Radioelectronic and Computer Systems, 2023, no. 1(105)
ISSN 1814-4225 (print)
ISSN 2663-2012 (online)

UDC 004.932.4
doi: 10.32620/reks.2023.1.09

Vladimir LUKIN[1], Ekaterina BATAEVA[2], Sergey ABRAMOV[1]

[1] *National Aerospace University "Kharkiv Aviation Institute", Kharkiv, Ukraine*
[2] *Open International University of Human Development Ukraine, Kyiv, Ukraine*

SALIENCY MAP IN IMAGE VISUAL QUALITY ASSESSMENT AND PROCESSING

Images are mainly viewed and analyzed by humans. Because of this, in the characterization of image quality and effectiveness of image processing, it is necessary to take into account the peculiarities of the human vision system and cognition that are very complex. Saliency maps as well as priority and meaning maps introduced recently are the attempts to incorporate specific features of human vision into image analysis and processing fields. Many authors that consider the aforementioned maps consider them from different viewpoints. Thus, the basic ***subject*** *of this paper is the factors that influence and determine these maps. Among such factors, there are low-level features as well as social and psychological ones such as emotions, age, and life values. The main* ***goal*** *of this paper is to give a brief survey of these factors and to consider how maps are already used in image quality assessment and processing as well as how they can be employed in the future. The* ***tasks*** *of the paper are to provide a definition of saliency, priority, and meaning maps, to analyze the factors that influence these maps, and to evaluate what improvement can be obtained due to taking maps into account in the assessment of image visual quality and such image processing operations as quality assessment, denoising, and lossy compression. The main* ***result*** *is that, by taking saliency maps into account, image quality assessment and processing efficiency can be sufficiently improved, especially for applications oriented on image viewing and analysis by observers or customers. This can be done by the simple weighting of local estimates of a given metric with further aggregation as well as by approaches based on neural networks. Using different quantitative criteria, we show what positive results can be got due to incorporating maps into quality assessment and image processing. As* ***conclusion****, we present possible directions of future research that are mainly related to an adaptation of denoising and lossy compression parameters to peculiarities of human attention.*

Keywords: *saliency map; quality assessment; image processing.*

1. Introduction

Imaging and images have become more and more widely employed in various areas such as agriculture, ecological monitoring, forestry, and advertising [1-3], etc. In some applications, including medicine and advertising, images are mostly subject to visual inspection or viewing by specialists and non-specialists, although computer methods and tools of data analysis are used as well [4, 5]. Finally, in applications such as remote sensing, images are mostly processed by computer means or specialized hardware/software, although visual analysis and interpreting are still used [6]. However, it is clear that, in any case, it is important how humans perceive and analyze images.

These studies intended for clarification of the main features of human perception of images were started long ago [7-9] by psychologists and neurobiologists. It was understood more than 25 years ago that primates (and humans in particular) have a wonderful ability to interpret natural scenes in real-time by "selecting a subset of the available sensory information before further processing [10], most likely to reduce the complexity of scene analysis [11]". Starting from the paper [12] cited by more than 10000 times, the aspects of visual attention and visual saliency have become a multidisciplinary topic intensively considered in image analysis and processing (see [13-15] and references therein). Note that, along with the saliency map concept, meaning and priority map concepts have been put forward and considered later (see, for example, [16, 17]). Despite several differences between them that will be discussed in the next two Sections, all these maps reflect the same fact that images are scanned and analyzed by humans, specifically where there are points (areas) attracting higher attention than other areas.

Studies on saliency and other maps have been divided into several directions. Eye gaze tracking devices have allowed getting statistics of gaze fixations in image areas [18, 19]. Significant efforts have been made to provide appropriate coincidence between true (obtained by processing data from humans) and predicted (obtained by automatic image processing) saliency maps (see, e.g., [20-22] and references therein). Special attention has been paid to the use of visual saliency maps in image processing including quality assessment [23, 24], denoising [25], lossy compression of images [26] and video [27], image deblurring [28] and inpainting [29]. The latter direction, in general, is explained by the fact that, since images and videos are mostly viewed by humans, processing has to be adapted to the peculiarities of the human vision system (HVS) and there are numerous ways to do so.

The goal of this paper is threefold. First, we would like to briefly introduce an interested reader to the theory and practice of visual saliency, give definitions of the considered types of maps, and discuss the types of features they take into account. Second, we consider the main applications of interest where saliency maps are already employed and analyze what improvements in image processing performance can be attained due to this. Third, we try to predict what can be further tendencies in incorporating saliency maps in image processing applications.

The paper is organized as follows. Definitions are presented in Section 2. Factors that might influence saliency are briefly considered in this Section as well. Section 3 deals with quality assessment taking the saliency map into account. In turn, other image processing applications where the use of a saliency map can be beneficial are discussed in Section 4. Finally, the conclusions and directions for further studies are given in Section 5.

2. Map definitions and factors influencing saliency

Let us differentiate some concepts in modern salience map theory – saliency map, meaning map, and priority map. The saliency map concept has been provided in psychological and computational theories and it refers to some physical, bottom-up properties of visual objects, such as color, luminance, motion [8, 11, 12], size, location [14], shape, brightness, oriented edges [16], and depth [30]. Numerous experiments have demonstrated that humans tend to primarily perceive larger details of visual images, focus on brighter objects that are in the foreground, and are well-illuminated. However, human attention is attracted not only by physical but also by the semantic, top-down features of images, and this process is influenced by life experiences, emotions, social values, age, gender, and profession of persons.

The priority map concept was developed later than the saliency map but earlier than the meaning map [13, 16, 31, 32]. J. H. Fecteau and D. P. Munoz recommend using the term priority map "to properly reflect the combined roles of salience and relevance in this selection process" [16]. It includes two main qualities – the physical, bottom-up, sensory-driven distinctiveness of visual objects, and the semantic, top-down, task- and goal-relevant information. The last aspect partially intersects with the meaning map concept because it can include different semantic constructs, not only the volitional tasks and cognitive goals of an observer but also the meaning attitudes of a viewer.

As has already been mentioned, there are many factors that may determine human attention. The saliency map is a two-dimensional map that quantifies attention. "Its amplitude at a given point represents how perceptually conspicuous the corresponding region is in visual space. And it does not matter what caused it to be conspicuous" [33].

Many authors have examined particular factors. The main findings are the following. Observers mostly look at the central part of the images [30]. The foreground is analyzed first, although some objects placed in the background might attract human attention as well [14]. Larger size objects usually attract attention first, whilst smaller size details can attract attention too [14]. More contrast, bright color, and better-illuminated objects are viewed first [34]. People mostly fixate on objects positioned closer, although the dependence of attention on depth is nonlinear [30]. In this sense, the findings are in good agreement with those in content-based image retrieval [35].

Meanwhile, there are other types of factors influencing attention. People tend to find human faces and bodies in images [36, 37]. Note that nowadays there are various approaches to face search in images [38]. Animals, flowers, buildings, and their particular parts as doors or windows can attract attention as well [36, 37]. These are objects inspired by experience and social life. Furthermore, emotions are important. For example, Humphrey *et al*. have examined how emotional objects can attract the attention of humans and guide the process of viewing the images along with or contrary to their physical features [39]. Interestingly, a person tends to focus attention on emotionally negative objects (for example, weapons) and remembers them better than positive or neutral ones.

Social factors play an important role as well. Nuthmann *et al*. have shown that viewers' age can influence the process of perception of the images. Older adults can longer focus on some image features and rate different objects as important and meaningful compared to younger viewers [40]. This aspect has a relation to the meaning map concept that precisely captures the features of goal-relevant, social-influenced perception of visual objects. This concept was developed by Henderson *et al*. [41] and it emphasizes the need to study not only the visual but also the semantic features of the images that simultaneously present in them and guide attention [17]. The authors stress that the semantic informativeness of images affects the attention of a person in the first moments of their viewing [17, 33, 39-41] alongside salience features. Obviously, such features are considerably harder to incorporate into computer-generated saliency or priority map since the

 ISSN 1814-4225 (print)
ISSN 2663-2012 (online)

corresponding analysis [42] is more complex than the analysis of low-level features.

As mentioned above, the priority map is goal - oriented. Below, we give an example of image analysis by a person asked to perform a special task – to decide whether or not denoising has improved the visual quality of images. For this purpose, an observer was shown three images simultaneously (Fig. 1) – a noisy and the corresponding filtered ones in the upper row as well as the noise-free (etalon, reference) image in the lower row. Analysis of gaze fixations in Fig. 1 shows the following. First, there are much more fixation points in noisy and filtered images than in the reference one. Second, the positions of the fixation points in noisy and filtered images coincide only partly. Third, attention is paid mostly to the object in the center and, basically, to near-edge points with relatively high contrasts. All these observations are in good agreement with the tendencies considered above.

Fig. 1. Illustration of gaze fixations in solving the task of filtering efficiency analysis

3. Quality assessment

The two main groups of visual quality metrics are full-reference (FR) and no-reference (NR). The former group employs a given and the corresponding reference images for metric calculation and is, in general, able to characterize image visual quality more accurately (adequately). The latter group does not need to have a reference in disposal but possesses less accuracy.

Different variants of saliency maps have been exploited in FR assessment of image visual quality. The paper [43] considers several known elementary quality metrics and several visual saliency maps (VSMs) where weighting that takes VSM into account is used. It is shown that due to this Spearman rank order correlation coefficient (SROCC) between the modified metrics and mean opinion score (MOS) for the modified database TID2013 [44] has been slightly (by a few percent) increased for some metrics as SSIM [45] and MAD [46]. Note here that VSM can be easily incorporated into assessment if a metric presumes its local calculation (e.g., in blocks) with further aggregation of local values.

One positive moment of the paper [43] is that it clearly shows by examples of VSMs that human attention is paid to humans (especially, faces), animals, flowers, and buildings.

The paper [24] takes into account that humans pay more attention to objects in the center of images in the design of center-oriented metrics. They analyze the designed metric performance for four known databases of distorted images and show that a small improvement in metric performance is provided for the database TID2013 compared to such an efficient elementary visual quality metric as MDSI [47].

The authors of the paper [48] considered the possibility of improving the performance of the metrics PSNR-HVS and PSNR-HVS-M [49] that are calculated in 8x8 blocks in the discrete cosine transform (DCT) domain. VSM has been taken into consideration in design modifications of the aforementioned metrics. It has been shown that performance can be sufficiently improved for such specific subsets of the database TID2008 as Exotic and Exotic2. The VSMs presented in the paper [48] show a special focus of humans on faces, elements of bodies, elements of buildings, and birds' heads.

A recent paper [50] has proposed a new pooling strategy based on saliency guiding. A thorough analysis conducted out for four databases shows that there is a rather large improvement in the designed metric in terms of rank correlation with MOS for some databases.

L. Zhang et al. [51] proposed VSI (Visual Saliency-Induced Index) in 2014. It provides aggregate SROCC close to 0.9 for the database TID2013, which is one of the best results. Even better SROCC values are produced by the visual saliency-based structural contrast quality index [52] proposed by Shahab Uddin et al. in 2019.

Certainly, the FR-metric performance depends on a used VSM. Different variants of VSMs are considered in [53] and it is shown that Ma's VSM [54] provides certain benefits compared to Itti's VSM.

The use of VSMs has gained popularity in NR quality metric design, especially based on neural networks [55-57]. J. Ryu has tested the metric [55] for the KADID-10K dataset and got SROCC equal to 0.834, which is considerably larger compared to many known NR metrics and close to good FR metrics. C. Charrier et al. have designed the SABIQ metric that has been tested for TID2013 and CSIQ databases. The results for particular types of distortions were either the best or among the best. The authors of the paper [57] have obtained excellent results, where it is demonstrated that SROCC and Pearson correlation both reached 0.88 for the KADID-10K dataset (http://database.mmsp-kn.de/kadid-10k-database.html). The presented results stress that one more time VSMs offer information useful for image quality assessment.

4. VSM in image processing

The peculiarities of HVS have also been widely exploited in image and video processing. We focus here on two basic operations such as image lossy compression and denoising.

Lossy compression is used to reduce the data size sufficiently (by several times) and the task is to minimize distortions introduced for a given compression ratio (CR). Although introduced distortions are still often characterized by conventional metrics such as mean square error (MSE) or peak signal-to-noise ratio (PSNR) [58], visual quality metrics are used more and more often [59-61]. Rate-distortion optimization is applied to obtain an appropriate compromise between CR and quality for a given image [62-64].

This quality can be perceived in different ways depending on the application and goals of image/video viewing (analysis) [65, 66]. In [65], it is shown that the central part of the image containing letters attracts the attention of observers and, due to this, the quality of its preservation at the lossy compression stage determines judgments of humans concerning what coder is the best. They show how this property can be incorporated into the coder design, leading to better results for a set of test images.

The first attempts to directly incorporate VSM into image and video compression were done about 20 years ago [67-69], where the authors modified JPEG format using region-dependent quantization [67], compressed salient and non-salient regions by separate algorithms [68], and applied saliency-based non-uniform compression [69]. Later region of interest (ROI) based lossy compression of images and video has become a hot topic [70-72] where the ROI concept is quite close to the VSM one. Such approaches can be implemented easier if compression is carried out using blocks. Then, orthogonal transform coefficients can be quantized adaptively taking into account belonging of a given block to ROI (or a fragment of high visual attention) or not. ROI-based compression has become especially popular in medical imaging [73], where it is important not to lose diagnostically valuable information.

In the image filtering theory, it was known 40 years ago that edges and small-sized details attract human attention and, thus, their preservation was paid special attention [74]. Nonlinear filter design was the first step [74-76] followed by intensive development of locally adaptive, transform-based, and non-local groups of techniques (see [77-79] and references therein, respectively). Scanning window nonlinear filters preserve valuable information due to nonlinear operations with local data; locally adaptive denoising techniques apply filters with different properties or parameters to locally active and passive areas; transform-based methods employ differences in spectral properties of a signal component in different regions; finally, non-local filters use the similarity of image fragments (patches, blocks) to improve denoising efficiency, in the first order, by better edge/detail/texture preservation. The most advanced denoising techniques such as BM3D [79] and its modifications [80] incorporate the useful properties of non-local and transform-based denoising.

This leads to the better visual quality of images denoised by the new generation of filters. To prove this, let us present the results of experiments carried out using an eye-tracker. An observer analyzed 45 frames (sets of three images simultaneously presented on the screen). Each frame contained three images: an original (noise-free) placed below, its noisy version placed on the top left, and a filtered one placed on the top right (see the examples in Figures 1 and 2). All test images were of equal size and displayed without any size changes (zooming). Additive i.i.d. Gaussian noise with variance values equal to 49, 100, and 225 was added to make noise visible in noisy images, at least, in homogeneous image regions. The example in Fig. 2 shows the case when noise is masked by texture in most fragments of the image.

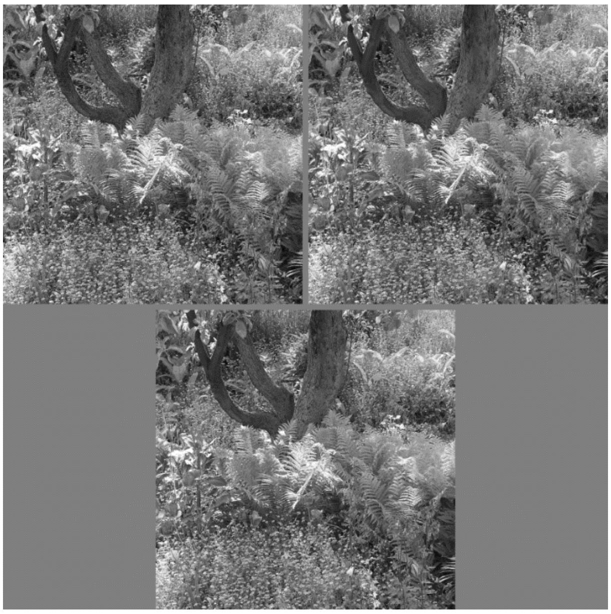

Fig. 2. Illustration of three images used in solving the task of filtering efficiency analysis

ISSN 1814-4225 (print)
ISSN 2663-2012 (online)

Five filters have been tested: 3x3 mean, 3x3 median, 5x5 sigma, 5x5 local statistic Lee (version for additive noise), and BM3D. The mean filter is an example of linear filters, and the median and sigma filters [74, 76] are examples of scanning window nonlinear denoising techniques. The local statistic Lee filter [75] (version for additive noise) is a simple locally adaptive filter, and BM3D [79] represents modern non-local denoisers.

Each image was processed by only one filter. Nine different images were processed by one of the listed filters. Each filter was applied to three images with high complexity (textural), three images of the middle, and three images of low (with a high percentage of homogeneous regions) complexity. Each filter was applied to three images with noise variance 49 (one complex, one simple, and one with middle complexity). The same with noise variances 100 and 225.

An observer was asked to compare noisy and filtered images having an opportunity to look at the corresponding original (noise-free) image placed in the lower row as well. For each frame, the observer was given up to 10 s for viewing and then he was asked to put marks to the “filtering effect” using 5 levels of opinions (-2 points if sufficient degradation of visual quality is observed due to filtering, -1 if degradation is observed, 0 if the visual quality of noisy and filtered images are approximately the same, +1 is an improvement of visual quality occurs, +2 if this improvement is essential). The observer did not know what filter was used for a given frame, he had only to say the opinion (score) that was fixed. The observer has not also seen the test images earlier (before experiments).

The observer’s head was fixed. This was done because of the two reasons. First, many known methodologies of performing visual experiments recommend doing so. Second, it was needed for performing an eye tracker calibration before performing experiments. The eye tracker is equipped with software that allows determining gaze fixations and the percentage ratio of attention attracted by each of the three images presented on the monitor. Recall that the noise-free image attracted very little attention. The main attention was paid to edge/detail fragments in both noisy and filtered images.

Filtering efficiency results (sums of points got by each filter for 9 images processed by it) are presented in Table 1. According to them, the BM3D filter is the best, and it has provided an improvement in image visual quality for most images processed by it. The sigma filter is also good enough. Other considered filters either remain the image visual quality almost the same (on average) as the Lee filter does or degrade it (as the mean and median filters).

These results are in good agreement with other experiments performed to study the filter efficiency and expedience of filter application for image quality improvement [81, 82].

Table 1

Comparison of filter performance

#	Filter type	Aggregate points
1	BM3D [79]	8
2	Sigma [76]	6
3	Lee [75]	1
4	Median [74]	-6
5	Mean [74]	-13

It has been shown in [81-83] that denoising is usually reasonable if images of low and middle complexity are contaminated by middle-intensity noise (and, of course, a filter is efficient enough). BM3D demonstrates sufficiently better efficiency than the DCT-based filter [83] in terms of visual quality and this is mainly due to better edge/detail preservation, which is paid special attention by observers.

5. Conclusions and perspective directions

In this paper, we have presented definitions of maps used in image analysis and processing. It is shown that many factors determine these maps and it is not an easy task to automatically generate SVM that agrees well with maps obtained from experiments with subjects that pay attention to both low-level and high-level features.

A brief analysis of VSM use in image quality assessment is given. It is demonstrated that VSMs help better assessment. Similarly, human attention is incorporated in image processing for such typical operations as lossy compression and denoising. Despite the improvement of PSNR by a few dB, processed images might be often assessed as having worse visual quality compared to original (noisy) ones [82].

Keeping this in mind, we can mention a few directions for further research that seems interesting and able to provide new benefits. First, there are new areas such as medical diagnostics [84, 85], where saliency maps can be useful. With the increased efficiency of computations, VSMs can be more widely used in video compression [86], including approaches based on artificial intelligence [87]. ROI-based denoising techniques might be of interest for some applications [88]. The prediction of filtering efficiency can be useful [89]. Most probably, convolutional neural networks will play an important role in these directions and applications.

Special tools for a thorough analysis of compression and denoising efficiency that allow the calculation of different metrics, including neural network-based ones can assist in better analysis and design.

Contributions of authors: conceptualization – **Vladimir Lukin, Ekaterina Bataeva**; methodology – **Vladimir Lukin**; formulation of tasks – **Ekaterina Bataeva**, **Vladimir Lukin**; analysis – **Sergey Abramov**; development of model – **Vladimir Lukin**, **Sergey Abramov;** software – **Sergey Abramov;** verification – **Sergey Abramov**; analysis of results – **Vladimir Lukin**, **Sergey Abramov**; visualization – **Sergey Abramov**; writing – original draft preparation – **Vladimir Lukin**, **Ekaterina Bataeva**; writing – review and editing – **Vladimir Lukin**.

All the authors have read and agreed to the published version of the manuscript.

References

1. Kussul, N., Lavreniuk, M., Shelestov, A., Skakun, S. Crop inventory at regional scale in Ukraine: Developing in season and end of season crop maps with multi-temporal optical and SAR satellite imagery. *European Journal of Remote Sensing*, 2018, vol. 51, no. 1, pp. 627-636. DOI: 10.1080/22797254.2018.1454265.

2. Zappavigna, M. Social media photography: construing subjectivity in Instagram images. *Visual Communication*, 2016, vol. 15, no. 3, pp. 271–292. DOI: 10.1177/1470357216643220.

3. Bataeva, E. V. Flanering and video mania: Modern and postmodern visual practices. *Voprosy Filosofii*, 2012, vol. 11, pp. 61–68. Available at: https://www.researchgate.net/publication/331037254_Flan-erstvo_i_videomania_modernye_i_postmodernye_vizualnye_praktiki (accessed 9.01.2023).

4. Lee, L.-K., Liew, S.-C. A Survey of Medical Image Processing Tools. *4th International Conference on Software Engineering and Computer Systems (ICSECS*), 2015, pp. 171-176. DOI: 10.13140/RG.2.1.3364.4241.

5. Daugherty, G. (ed.) *Medical Image Processing: Techniques and Applications*. Springer New York, NY, 2011. 380 p. DOI: 10.1007/978-1-4419-9779-1.

6. Safridatul, A. et al. Visual Analysis of Satellite Landsat Images Multitemporal and GPS as a Geographic Information System for Mapping of Nugmet Plantations in Tapaktuan. *1st South Aceh International Conference on Engineering and Technology*, 2019. 13 p. DOI: 10.1088/1757-899X/506/1/012037.

7. Treisman, A. M., Gelade, C. A feature-integration theory of attention. *Cognitive Psychology*, 1980, vol. 12, no 1, pp. 97-136. DOI: 10.1016/0010-0285(80)90005-5.

8. Koch, C., Ullman, S. Shifts in selective visual attention: towards the underlying neural circuitry. *Human Neurobiology*, 1985, vol. 4, pp. 219-227. DOI: 10.1007/978-94-009-3833-5_5.

9. Olshausen, B. A., Anderson, C. H. Van Essen, D. C. A neurobiological model of visual attention and invariant pattern recognition based on dynamic routing of information. *J Neuroscience*, 1993, vol. 13, iss. 11, pp. 4700–4719. DOI: 10.1523/JNEUROSCI.13-11-04700. 1993.

10. Tsotsos, J. K. et al. Modeling visual attention via selective tuning. *Artificial Intelligence*, 1995, vol. 78, iss. 1-2, pp. 507-545. DOI: 10.1016/0004-3702(95)00025-9.

11. Itti, L., Koch, C., Niebur, E. Computational modelling of visual attention. *Nature Reviews Neuroscience*, 2001, vol. 2, pp. 194–203. DOI: 10.1038/35058500.

12. Itti, L., Koch, C., Niebur, E. A model of saliency-based visual attention for rapid scene analysis. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 1998, no. 20, pp. 1254–1259. DOI: 10.1109/34.730558.

13. Machner, B. et al. Unbalancing the Attentional Priority Map via Gaze-Contingent Displays Induces Neglect-Like Visual Exploration. *Frontiers in Human Neuroscience*, 2020, vol. 14, Article no. 41. DOI: 10.3389/fnhum.2020.00041.

14. Li, J. Y. *Visual Attention is Beyond One Single Saliency Map*, 2018. Available at: https://arxiv.org/abs/1811.02650 (accessed 9.01.2023).

15. Cong, R. et al. Review of Visual Saliency Detection with Comprehensive Information. *IEEE Transactions on Circuits and Systems for Video Technology*, 2018, vol. 29, pp. 2941-2959. DOI: 10.1109/TCSVT.2018.2870832.

16. Fecteau, J. H., Munoz, D. P. Salience, relevance, and firing: a priority map for target selection. *Trends in Cognitive Sciences*, 2006, vol. 10, iss. 8, pp. 382-390. DOI: 10.1016/j.tics.2006.06.011.

17. Henderson, J. M., Hayes, T. R. Meaning-based guidance of attention in scenes as revealed by meaning maps. *Nature Human Behaviour*, 2017, no. 1, pp. 743–747. DOI: 10.1038/s41562-017-0208-0.

18. Rahman, S. et al. Classifying Eye-Tracking Data Using Saliency Maps. *25th International Conference on Pattern Recognition (ICPR*), 2020, pp. 9288-9295. DOI: 10.1109/ICPR48806.2021.9412308.

19. Ramanathan, S. et al. An Eye Fixation Database for Saliency Detection in Images. *Lecture Notes in Computer Science*, 2010, vol. 6314, pp. 30-43. DOI: 10.1007/978-3-642-15561-1_3.

20. Liu, T. et al. Learning to detect a salient object. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2011, no. 33, iss. 2, pp. 353-367. DOI: 10.1109/TPAMI.2010.70.

ISSN 1814-4225 (print)
ISSN 2663-2012 (online)

21. Obeso, A. M. et al. Comparative study of visual saliency maps in the problem of classification of architectural images with Deep CNNs. *Eighth International Conference on Image Processing Theory, Tools and Applications (IPTA)*, 2018, pp. 1-6. DOI: 10.1109/IPTA.2018.8608125.

22. Ullah, I. et al. A brief survey of visual saliency detection. *Multimedia Tools and Applications*, 2020, vol. 79, pp. 34605-34645. DOI: 10.1007/s11042-020-08849-y.

23. Liu, Z., Huang, Y., Niu, Y., Lin, L. Saliency-aware Image Quality Assessment. *Proceedings of the 2016 2nd Workshop on Advanced Research and Technology in Industry Application*, 2016, pp. 1352-1358. DOI: 10.2991/wartia-16.2016.280.

24. Layek, M. A. et al. Center-Emphasized Visual Saliency and a Contrast-Based Full Reference Image Quality Index. *Symmetry*, 2019, vol. 11, iss. 3, article no. 296. DOI: 10.3390/sym11030296.

25. Kim, C., Milanfar, P. Visual saliency in noisy images. *Journal of Vision*, 2013, vol. 13, pp. 1-14. DOI: 10.1167/13.4.5.

26. Yu, S. X., Lisin, D. A. Image Compression Based on Visual Saliency at Individual Scales. *Advances in Visual Computing. Lecture Notes in Computer Science*, 2009, vol. 5875, pp. 157-166. DOI: 10.1007/978-3-642-10331-5_15.

27. Gupta, R., Khanna, M. T., Chaudhury, S. Visual saliency guided video compression algorithm. *Signal Processing: Image Communication*, 2013, vol. 28, iss. 9, pp. 1006-1022. DOI: 10.1016/j.image.2013.07.003.

28. Zhou, B., Zhang, Z., Hao, W., Wang, Z. Image deblurring based on visual saliency. *International Conference on Systems and Informatics*, 2012, pp. 1919-1922. DOI: 10.1109/ICSAI.2012.6223423.

29. Hu, C., Jia, S., Zhang, F., Li, X. A saliency-guided street view image inpainting framework for efficient last-meters wayfinding. *ISPRS Journal of Photogrammetry and Remote Sensing*, 2023, vol. 195, pp. 365-379. DOI: 10.1016/j.isprsjprs.2022.11.009.

30. Desingh, K., Krishna, K.M., Rajan, D., Jawahar, C.V. Depth really Matters: Improving Visual Salient Region Detection with Depth. *Proceedings British Machine Vision Conference*, 2013, pp. 98.1-98.11. DOI: 10.5244/C.27.98.

31. Einhäuser, W., Rutishauser, U., Koch, C. Task-demands can immediately reverse the effects of sensory-driven saliency in complex visual stimuli. *Journal of Vision*, 2008, vol. 8, iss. 2, pp. 1–19. DOI: 10.1167/8.2.2.

32. Serences, J. T., Yantis, S. Selective visual attention and perceptual coherence, *Trends in Cognitive Sciences*, 2006, vol. 10, iss. 1, pp. 38-45. DOI: 10.1016/j.tics.2005.11.008.

33. Veale, R., Hafed, Z. M., Yoshida, M. How is visual salience computed in the brain? Insights from behavior, neurobiology and modeling. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 2017, vol. 372, iss. 1714. DOI: 10.1098/rstb.2016.0113.

34. Henderson, J. M., Hayes, T. R., Peacock, C. E., Rehrig, G. Meaning and attentional guidance in scenes: A review of the meaning map approach. *Vision*, 2019, vol. 3, iss. 2, pp. 1-19. DOI: 10.3390/vision3020019.

35. Dharani, T., Aroquiaraj, I. L. A survey on content based image retrieval. *International Conference on Pattern Recognition, Informatics and Mobile Engineering*, 2013, pp. 485-490. DOI: 10.1109/ICPRIME.2013.6496719.

36. Yang, H. et al. Blind Image Quality Assessment of Natural Distorted Image Based on Generative Adversarial Networks. *IEEE Access*, 2019, vol. 7, pp. 179290-179303. DOI: 10.1109/ACCESS.2019.2957235.

37. Wang, H., Xiao, S. An Object IQA Framework Based on Saliency Map Similarity, *IEEE 6th International Conference on Computer and Communications (ICCC*), 2020, pp. 1985-1989. DOI: 10.1109/ICCC51575.2020.9345144.

38. Lu, D., Yan, L. Face Detection and Recognition Algorithm in Digital Image Based on Computer Vision Sensor. *Journal of Sensors*, 2021, vol. 2021, pp. 1-16. DOI: 10.1155/2021/4796768.

39. Humphrey, K., Underwood, G., Lambert, T. Salience of the lambs: A test of the saliency map hypothesis with pictures of emotive objects. *Journal of Vision*, 2012, vol. 12, iss. 1, article no. 22, pp. 1-15. DOI: 10.1167/12.1.22.

40. Nuthmann, A., Schütz, I., Einhäuser, W. Salience-based object prioritization during active viewing of naturalistic scenes in young and older adults. *Scientific Reports*, 2020, vol. 10, iss. 1, article no. 22057. DOI: 10.1038/s41598-020-78203-7.

41. Henderson, J. M., Hayes, T. R., Peacock, C. E., Rehrig, G. Meaning maps capture the density of local semantic features in scenes: A reply to Pedziwiatr, Kümmerer, Wallis, Bethge & Teufel. *Cognition*, 2021, vol. 214, article no. 104742. DOI: 10.1016/ j.cognition.2021.104742.

42. Kulkarni, S. S., Reddy, N. P., Hariharan, S. Facial expression (mood) recognition from facial images using committee neural networks. *BioMedical Engineering OnLine*, 2009, vol. 8, pp. 1-16. DOI: 10.1186/1475-925X-8-16.

43. Liu, Z., Huang, Y., Niu, Y., Lin, L. Saliency-aware Image Quality Assessment, *2nd Workshop on Advanced Research and Technology in Industry Applications*, 2016, pp. 1354-1360. DOI: 10.2991/wartia-16.2016.280.

44. Ponomarenko, N. et al. Image database TID2013: Peculiarities, results and perspectives. *Signal Processing: Image Communication*, 2015, vol. 30, pp. 57-77. DOI: 10.1016/j.image.2014.10.009.

45. Wang, Z., Bovik, A. C., Sheikh, H. R., Simoncelli, E. P. Image quality assessment: from error visibility to structural similarity. *IEEE Transactions on Image Processing,* 2004, vol. 13, iss. 4, pp. 600-612. DOI: 10.1109/TIP.2003.819861.

46. Larson, E. C., Chandler, D. Most apparent distortion: Full-reference image quality assessment and the role of strategy. *Journal of Electronic Imaging,* 2010, vol. 19, iss. 1, article no. 011006. DOI: 10.1117/1.3267105.

47. Nafchi, H. Z., Shahkolaei, A., Hedjam, R., Cheriet, M. Mean Deviation Similarity Index: Efficient and Reliable Full-Reference Image Quality Evaluator. *IEEE Access,* 2016, vol. 4, pp. 5579-5590. DOI: 10.1109/ACCESS.2016.2604042.

48. Tong, Y., Konik, H, Cheikh, F.A., Tremeau, A. Full Reference Image Quality Assessment Based on Saliency Map Analysis. *Journal of Imaging Science and Technology,* 2010, vol. 54, iss. 3, pp. 1-21. DOI: 10.2352/J.ImagingSci.Technol.2010.54.3.030503.

49. Ponomarenko, N. et al. On between-coefficient contrast masking of DCT basis functions. *International Conference on VPQM, USA,* 2007, pp. 1-4. Available at: https://ponomarenko.info/vpqm07_p.pdf (accessed 9.01.2023).

50. Varga, D. Saliency-Guided Local Full-Reference Image Quality Assessment. *Signals,* 2022, vol. 3, iss. 3, pp. 483-496. DOI: 10.3390/signals3030028.

51. Zhang, L., Shen, Y., Li, H. VSI: A Visual Saliency-Induced Index for Perceptual Image Quality Assessment. *IEEE Transactions on Image Processing,* 2014, vol. 23, pp. 4270-4281. DOI: 10.1109/TIP.2014.2346028.

52. Uddin, A. F. M. S., Chung, T. C., Bae, S.-H. Visual saliency based structural contrast quality index. *Electronics Letters,* 2019, vol. 55, iss. 4, pp. 194-196. DOI: 10.1049/el.2018.6435.

53. Adrizzone, E., Bruno, A. Image Quality Assessment by saliency maps. *VISAPP,* 2012, vol. 1, pp. 479-483. Available at: https://www.scitepress.org/papers/2012/38677/38677.pdf (accessed 9.01.2023).

54. Ma, L., Li, S., Ngan, K. N. Visual Horizontal Effect for Image Quality Assessment. *IEEE Signal Processing Letters,* 2010, vol. 17, iss. 7, pp. 627-630. DOI: 10.1109/LSP.2010.2048726.

55. Ryu, J. A Visual Saliency-Based Neural Network Architecture for No-Reference Image Quality Assessment. *Applied Sciences,* 2022, vol. 12, iss. 19, article no. 9567. DOI: 10.3390/app12199567.

56. Charrier, C., Saadane, A., Fernandez-Maloigne, C. Blind Image Quality Assessment based on the use of Saliency Maps and a Multivariate Gaussian Distribution. *20th International Conference on Image Analysis and Processing (ICIAP),* 2019, pp. 1-12. DOI: 10.1007/978-3-030-30645-8_13.

57. Lamichhane, K., Carli, M., Battisti, F. A CNN-based no reference image quality metric exploiting content saliency. *Signal Processing: Image Communication,* 2023, vol. 111, article no. 116899. DOI: 10.1016/j.image.2022.116899.

58. Li, F., Krivenko, S., Lukin, V. Two-step providing of desired quality in lossy image compression by SPIHT. *Radioelectronic and computer systems,* 2020, no. 2(94), pp. 22-32. DOI: 10.32620/reks.2020.2.02.

59. Oh, H., Bilgin, A., Marcellin, M. Visually lossless JPEG 2000 for remote image browsing. *Information,* 2016, vol. 7, no. 3, article no. 45. DOI: 10.3390/info7030045.

60. Bondžulić, B. et al. Efficient prediction of the first just noticeable difference point for JPEG compressed images. *Acta Polytechnica Hungarica,* 2021, vol. 18, iss. 8, pp. 201-220. DOI: 10.12700/APH.18.8.2021.8.11.

61. Yang, K., Jiang, H. Optimized-SSIM based quantization in optical remote sensing image compression. *Proceedings of Sixth International Conference on Image and Graphics,* 2011, pp. 117-122. DOI: 10.1109/ICIG.2011.38.

62. Jeong, Y. W. et al. *Rate distortion optimization encoding system and method of operating the same*. Patent US, no. 10,742,995 B2. 2020. Available at: https://patents.justia.com/patent/10742995 (accessed 9.01.2023).

63. Ortega, A., Ramchandran, K. Rate-distortion methods for image and video compression. *IEEE Signal Processing Magazine,* 1998, vol. 15, no. 6, pp. 23-50. DOI: 10.1109/79.733495.

64. Jin, L. et al. Statistical study on perceived JPEG image quality via MCL-JCI dataset construction and analysis. *Proceedings of IS&T International Symposium on Electronic Imaging: Image Quality and System Performance XIII,* 2016, vol. 28, pp. 1-9. DOI: 10.2352/ISSN.2470-1173.2016.13.IQSP-222.

65. Patel, Y., Appalaraju, S., Manmatha, R. Saliency Driven Perceptual Image Compression. *IEEE Winter Conference on Applications of Computer Vision (WACV),* 2021, pp. 227-236. DOI: 10.1109/WACV48630.2021.00027.

66. Hu, S. et al. Compressed image quality metric based on perceptually weighted distortion. *IEEE Transactions on Image Processing,* 2015, vol. 24, iss. 12, pp. 5594-5608. DOI: 10.1109/TIP.2015.2481319.

67. Golner, M., Mikhael, W., Krishnang, V. Modified JPEG Image Compression With Region-dependent Quantization. *Circuits, Systems and Signal Processing,* 2002, vol. 21, pp. 163–180. DOI: 10.1007/s00034-002-2004-x.

68. Lee, S.-H., Shin, J.-K., Lee, M. Non-uniform image compression using biologically motivated saliency map model. *Proceedings of the 2004 Intelligent Sensors,*

 ISSN 1814-4225 (print)
ISSN 2663-2012 (online)

Sensor Networks and Information Processing Conference, 2004, pp. 525-530. DOI: 10.1109/ISSNIP. 2004.1417516.

69. Itti, L. Automatic foveation for video compression using a neurobiological model of visual attention. *IEEE Transactions on Image Processing,* 2004, vol.13, pp. 1304-1318. DOI: 10.1109/TIP.2004.834657.

70. Hadizadeh, H., Bajić, I. V. Saliency-aware video compression. *IEEE Trans. Image Process,* 2014, vol. 23, iss. 1, pp. 19-33. DOI: 10.1109/TIP. 2013.2282897.

71. Li, Z., Qin, S., Itti, L. Visual attention guided bit allocation in video compression. *Image and Vision Computing,* 2011, vol. 29, iss. 1, pp. 1-14. DOI: 10.1016/j.imavis.2010.07.001.

72. Cai, C., Chen, L., Zhang, X., Gao, Z. End-to-End Optimized ROI Image Compression. *IEEE Transactions on Image Processing,* 2020, vol. 29, pp. 3442-3457. DOI: 10.1109/TIP.2019.2960869.

73. Kaur, R., Rani, R. ROI and Non-ROI based Medical Image Compression Techniques: A Survey and Comparative Review. *First International Conference on Secure Cyber Computing and Communication (ICSCCC),* 2018, pp. 550-555. DOI: 10.1109/ICSCCC. 2018.8703337.

74. Astola, J., Kuosmanen, P. *Fundamentals of Nonlinear Digital Filtering.* Boca Raton, CRC Press Publ., 1997. 288 p. DOI: 10.1201/9781003067832.

75. Lee, J.-S. Refined filtering of image noise using local statistics. *Computer Graphics and Image Processing,* 1981, vol. 15, iss. 4, pp. 380-389. DOI: 10.1016/S0146-664X(81)80018-4.

76. Lee, J. S. Digital image smoothing and the sigma filter. *Computer Vision, Graphics, and Image Processing,* 1983, vol. 24, iss. 2, pp. 255-269. DOI: 10.1016/0734-189X(83)90047-6.

77. Lukin, V. V. et al. Digital adaptive robust algorithms for radar image filtering. *Journal of Electronic Imaging,* 1996, vol. 5, iss. 3, pp. 410-421. DOI: 10.1117/12.240715.

78. Mallat, S. A *Wavelet Tour of Signal Processing,* Academic Press, 2008. 832 p. Available at: https://www.sciencedirect.com/book/9780123743701/a-wavelet-tour-of-signal-processing (accessed 9.01.2023).

79. Dabov, K., Foi, A., Katkovnik, V., Egiazarian, K. Image Denoising by Sparse 3-D Transform-Domain Collaborative Filtering. *IEEE Transactions on Image Processing,* 2007, vol. 16, iss. 8, pp. 2080-2095. DOI: 10.1109/TIP.2007.901238.

80. Mäkitalo, M., Foi, A., Fevralev, D., Lukin, V. Denoising of single-look SAR images based on variance stabilization and nonlocal filters. *International Conference on Mathematical Methods in Electromagnetic Theory,* 2010, pp. 1-4. DOI: 10.1109/MMET.2010.5611418.

81. Rubel, A., Rubel, O, Lukin, V. On Prediction of Image Denoising Expedience Using Neural Networks. *Proceedings of PICST,* 2018, pp. 629-634. DOI: 10.1109/INFOCOMMST.2018.8632050.

82. Zemliachenko, A. et al., Analysis of Opportunities to Improve Image Denoising Efficiency for DCT-based Filters. *Radioelectronic and Computer Systems,* 2018, no. 2, pp. 4-12. DOI: 10.32620/reks. 2018.2.01.

83. Pogrebnyak, O., Lukin, V. Wiener discrete cosine transform-based image filtering. *Journal of Electronic Imaging,* 2012, no 4, pp. 1-14. DOI: 10.1117/1.JEI.21.4.043020.

84. Arun, N. et al. Assessing the Trustworthiness of Saliency Maps for Localizing Abnormalities in Medical Imaging. *Radiology: Artificial Intelligence,* 2021, vol. 3, pp. 1-12. DOI: 10.1148/ryai.2021200267.

85. Hosokawa, S. et al. Validation of a Saliency Map for Assessing Image Quality in Nuclear Medicine: Experimental Study Outcomes. *Radiation,* 2022, vol. 2, pp. 248-258. DOI: 10.3390/radiation2030018.

86. Lyudvichenko, V., Erofeev, M., Ploshkin, A., Vatolin, D. Improving Video Compression with Deep Visual-attention Models. *IMIP '19: Proceedings of the 2019 International Conference on Intelligent Medicine and Image Processing,* 2019, pp. 88–94. DOI: 10.1145/3332340.3332358.

87. Pelurson, S. et al. AI-Based Saliency-Aware Video Coding. *SMPTE Motion Imaging Journal,* 2022, vol. 131, iss. 4, pp. 21-29. DOI: 10.5594/JMI.2022. 3160541.

88. Chahid, A., Serrai, H., Achten, E., Laleg-Kirati, T.-M. A New ROI-Based performance evaluation method for image denoising using the Squared Eigenfunctions of the Schrödinger Operator. *40th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC),* 2018, pp. 5579-5582. DOI: 10.1109/EMBC.2018.8513615.

89. Rubel, O., Rubel, A., Lukin, V., Egiazarian, K. Selection of Lee Filter Window Size Based on Despeckling Efficiency Prediction for Sentinel SAR Images. *Remote Sensing,* 2021, vol. 13, article no. 1887. DOI: 10.3390/rs13101887.

Надійшла до редакції 19.09.2022, розглянута на редколегії 20.02.2023

КАРТИ ЗНАЧУЩОСТІ В ОЦІНЦІ ВІЗУАЛЬНОЇ ЯКОСТІ ТА ОБРОБЦІ ЗОБРАЖЕНЬ

Володимир Лукін, Катерина Батаєва, Сергій Абрамов

Зображення в основному переглядають і аналізують люди. Через це при характеристиці якості зображення та ефективності обробки зображень необхідно враховувати особливості системи зору та пізнання людини, які є дуже складними. Карта значущості, а також карти пріоритетів і смислів, що нещодавно запропоновані, є спробами врахувати специфічні особливості людського зору до застосувань аналізу та обробки зображень. Багато авторів, які розглядають вищезгадані карти, аналізують їх з різних точок зору. Таким чином, основним **предметом** даної статті є фактори, які впливають і визначають ці карти. Серед таких факторів є ознаки низького рівня, а також соціально-психологічні фактори, такі як емоції, вік, життєві цінності. Основна **мета** статті полягає в тому, щоб дати короткий огляд цих факторів і розглянути, як карти вже використовуються для оцінки якості зображення та обробки, а також як вони можуть бути використані в майбутньому. **Завдання** роботи полягають у тому, щоб дати визначення картам значущості, пріоритету та смислу, проаналізувати фактори, що впливають на ці карти, оцінити, яке покращення можна отримати за рахунок врахування карт при оцінці візуальної якості зображення та інших операцій обробки зображень, як-то придушення шумів і стиснення з втратами. Основний **результат** полягає в тому, що, беручи до уваги карти значущості, можна значно покращити оцінку якості зображення та ефективність обробки, особливо для програм, орієнтованих на перегляд та аналіз зображень спостерігачами або клієнтами. Це можна зробити за допомогою простого зважування локальних оцінок заданої метрики з подальшим агрегуванням, а також за допомогою підходів на основі нейронних мереж. Використовуючи різні кількісні критерії, ми показуємо, який позитивний результат можна отримати за рахунок включення карт в оцінку якості та обробку зображень. У якості **висновків** ми представляємо можливі напрямки майбутніх досліджень, які, головним чином, стосуються адаптації параметрів шумозаглушення та стиснення з втратами до особливостей людської уваги.

Ключові слова: карта значущості; оцінка якості; обробка зображень.

Лукін Володимир Васильович – д-р техн. наук, проф., зав. каф. інформаційно-комунікаційних технологій ім. О. О. Зеленського, Національний аерокосмічний університет ім. М. Є. Жуковського «Харківський авіаційний інститут», Харків, Україна.

Батаєва Катерина Вікторівна – д-р філос. наук, проф., проф. каф. соціально-реабілітаційних технологій, Відкритий міжнародний університет розвитку людини "Україна", Київ, Україна.

Абрамов Сергій Клавдійович – канд. техн. наук, доц., доц. каф. інформаційно-комунікаційних технологій ім. О. О. Зеленського, Національний аерокосмічний університет ім. М. Є. Жуковського «Харківський авіаційний інститут», Харків, Україна.

Vladimir Lukin – Doctor of Technical Sciences, Professor, Head of Information-Communication Technologies named after O. O. Zelensky Department, National Aerospace University "Kharkiv Aviation Institute", Kharkiv, Ukraine,
e-mail: lukin@ai.kharkov.com, ORCID: 0000-0002-1443-9685.

Ekaterina Bataeva – Doctor of Philosophic Sciences, Professor, Professor of Social Rehabilitation Technologies Department, Open International University of Human Development Ukraine, Kyiv, Ukraine,
e-mail: bataevaekaterina72@yahoo.com, ORCID: 0000-0002-4628-4817.

Sergey Abramov – Candidate of Technical Sciences, Assoc. Prof., Associate Professor of Information-Communication Technologies named after O. O. Zelensky Department, National Aerospace University "Kharkiv Aviation Institute", Kharkiv, Ukraine,
e-mail: s.abramov@khai.edu, ORCID: 0000-0002-8295-9439.