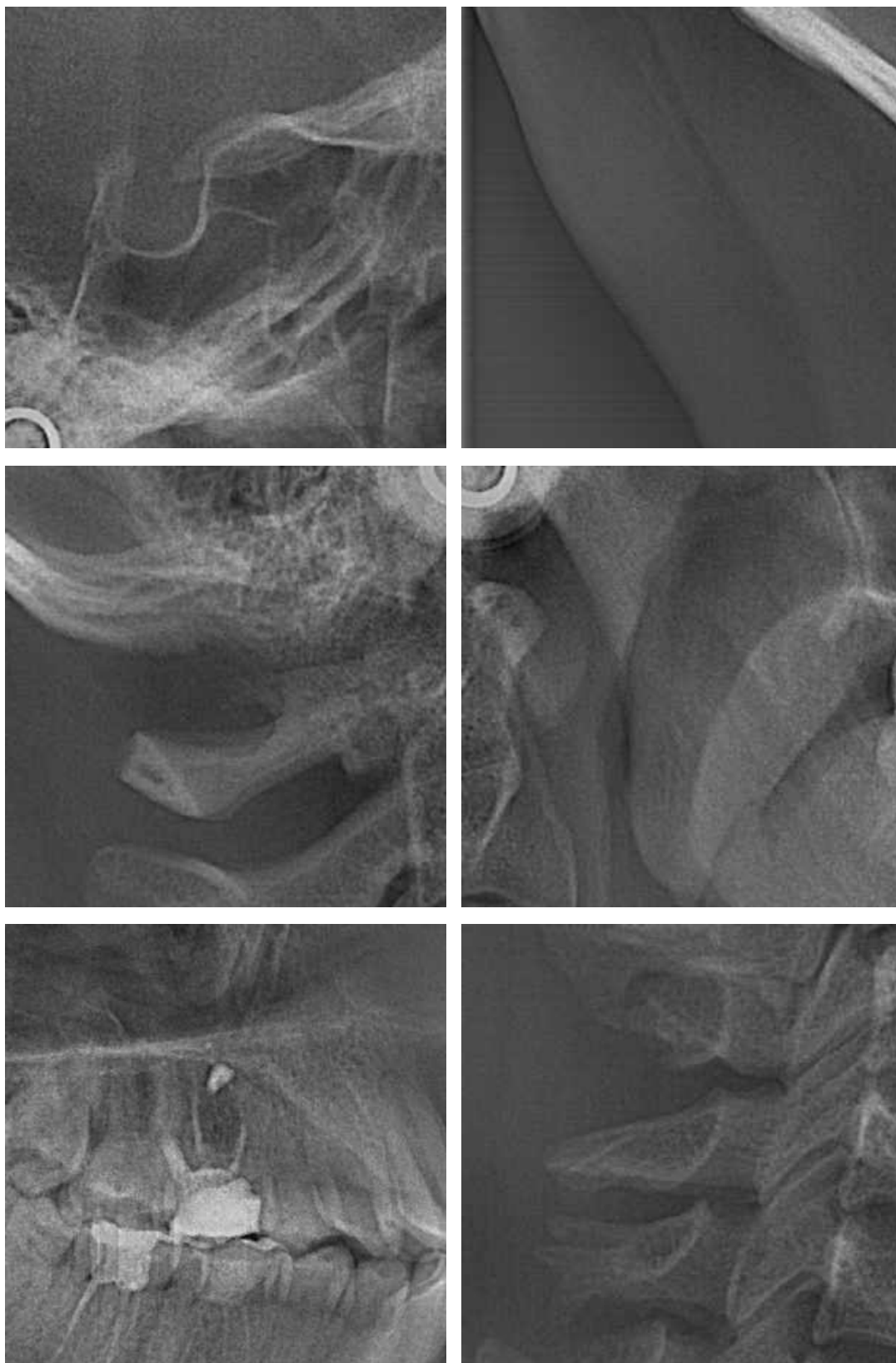


Рисунок Е.4 – Збільшені фрагменти тестових зображень Dental1 та Dental2



Рисунок Е.5 – Реальне РСА-зображення ДЗ великого розміру та його збільшені фрагменти

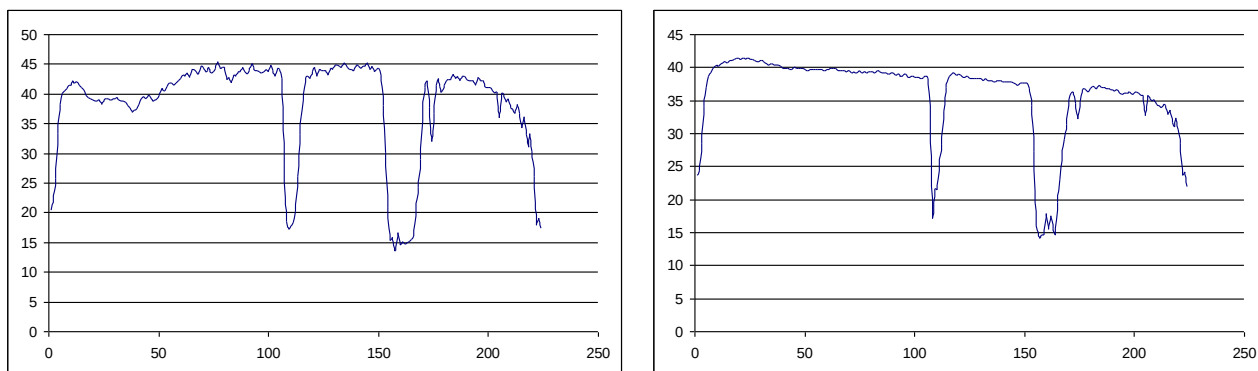


Рисунок Е.6 – Оцінені PSNR для піддіапазонів для двох зображень (система AVIRIS), 224 канали



Рисунок Е.7 – Приклад цифрового зображення, сформованого системою ДЗ міліметрового діапазону хвиль

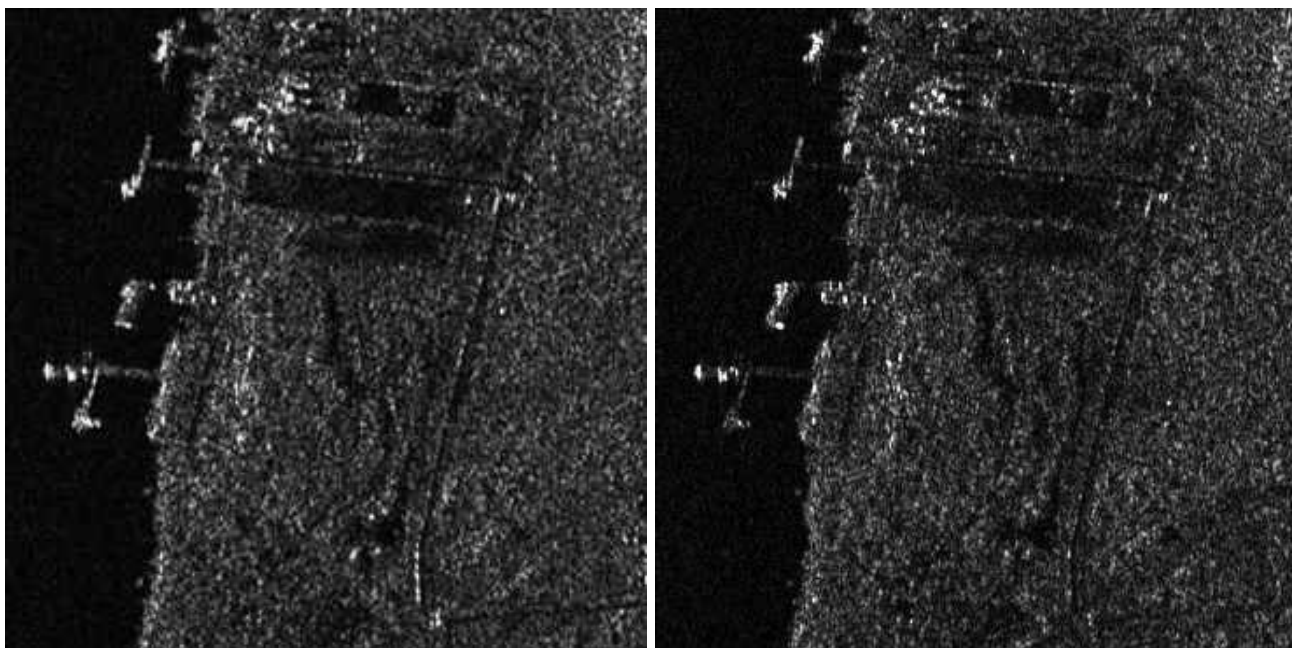


Рисунок Е.8 – Фрагменти спотвореного Релєєвським шумом цифрового зображення, сформованого системою ДЗ TerraSAR-X (два останні фрагменти – вертикальна та горизонтальна поляризація)



Рисунок Е.9 – Зображення Леппа, спотворене спекл-шумом з дисперсією $\sigma_{\mu}^2 = 0,04$



Рисунок Е.10 – Зображення House, спотворене гаусовим та пуассонівським шумами

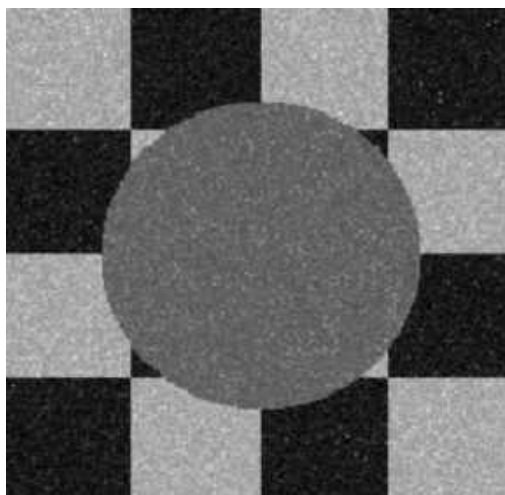


Рисунок Е.11 – Зображення, спотворене Релеєвським шумом

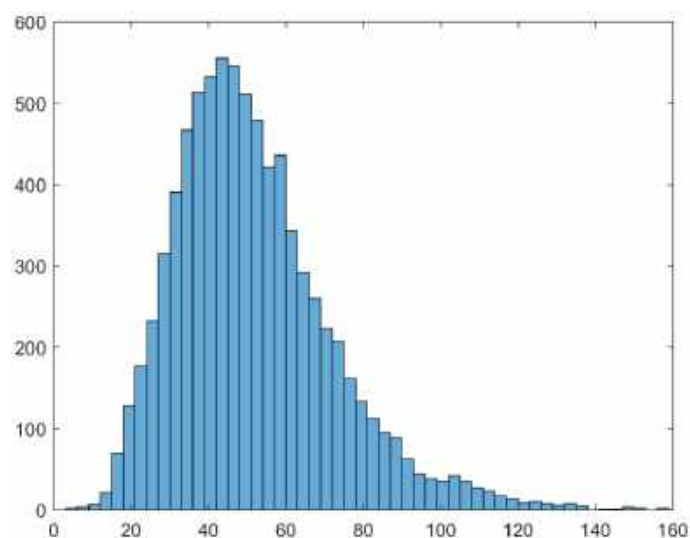
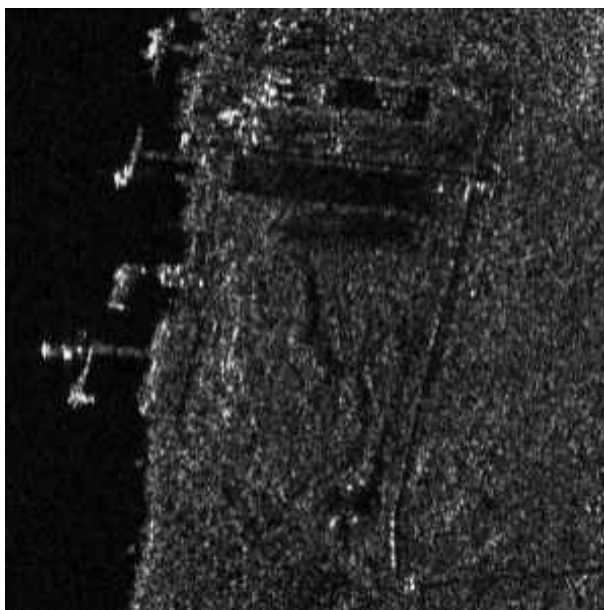


Рисунок Е.12 – Приклад однопоглядового РСА зображення, отриманого системою TerraSAR-X і гістограми однорідної ділянки цього зображення

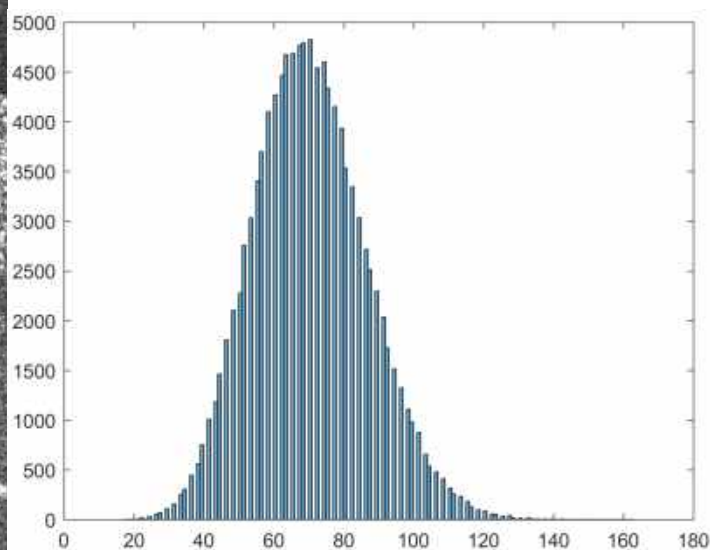
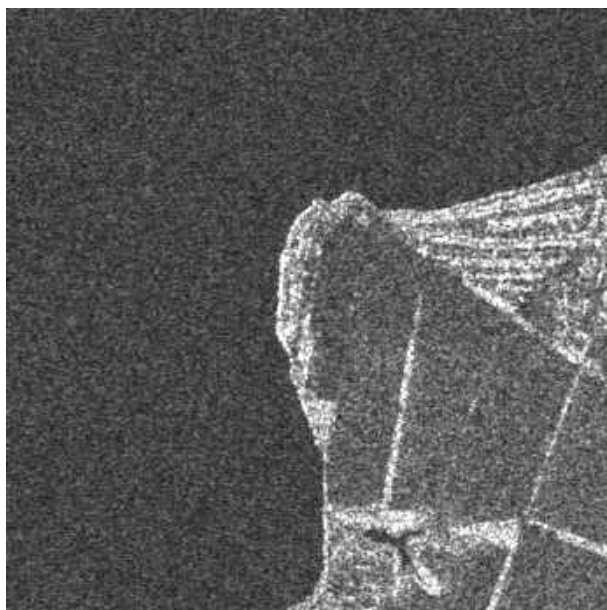


Рисунок Е.13 – Приклад фрагменту багатопоглядового РСА зображення, отриманого системою Sentinel-1, і гістограми однорідної ділянки цього зображення

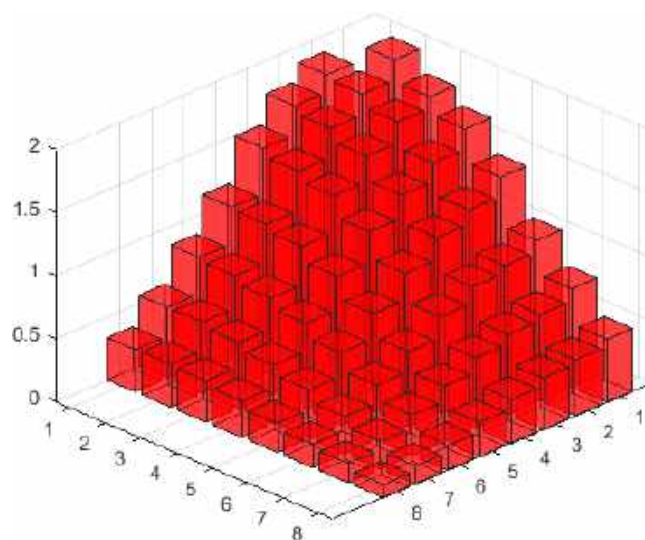
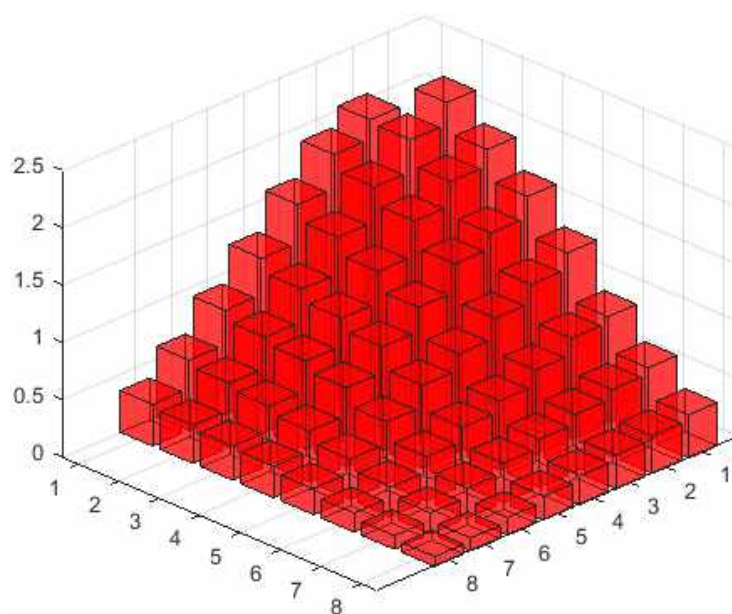
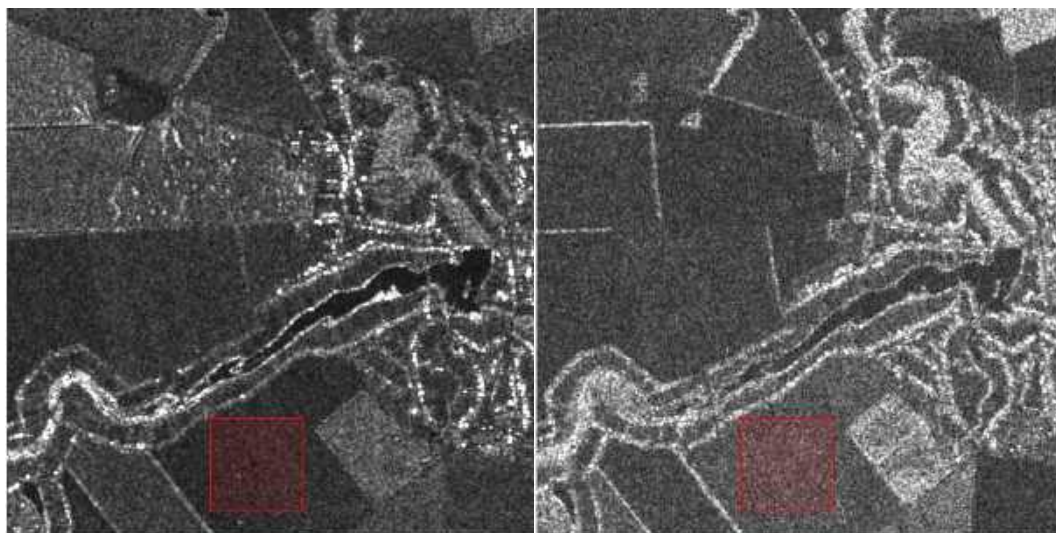
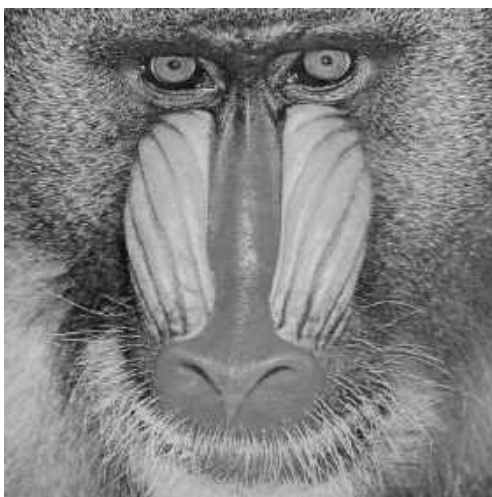


Рисунок Е.14 – Приклади РСА зображень системи Sentinel-1 і відповідних двовимірних ДКП-спектрів для вертикальної (зліва та зверху) та горизонтальної поляризацій



Baboon



Barbara



Goldhill



Lenna



Peppers



MRT_Prepared

Рисунок Е.15 – Тестові оптичні зображення (включно з одним медичним)



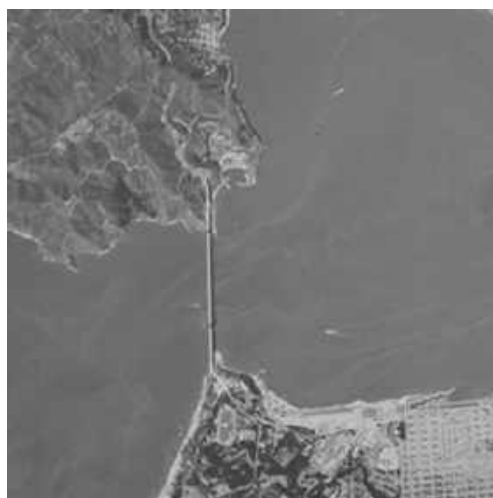
Aerial



Airfield



Diego



Frisco



FR01



FR02

Рисунок Е.16 – Тестові зображення ДЗЗ

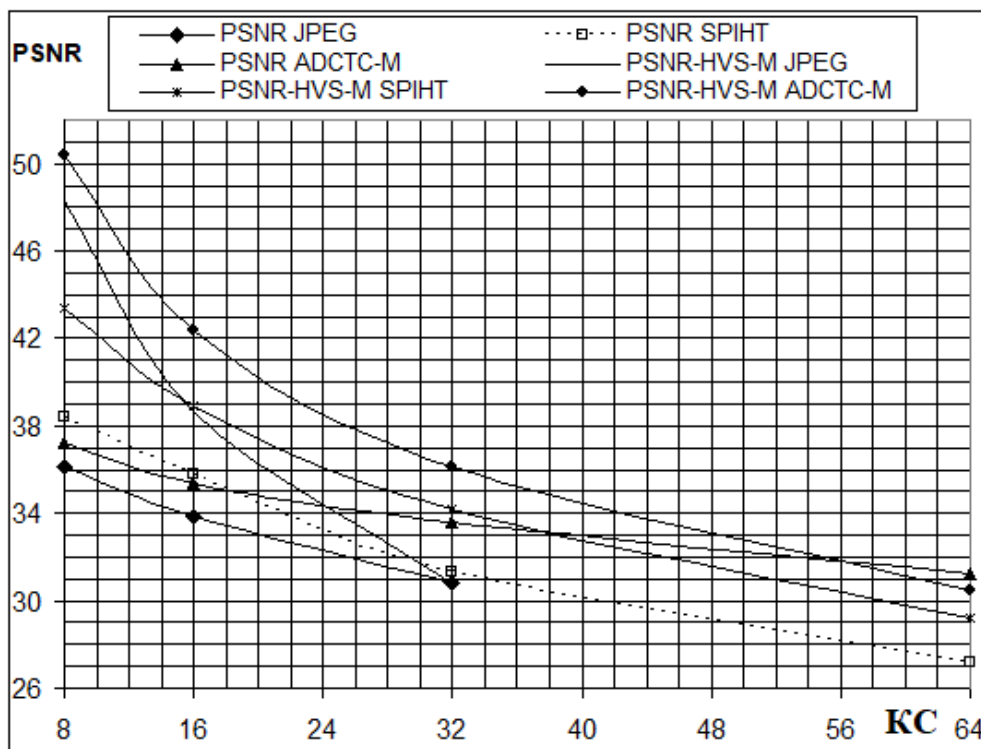


Рисунок Е.17 – Залежності PSNR і PSNR-HVS-M від КС для трьох кодерів для тестового зображення Peppers

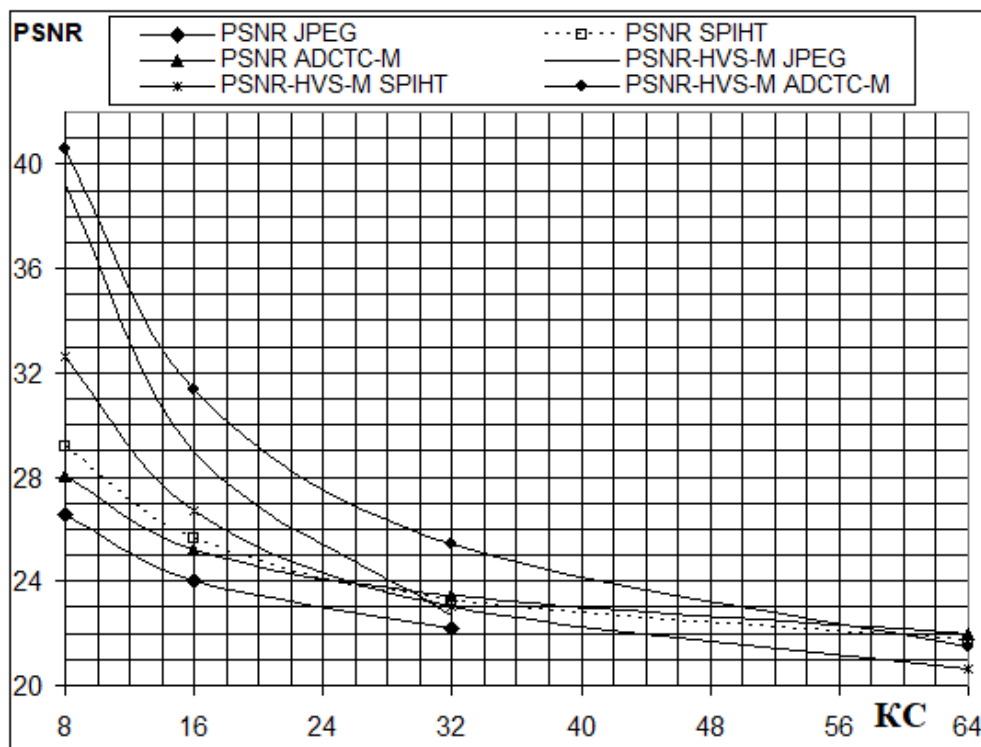


Рисунок Е.18 – Залежності PSNR і PSNR-HVS-M від КС для трьох кодерів для тестового зображення Baboon

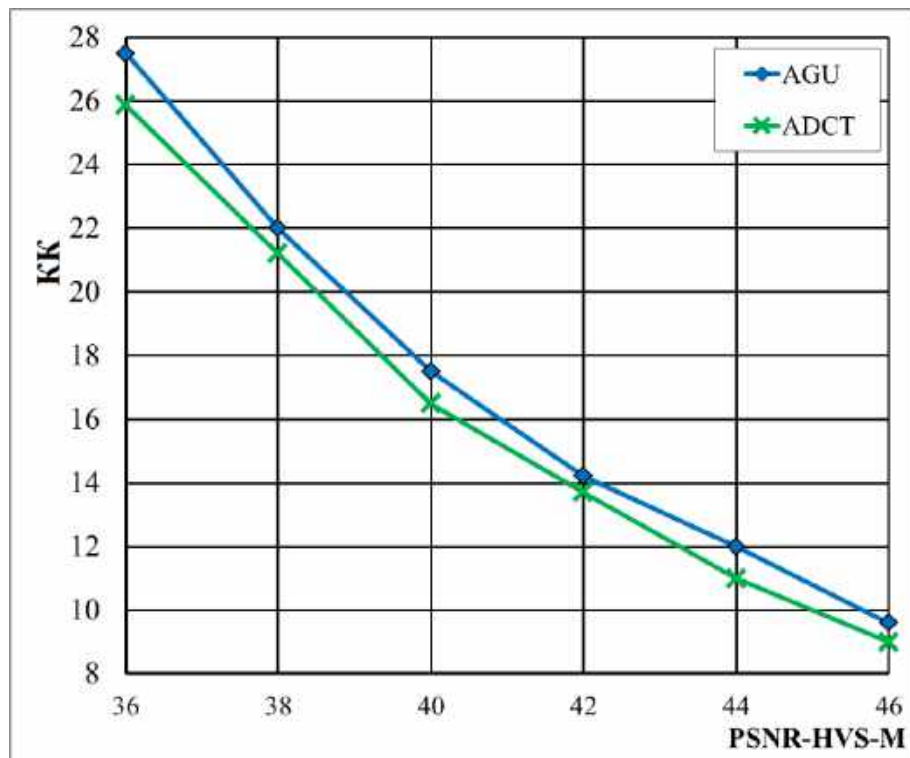


Рисунок Е.19 – Усереднена (для багатьох 8-бітних тестових зображень) залежність КК від PSNR-HVS-M для кодерів AGU і ADCT

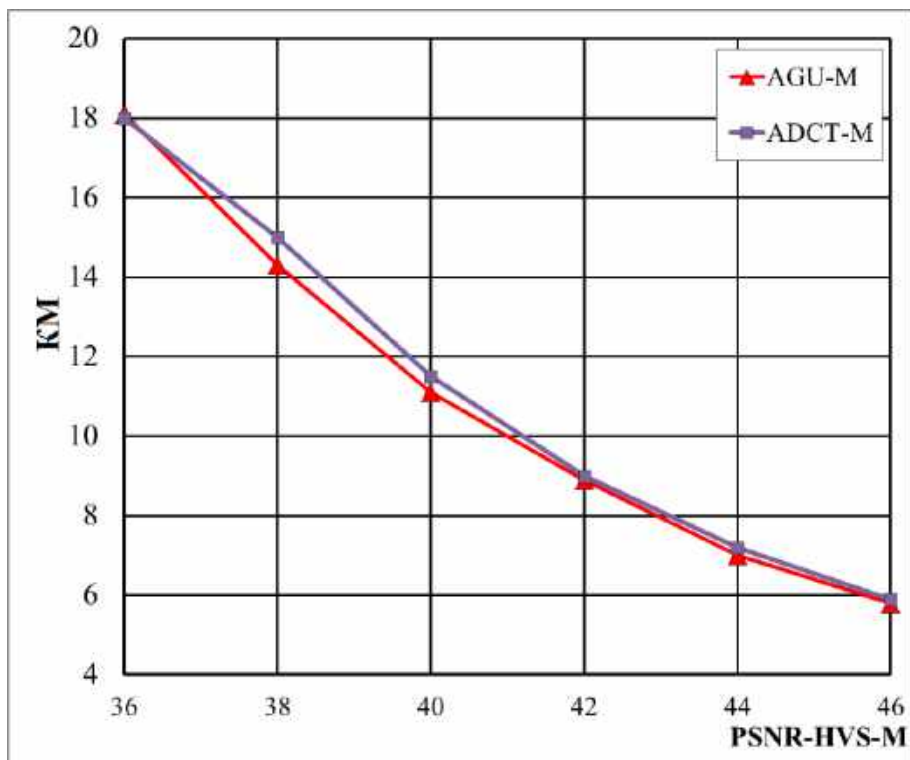
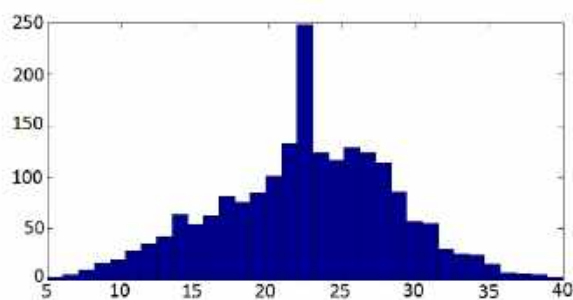


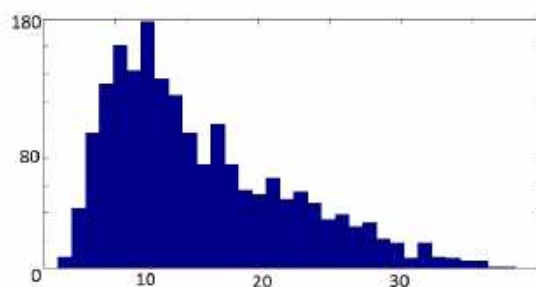
Рисунок Е.20 – Усереднена (для багатьох 8-бітних тестових зображень) залежність КМ від PSNR-HVS-M для кодерів AGU-M і ADCT-M

ДОДАТОК Ж

ДОДАТКОВІ МАТЕРІАЛИ ДО РОЗДІЛУ 2

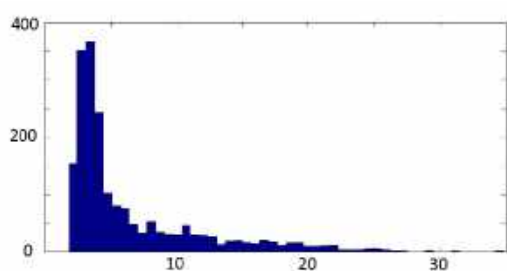


а

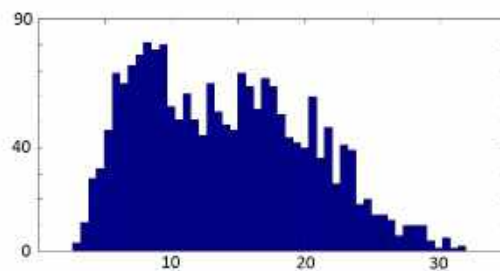


б

Рисунок Ж.1 – Гістограми значень елементарного детектора з урахуванням ДКП:
а – не текстури; б – текстури

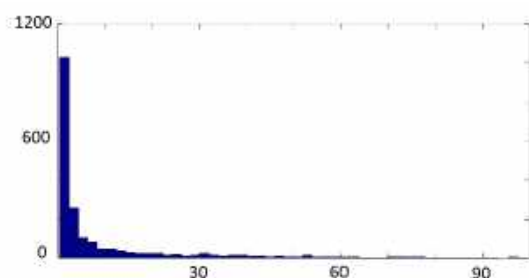


а

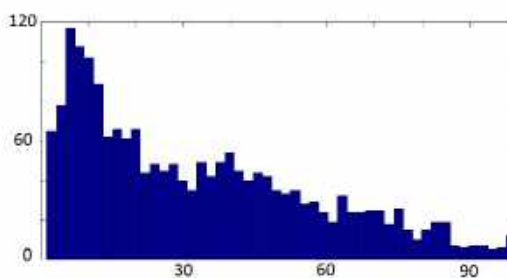


б

Рисунок Ж.2 – Гістограми значень елементарного детектора з урахуванням КР: а
– не текстури; б – текстури



а



б

Рисунок Ж.3 – Гістограми значень елементарного детектора з урахуванням ВЛД:
а – не текстури; б – текстури

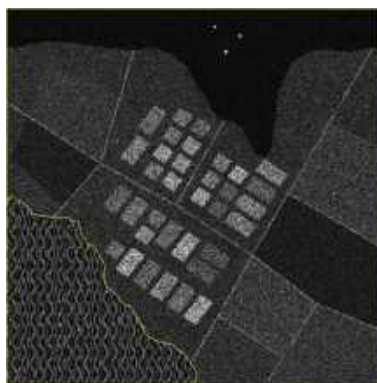
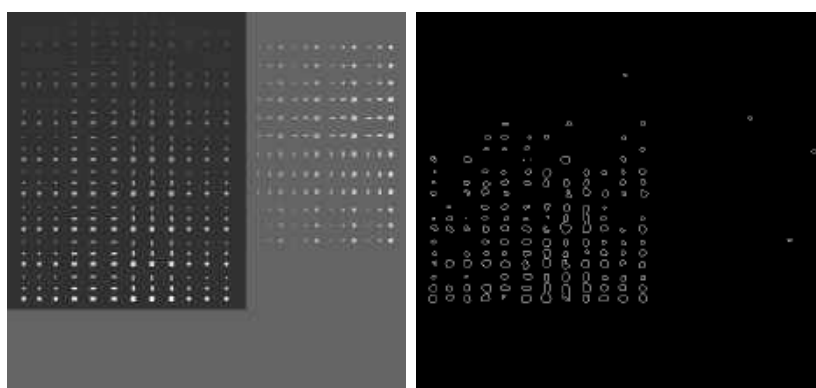


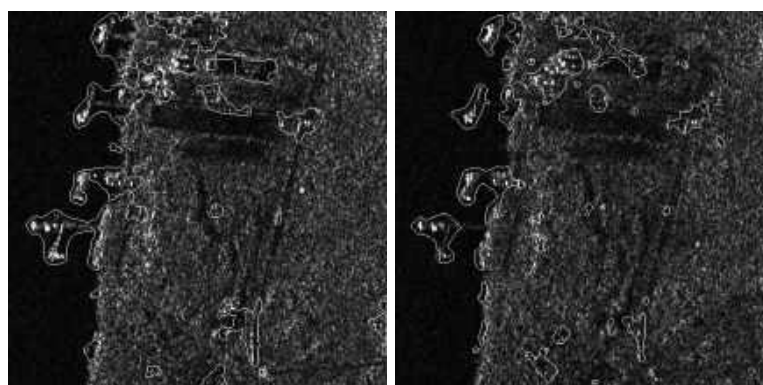
Рисунок Ж.4 – Зашумлене тестове зображення з п'ятнадцятьма об'єктами малого розміру та результати виявлення (межу показано жовтою лінією)



а

б

Рисунок Ж.5 – Синтетичне зображення: тестове зображення без шумів (а); виявлення об'єктів на версії зображення, спотвореного спеклом (б)



а

б

Рисунок Ж.6 – HH (а) і VV (б) поляризаційні однопоглядові зображення TerraSAR-X регіону в Індонезії з неоднорідними областями, виявленими методом Level Set.

Таблиця Ж.1 – Ефективність придушення спеклу на основі ДКП у поєднанні з частотно-залежним порогом, блоки 16x16 пікселів

β	FR01			FR02		
	MSE_{out}	MSE_{het}	MSSIM	MSE_{out}	MSE_{het}	MSSIM
4,0	527	721	0,809	547	755	0,798
4,2	510	713	0,809	530	758	0,799
4,4	498	712	0,809	520	766	0,798
4,6	492	717	0,807	515	780	0,796
4,8	489	725	0,805	513	796	0,793
5,0	490	736	0,801	514	816	0,790

Таблиця Ж.2 – Ефективність адаптивних методів придушення спеклу

Опис комбінованого фільтру та його параметри	FR01			FR02		
	MSE_{out}	MSE_{het}	MSSIM	MSE_{out}	MSE_{het}	MSSIM
ДКП: 4.9 (комб.) + 4.0 (жорсткий, незалеж. від частоти)	449	666	0,823	470	696	0,814
ДКП 4.9 (комб.) + 3.0 вейвлет Хаара (жорсткий)	445	667	0,823	466	688	0,817
ДКП +Frost	454	648	0,819	474	699	0,809
ДКП +Vozel9	429	617	0,825	457	654	0,814



а

б

Рисунок Ж.7 – Вихідні зображення FR01 для адаптивних фільтрів на основі ДКП з перемиканням параметрів у блоках залежно від карти гетерогенності: $\beta=4,0$ (а), $\beta=3,0$ (б)



а

б

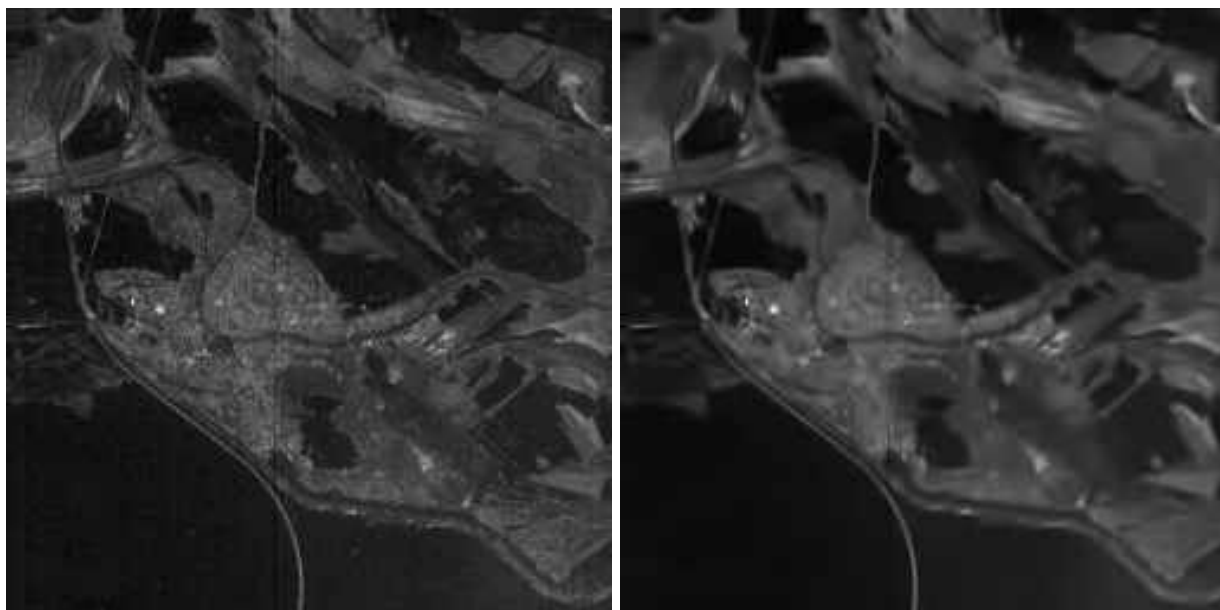
Рисунок Ж.8 – Вихідні зображення для адаптивних фільтрів: ДКП+Хаар (а), ДКП+інтегро-диференціальний фільтр [228] (б)

Таблиця Ж.3 – Параметри шуму та результати прогнозування для зображень каналів з індексами 13...22 набору даних сенсора Hyperion EO1H2010262004157110KP

Індекс каналу	$\hat{\sigma}_0(n)$	$\hat{k}(n)$	Динам. діапазон	$P_{2\sigma}$	Прогноз IPSNR на основі $P_{2\sigma}$	Прогноз IPSNR на основі $P_{2.7\sigma}$	Прогноз IPSNR на основі $P_{0.5\sigma}$	Прогноз IPSNR – Н на основі $P_{2\sigma}$
13	63,61	0,78	4166,00	0,85	5,60	4,72	5,90	3,45
14	41,54	0,62	4098,00	0,78	3,84	3,31	4,09	2,19
15	27,38	0,49	4316,00	0,71	2,42	2,08	2,65	1,25
16	20,01	0,42	4411,00	0,64	1,70	1,44	1,92	0,82
17	17,45	0,39	4418,00	0,63	1,56	1,32	1,77	0,74
18	16,64	0,37	4615,00	0,62	1,44	1,23	1,64	0,67
19	14,09	0,34	4741,00	0,59	1,27	1,07	1,47	0,57
20	13,29	0,31	4885,00	0,57	1,10	0,91	1,30	0,48
21	12,05	0,31	5007,00	0,54	0,96	0,78	1,15	0,41
22	10,98	0,30	4872,00	0,52	0,83	0,65	1,00	0,34

Таблиця Ж.4 – Параметри шуму та результати прогнозування для зображень каналів з індексами 13...22 набору даних сенсора Hyperion EO1H1800252002116110KZ

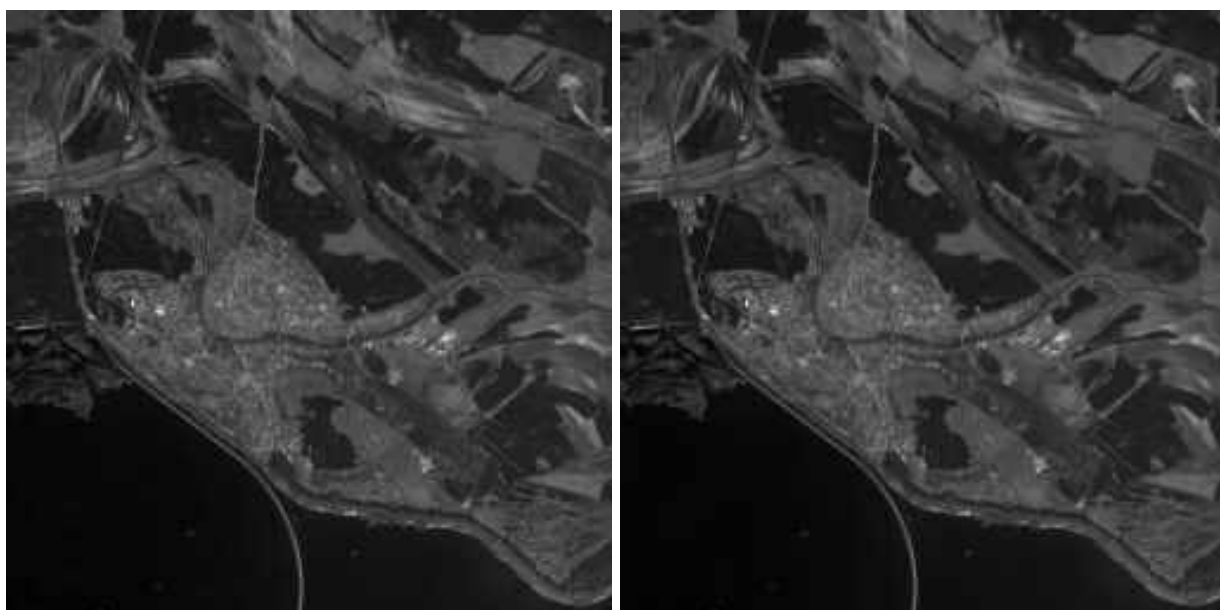
Індекс каналу	$\hat{\sigma}_0(n)$	$\hat{k}(n)$	Динам. діапазон	$P_{2\sigma}$	Прогноз IPSNR на основі $P_{2\sigma}$	Прогноз IPSNR на основі $P_{2.7\sigma}$	Прогноз IPSNR на основі $P_{0.5\sigma}$	Прогноз IPSNR – Н на основі $P_{2\sigma}$
13	54,44	0,78	1793,00	0,93	9,06	7,33	9,43	6,17
14	36,39	0,45	1835,00	0,88	6,85	5,80	7,20	4,40
15	23,74	0,27	1894,00	0,81	4,35	3,86	4,68	2,54
16	16,85	0,20	1985,00	0,73	2,80	2,53	3,11	1,49
17	14,48	0,17	1943,00	0,70	2,41	2,16	2,71	1,24
18	14,83	0,14	2118,00	0,67	2,02	1,80	2,30	1,01
19	11,84	0,13	2190,00	0,63	1,58	1,38	1,84	0,75
20	11,91	0,10	2265,00	0,60	1,33	1,13	1,57	0,61
21	10,90	0,10	2321,00	0,58	1,14	0,94	1,35	0,50
22	9,77	0,09	2226,00	0,55	1,00	0,80	1,15	0,43



а

б

Рисунок Ж.9 – Вхідні (а) і вихідні (б) зображення EO1H1800252002116110KZ в 13-му каналі даних Hyperion



а

б

Рисунок Ж.10 – Вхідні (а) і вихідні (б) зображення EO1H1800252002116110KZ в 22-му каналі даних Hyperion

ДОДАТОК К

ДОДАТКОВІ МАТЕРІАЛИ ДО РОЗДІЛУ 3

Таблиця К.1 – Статистика та параметри забезпечення $\text{PSNR}_{\text{des}}=34$ дБ

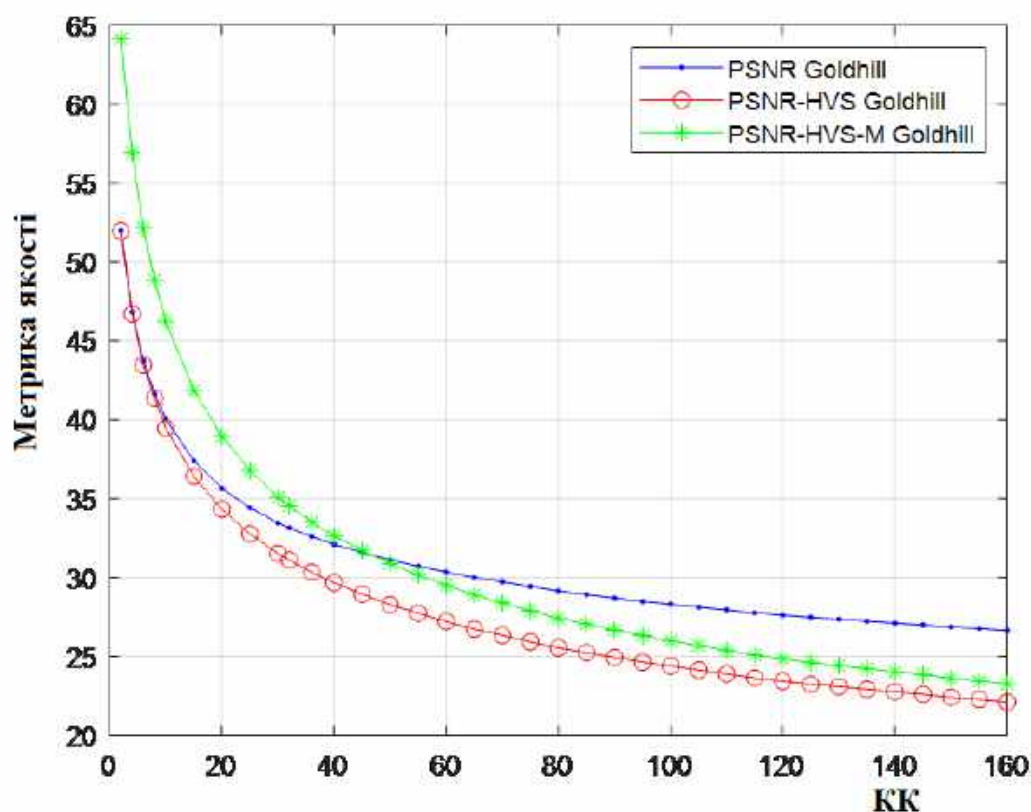
Тестове зображення	KK_{init}	$\text{PSNR}_{\text{init}}$	ΔKK	KK_{rec}	$\text{PSNR}_{\text{prov}}$
Baboon	29,45	31,18	-14,72	14,73	36,23
Barbara	29,45	34,96	5,07	34,53	34,08
Goldhill	29,45	33,57	-2,23	27,22	34,43
Lenna	29,45	35,66	8,69	38,15	34,59
Patterns	29,45	36,95	15,50	44,96	33,86
Mountain	29,45	30,43	-18,67	10,78	38,35
Zelda	29,45	36,27	11,91	41,36	34,53
Boat	29,45	32,89	-5,77	23,68	34,03
Pentagon	29,45	34,03	0,20	29,66	34,01
Дисперсія	4,46		2,79		

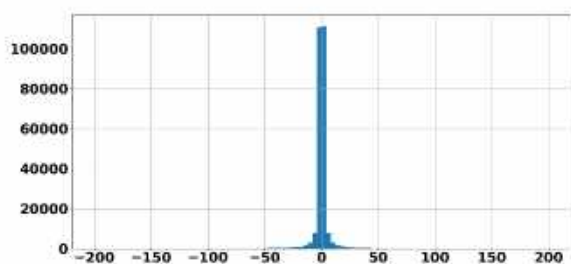
Таблиця К.2 – Статистика та параметри забезпечення $\text{PSNR}_{\text{des}}=37$ дБ

Тестове зображення	KK_{init}	$\text{PSNR}_{\text{init}}$	ΔKK	KK_{rec}	$\text{PSNR}_{\text{prov}}$
Baboon	17,30	35,07	-5,46	11,84	38,06
Barbara	17,30	37,92	2,63	19,94	37,1
Goldhill	17,30	36,59	-1,14	16,16	36,99
Lenna	17,30	37,92	2,62	19,92	37,28
Patterns	17,30	39,33	6,64	23,95	38,025
Mountain	17,30	34,35	-7,53	9,77	39,11
Zelda	17,30	38,25	3,56	20,87	37,55
Boat	17,30	35,87	-3,20	14,10	37,19
Pentagon	17,30	37,66	1,88	19,19	36,66
Дисперсія	2,36		0,96		

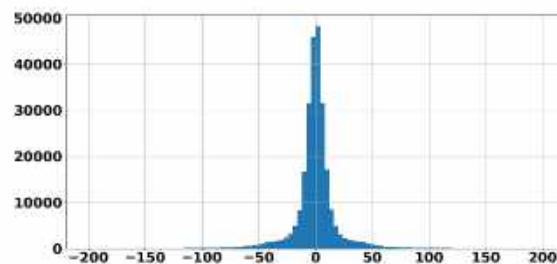
Таблиця К.3 – Статистика та параметри забезпечення $\text{PSNR}_{\text{des}}=40$ дБ

Тестове зображення	KK_{init}	$\text{PSNR}_{\text{init}}$	ΔKK	KK_{rec}	$\text{PSNR}_{\text{prov}}$
Baboon	10,44	38,93	-2,21	8,23	40,82
Barbara	10,44	40,79	1,64	12,09	40,02
Goldhill	10,44	39,82	-0,37	10,07	40,02
Lenna	10,44	40,39	0,81	11,26	40,05
Patterns	10,44	41,06	2,22	12,67	40,51
Mountain	10,44	38,58	-2,95	7,49	41,40
Zelda	10,44	40,38	0,80	11,25	40,09
Boat	10,44	40,52	-2,87	7,67	40,43
Pentagon	10,44	40,79	1,65	12,09	40,41
Дисперсія	0,68		0,38		

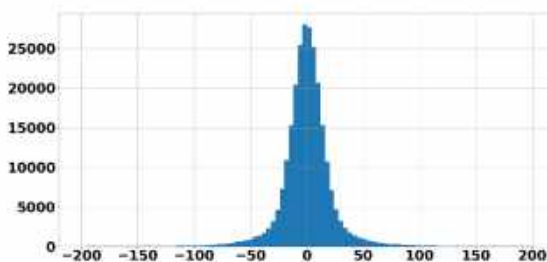
Рисунок К.1 – Залежність метрики візуальної якості від KK



а

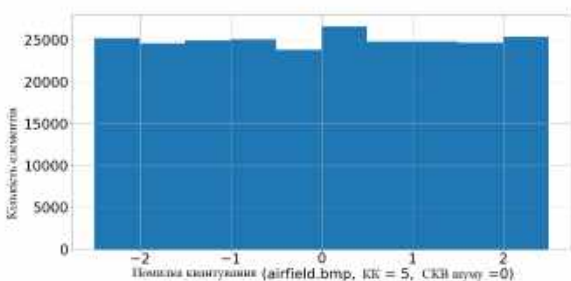


б

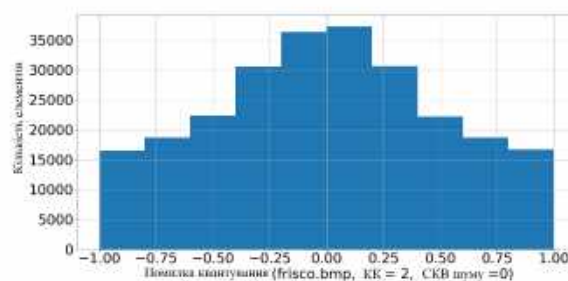


в

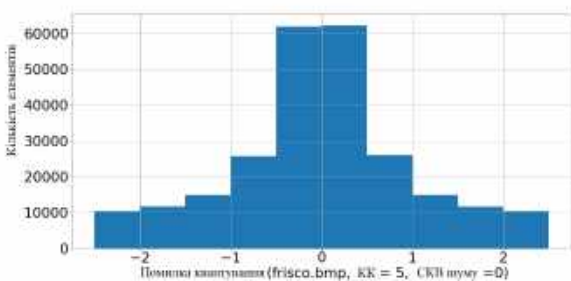
Рисунок К.2 – Розподіл коефіцієнтів АС ДКП для безшумного зображення Frisco (а), безшумного зображення Airfield (б) і зашумленого зображення Airfield (в), межі від -200 до 200



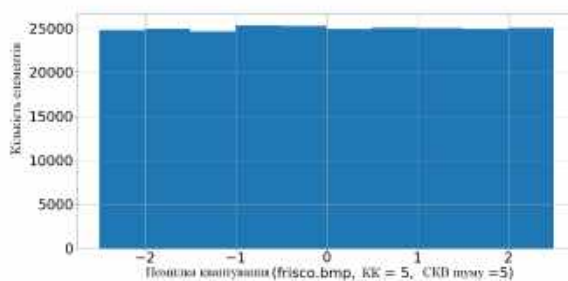
а



б

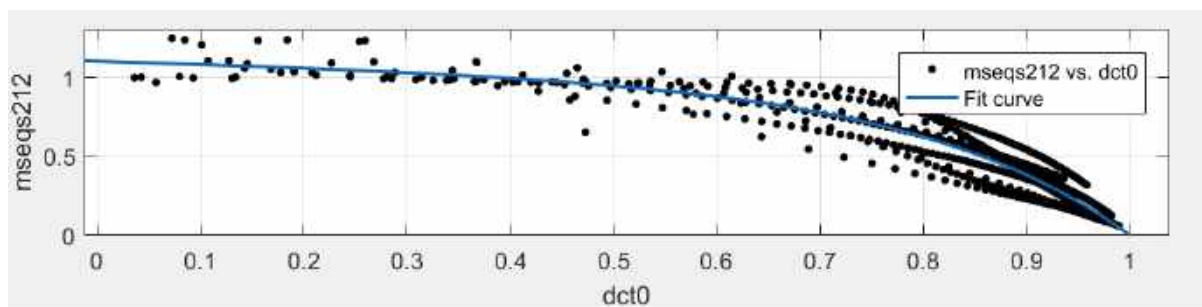


в

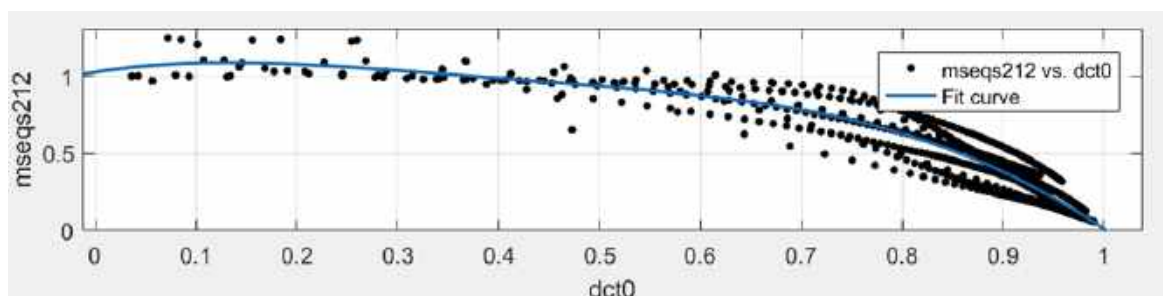


д

Рисунок К.3 – Приклади гістограм помилки квантування для коефіцієнтів ДКП



а



б

Рисунок К.4 – Діаграма розсіювання $12MSE/KK^2$ від P_0 для безшумних зображень і двох варіантів підігнаних кривих, експоненційна (а) та поліноміальна (б), кодер ADCT

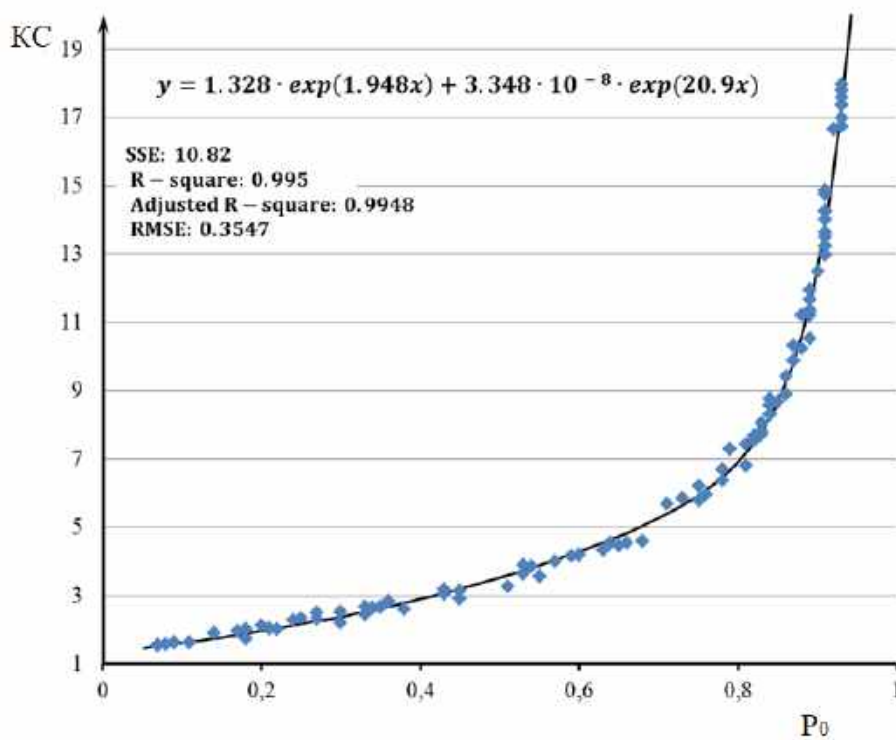


Рисунок К.5 – Діаграма розсіювання KC на P_0 та підігнана крива

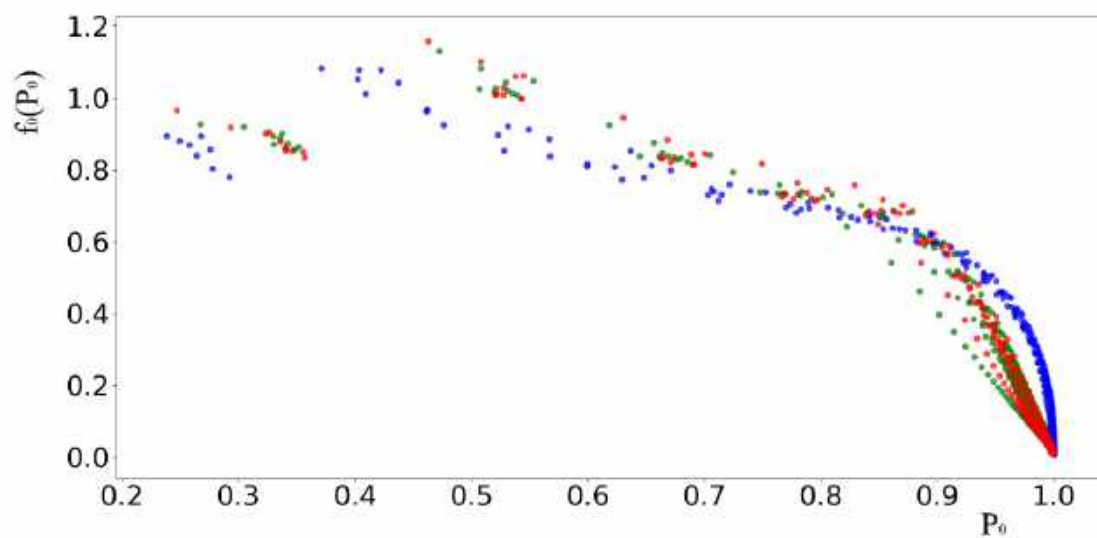
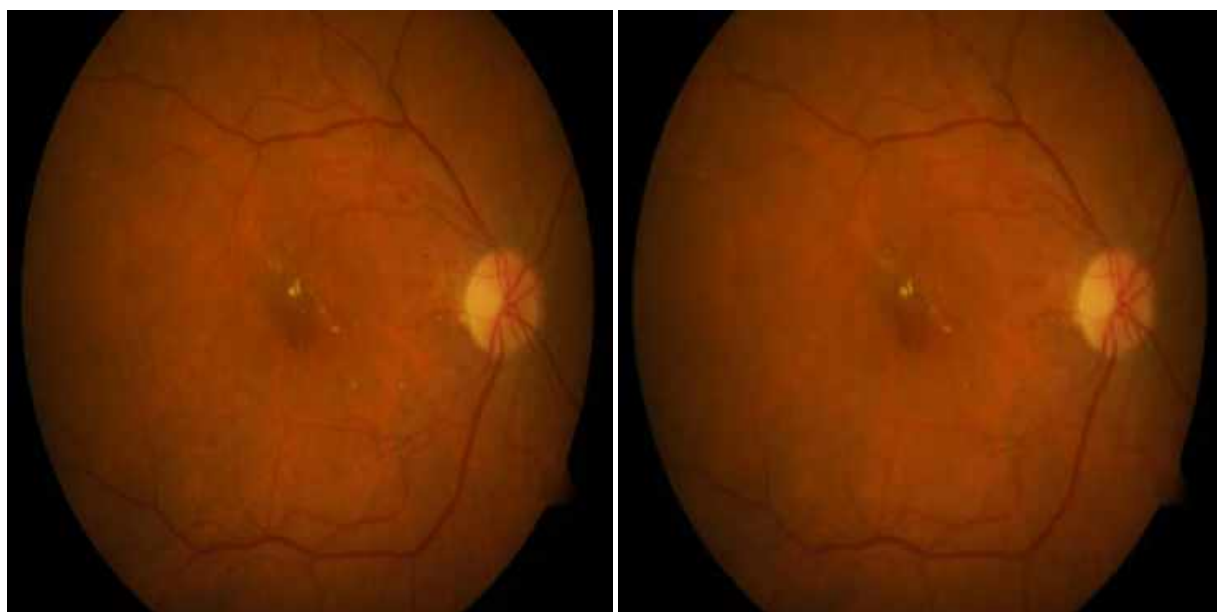


Рисунок К.6 – Діаграма розсіювання $f_0(P_0)$ для окремих компонент кольорових зображень сітківки, кодер AGU



а

б

Рисунок К.7 – Оригінальне (а) і стиснуте (б) зображення сітківки



Рисунок К.8 – Результати стиснення для тестового зображення Goldhill: безшумне нестиснуте зображення (а); зображення, стиснуте з $MSE=20$ ($KC=11,5$, $PSNR-HVS-M=38,1$ дБ) (б); зображення, стиснуте з $MSE=40$ ($KC=23,9$, $PSNR-HVS-M=32,6$ дБ) (в); зображення, стиснуте з $MSE=60$ ($KC=40,1$, $PSNR-HVS-M=29,5$ дБ) (г)

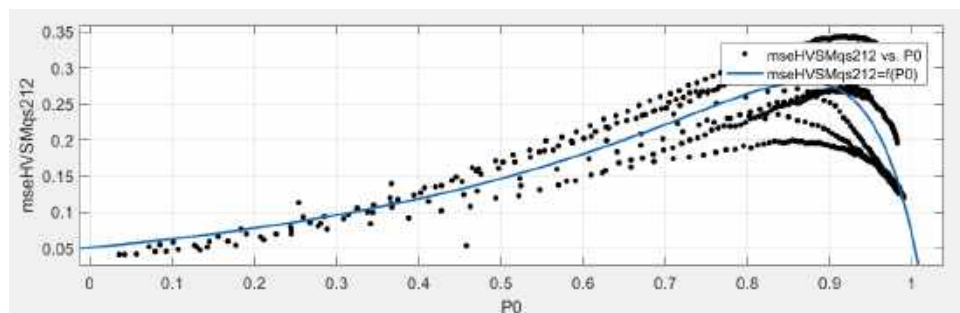


Рисунок К.9 – Діаграма розсіювання $MSE_{HVS-M}/(KC^2/12)$ від P_0 для зображень без шуму та підігнаної кривої, кодер AGU

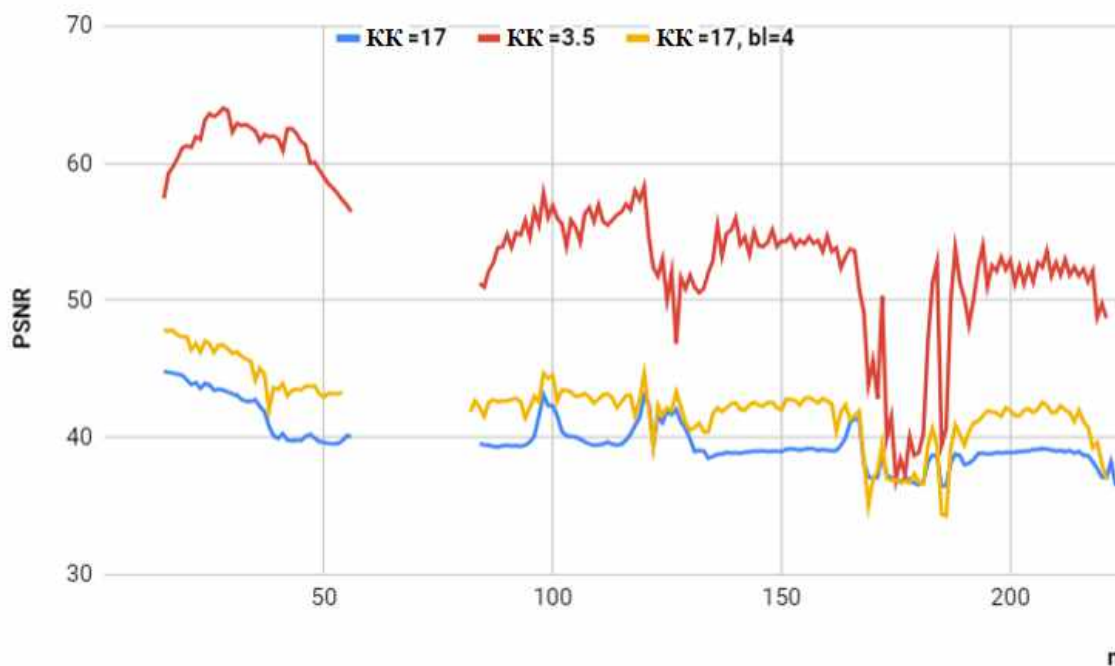


Рисунок К.10 – PSNR для покомпонентного стиснення за методом ([165], $KK=3,5$), запропонованим покомпонентним підходом ($KK=17$) і запропонованим методом 3D стиснення ($KK=17, bl=4$)

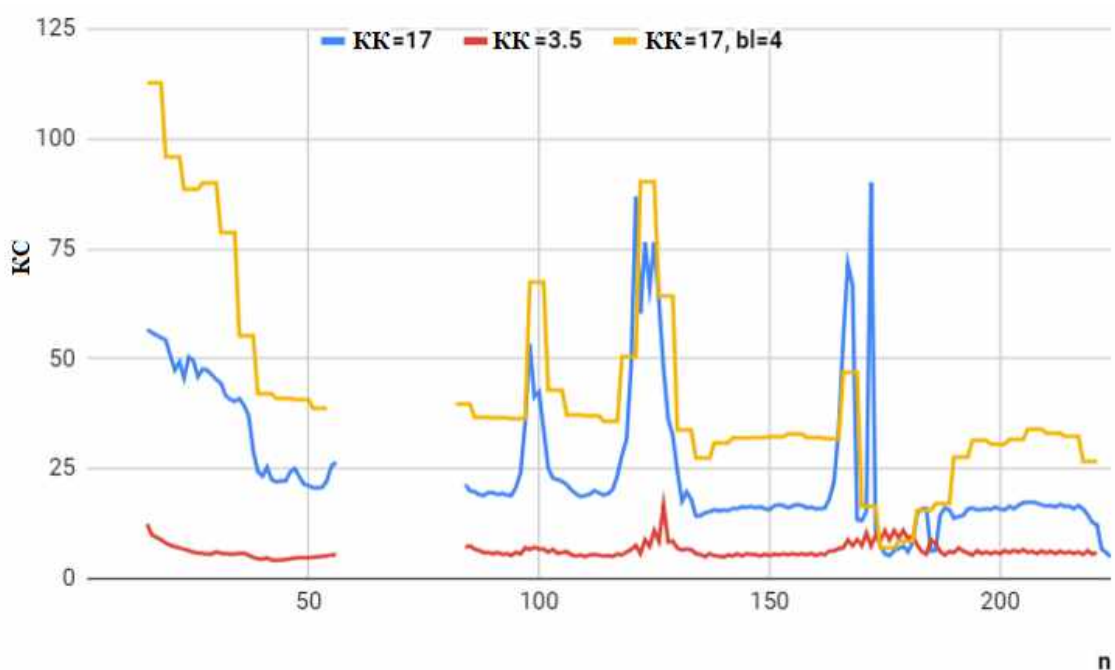
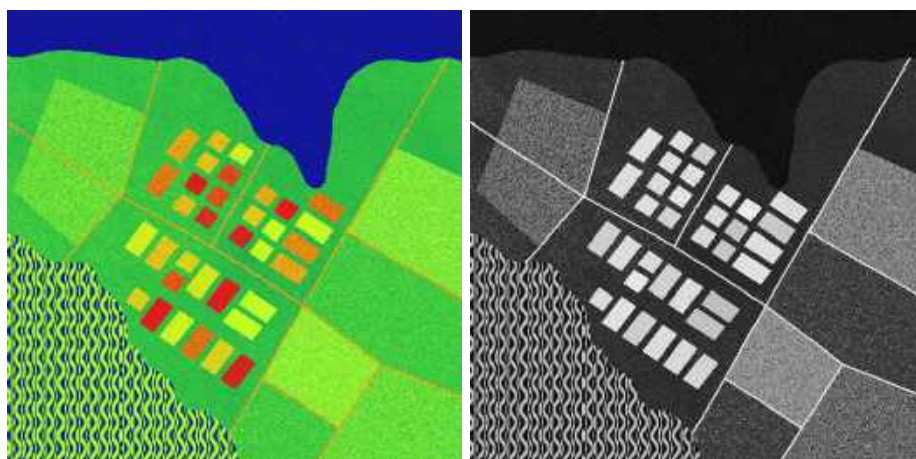
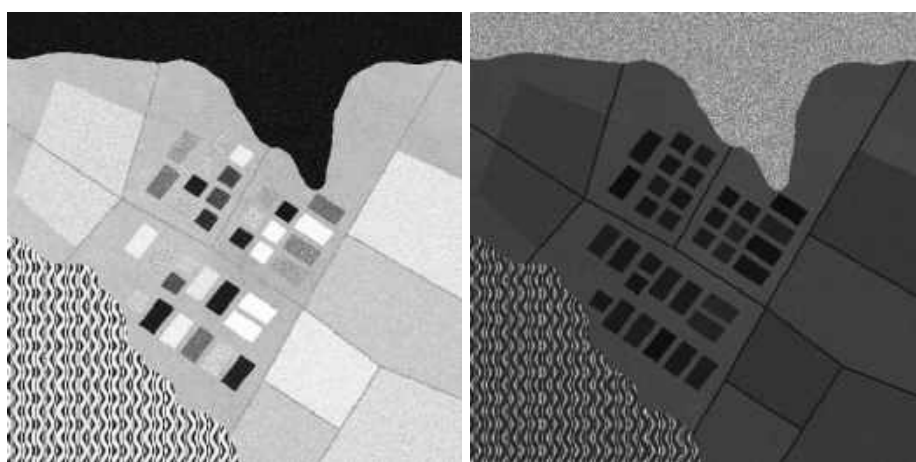


Рисунок К.11 – KS для покомпонентного стиснення за методом ([165], $KK=3,5$), запропонованим покомпонентним підходом ($KK=17$) і запропонованим методом 3D стиснення ($KK=17, bl=4$)



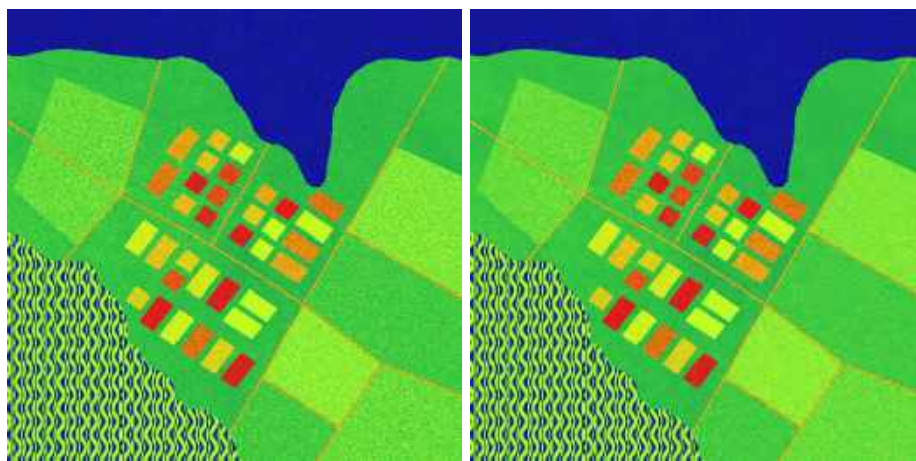
а

б



в

г



д

е

Рисунок К.12 – Триканальне тестове зображення: оригінал (а), його три компоненти (б–г), стиснене із забезпеченням PSNR-HVS-M = 36 дБ (д), стиснене із забезпеченням PSNR-HVS-M = 30 дБ (е)

Таблиця К.4 – Інформація про класи зображень

#	Карта класу (зразок об'єкта)	Червоний (гістограма та наближення)	Зелений (гістограма та наближення)	Синій (гістограма та наближення)
1				
2				
3				
4				
5				
6				
7				

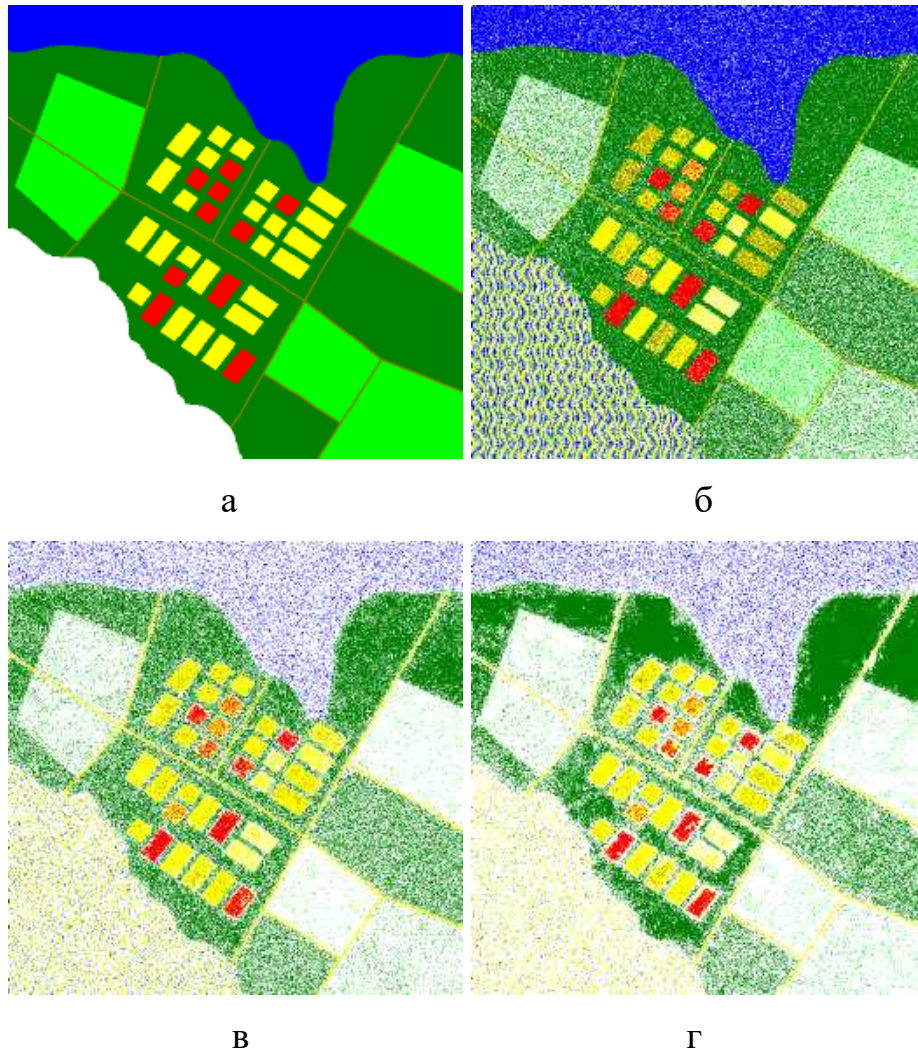
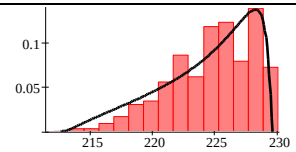
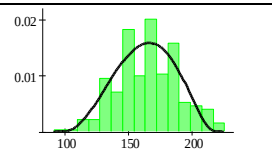
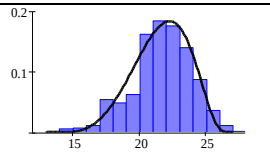
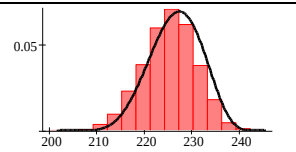
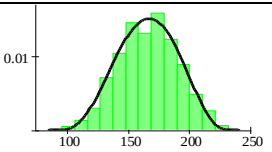
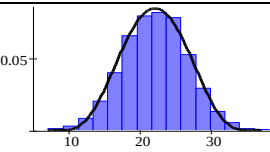
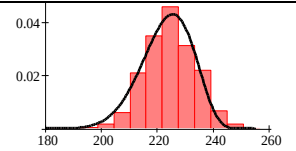
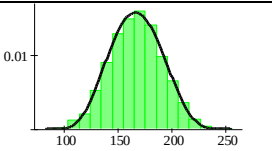
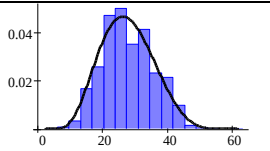
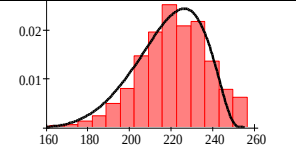
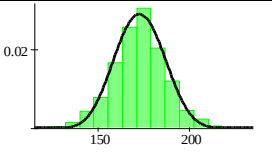
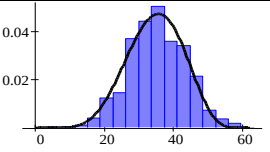


Рисунок К.13 – Результати класифікації: істинна карта (а); карта класифікації для вихідного зображення (б); карта класифікації для зображення, стисненого із забезпеченням PSNR-HVS-M = 36 дБ (в); карта класифікації для зображення, стисненого із забезпеченням PSNR-HVS-M = 30 дБ (г)

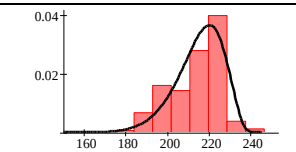
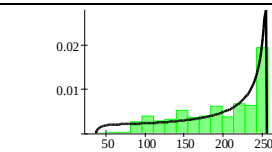
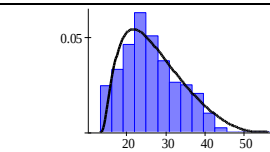
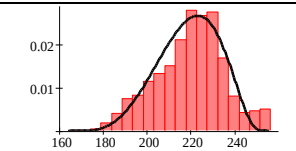
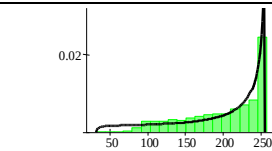
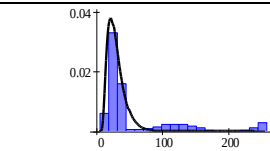
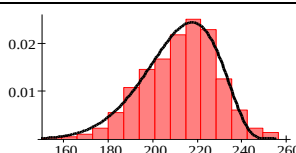
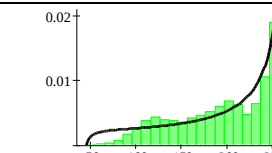
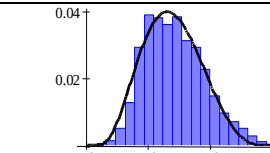
Таблиця К.5 – Імовірності правильної класифікації для окремих класів залежно від якості зображення

PSNR-HVS-M, дБ	P ₁₁	P ₂₂	P ₃₃	P ₄₄	P ₅₅	P ₆₆	P ₇₇
∞ (оригінал)	0,475	0,667	0,773	0,854	0,193	0,822	0,415
48	0,190	0,781	0,686	0,384	0,093	0,809	0,579
45	0,147	0,795	0,652	0,308	0,077	0,777	0,623
42	0,101	0,804	0,615	0,253	0,058	0,712	0,665
36	0,031	0,802	0,568	0,163	0,040	0,545	0,777
30	0,001	0,753	0,478	0,177	0,042	0,556	0,843

Таблиця К.6 – Ілюстрація змін розподілу через стиснення з втратами для класу 1

PSNR-HVS-M, дБ	Червоний (гістограма та наближення)	Зелений (гістограма та наближення)	Синій (гістограма та наближення)
∞ (оригінал)			
42			
36			
30			

Таблиця К.7 – Ілюстрація змін розподілу через стиснення з втратами для класу 2

PSNR-HVS-M, дБ	Червоний (гістограма та наближення)	Зелений (гістограма та наближення)	Синій (гістограма та наближення)
∞ (оригінал)			
36			
30			

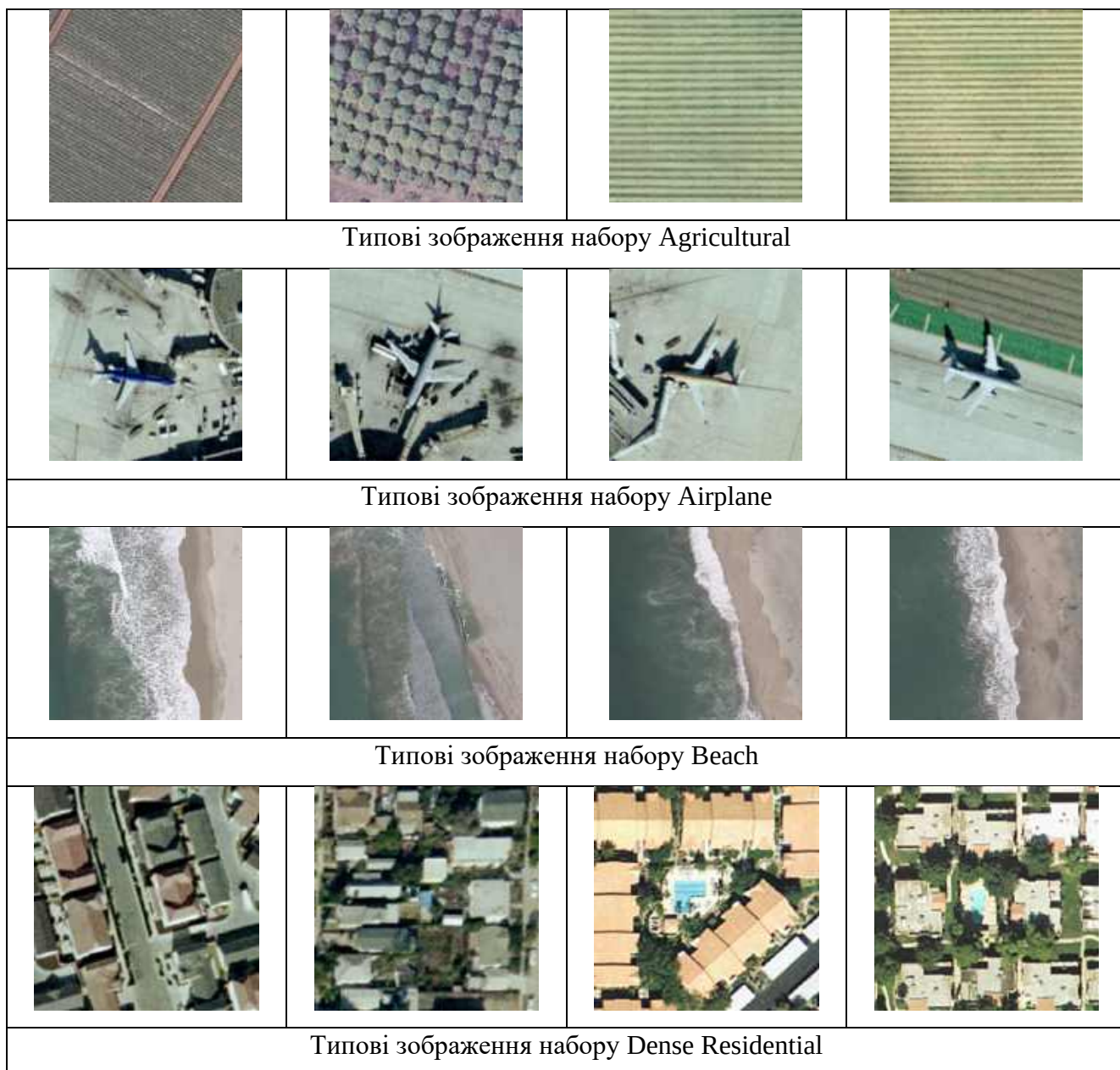


Рисунок К.14 – Приклади зображень у наборах даних

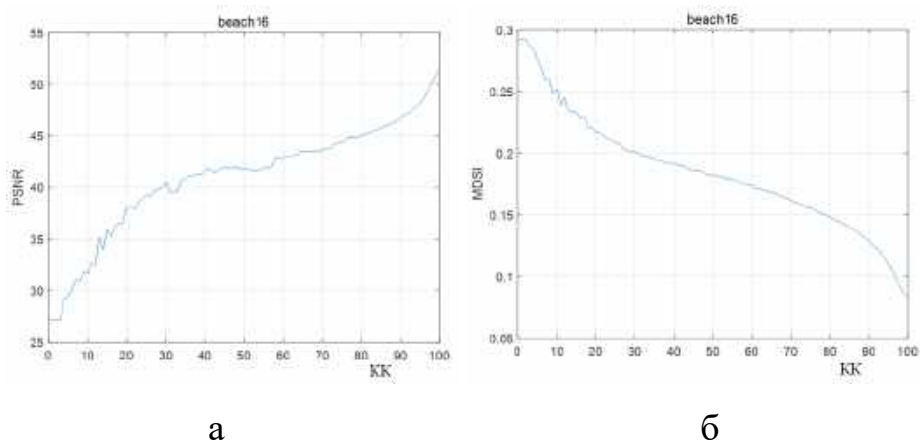


Рисунок К.15 – Приклади КІШС PSNR(KK) (а) та MDSI(KK) (б) для зображення Beach16

ДОДАТОК Л

ДОДАТКОВІ МАТЕРІАЛИ ДО РОЗДІЛУ 4

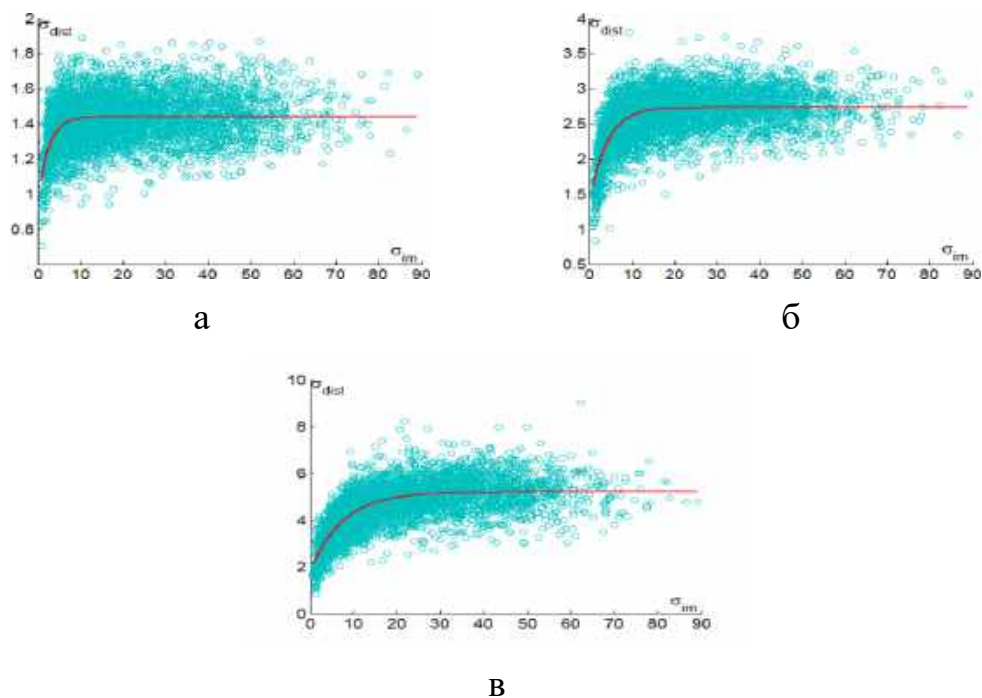


Рисунок Л.1 – Діаграма розсіювання σ_{dist} проти σ_{im} для тестового зображення FR01, КК дорівнює: 5 (а), 10 (б) і 20 (в)

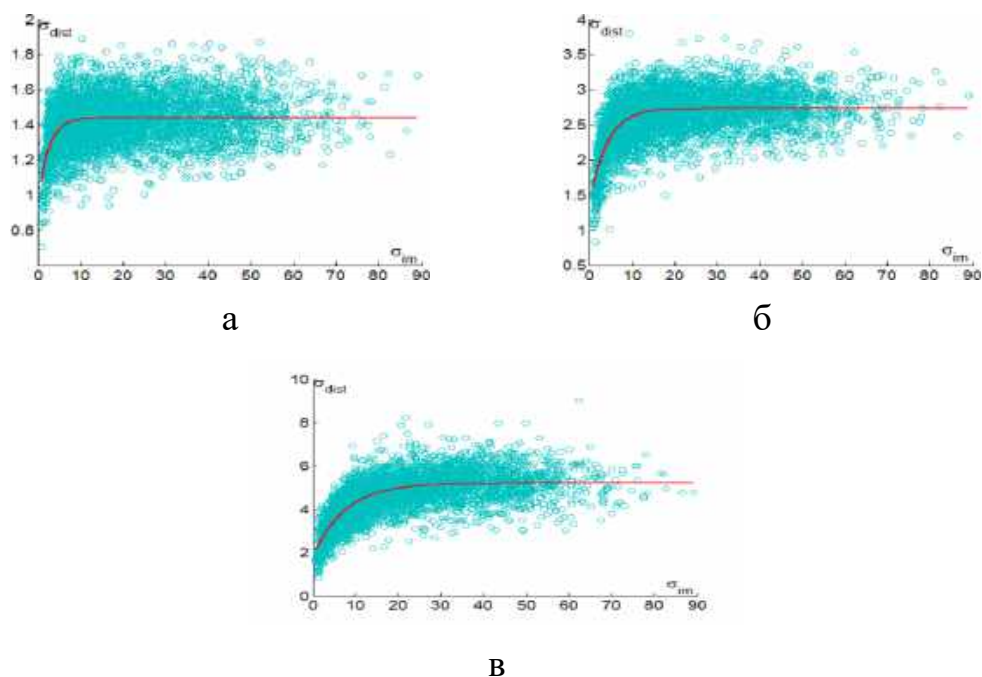


Рисунок Л.2 – Діаграма розсіювання σ_{dist} проти σ_{im} для тестового зображення FR02, КК дорівнює: 5 (а), 10 (б) і 20 (в)

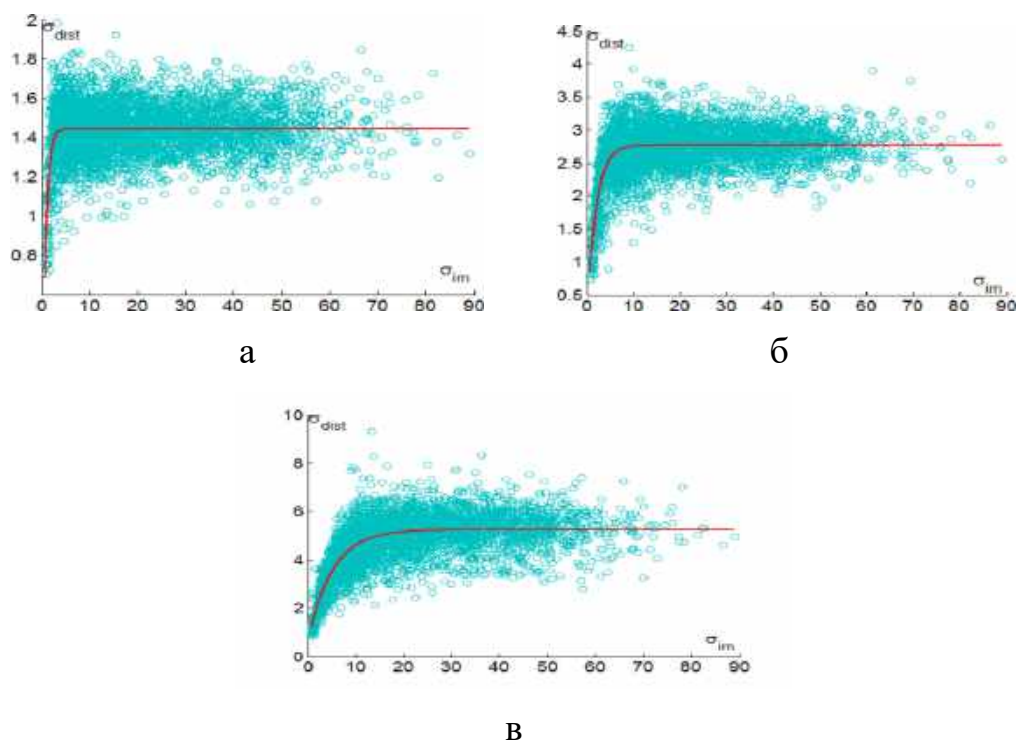


Рисунок Л.3 – Діаграма розсіювання σ_{dist} проти σ_{im} для тестового зображення FR01, КК дорівнює: 5 (а), 10 (б) і 20 (в)

Таблиця Л.1 – Параметри підігнаних кривих для кодера AGU

Зображення	КК=5	КК=10	КК=20
FR01	$a = -0,50$	$a = -1,37$	$a = -3,41$
	$b = -0,40$	$b = -0,25$	$b = -0,13$
	$c = 1,44$	$c = 2,73$	$c = 5,21$
FR02	$a = -0,91$	$a = -2,04$	$a = -4,03$
	$b = -0,57$	$b = -0,33$	$b = -0,16$
	$c = 1,44$	$c = 2,74$	$c = 5,24$
FR03	$a = -0,74$	$a = -1,28$	$a = -3,35$
	$b = -0,63$	$b = -0,25$	$b = -0,14$
	$c = 1,44$	$c = 2,77$	$c = 5,31$
FR04	$a = -1,01$	$a = -2,41$	$a = -4,97$
	$b = -0,55$	$b = -0,33$	$b = -0,17$
	$c = 1,45$	$c = 2,80$	$c = 5,46$

Таблиця Л.2 – Порівняння справжніх і прогнозованих значень MSE для кодера AGU

Зображення	КК	MSE	MSE _{pred}
FR01	5	2,03	1,88
FR02	5	1,99	1,85
FR03	5	2,07	1,90
FR04	5	1,96	1,78
FR01	10	6,91	6,97
FR02	10	6,85	6,85
FR03	10	7,24	7,12
FR04	10	6,74	6,55
FR01	20	21,84	23,88
FR02	20	21,95	23,67
FR03	20	23,30	24,39
FR04	20	22,02	22,24

Таблиця Л.3 – Порівняння справжніх і прогнозованих значень MSE для кодера AGU

Зображення	КК	MSE	MSE _{pred}
FR01	5	2,08	1,94
FR02	5	2,03	1,90
FR03	5	2,10	1,95
FR04	5	1,98	1,84
FR01	10	7,34	7,30
FR02	10	7,13	7,15
FR03	10	7,71	7,45
FR04	10	7,03	6,83
FR01	20	23,03	25,51
FR02	20	23,17	25,12
FR03	20	25,38	26,23
FR04	20	23,92	23,66

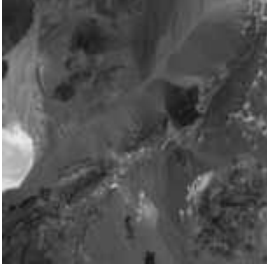
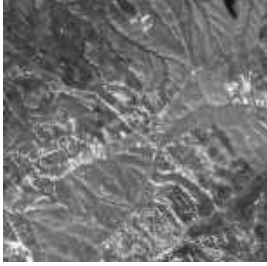
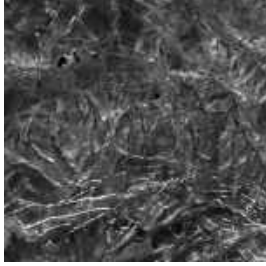
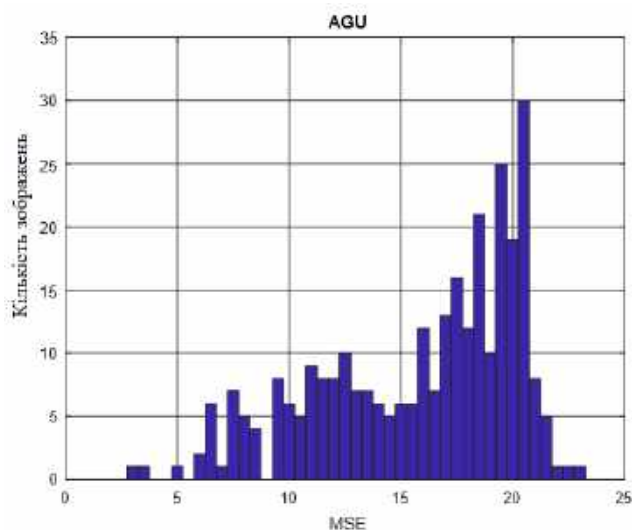
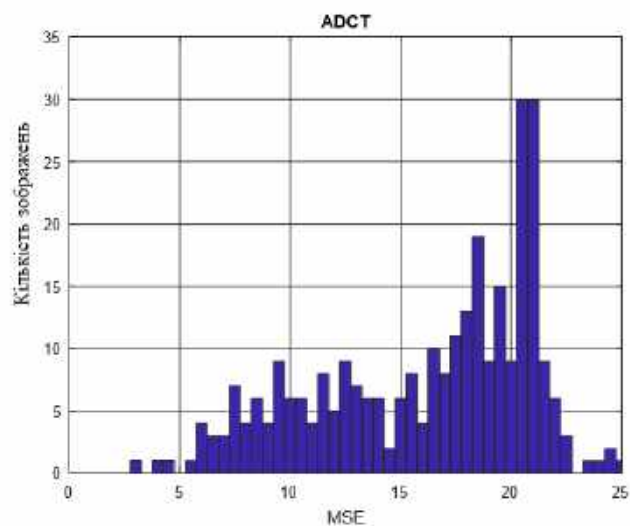
		
A13sm	Aerial	Baboon
		
Barbara	Bikes	Boat
		
Goldhill	Grass	Lena
		
Lu01	Man	Peppers
		
SanDiego	Sent01	Sent02

Рисунок Л.4 – Зменшені копії тестових зображень, використаних в експериментах



а



б

Рисунок Л.5 – Гістограми значень MSE, отриманих для 300 зображень із набору даних TAMPERE17: за допомогою кодера AGU (а), за допомогою кодера ADCT (б)

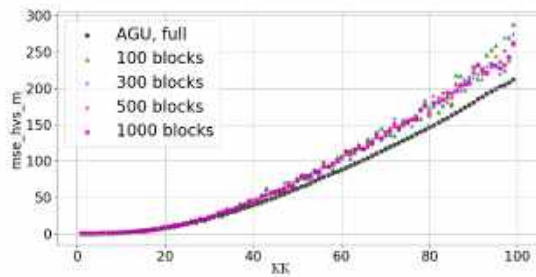


а

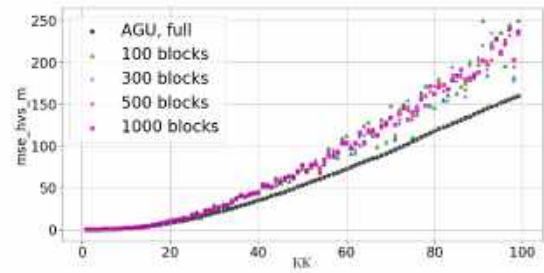


б

Рисунок Л.6 – Два зразки зображень у градаціях сірого з набору даних TAMPERE17 із найменшими отриманими значеннями MSE: зображення № 257 (а), зображення № 263 (б)

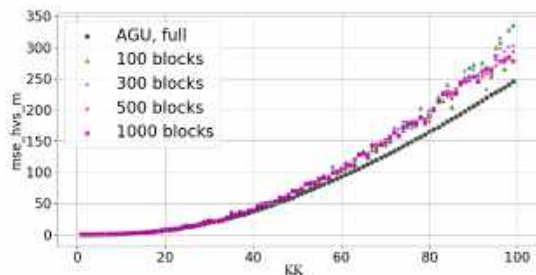


а

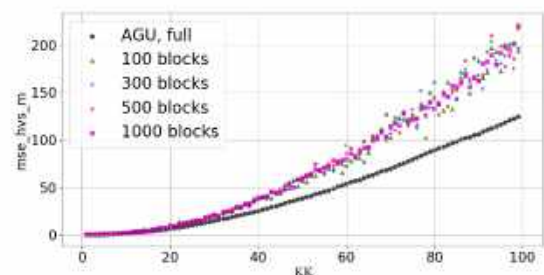


б

Рисунок Л.7 – Діаграми розсіювання оцінок $MSE_{HVS-M \Sigma}$ для різної кількості проаналізованих блоків і $MSE_{HVS-M AGU}$ для стиснених тестових зображень Airfield (а) і Goldhill (б), кодер AGU



а



б

Рисунок Л.8 – Діаграми розсіювання оцінок $MSE_{HVS-M \Sigma}$ для різної кількості проаналізованих блоків і $MSE_{HVS-M AGU}$ для стиснених тестових зображень Baboon (а) і Barbara (б), кодер AGU

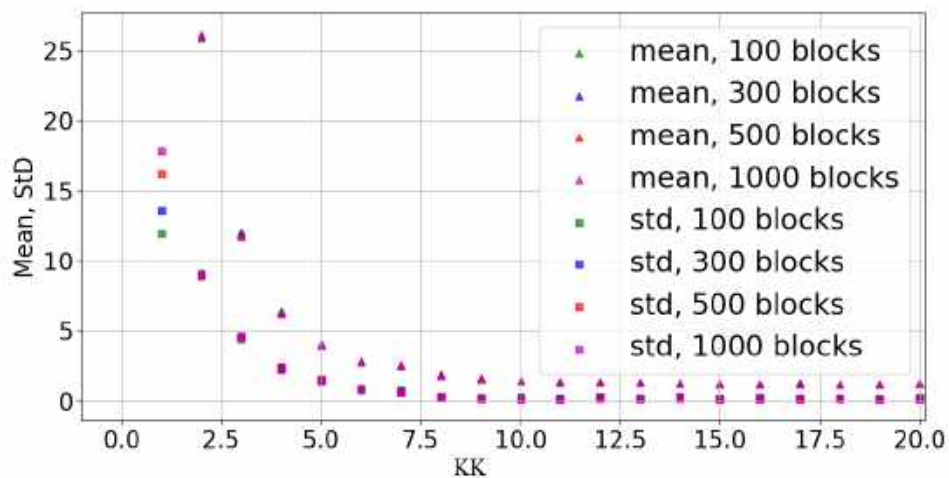


Рисунок Л.9 – Середнє значення та стандартні відхилення співвідношення $MSE_{HVS-M \Sigma} / MSE_{HVS-M AGU}$ для всіх тестових зображень, стиснутих AGU

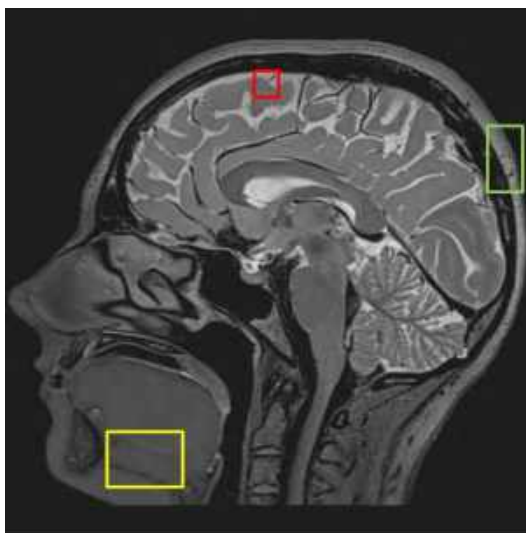
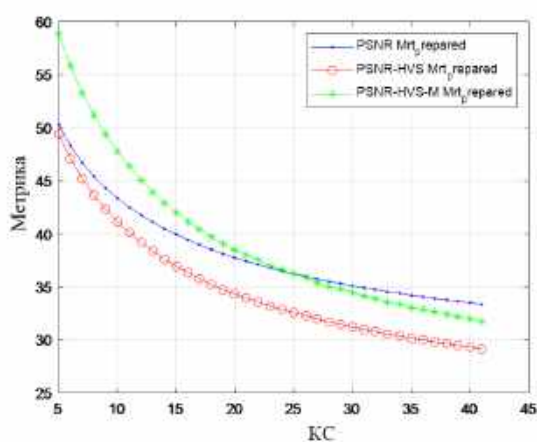


Рисунок Л.10 – Тестове зображення $Mrt_{prepared}$ із зазначенням характерних фрагментів

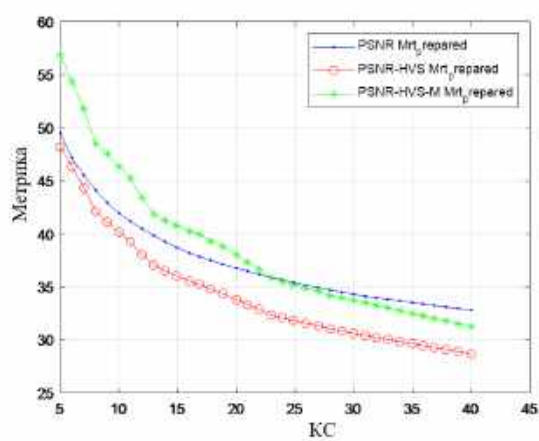


а б

Рисунок Л.11 – Фрагменти стиснених зображень



а



б

Рисунок Л.12 – Залежності розглянутих метрик від КС для тестового зображення $MRT_{prepared}$, кодер AGU (а) та SPIHT (б)

Таблиця Л.4 – Набори використаних БНП для кодера SPIHT для розглянутих тестових зображень і дисперсій шуму

Зображення	σ^2	b						
		0,5	1,5	2,5	3,5	4,5	5,5	6,5
Baboon	50	4,41	2,87	2,17	1,72	1,35	1,07	0,88
	100	4,09	2,54	1,86	1,40	1,02	0,76	0,59
Barbara	50	3,71	2,19	1,52	1,03	0,63	0,48	0,39
	100	3,54	2,03	1,38	0,86	0,48	0,35	0,27
GoldHill	50	3,72	2,20	1,53	1,03	0,64	0,44	0,33
	100	3,53	2,02	1,36	0,84	0,44	0,29	0,21

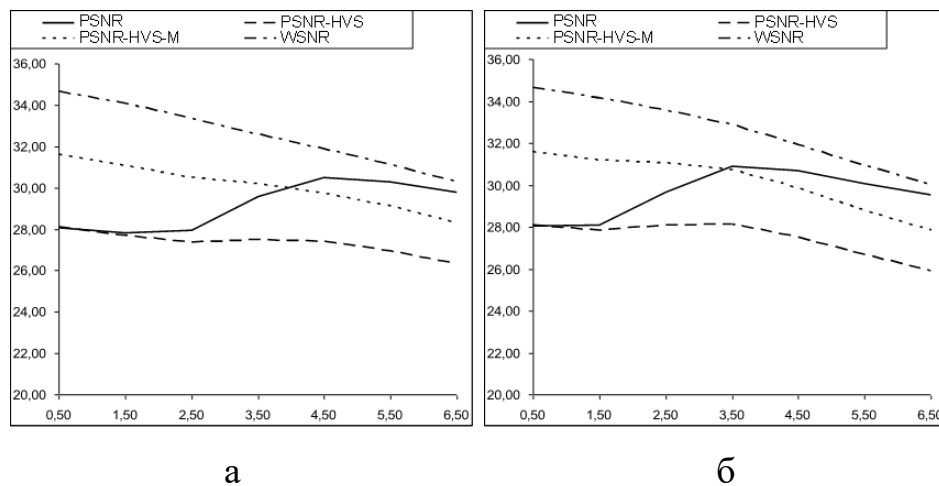


Рисунок Л.13 – Залежності метрик від b для зображення Goldhill, $\sigma^2=100$, для кодерів AGU (а) та ADCT (б)

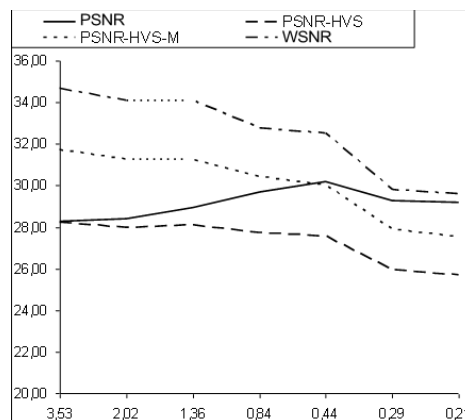


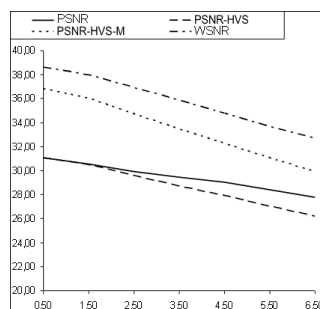
Рисунок Л.14 – Залежності метрик від БНП для зображення Goldhill, $\sigma^2=100$, для кодера SPIHT

Таблиця Л.5 – Залежності метрик від b для зображення Goldhill, $\sigma^2=50$, кодер ADCT

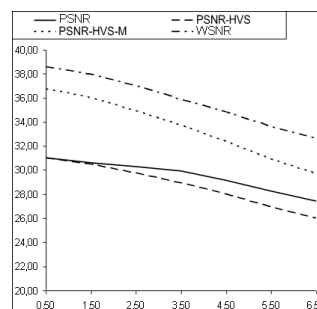
b	PSNR, дБ	PSNR-HVS, дБ	PSNR-HVS-M, дБ	WSNR, дБ	MS-SSIM
0,5	31,06	31,06	35,19	37,64	0,9724
1,5	30,92	30,69	34,67	37,07	0,9709
2,5	31,84	30,59	34,25	36,42	0,9733
3,5	32,58	30,33	33,58	35,48	0,9741
4,5	32,22	29,59	32,51	34,47	0,9701
5,5	31,55	28,72	31,39	33,47	0,9635
6,5	30,91	27,86	30,31	32,48	0,9562

Таблиця Л.6 – Залежності метрик від b для зображення Goldhill, $\sigma^2=100$, кодер ADCT

b	PSNR, дБ	PSNR-HVS, дБ	PSNR-HVS-M, дБ	WSNR, дБ	MS-SSIM
0,5	28,07	28,04	32,75	35,57	0,9701
1,5	27,73	27,53	32,03	34,93	0,9674
2,5	27,85	27,06	31,25	34,02	0,9666
3,5	27,76	26,40	30,22	33,05	0,9639
4,5	27,11	25,49	28,96	31,91	0,9563
5,5	26,29	24,52	27,67	30,85	0,9450
6,5	25,51	23,64	26,52	29,94	0,9330



а



б

Рисунок Л.15 – Залежності метрик від b для зображення Baboon, $\sigma^2=50$, для кодерів AGU (а) та ADCT (б)

Таблиця Л.7 – Залежності метрик від b для зображення Barbara, $\sigma^2=50$, кодер AGU

b	PSNR, дБ	PSNR-HVS, дБ	PSNR-HVS-M, дБ	WSNR, дБ	MS-SSIM
0,5	31,06	31,00	35,09	37,61	0,9697
1,5	30,78	30,60	34,58	37,08	0,9680
2,5	30,85	30,25	34,01	36,27	0,9675
3,5	32,19	30,34	33,82	35,53	0,9739
4,5	32,83	30,18	33,41	34,76	0,9768
5,5	32,50	29,72	32,81	34,05	0,9761
6,5	31,93	29,05	31,90	33,16	0,9735

Таблиця Л.8 – Залежності метрик від b для зображення Barbara, $\sigma^2=100$, кодер SPIHT

БНП	PSNR, дБ	PSNR-HVS, дБ	PSNR-HVS-M, дБ	WSNR, дБ	MS-SSIM
3,5435	28,29	28,19	31,71	34,64	0,9482
2,0322	28,48	28,01	31,37	34,12	0,9481
1,3768	29,01	28,25	31,56	34,10	0,9526
0,8567	30,07	28,23	31,15	33,06	0,9613
0,4779	29,70	27,65	30,38	32,50	0,9616
0,3445	28,68	26,25	28,55	30,35	0,9490
0,272	27,88	25,60	27,82	30,06	0,9401

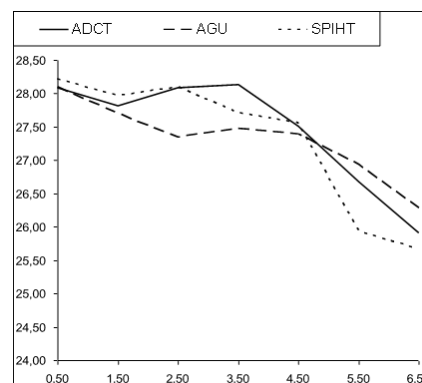


Рисунок Л.16 – Залежності PSNR-HVS від b для різних кодери, $\sigma^2=100$, зображення Goldhill



а

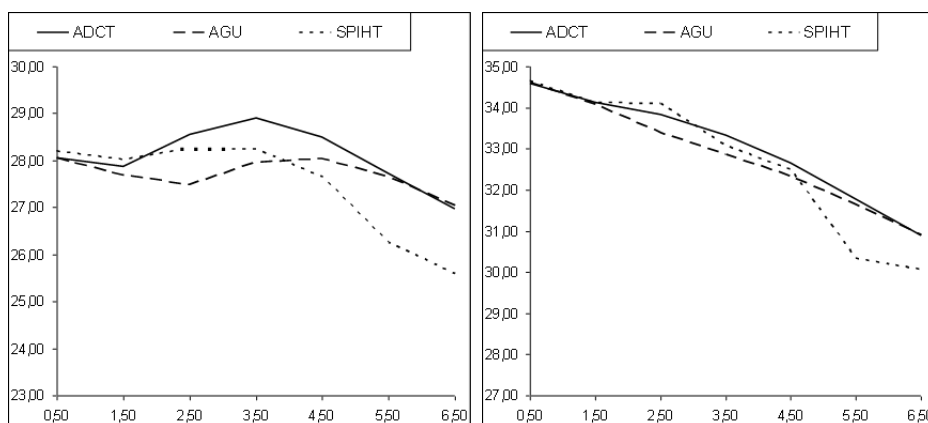
б



в

г

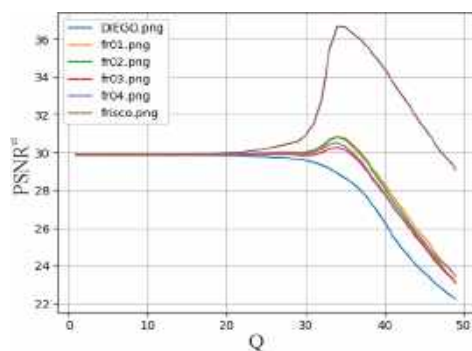
Рисунок Л.17 – Зображення Goldhill: безшумне (а); зашумлене, $\sigma^2=100$ (б); стиснуте ADCT, $b=3,5$ (в); стиснуте AGU, $b=3,5$ (г)



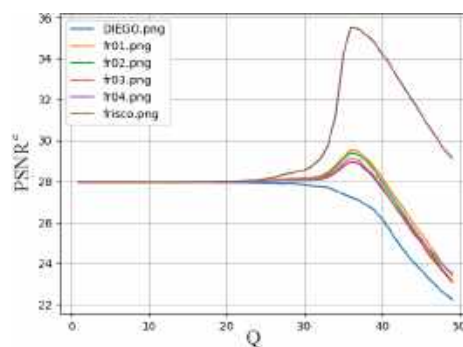
а

б

Рисунок Л.18 – Залежності PSNR-HVS (а) та WSNR (б) від b для різних кодерів, $\sigma^2=100$, зображення Barbara

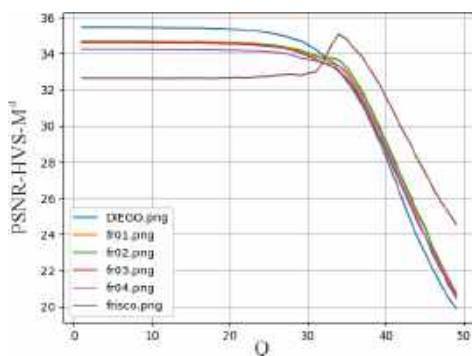


а

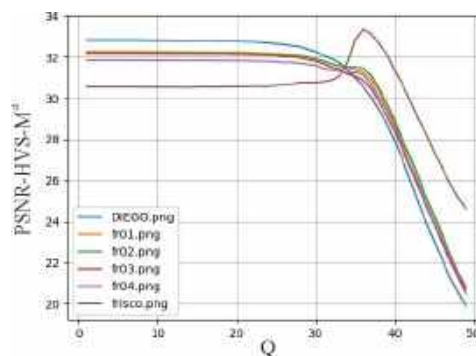


б

Рисунок Л.19 – Залежності PSNR від Q для кодера BPG для шести тестових зображень, спотворених АБГШ, з дисперсією, що дорівнює 64 (а) і 100 (б)

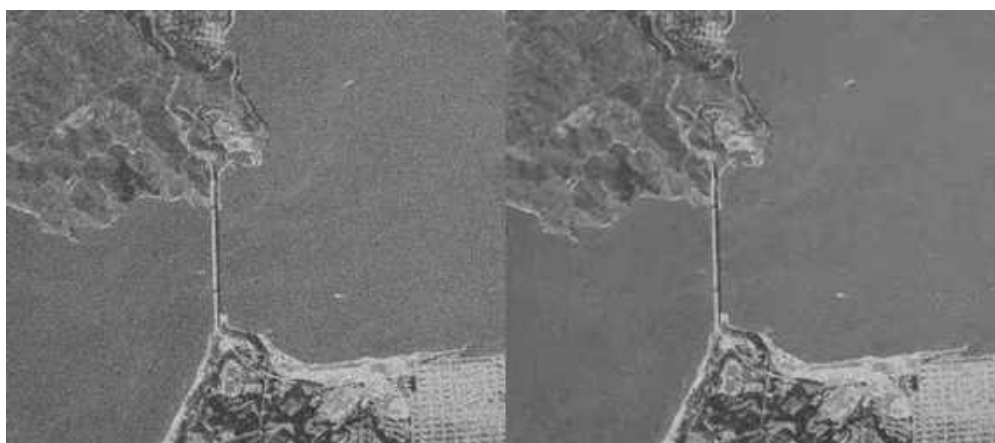


а



б

Рисунок Л.20 – Залежності PSNR-HVS-M від Q для кодера BPG для шести тестових зображень, спотворених АБГШ, з дисперсією, що дорівнює 64 (а) та 100 (б)



а

б

Рисунок Л.21 – Зашумлене зображення Frisco (а) і те саме зображення, стиснене в ОПТ (б), дисперсія шуму дорівнює 196

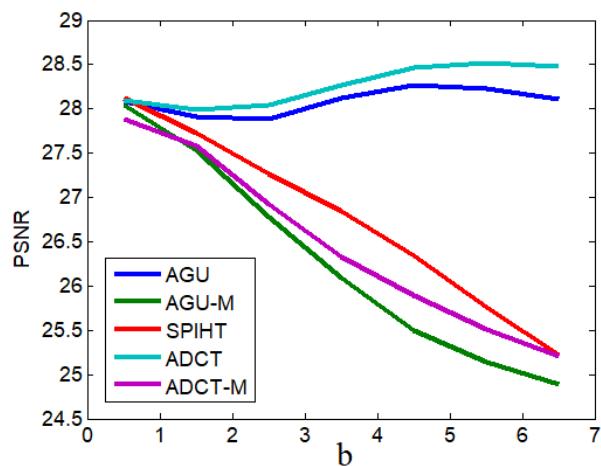


Рисунок Л.22 – Залежності PSNR(b) для досліджуваних кодерів, Barbara, спотворено просторово-корельованим шумом з дисперсією рівною 100

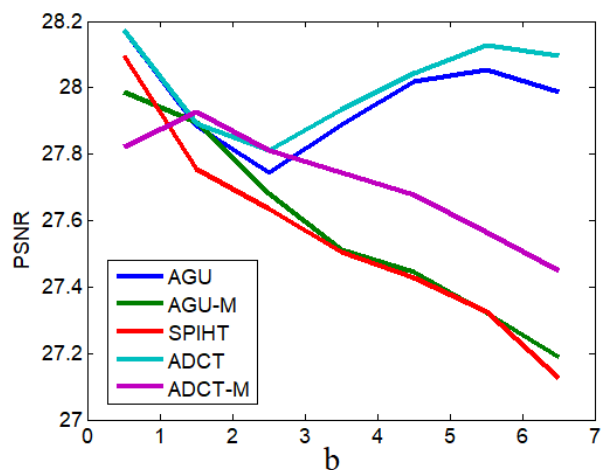


Рисунок Л.23 – Залежності PSNR(b) для досліджуваних кодерів, Goldhill, спотворено просторово-корельованим шумом з дисперсією рівною 100

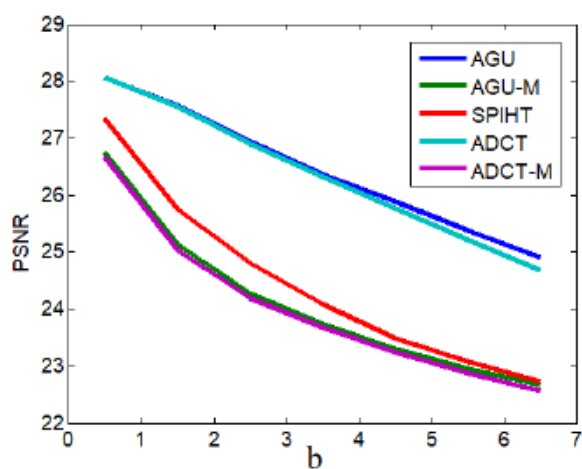


Рисунок Л.24 – Залежності PSNR(b) для досліджуваних кодерів, Baboon, спотворено просторово-корельованим шумом з дисперсією рівною 100

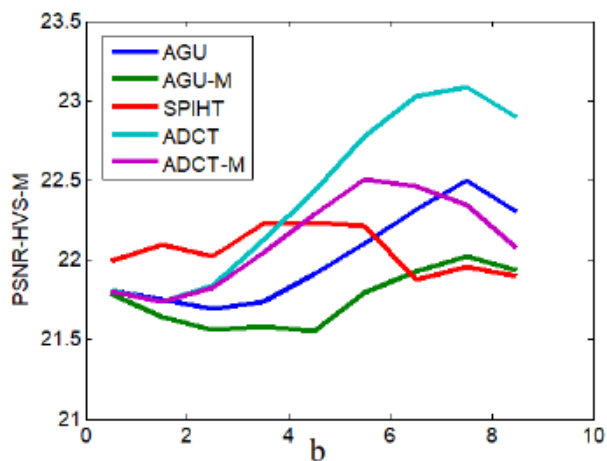


Рисунок Л.25 – Залежності PSNR-HVS-M(b) для досліджуваних кодерів, Goldhill, спотворено просторово-корельованим шумом з дисперсією рівною 200

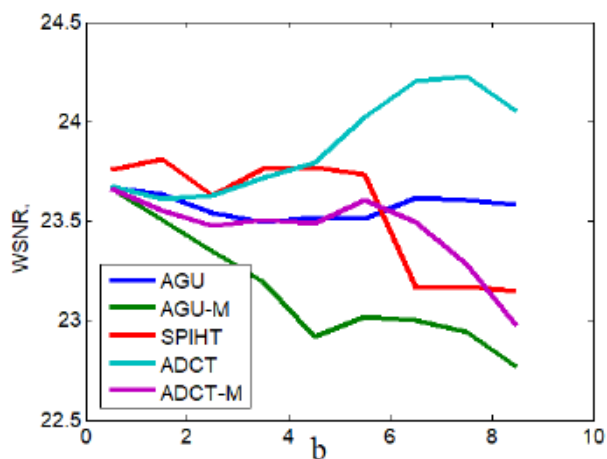


Рисунок Л.26 – Залежності WSNR(b) для досліджуваних кодерів, Goldhill, спотворено просторово-корельованим шумом з дисперсією рівною 200

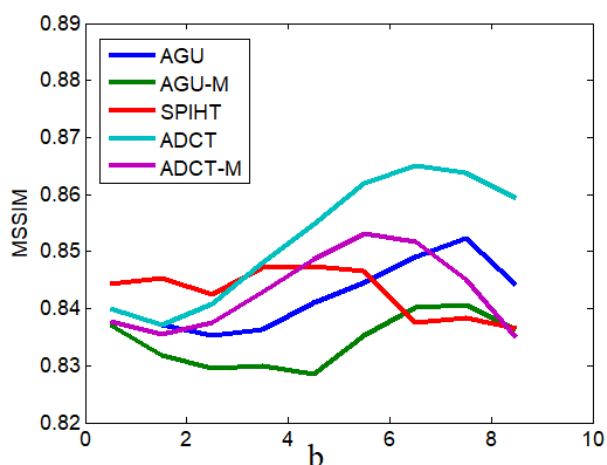
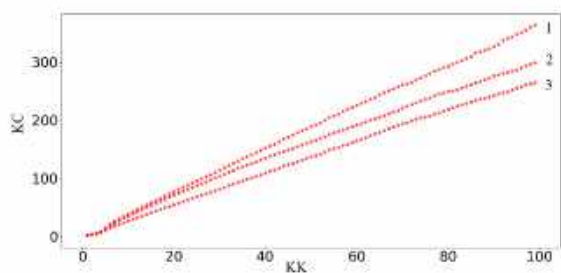


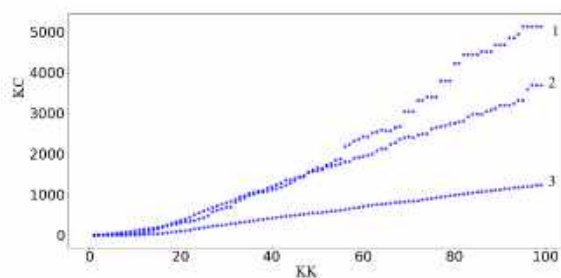
Рисунок Л.27 – Залежності MSSIM(b) для досліджуваних кодерів, Goldhill, спотворено просторово-корельованим шумом з дисперсією рівною 200

ДОДАТОК М

ДОДАТКОВІ МАТЕРІАЛИ ДО РОЗДІЛУ 5

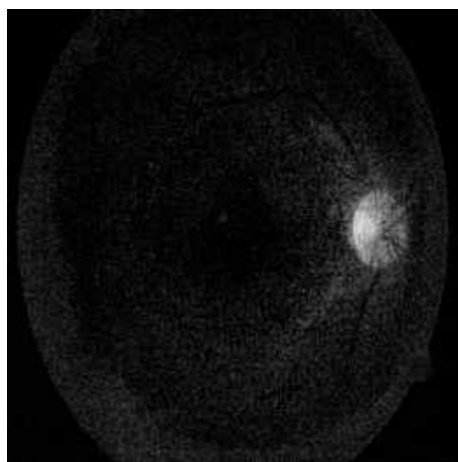


а

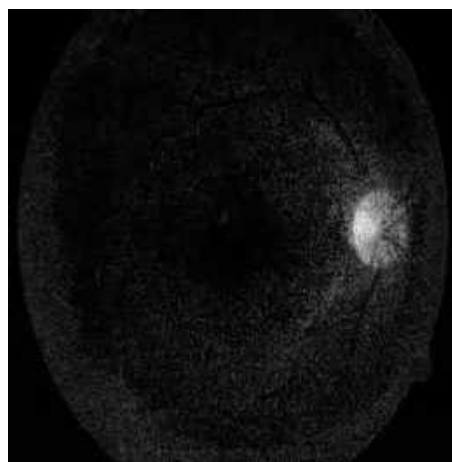


б

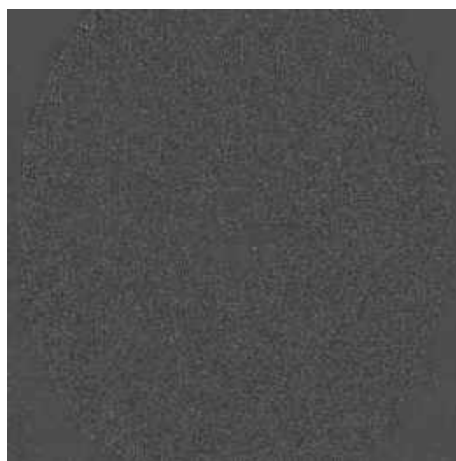
Рисунок М.1 – Залежності КС від КК для трьох різних тестових зображень (1, 2, 3) для R (а) і B (б) компонентів кольору



а

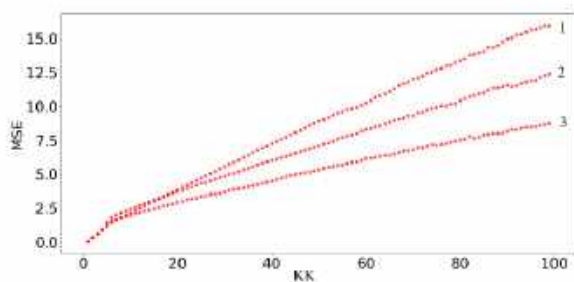


б

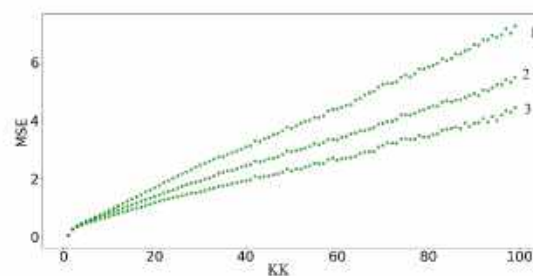


в

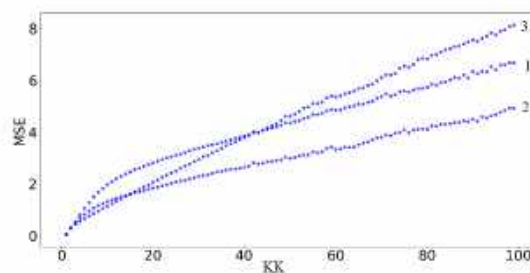
Рисунок М.2 – Оригінал (а), стиснене з КК=40 (б) і різницеве (в) зображення сітківки



а

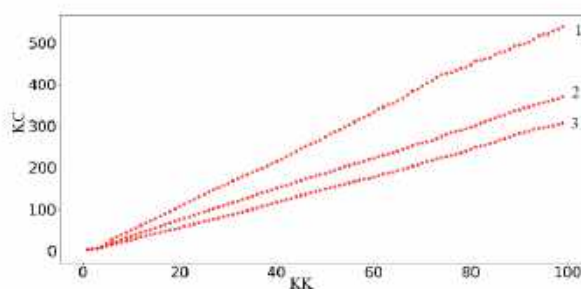


б

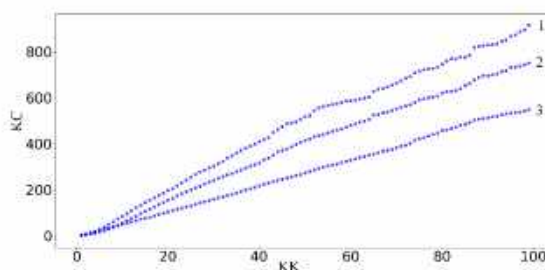


в

Рисунок М.3 – Залежності MSE від КК для трьох різних тестових зображень (1, 2, 3) для Y (а), Cr (б) і Cb (в) компонентів кольору

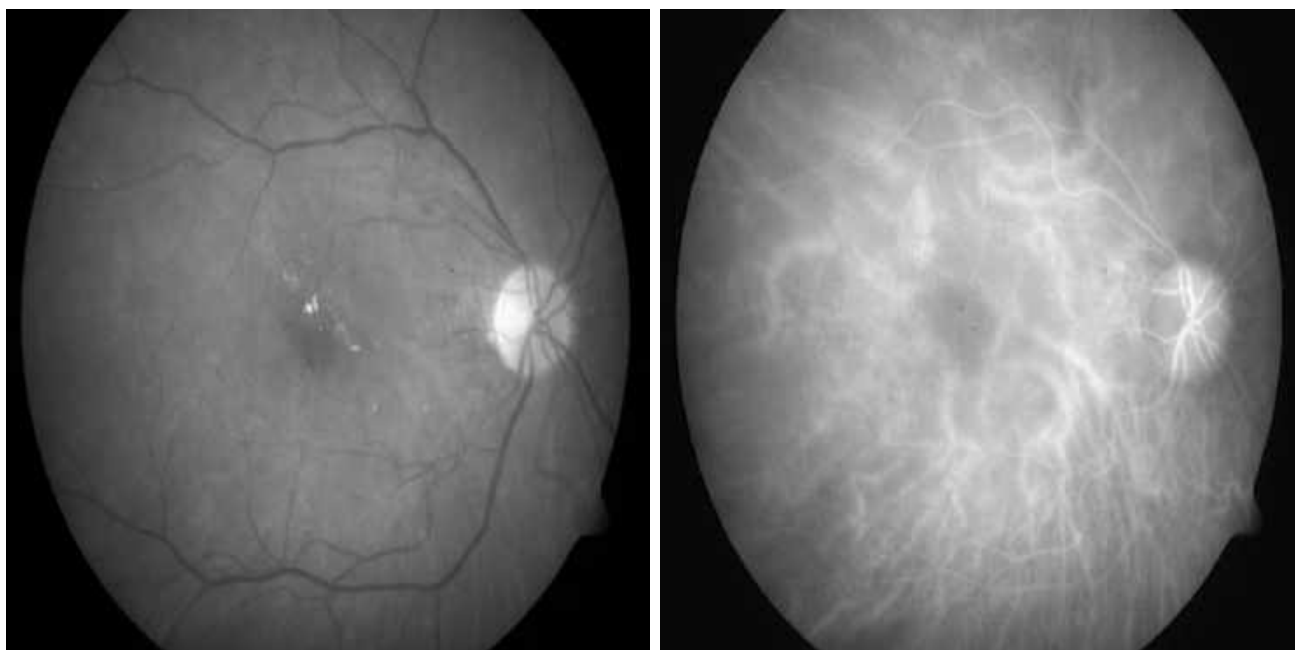


а



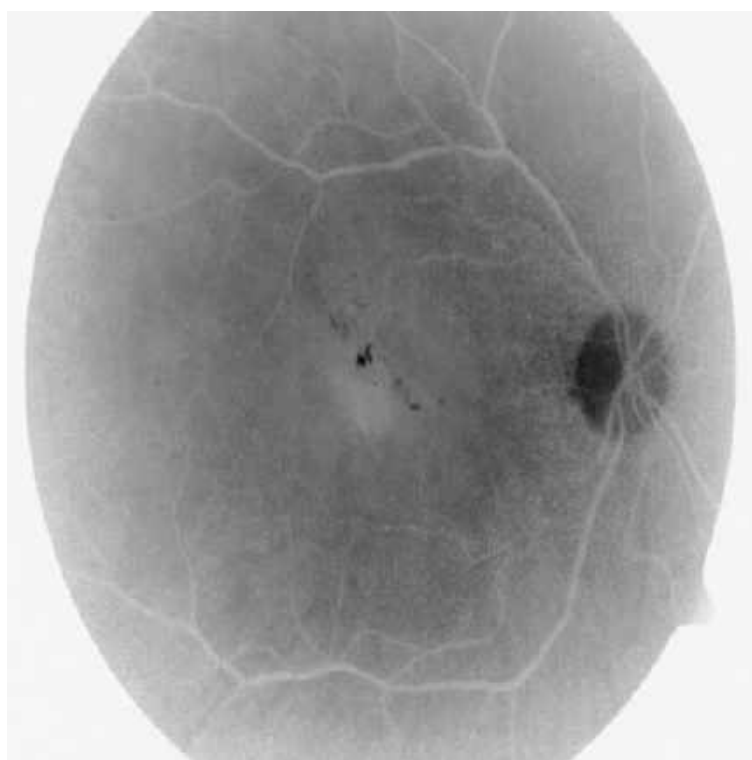
б

Рисунок М.4 – Залежності КС від КК для трьох різних тестових зображень (1, 2, 3) для Y (а) і Cb (б) компонентів кольору



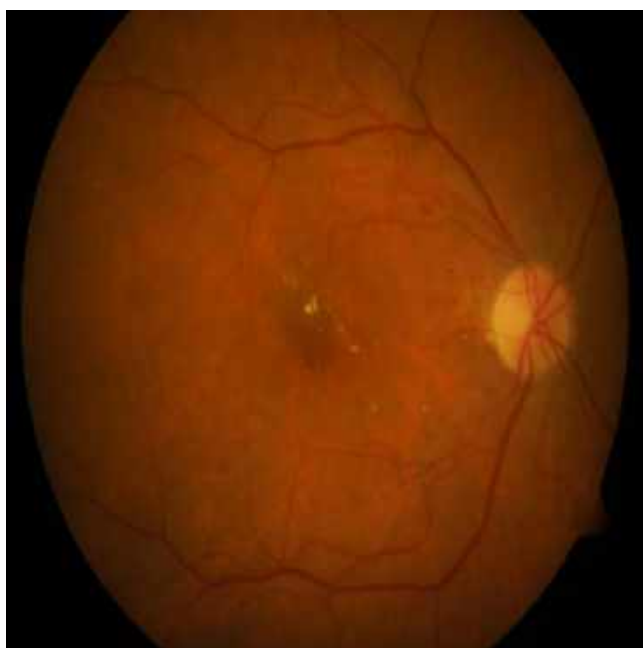
а

б

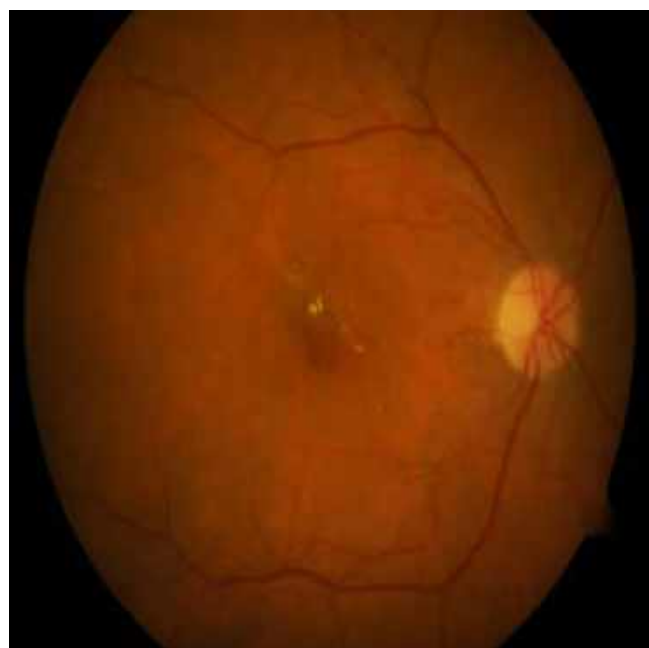


в

Рисунок М.5 – Y (а), Cr (б) і Сb (в) компоненти кольорового зображення сітківки



а



б

Рисунок М.6 – Оригінальне (а) і стиснуте (б) зображення сітківки

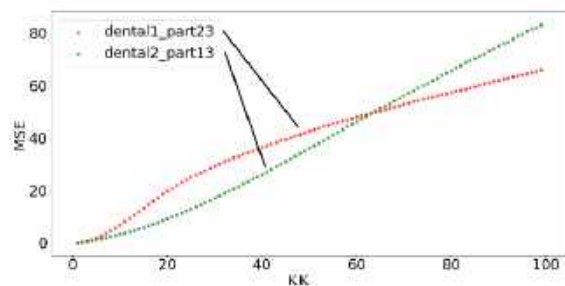


а

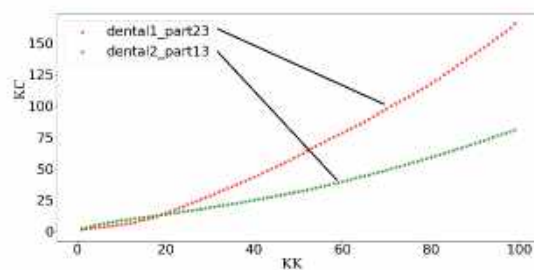


б

Рисунок М.7 – Два тестових фрагменти різної складності: dental1_part23 (а), dental2_part13 (б)

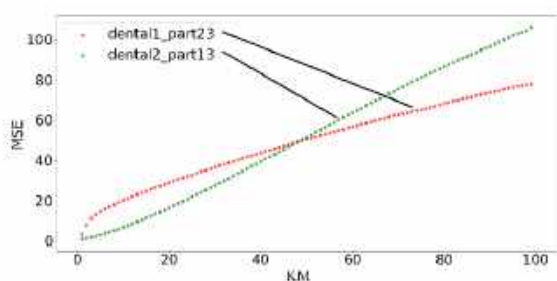


а

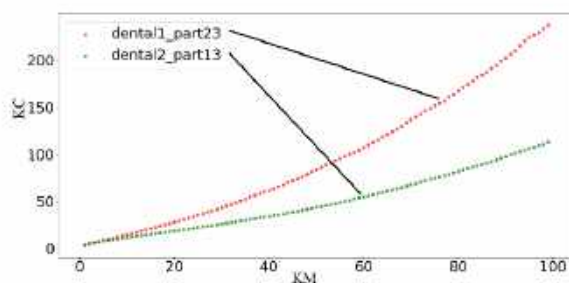


б

Рисунок М.8 – Залежності MSE (а) та KC (б) від KK для двох тестових зображень, кодер AGU

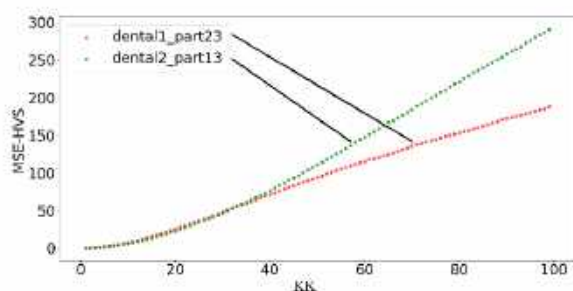


а

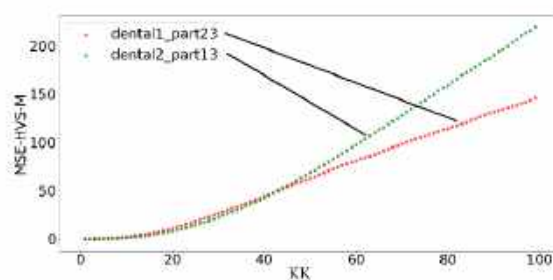


б

Рисунок М.9 – Залежності MSE (а) та KC (б) від KM для двох тестових зображень, кодер AGU-M



а



б

Рисунок М.10 – Залежності MSE_{HVS} (а) і MSE_{HVS-M} (б) від KK для двох тестових зображень, кодер AGU

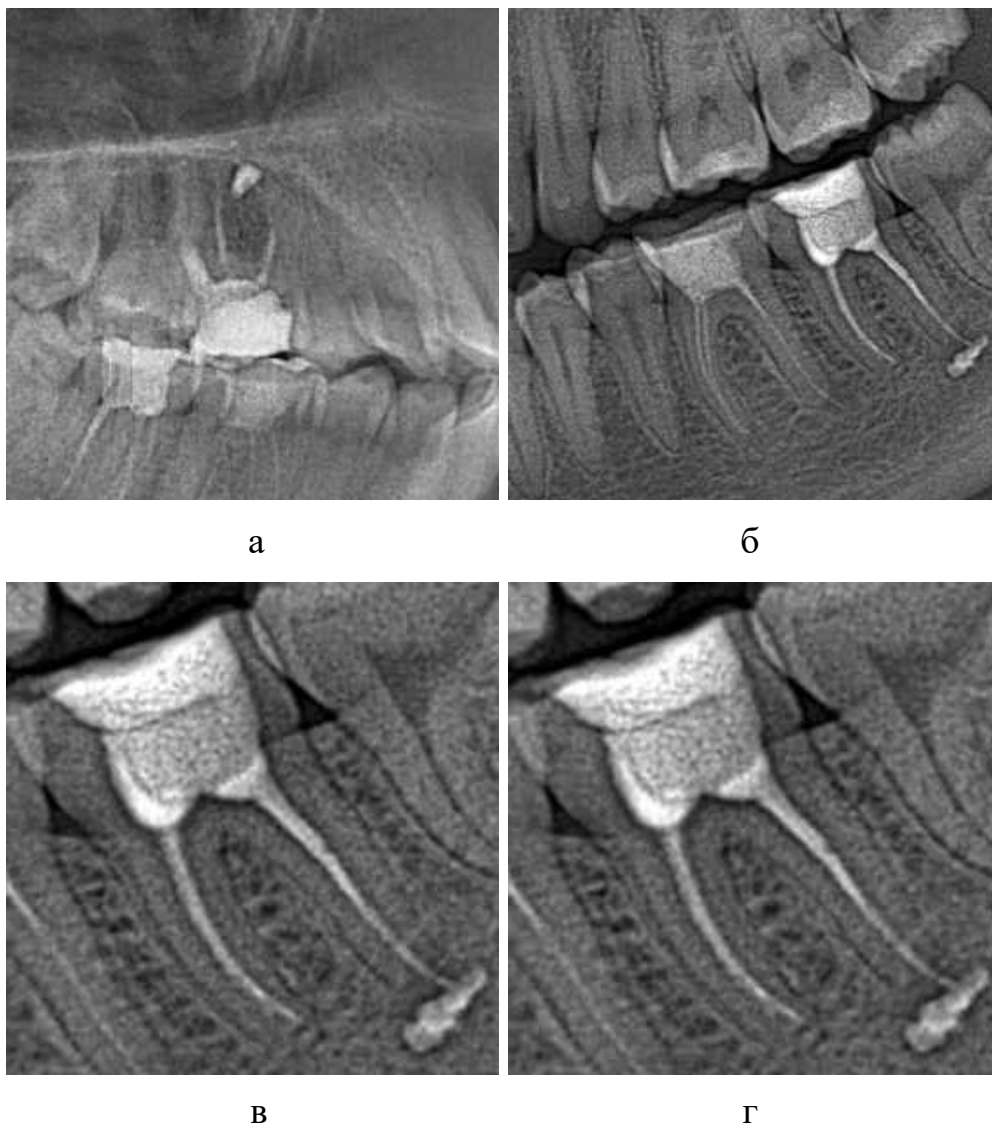


Рисунок М.11 – Тестові фрагменти, для яких $KM=8$, $KC=12,5$ (а), $KM=8$, $KC=10,6$ (б); збільшений фрагмент пролікованого зуба до (в) і після (г) компресії

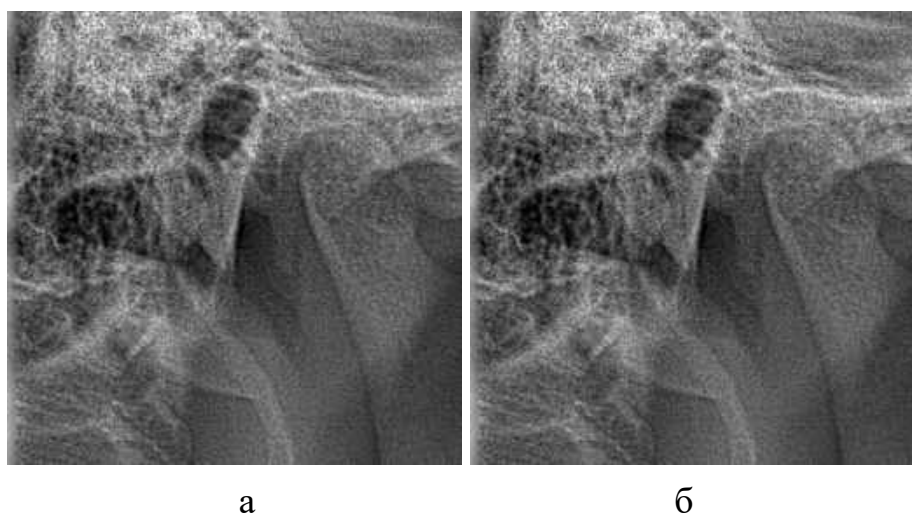
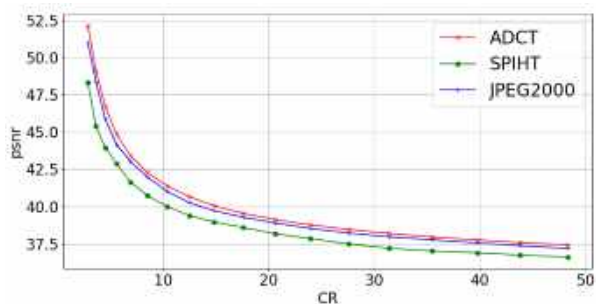
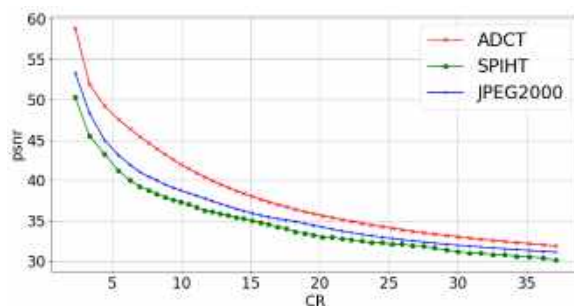


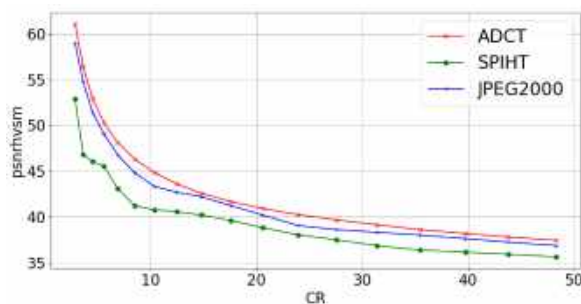
Рисунок М.12 – Фрагмент Dental2_part00: оригінал (а), $KM=8$, $KC=8,9$ (б)



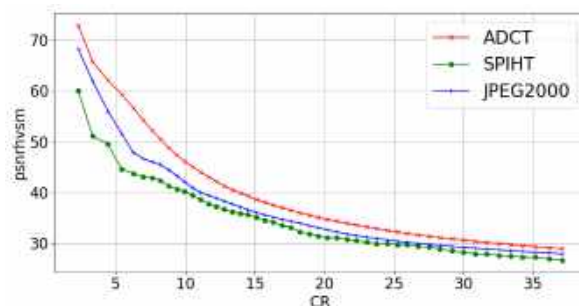
а



б

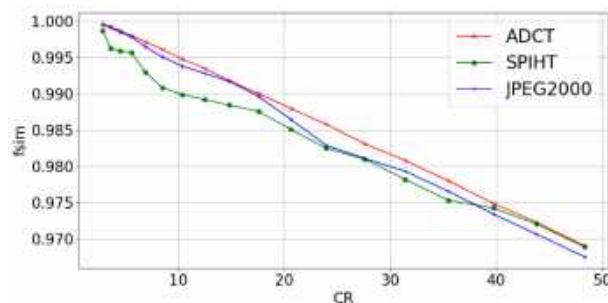


в

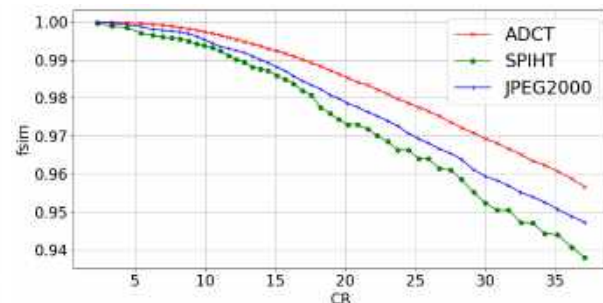


г

Рисунок М.13 – Залежності PSNR та PSNR-HVS-M від КС для фрагментів зубів складної структури (а, в) і простої структури (б, г)



а



б

Рисунок М.14 – Залежності FSIM від КС для фрагментів зубів складної структури (а) і простої структури (б)

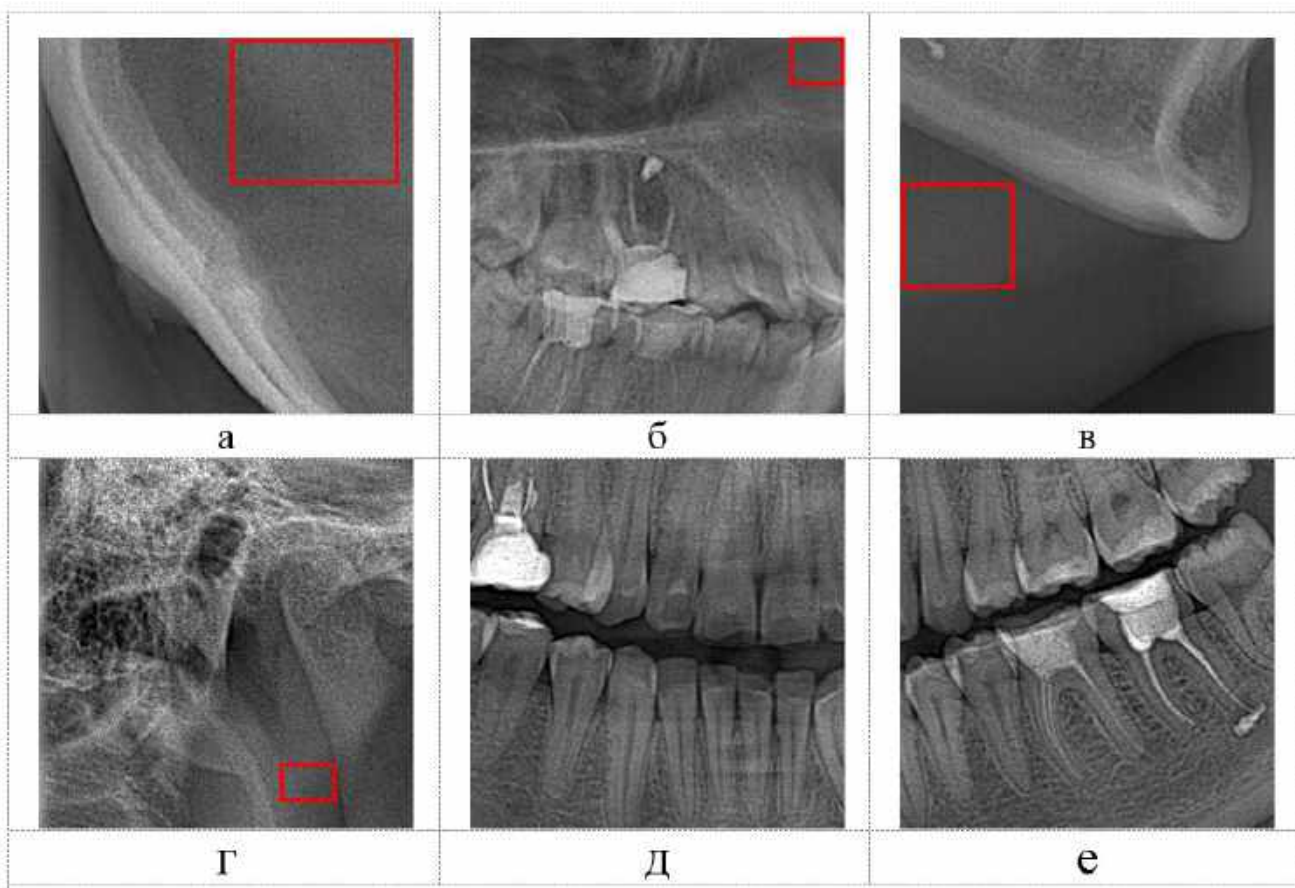


Рисунок М.15 – Фрагменти панорамного знімку зубів з індексами 2 (а), 9 (б), 12 (в), 13 (г), 19 (д), 20 (е)

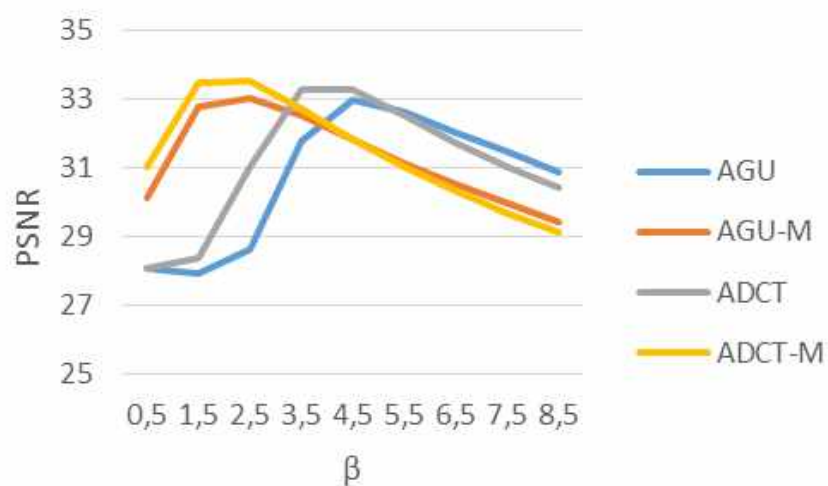


Рисунок М.16 – Залежності $PSNR_{tc}(\beta)$ для чотирьох методів стиснення тестового зображення MRT_Prepared, АБГШ $\sigma=10$

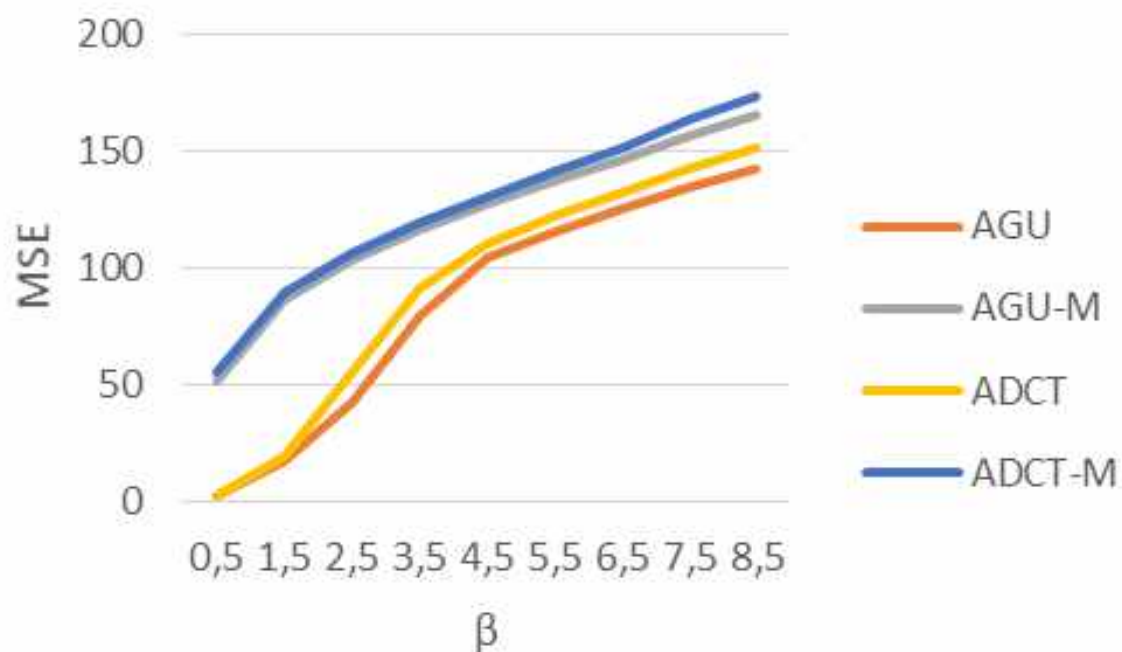


Рисунок М.17 – Залежності $MSE_{cn}(\beta)$ для чотирьох методів стиснення тестового зображення MRT_Prepared, АБГШ $\sigma=10$

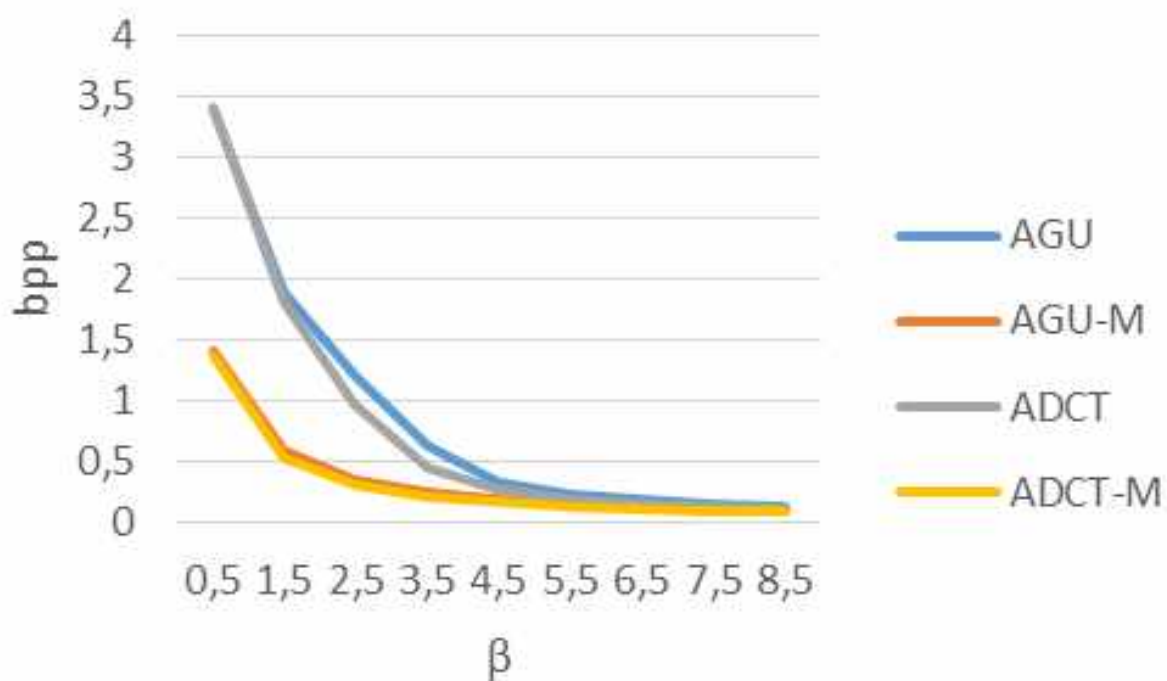


Рисунок М.18 – Залежності БНП(β) для чотирьох методів стиснення тестового зображення MRT_Prepared, АБГШ $\sigma=10$

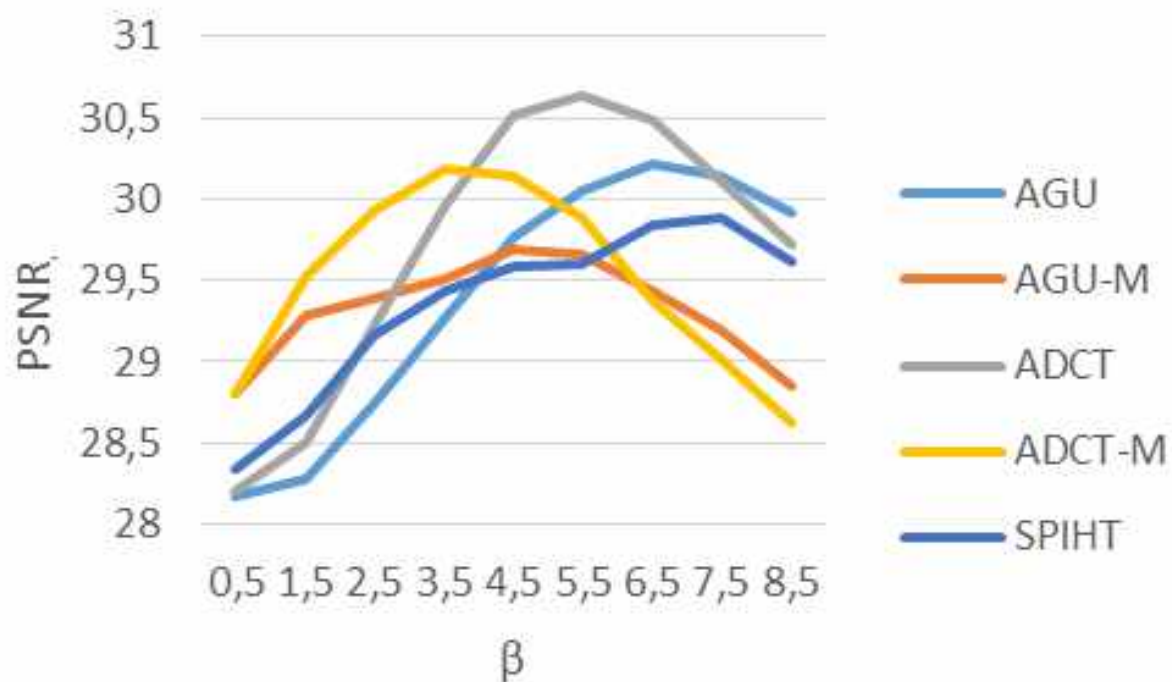


Рисунок М.19 – Залежності $PSNR_{ic}(\beta)$ для п'яти методів стиснення тестового зображення MRT_Prepared, СКШ, $\sigma=10$

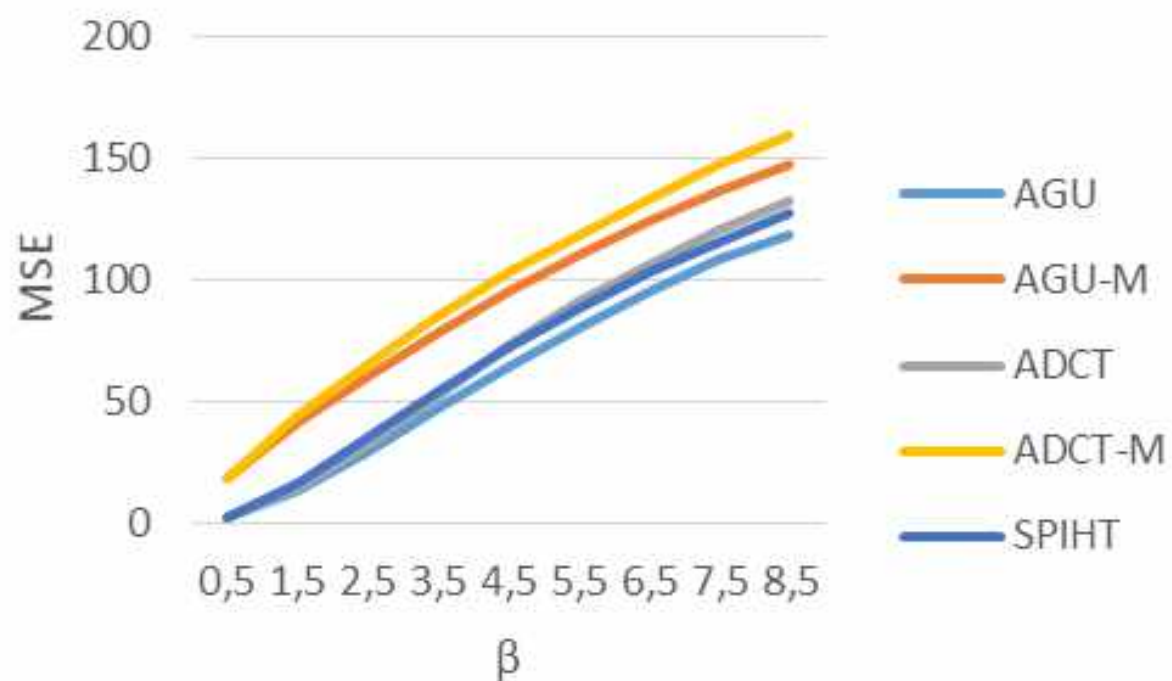


Рисунок М.20 – Залежності $MSE_{cn}(\beta)$ для п'яти методів стиснення тестового зображення MRT_Prepared, СКШ, $\sigma=10$

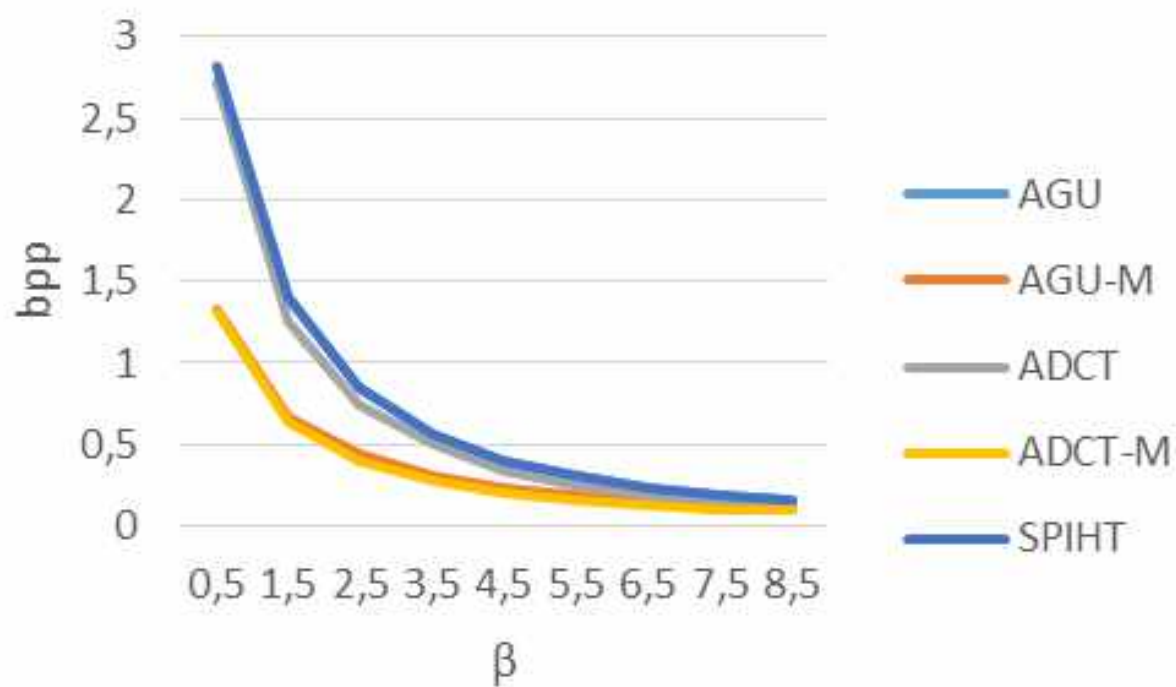


Рисунок М.21 – Залежності БПП(β) для п'яти методів стиснення тестового зображення MRT_Prepared, СКШ, $\sigma=10$

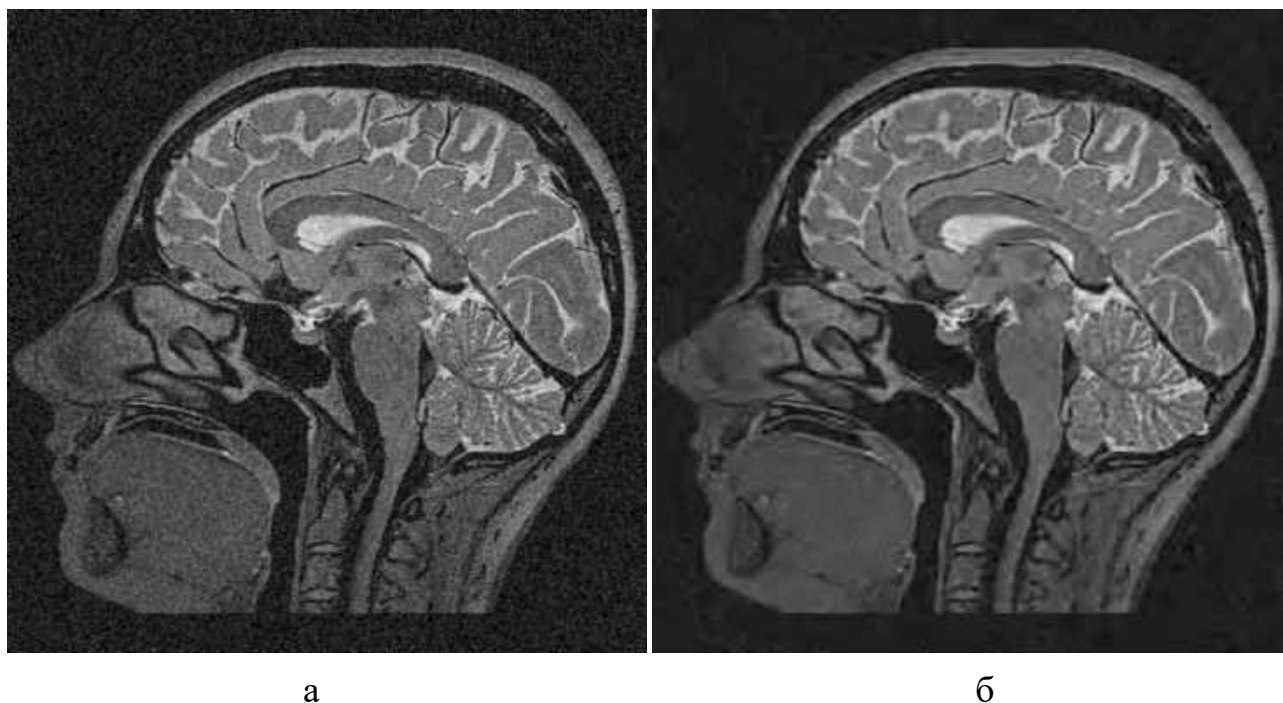


Рисунок М.22 – Зашумлене (а) та стиснене (б) зображення MRT_Prepared

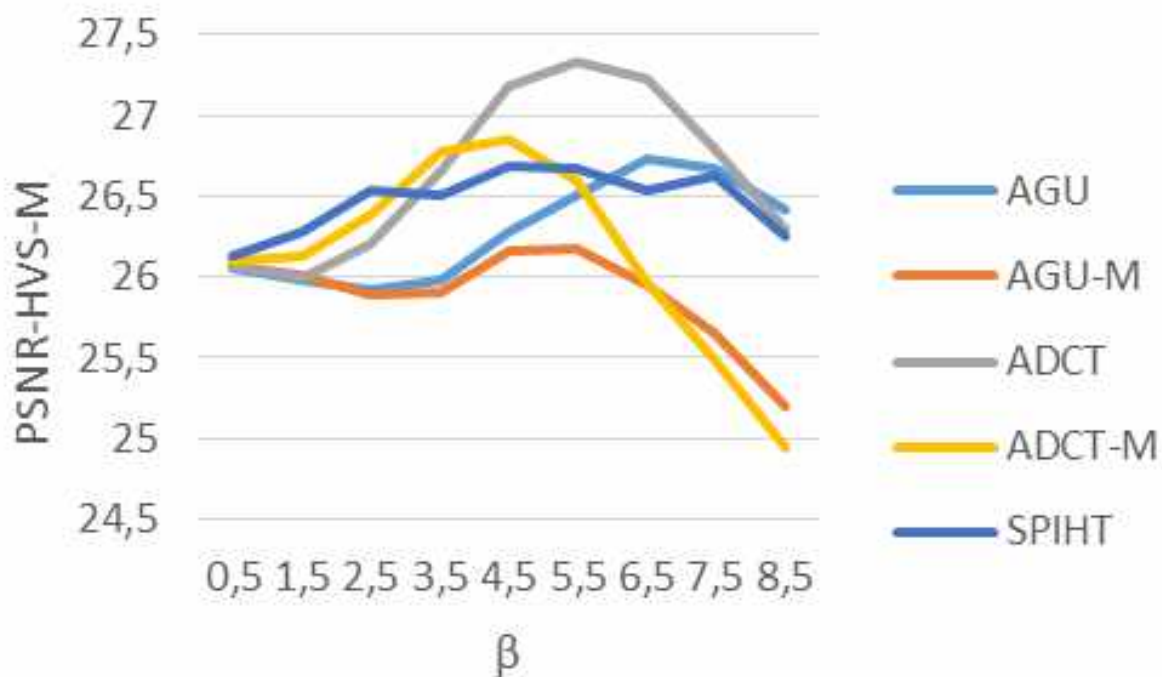


Рисунок М.23 – Залежності $PSNR-HVS-M_{tc}(\beta)$ для п'яти методів стиснення тестового зображення MRT_Prepared, СКШ, $\sigma=10$

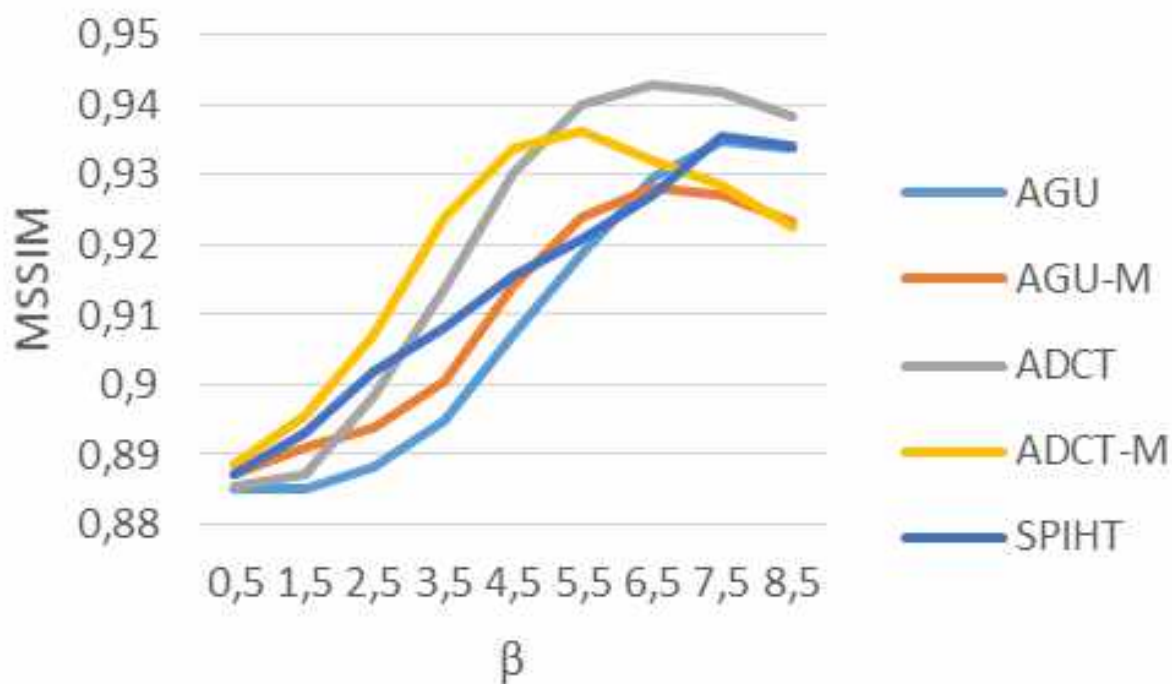


Рисунок М.24 – Залежності $MSSIM_{tc}(\beta)$ для п'яти методів стиснення тестового зображення MRT_Prepared, СКШ, $\sigma=10$

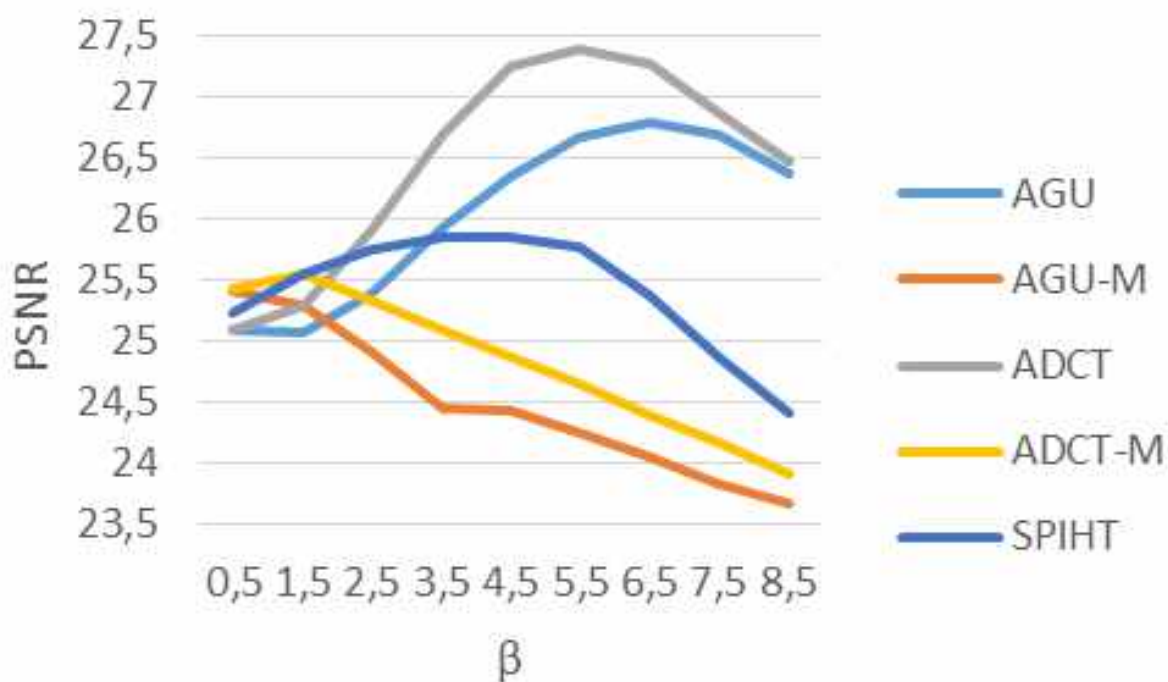


Рисунок М.25 – Залежності $PSNR_{tc}(\beta)$ для п'яти методів стиснення тестового зображення Barbara, СКШ, $\sigma^2=200$

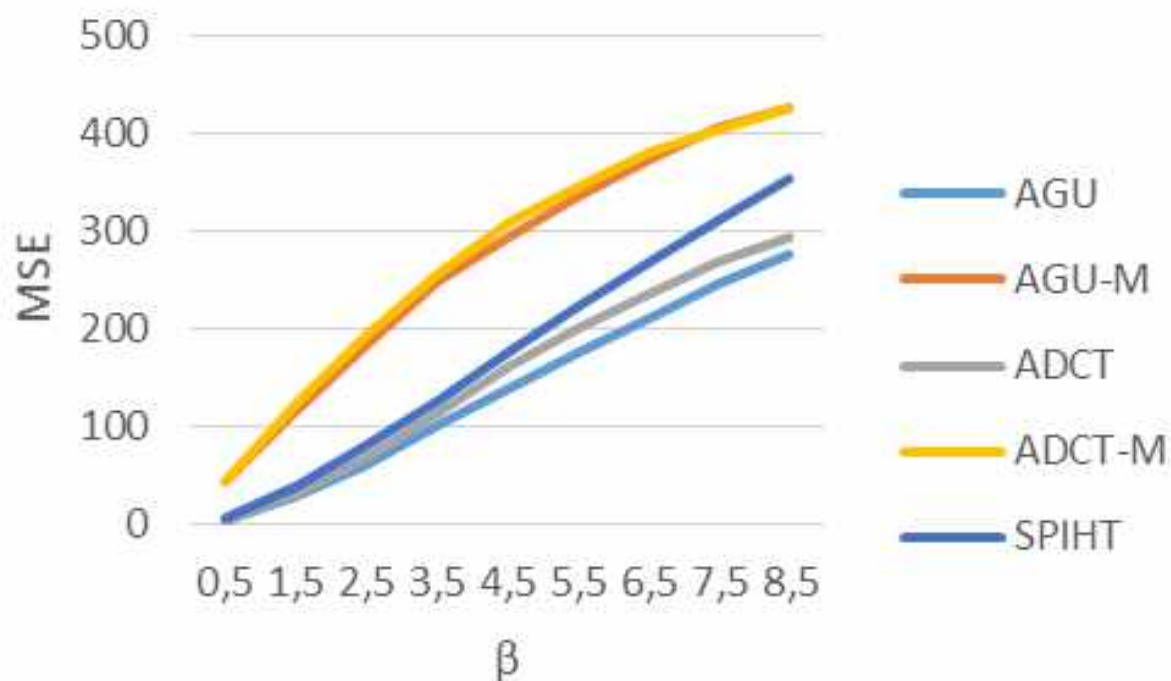


Рисунок М.26 – Залежності $MSE_{cn}(\beta)$ для п'яти методів стиснення тестового зображення Barbara, СКШ, $\sigma^2=200$

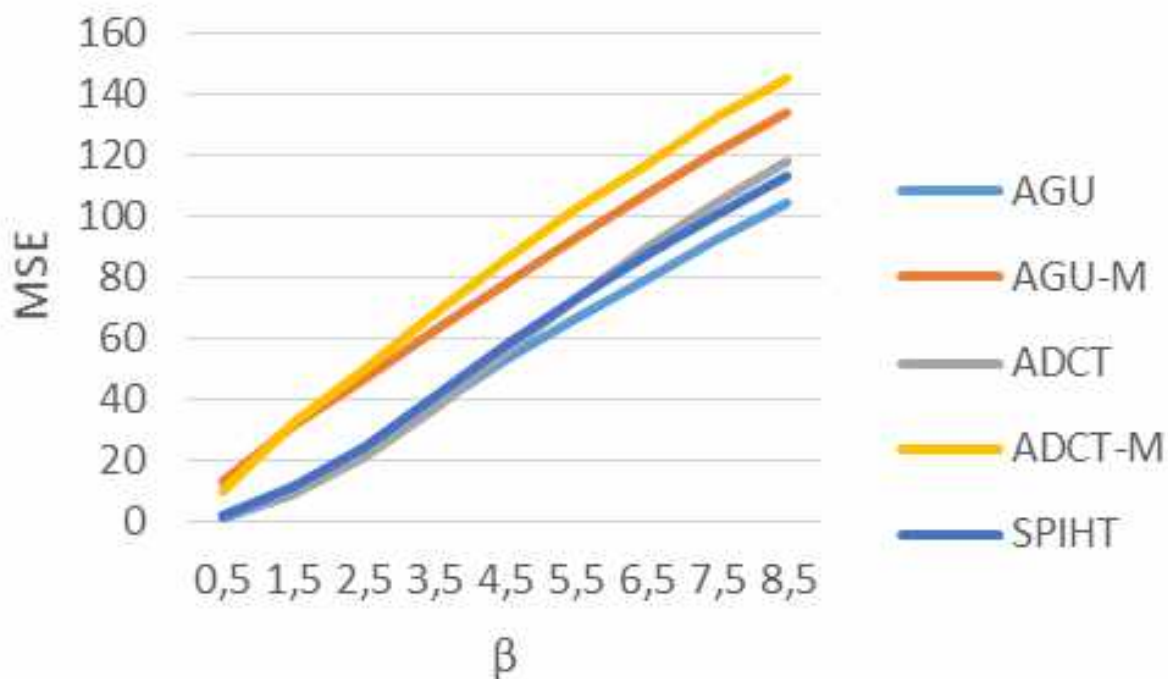


Рисунок М.27 – Залежності $MSE_{\text{сн}}(\beta)$ для п'яти методів стиснення тестового зображення Barbara, ВКШ, $\sigma=10$

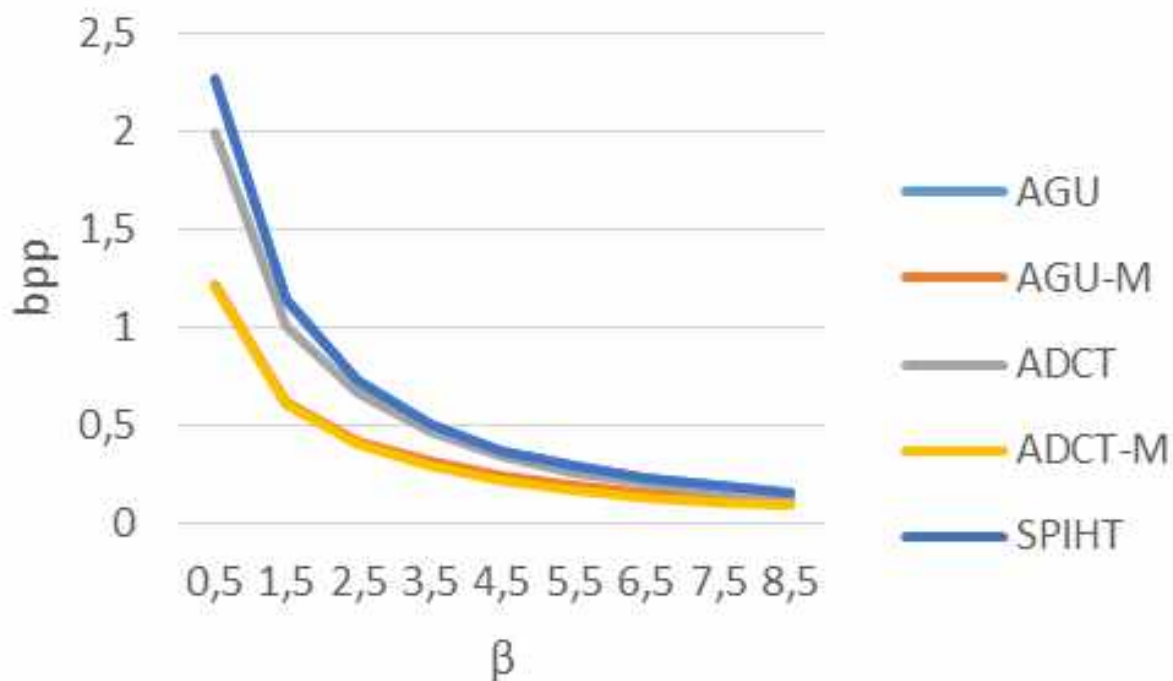


Рисунок М.28 – Залежності БНП(β) для п'яти методів стиснення тестового зображення Barbara, ВКШ, $\sigma=10$ (криві для кодерів AGU і ADCT, а також кодерів AGU-M і ADCT-M практично збігаються)

ДОДАТОК Н
АКТ ПРО ВПРОВАДЖЕННЯ РЕЗУЛЬТАТІВ ДОКТОРСЬКОЇ ДИСЕРТАЦІЇ
В ІНСТИТУТІ РАДІОФІЗИКИ ТА ЕЛЕКТРОНІКИ ІМ. О.Я. УСИКОВА
НАН УКРАЇНИ

ЗАТВЕРДЖУЮ

В.о. директора Інституту радіофізики
та електроніки ім. О.Я. Усикова НАН України

Юрій Логвінов

«25» грудня 2024 р.

АКТ

науково-технічної комісії щодо впровадження наукових положень і
результатів дисертації на здобуття ступеня доктора технічних наук
докторанта кафедри інформаційно-комунікаційних
технологій ім. О.О. Зеленського
Національного аерокосмічного університету «ХАІ»
Кривенка Сергія Станіславовича

Науково-технічна комісія Інституту радіофізики та електроніки ім. О.Я. Усикова НАН України у складі голови комісії виконуючого обов'язки заступника директора з наукової роботи, доктора фіз.-мат. наук Когута Олександра Євгеновича та членів комісії завідуючого відділом фізичних основ радіолокації канд. фіз.-мат. наук Могилі Анатолія Андрійовича та старшого наукового співробітника відділу фізичних основ радіолокації, доктора фіз.-мат. наук Миценко Ігоря Михайловича
склала даний акт про те, що в результаті виконання спільних науково-дослідних робіт, були використані та впроваджені наступні положення та результати, що отримані докторантом Кривенком С.С.:

- методи фільтрації радіолокаційних зображень, що спотворені сигнально-залежними та/або просторово-корельованими завадами;
- метод прогнозування ефективності фільтрації радіолокаційних зображень для кількох метрик якості.

Розроблені в дисертації методи та програмно-алгоритмічні засоби дозволили:

- 1) підвищити якість радіолокаційних даних у відповідності до різних критеріїв;
- 2) забезпечити прогнозування характеристик фільтрації та вибір найкращих типу та параметрів фільтрів.

Голова комісії:

Олександр Когут

Члени комісії:

Анатолій Могила

Ігор Миценко

ДОДАТОК П
АКТ ПРО ВПРОВАДЖЕННЯ РЕЗУЛЬТАТІВ ДОКТОРСЬКОЇ ДИСЕРТАЦІЇ
В КНП ХОР «ОБЛАСНА ДИТЯЧА КЛІНІЧНА ЛІКАРНЯ»

«ЗАТВЕРДЖУЮ»
Директор
КНП ХОР «ОБЛАСНА ДИТЯЧА
КЛІНІЧНА ЛІКАРНЯ»
к.мед.н. Муратов Г.Р.
« 3 » 2025 р.



АКТ
комісії щодо впровадження результатів
дисертації на здобуття ступеня доктора технічних наук докторанта
Національного аерокосмічного університету ім. М.Є. Жуковського «ХАІ»
Кривенка Сергія Станіславовича

Локальна етична комісія у складі: голова комісії: Халтуріна Т.О., член комісії: Білогуров С.М., секретар: Койда М.Л., склала даний акт про те, що в результаті виконання спільних науково-дослідних робіт були використані та впроваджені наступні результати, що отримані докторантом Кривенком С.С.:

- метод прогнозування показників якості первинних рентгенівських зображень, що забезпечує стиснення зображень без втрати діагностично важливої інформації (текст додається);

- метод стиснення медичних зображень із урахування спотворень, що вносяться апаратурою формування зображень, та забезпеченням найкращих з точки зору візуальної якості стиснення (текст додається).

Місце впровадження: рентгенологічне відділення КНП ХОР «Обласна дитяча клінічна лікарня».

Джерела інформації: 1) Lukin, V., Krivenko, S., Kaluzhinov, I., Krylova, O., Kryvenko, L. (2022). Lossy Compression of Remote Sensing and Dental Images Corrupted by Spatially Correlated Noise. In: Nechyporuk, M., Pavlikov, V., Kritskiy, D. (eds) Integrated Computer Technologies in Mechanical Engineering - 2021. ICTM 2021. Lecture Notes in Networks and Systems, vol 367. Springer, Cham. https://doi.org/10.1007/978-3-030-94259-5_77. 2) L. Kryvenko, O. Krylova, V. Lukin, and S. Kryvenko, "Intelligent visually lossless compression of dental images," Adv. Opt. Technol., vol. 13, p. 1306142, Feb. 2024, doi: 10.3389/aot.2024.1306142.

Розроблені в дисертації методи та програмно-алгоритмічні засоби було впроваджено у діагностичну діяльність КНП ХОР «Обласна дитяча клінічна лікарня», що дозволило покращити лікувально-діагностичний процес пацієнтів дитячого віку за рахунок оптимізації рентгенологічних методів дослідження пацієнта.

Голова комісії

Члени комісії

Секретар



Тетяна ХАЛТУРІНА

Сергій БІЛОГУРОВ

Марія КОЙДА

ДОДАТОК Р
АКТ ПРО ВПРОВАДЖЕННЯ РЕЗУЛЬТАТІВ ДОКТОРСЬКОЇ ДИСЕРТАЦІЇ
В ТОВ НК «КОЛІБРІ»



Директорка ТОВ НК КОЛІБРИ
Щапохіна Ольга Анатоліївна

«29» січня 2025 р.

АКТ

НАУКОВО-ТЕХНІЧНОЇ КОМІСІЇ ЩОДО ВПРОВАДЖЕННЯ НАУКОВИХ ПОЛОЖЕНЬ І РЕЗУЛЬТАТІВ ДОКТОРСЬКОЇ ДИСЕРТАЦІЇ ДОКТОРАНТА НАЦІОНАЛЬНОГО АЕРОКОСМІЧНОГО УНІВЕРСИТЕТУ ІМ. М.С. ЖУКОВСЬКОГО «ХАІ» КРИВЕНКА СЕРГІЯ СТАНІСЛАВОВИЧА

Науково-технічна комісія у складі: голова комісії провідний спеціаліст Щебликін А. О., члени комісії: провідний спеціаліст Колеснік О. А., провідний спеціаліст Семененко К. П. склала даний акт про те, що в результаті виконання спільних науково-дослідних робіт були використані та впроваджені наступні наукові положення та результати, що отримані докторантом Кривенком С. С.:

- методи оцінки якості медичних сигналів за допомогою методів стиснення оптичних зображень;

- двоетапні процедури стиснення зображень з втратами для забезпечення бажаної якості стиснення зображення з втратами відповідно до метрик візуальної якості;

Розроблені в дисертації методи та програмно-алгоритмічні засоби дозволили:

- 1) суттєво спростити процедуру оцінки ефективності фільтрації та стиснення зображень, що спотворені шумами складної природи;

- 2) прискорити процедуру стиснення медичних та оптичних зображень із збереженням інформаційних ознак, що дозволяє використовувати наявні автоматичні методи класифікації зображень;

Отримані результати було використано під час обробки великих обсягів даних пацієнтів з публічних та приватних баз даних, що використовуються для дослідження роботи приладу персональної медичної діагностики KOLIBRI.

ГОЛОВА КОМІСІЇ

ЧЛЕНИ КОМІСІЇ:

ЩЕБЛИКІН А.О.

КОЛЕСНИК О.А.

СЕМЕНЕНКО К.П.

ДОДАТОК С
АКТ ПРО ВПРОВАДЖЕННЯ РЕЗУЛЬТАТІВ ДОКТОРСЬКОЇ ДИСЕРТАЦІЇ
В ТОВ «БІОПРОМІНЬ»



«Затверджую»

Директор ТОВ БІОПРОМІНЬ

А. А. Пулавський

27 грудня 2024 року

(дата)

АКТ

впровадження результатів наукових досліджень докторанта Національного аерокосмічного університету ім. М.Є. Жуковського «ХАІ»

Кривенка Сергія Станіславовича

Комісія в складі: голова – Сидора В.Л., члени комісії – Ахтемійчук М.І., Баранніков С.О. склала цей акт про те, що такі наукові положення і результати, одержані Кривенком Сергієм Станіславовичем у його статті «Blind denoising of dental X-Ray images», яку використано в його дисертації на здобуття наукового ступеня доктора технічних наук, а саме

- двоетапний підхід до придушення шуму - метод, що складається з двох нейронних мереж (NSPNet і GDNet). Перша оцінює спектр і рівень шуму для кожного пікселя, створюючи карти стандартних відхилень на трьох масштабах. Друга мережа використовує ці карти для ефективного придушення шуму. Це дозволяє значно підвищити ефективність фільтрації зображень в складних умовах;
- метод адаптації до просторово корельованого шуму. Ця модель ефективно працює із шумом зі змінним спектром і рівнем, що складно оцінити заздалегідь. Це досягнуто через навчання з використанням згенерованих шумових даних різної природи. Це дозволяє працювати в умовах шумів складних типів з невідомими ап-ріорі параметрами;

були використані при розробці програмного забезпечення медичних приладів для підвищення ефективності вторинної обробки і аналізу зашумлених даних, отриманих з медичних датчиків.

Голова комісії

Члени комісії

Сидора В.Л.

Ахтемійчук М.І.

Баранніков С.О.