

УДК 681.34

АВТОМАТИЧЕСКОЕ ОЦЕНИВАНИЕ ДИСПЕРСИИ ОШИБОК ИЗМЕРЕНИЙ ПРИ ОБРАБОТКЕ ОДНОМЕРНЫХ ИНФОРМАЦИОННЫХ СИГНАЛОВ*С.К. Абрамов, А.А. Зеленский, д-р техн. наук, В.В.Лукин, канд. техн. наук**Национальный аэрокосмический университет им. Н.Е.Жуковского "ХАИ"*

Предложен метод автоматического оценивания дисперсии ошибок измерения информационных параметров, изменяющихся во времени. Метод надежно функционирует в условиях ограниченных априорных сведений о характеристиках ошибок измерения и особенностях изменения оцениваемых информационных параметров.

* * *

Запропоновано метод автоматичного оцінювання дисперсії похибок вимірювання інформаційних параметрів, що змінюються. Метод надійно функціонує в умовах обмежених априорних відомостей про характеристики похибок вимірювання та особливості зміни інформаційних параметрів, що оцінюються.

* * *

A method of automatic estimation of variance of measuring errors for informative parameters changing in time is proposed. This method reliably performs in case of limited a priori knowledge about measurement error characteristics and peculiarities of information parameter changing.

Обычно в измерительных системах в целях улучшения их точностных характеристик и повышения надежности выделения полезной информации обработка данных проводится в два этапа [1]. В ходе первичной обработки формируется последовательность оценок контролируемого информационного параметра. При этом для каждого первичного измерения вследствие различных причин возникают ошибки, в результате чего получаемая последовательность оценок информационного параметра представляет собой дискретизированный во времени процесс, содержащий как информационную, так и шумовую составляющую.

В ходе вторичной обработки, которая в современных системах обычно реализуется в цифровом виде, наблюдаемый процесс подвергается фильтрации, целью которой является повышение точности первичных измерений, т.е. подавление шумовой составляющей при одновременном сохранении информационной [1].

Для решения этой задачи применяются оптимальные фильтры Калмана и Винера [1], неадаптивные и локально-адаптивные нелинейные фильтры

[2], методы обработки на основе локальных ортогональных преобразований [3] и т.п. При использовании многих из них необходимо знание характеристик шумовой составляющей обрабатываемого процесса [1,2,3], в частности дисперсии.

Во многих практических ситуациях дисперсия неизвестна, и, соответственно, ее необходимо оценить, причем сделать это непосредственно по наблюдаемой зашумленной реализации процесса, априорная информация о поведении информационной составляющей которого обычно ограничена [2,3]. Задача усложняется, если дисперсия шумовой составляющей обрабатываемого процесса изменяется во времени, как это, например, имеет место в системах тракторных измерений [1]. Таким образом, актуальной является задача автоматической оценки дисперсии шумовой составляющей с учетом ее возможного медленного изменения во времени при ограниченных априорных сведениях о характере информационной компоненты наблюдаемого процесса.

Положим, что характеристики шумовой составляющей в течение некоторого промежутка времени (интервала квазистационарности) можно в первом

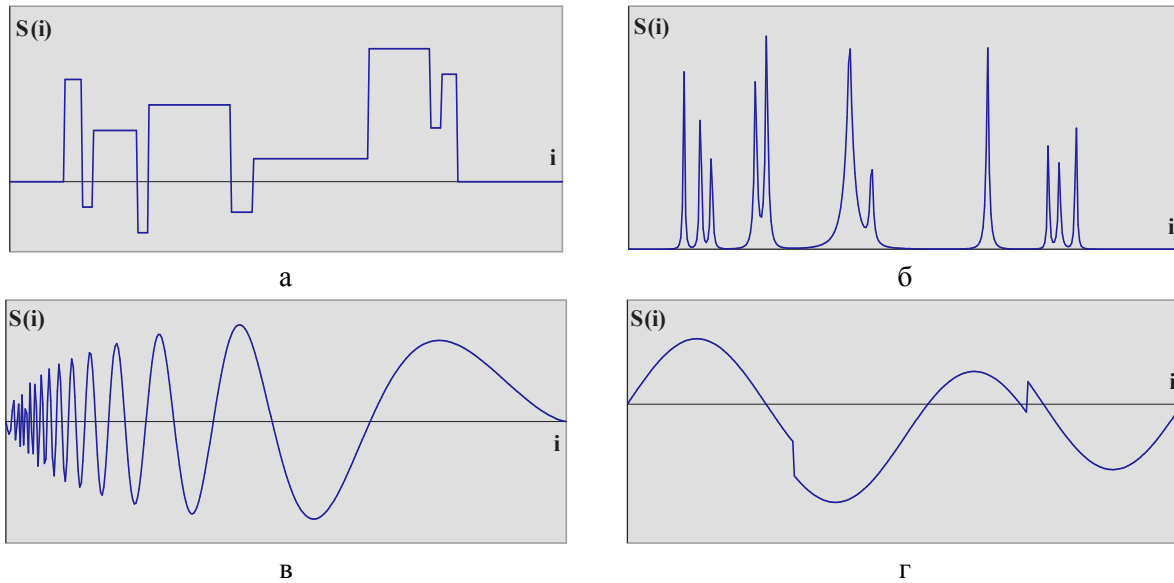


Рис. 1. Тестовые сигналы: а - “Blocks”; б - “Bumps”; в - “Doppler”; г - “HeaviSine”

приближении считать постоянными, а сама шумовая компонента является аддитивной, δ -коррелированной и хорошо описывается гауссовским процессом с нулевым средним и дисперсией σ_a^2 . На практике интервал квазистационарности может быть заранее известен или предварительно оценен с учетом характерных особенностей измерительной системы и обрабатываемых ею информационных сигналов.

Для описания наблюдаемого процесса X_i на интервале квазистационарности дисперсии первичных измерений воспользуемся следующей моделью: $X_i = S_i + \eta_i$, где S_i - информационный сигнал; η_i - аддитивная шумовая компонента. При проведении исследований использовались типовые тестовые сигналы [3]: “Blocks”, “Bumps”, “Doppler”, “HeaviSine” длительностью 1024 отсчета (рис.1).

Предлагаемый алгоритм автоматической оценки дисперсии σ_a^2 состоит из следующих основных этапов:

1) информационный сигнал разбивается на участки фиксированной или переменной длины L , на которых η_i можно считать квазистационарными;

2) каждый из таких участков, в свою очередь,

разбивается на N_Δ перекрывающихся или неперекрывающихся подынтервалов с фиксированным размером L_Δ ;

3) для каждого подынтервала рассчитывается оценка локальной дисперсии шумовой компоненты $\sigma_{лок_i}^2$; при этом в соответствии с традиционной методикой

$$\sigma_{лок_i}^2 = \frac{1}{L_\Delta - 1} \sum_{j=m_i - M_\Delta}^{m_i + M_\Delta} (X_j - \bar{X}_i)^2, \quad (1)$$

где m_i - индекс центрального отсчета i -го подынтервала, для которого рассчитывается локальная дисперсия; X_j - отсчеты зашумленного информационного сигнала в пределах i -го подынтервала; \bar{X}_i - среднее значение зашумленного информационного сигнала на i -м подынтервале; $M_\Delta = (L_\Delta - 1)/2$;

4) с использованием описанной ниже процедуры на основе мириадной операции [4] находится мода полученной выборки оценок $\{\sigma_{лок_i}^2\}$, $i = 1, \dots, N_\Delta$, координата которой и принимается в качестве оценки дисперсии шумовой компоненты.

В основе предложенного метода лежит предпо-

ложение о том, что большинство оценок $\sigma_{лок_i}^2$, полученных для подынтервалов, соответствует "плоским" участкам информационного сигнала и группируется в окрестности истинного значения σ_a^2 , формируя моду выборки $\{\sigma_{лок_i}^2\}$ [5].

Существенное влияние на точностные характеристики метода оказывает выбор соответствующих параметров на каждом из перечисленных выше этапов. Параметр L (этап 1) выбирается таким образом, чтобы он, с одной стороны, не превосходил интервал квазистационарности шумовой компоненты. С другой стороны, L должен быть достаточно большим, чтобы количество оценок N_Δ , получаемых для этого интервала, было достаточным для формирования моды распределения $\{\sigma_{лок_i}^2\}$ (этап 4). При моделировании L полагалось равным длительности тестовых сигналов ($L=1024$).

На этапе 2 необходимо правильно задать размер подынтервалов L_Δ , а также способ их формирования - с перекрытием или без перекрытия. Использование перекрывающихся подынтервалов при фиксированных L и L_Δ позволяет приблизительно в L_Δ раз увеличить количество оценок N_Δ .

Для обоснования выбора L_Δ и способа формирования подынтервалов был сформирован участок сигнала $S = const$ ($L=1024$), искаженного аддитив-

ным гауссовским шумом с нулевым средним и дисперсией $\sigma_a^2 = 0,5$. По полученному таким образом сигналу были сформированы выборки $\{\sigma_{лок_i}^2\}$.

Гистограммы полученных выборок (количество $N_{hst}(h_i)$ оценок $\sigma_{лок_i}^2$, попадающих в различные интервалы значений $h_i \pm \frac{h_i + h_{i+1}}{2}$) для различных L_Δ и различных способов их формирования показаны на рис. 2. Как видно, при использовании неперекрывающихся подынтервалов даже небольшого размера ($L_\Delta=7$) получаемая гистограмма является изрезанной и ее мода выражена не явно (рис. 2,а). При использовании перекрывающихся подынтервалов мода гистограмм носит более явно выраженный характер (рис. 2,б), поэтому их применение более целесообразно.

Исследование гистограмм при большем числе оценок $\sigma_{лок_i}^2$ ($N_\Delta=10^6$) показало, что координата моды λ_{mod} смещена относительно σ_a^2 для рассмотренных значений $L_\Delta=7, 9, 11$. При этом $\lambda_{mod} < \sigma_a^2$, а $|\lambda_{mod} - \sigma_a^2|$ уменьшается при возрастании L_Δ . Поэтому, если алгоритм оценки дисперсии основывается на определении координаты моды распределения $\{\sigma_{лок_i}^2\}$, то представляется разумным увеличивать L_Δ для уменьшения $\sigma_a^2 - \lambda_{mod}$.

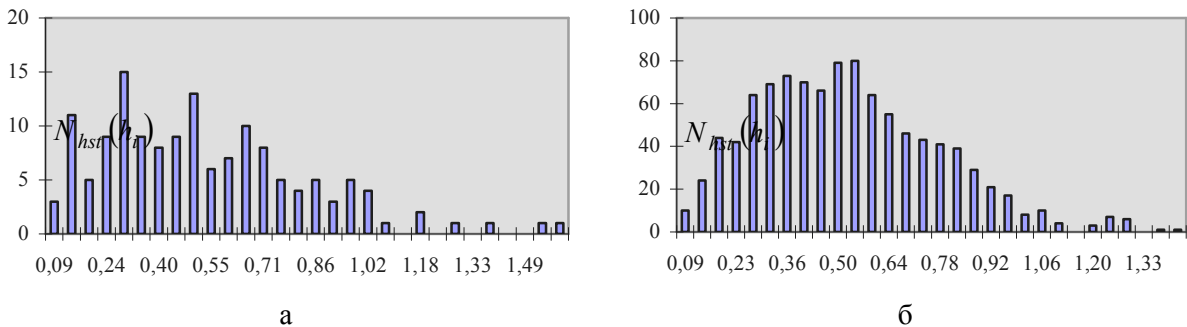


Рис. 2. Гистограммы выборок оценок локальной дисперсии:
 а - для неперекрывающихся подынтервалов размером 7 отсчетов (146 оценок);
 б - для перекрывающихся подынтервалов размером 9 отсчетов (1016 оценок)

h_i

При расчете оценок локальной дисперсии в соответствии с (1) для $S = \text{const}$ на них не оказывала влияния динамика информационного сигнала. Изменение значений S_j в пределах L_Δ приводит к возрастанию $\sigma_{лок_i}^2$ (1). Действительно, при условии некоррелированности сигнальной и шумовой составляющих имеем

$$\sigma_{лок_i}^2 \approx \text{Var}(\eta_i) + \text{Var}(S_i), \quad (2)$$

$$\text{Var}(S_i) = \frac{1}{L_\Delta - 1} \sum_{j=m_i-M_\Delta}^{m_i+M_\Delta} (S_j - \bar{S}_i)^2,$$

$$\text{Var}(\eta_i) = \frac{1}{L_\Delta - 1} \sum_{j=m_i-M_\Delta}^{m_i+M_\Delta} (\eta_j - \bar{\eta}_i)^2,$$

где \bar{S}_i и $\bar{\eta}_i$ - средние значения соответственно сигнальной и шумовой составляющих для i -го подынтервала. Поэтому, если для большого числа подынтервалов $\text{Var}(S_i)$ оказывается сравнимой или большей, чем $\text{Var}(\eta_i)$, то распределение оценок $\{\sigma_{лок_i}^2\}$ существенно отличается от распределения $\{\sigma_a^2\}$ для $S = \text{const}$. В итоге может наблюдаться существенная положительная смещенность оценок (положения моды λ_{mod} относительно истинного значения σ_a^2), т.е. $\lambda_{mod} > \sigma_a^2$, а разность $\lambda_{mod} - \sigma_a^2$ оказывается тем больше, чем больше L_Δ .

Ниже рассматриваются два способа уменьшения негативного влияния динамики сигнала: ограничение L_Δ сверху и модификация методики получения оценок локальной дисперсии $\sigma_{лок_i}^2$.

На этапе 4 для получения оценки координаты моды распределения λ_{mod} , принимаемой затем в качестве оценки σ_a^2 , могут использоваться различные методы [5]. Нами для поиска λ_{mod} предлагается применять мириадную операцию [4], которая при соответствующем выборе величины настроечного

параметра k способна определять координату моды обрабатываемой выборки. Оптимальная величина настроечного параметра k мириадного метода, при которой он обладает наилучшими точностными характеристиками, зависит от оцениваемой дисперсии, причем оптимальное k прямо пропорционально σ_a^2 [5]. Тогда, зная оптимальное соотношение

$(k/\sigma_a^2)_{onm}$, можно определить

$$k_{onm} = \hat{\sigma}_{предв}^2 \left(\frac{k}{\sigma_a^2} \right)_{onm}, \quad \text{где } \hat{\sigma}_{предв}^2 - \text{предварительная оценка } \sigma_a^2.$$

Для получения $\hat{\sigma}_{предв}^2$ предлагается использовать квантильный метод оценки дисперсии [5,6], предусматривающий нахождение определенной порядковой статистики, n -процентного квантиля Q_n [7] полученной выборки $\{\sigma_{лок_i}^2\}$. При правильном выборе n этот квантиль примерно соответствует координате моды выборки [6,7]. Таким образом, частная задача состоит в определении $(k/\sigma_a^2)_{onm}$ для мириадного метода (в предположении, что σ_a^2 известна) и n_{onm} - для квантильного. Затем найденные оптимальные значения используются в качестве параметров автоматического двух-этапного квантильно-мириадного метода оценки дисперсии.

Рассмотрим первый способ уменьшения влияния динамики сигнала. Ограничим $L_\Delta = 9$, а для расчета $\sigma_{лок_i}^2$ воспользуемся выражением (1).

В качестве количественных критериев точности оценивания дисперсии шумовой компоненты использовалась смещенность получаемых оценок относительно истинного значения $\delta = \langle \hat{\sigma}^2 \rangle - \sigma_a^2$, их дисперсия $\theta^2 = \left\langle \left(\hat{\sigma}^2 - \langle \hat{\sigma}^2 \rangle \right)^2 \right\rangle$, а также приведенная ошибка $\varepsilon = \delta^2 + \theta^2$. Усреднение $\langle \bullet \rangle$ проводи-

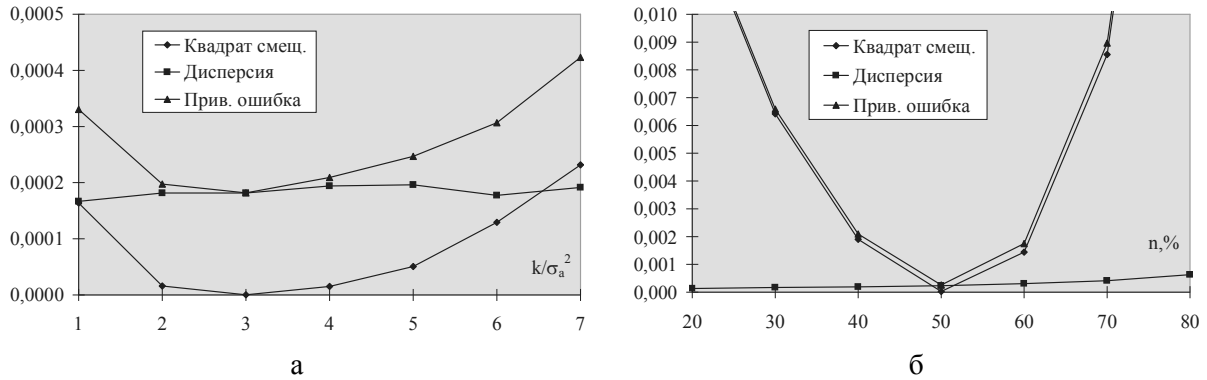


Рис.3. Характеристики точности оценивания для а - мириадного и б - квантильного методов в зависимости от их параметров (тестовый сигнал “Blocks”; $\sigma_a^2=0,3$)

лось по 1000 независимых реализаций.

График приведенной ошибки ϵ для мириадного метода (рис.3,а) носит экстремальный характер, который обусловлен в основном изменением смещенности δ получаемых оценок. Значение δ^2 увеличивается по мере удаления величины параметра k/σ_a^2 от оптимальной (соответствующей минимуму ϵ), а дисперсия получаемых оценок θ медленно возрастает при увеличении k/σ_a^2 .

Схожий характер поведения зависимостей δ^2 , θ и ϵ можно наблюдать и для квантильного метода (рис.3,б). Здесь также существует оптимальный номер порядковой статистики n_{onm} , при котором имеет место наименьшая приведенная ошибка ϵ . Абсолютная величина смещенности δ при отходе от n_{onm} увеличивается, а дисперсия оценивания θ при увеличении n медленно возрастает.

Аналогичные тенденции в поведении количественных критериев наблюдаются и для остальных тестовых сигналов. Величины оптимальных значений $(k/\sigma_a^2)_{onm}$ и n_{onm} , при которых имеют место наименьшие значения ϵ (и соответствующие этим параметрам значения δ , θ), для всех тестовых сигналов при дисперсии помех $\sigma_a^2=0,3$ приведены в табл. 1. Как видно, мириадный метод оценивания дисперсии при выборе $(k/\sigma_a^2)_{onm}$ обеспечивает более высокую точность (меньшие значения ϵ) в сравнении с квантильным методом для всех рассматриваемых тестовых сигналов, за исключением сигнала “Doppler”. Для остальных тестовых сигналов мириадный метод позволяет обеспечить относительную ошибку $\Delta\epsilon = (\epsilon^{0,5} / \sigma_a^2) \cdot 100\%$ порядка 4-6%, а квантильный - порядка 5-7%.

Общим недостатком применения как мириад-

Таблица 1

Точность оценок σ_a^2 при оптимальных значения параметров $(k/\sigma_a^2)_{onm}$ для мириадного и n_{onm} для квантильного методов

Сигнал	Мириадный метод				Квантильный метод			
	$(\frac{k}{\sigma_a^2})_{onm}$	$\delta, \times 10^{-3}$	$\theta^2, \times 10^{-4}$	$\epsilon, \times 10^{-4}$	$n_{onm}, \%$	$\delta, \times 10^{-3}$	$\theta^2, \times 10^{-4}$	$\epsilon, \times 10^{-4}$
Blocks	2,5	-2,54	1,73	1,79	50	-5,16	2,33	2,60
Bumps	0,5	-4,18	2,66	2,84	40	8,34	3,12	3,82
Doppler	0,1	31,41	8,80	18,67	20	-34,46	3,09	14,96
HeaviSine	0,2	4,29	2,68	2,87	40	-16,15	1,77	4,38

ного, так и квантильного методов является то, что наименьшие значения ϵ для рассмотренных тестовых сигналов наблюдаются при существенно различных величинах настроечных параметров.

Проведенный нами анализ гистограмм распределений оценок $\{\sigma_{лок_i}^2\}$ для рассматриваемых тестовых сигналов показал, что разброс оптимальных значений $\left(k/\sigma_a^2\right)_{onm}$ и n_{onm} обусловлен существенным различием в структуре формируемых выборок $\{\sigma_{лок_i}^2\}$. В частности, для тестовых сигналов “Bumps”, “HeaviSine” и “Doppler” по сравнению с сигналом “Blocks” гораздо большая часть оценок локальной дисперсии приходится на участки информационного сигнала со значительными динамическими изменениями, поэтому большинство оценок $\{\sigma_{лок_i}^2\}$ оказываются завышенными вследствие влияния $Var(S_i)$ в (2) и их распределение имеет тяжелый “хвост”. В связи с этим при отсутствии априорных сведений о характеристиках информационной составляющей и использовании (1) для расчета $\sigma_{лок_i}^2$ сложно выбрать такие фиксированные параметры $\left(k/\sigma_a^2\right)$ и n комбинированного квантильно-мириадного метода автоматической оценки дисперсии шумовой компоненты, при которых для всех типов сигналов точность оценок дисперсии шумовой компоненты была бы приемлемой.

Рассмотрим второй метод уменьшения влияния динамики сигнала. Модифицируем методику получения оценок локальной дисперсии следующим образом:

$$\sigma_{лок_i}^2 = \frac{1}{L_{\Delta} - 1} \sum_{j=m_i-M_{\Delta}}^{m_i+M_{\Delta}} \left[(X_j - X_j^f) - \bar{X}_{m_i} \right]^2, \quad (3)$$

где X_j^f - отсчеты фильтрованного процесса в пределах i -го подынтервала;

$$\bar{X}_{m_i} = \frac{1}{L_{\Delta}} \sum_{j=m_i-M_{\Delta}}^{m_i+M_{\Delta}} (X_j - X_j^f) - \text{среднее значение раз-$$

ностного сигнала на i -м подынтервале. В качестве предварительного фильтра предлагается использовать α -урезанный фильтр [7] с размером скользящего окна $N^{\alpha-tr} = 9$ и параметром усечения $\alpha=0,2$. Благодаря использованию предложенной модификации удастся заметно уменьшить вес “хвоста” распределения и увеличить концентрацию оценок $\sigma_{лок_i}^2$ в окрестности истинного значения дисперсии шумовой компоненты.

Результаты численного моделирования при использовании модифицированных мириадного и квантильного методов при оптимальных значениях настроечных параметров $\left(k/\sigma_a^2\right)_{onm}$ и n_{onm} для всех рассматриваемых тестовых сигналов приведены в табл. 2. Очевидно, что точность оценок дисперсии для сигналов “Bumps”, “HeaviSine” и “Doppler” существенно повысилась (сравните данные в табл. 2 и 1), некоторое ухудшение точности оценок (увеличение ϵ) произошло лишь для сигнала “Blocks”. При этом важным положительным моментом предложенной модификации является то, что оптимальные значения настроечных параметров $\left(k/\sigma_a^2\right)_{onm}$ и n_{onm} для всех рассматриваемых тестовых сигналов отличаются незначительно. Это создает предпосылки для применения автоматического квантильно-мириадного метода, который предусматривает: 1) проведение предварительной фильтрации наблюдаемого процесса на анализируемом интервале квазистационарности; 2) получение оценок локальной дисперсии в соответствии с (3); 3) получение предварительной оценки дисперсии $\hat{\sigma}_{предв}^2$ с использованием квантильной операции; 4) получение уточненной оценки дисперсии шумовой компоненты с помощью мириадной операции.

Таблица 2

Точность оценок σ_a^2 для модифицированных мириадного, квантильного и комбинированного методов ($\sigma_a^2=0,3$)

Сигнал	Мириадный метод		Квантильный метод		Комбинированный метод			
	$(k/\sigma_a^2)_{opt}$	$\epsilon, \times 10^{-4}$	$n_{opt}, \%$	$\epsilon, \times 10^{-4}$	$n, \%$	k/σ_a^2		
						1,5	2,0	2,5
Blocks	2,0	2,57	50	2,97	50	2,70	3,26	4,77
Bumps	2,0	2,22	50	3,31	50	2,63	2,64	3,20
Doppler	2,2	2,13	50	5,39	50	3,87	2,71	2,63
HeaviSine	4,0	1,93	60	4,31	50	4,14	3,26	2,65
	2,0	2,61	50	10,18				

Для комбинированного метода (см. правую часть табл. 2) при получении оценки $\hat{\sigma}_{предв}^2$ рекомендуется использовать квантиль $n=50\%$, а при получении уточненной оценки - $(k/\sigma_a^2)=2$. Величины ϵ для комбинированного метода при этом оказываются в 1,15-1,4 раза больше, чем для модифицированного мириадного при оптимальном выборе $(k/\sigma_a^2)_{opt}$. Однако на практике оптимальная величина (k/σ_a^2) зависит от типа обрабатываемых сигналов и реализовать высокую точность оценивания достаточно трудно. Поэтому более целесообразным оказывается применение автоматического комбинированного метода, обеспечивающего приемлемую точность для любых сигналов.

Заклучение

В работе предложен эффективный метод автоматической оценки дисперсии помех, успешно функционирующий при минимальных априорных сведениях о характере шумовой компоненты. Метод применим для обработки измерительных сигналов различных типов и обеспечивает относительную ошибку оценивания менее 6%.

Литература

1. Фалькович С.Е., Хомяков Э.Н. Статистическая теория измерительных радиосистем. - М.: Ра-

дио и связь, 1981.

2. Лукин В.В. Анализ поведения показателей локальной активности для нелинейных адаптивных фильтров // Радиофизика и электроника. – Х.: ИРЭ НАНУ. – 1998. - Вып. 3. - N 2. - С.80-89.

3. Oktem R. Transform Domain Algorithms for Image Compression and Denoising // Thesis for the Degree of Doctor of Technology. Tampere University of Technology, Tampere, Finland, 2000. - 142 p.

4. Kalluri S., Arce G. Adaptive Weighted Myriad Filter Algorithms for Robust Signal Processing in α -stable Noise Environments // IEEE Trans. on Signal Processing. - 1998. - Vol.46. - No.2. - P.322-334.

5. Абрамов С.К., Лукин В.В. Мириадный метод автоматической оценки дисперсии помех на изображениях и его оптимизация // Радиотехника. - 2001. - Вып.123. - С.177-184.

6. Zelensky A., Melnik V., Lukin V. et al. Airborne Multichannel Remote Sensing Data Processing Techniques and Software. Proc. of ERIM, USA. - 1996. - P. III-151-160.

7. Astola J., Kuosmanen P. Fundamentals of Non-linear Digital Filtering, CRC Press LLC, USA, 1997. - 276 p.

Поступила в редакцию 27.02.03

Рецензенты: д-р техн. наук, профессор Кулемин Г.П., ИРЭ НАН Украины, г. Харьков; д-р техн. наук, профессор, Батаев В.А., Национальный аэрокосмический университет им. Н.Е. Жуковского "ХАИ", г. Харьков.