

ВЫБОР МЕРЫ РАЗЛИЧИЯ ОБРАЗЦОВОГО И ОТФИЛЬТРОВАННОГО ПЕРИОДИЧЕСКИХ СИГНАЛОВ В ЗАДАЧЕ ПОИСКА ОПТИМАЛЬНЫХ ЗНАЧЕНИЙ ПАРАМЕТРА ЦИФРОВОГО ФИЛЬТРА ПРИ ВЫСОКИХ УРОВНЯХ ФЛУКТУАЦИОННОГО ШУМА

В статье рассмотрен выбор меры различия результатов фильтрации искажённых дискретных квантованных периодических сигналов в задаче поиска оптимальных значений параметра цифрового фильтра нижних частот «простое скользящее среднее», представленного в рекурсивной форме. Для сравнения обработанных сигналов были выбраны меры различия — среднеквадратичная ошибка, корень среднеквадратичной ошибки, средняя абсолютная ошибка. Для выполнения экспериментов создана программная библиотека для генерирования периодических дискретных квантованных сигналов, их искажения при помощи флуктуационной помехи (шума) с нормальным распределением, последующей фильтрации искажённых сигналов. Выполнена серия экспериментов по генерированию искажённых периодических сигналов с различными уровнями шума, найдены значения мер различия для отфильтрованных сигналов с различными значениями параметра цифрового фильтра. В результате сравнения выбрана мера, для которой характерна высокая различимость оптимальных значений параметра цифрового фильтра при высоких уровнях шума.

Ключевые слова: дискретный сигнал, цифровой фильтр, скользящая фильтрация, мера различия, оптимальный параметр

O.O. SERHEIEV-HORCHYNSKYI
NTUU «Igor Sikorsky Kyiv Polytechnic Institute»

SELECTION OF ERROR MEASURE FOR REFERENCE AND FILTERED PERIODIC SIGNALS IN THE PROBLEM OF SEARCHING FOR OPTIMUM VALUES OF DIGITAL FILTER PARAMETERS AT HIGH LEVELS OF FLUCTUATING NOISE

The article covers selection of a error measure for the results of filtering distorted periodic signals when searching for optimum values of low band digital filter parameters represented by a "simple moving average". The following error measures were selected for comparison — mean squared error, root mean squared error, mean absolute error. A series of experiments was carried out for the generation and filtering of distorted periodic signals with different noise levels. Ultimately, a measure was selected based on comparison results, representing high sensitivity to the optimum values of the digital filter parameter at high noise levels.

Keywords: discrete signal, digital filter, moving average, optimal parameter, error measure.

Введение

В различных технических задачах требуется сравнение данных, которые в общем случае имеют различное представление. Одним из типов представления данных является представление в виде последовательности значений дискретных квантованных сигналов (временных рядов). В цифровой обработке сигналов при проектировании цифровых фильтров, осуществляющих подстройку по образцовым сигналам, требуется расчёт меры различия значений образцового и отфильтрованного временных рядов [1].

В [2] поиск оптимального значения параметра цифрового фильтра нижних частот (ФНЧ) «простое скользящее среднее» (ПСС) осуществлялся при сравнении отфильтрованного и аппроксимированного временных рядов. Для оценки различия временных рядов была применена мера «среднеквадратичная ошибка» (СКО). В прикладных методах анализа данных применяются и другие численные показатели различия временных рядов, которые называются «мерами ошибки» (error measures) [3], либо «метриками расстояния» (distance metrics) [4].

Чтобы проверить целесообразность применения меры СКО для оценки оптимальности фильтрации была создана программная библиотека для цифровой обработки сигналов. Для оценки различия были сгенерированы исходные неискаженные сигналы различной формы и их модифицированные копии, полученные путём добавления к исходным сигналам случайной составляющей (помехи) и последующей фильтрации искажённых сигналов. Информация о модулях, вошедших в разработанную программную библиотеку, приведена в разделе «Создание тестовой среды».

В разделе «Выбор мер различия» описаны математические выражения, которые определяют выражения выбранных для сравнения мер различия значений образцового и отфильтрованного временных рядов.

В разделе «Постановка эксперимента» проведён анализ характеристик целевых функций мер различия временных рядов при различных уровнях шума.

В разделе «Сравнение мер различия» описаны эксперименты, которые были проведены с целью выбора меры, для которой характерна высокая различимость оптимальных значений параметра цифрового фильтра при высоких уровнях шума.

В разделе «Выводы» обобщены результаты, полученные в разделе «Сравнение мер различия».

Создание тестовой среды

В процессе исследований была создана библиотека программных модулей для генерирования и обработки дискретных квантованных периодических сигналов с различными характеристиками. Структура библиотеки изображена на рисунке 1.

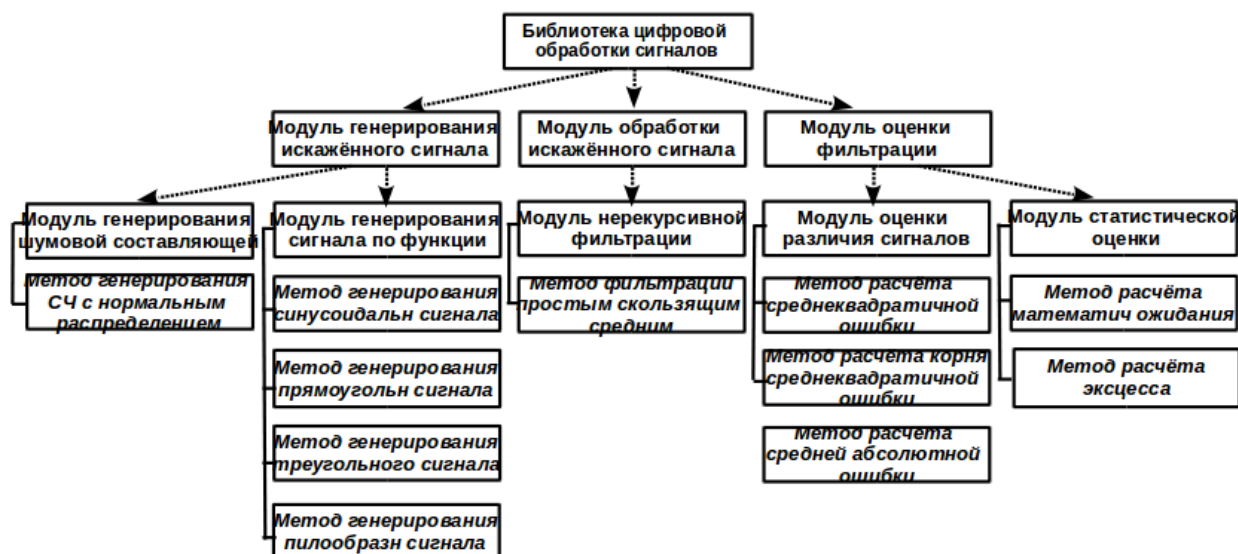


Рис. 1 Структура созданной библиотеки для цифровой обработки сигналов

Функции программных модулей следующие. «Модуль генерирования искажённого сигнала» предназначен для генерирования значений дискретных квантованных периодических сигналов и шумовой составляющей (помехи).

«Модуль генерирования сигнала по функции» реализует программное генерирование значений дискретных квантованных сигналов следующих форм: синусоидального, прямоугольного, пилообразного, треугольного. Программная реализация модуля включает программный метод, который предоставляет возможность настройки амплитуды, частоты и фазы исходного неискаженного сигнала.

«Модуль генерирования шумовой составляющей» добавлен для моделирования различных условий передачи неискаженного сигнала. В модуле реализован программный метод генерирования случайной составляющей искажённого сигнала с нормальным распределением и возможностью задания значений математического ожидания и среднеквадратичного отклонения.

Для минимизации уровня шума был создан «Модуль нерекурсивной фильтрации сигнала». В модуле реализован программный метод фильтрации простым скользящим средним (ПСС). В фильтре ПСС обработанное значение $y[k]$ в момент времени k определяется выражением [2]:

$$y_{\text{отфильтр}}[k] = \frac{1}{m_1 + m_2 + 1} \cdot \sum_{i=-m_1}^{m_2} x_{\text{искаж}}[k-i] \quad (1)$$

где $m_1 + m_2 + 1$ — параметр фильтрации (количество искажённых значений сигнала, участвующих в расчёте отфильтрованного значения для заданного момента времени k), $x_{\text{искаж}}[k-i]$ — значение искажённого сигнала в момент времени дискретизации $k-i$, $y_{\text{отфильтр}}[k]$ — сигнал на выходе фильтра в текущий момент времени k .

В компоненте «Модуль оценки фильтрации» были реализованы программные методы оценки различия сигналов (меры различия): среднеквадратичная ошибка, корень среднеквадратичной ошибки, средняя абсолютная ошибка. Перечисленные программные методы были объединены в «Модуль оценки различия сигналов». Для дополнительной оценки значений числовых последовательностей был реализован «Модуль статистической оценки».

Созданная программная библиотека была применена для формирования искажённых сигналов и их последующей цифровой обработки. Цель выполненных экспериментов была направлена на определение меры различия сигналов, для которой характерна высокая различимость оптимальных значений параметра цифрового фильтра при высоких уровнях шума.

Выбор мер различия

Как было отмечено во введении, при решении технических задач часто возникает необходимость сравнения временных рядов с целью определения степени их подобия. Подобие временных рядов можно оценить при помощи мер различия значений анализируемых сигналов. К задачам, в которых применяют меры различия относятся задачи адаптивной фильтрации, классификации по образцу, обучения по образцу и некоторые другие.

В публикациях, касаючихся изучаеюой в данной статье тематики, предложены различные математические выражения для расчёта меры различия числовых последовательностей, которые можно разделить на следующие группы: меры расстояния, угловые меры, корреляционные меры [3, 4, 5]. Из трёх упомянутых групп мер были выбраны следующие. Мера «среднеквадратичная ошибка» (СКО, Mean Square Error, MSE), которая определяется выражением [2]:

$$M_{\text{СКО}}(\{x_{\text{неискаж}}[k]\}, \{y_{\text{отфильтр}}[k]\}) = \frac{1}{N} \cdot \sum_{k=0}^{N-1} (x_{\text{неискаж}}[k] - y_{\text{отфильтр}}[k])^2 \quad (2)$$

где k — номер отсчёта, $x_{\text{неискаж}}[k]$ — значение неискажённого сигнала в момент времени дискретизации k , $y_{\text{отфильтр}}[k]$ — значение отфильтрованного сигнала в момент времени дискретизации k , N — общее количество отсчётов искажённого сигнала. Мера «корень среднеквадратичной ошибки» (КСКО, Root Mean Square Error, RMSE), которая определяется выражением [3]:

$$M_{\text{КСКО}}(\{x_{\text{неискаж}}[k]\}, \{y_{\text{отфильтр}}[k]\}) = \sqrt{\frac{1}{N} \cdot \sum_{k=0}^{N-1} (x_{\text{неискаж}}[k] - y_{\text{отфильтр}}[k])^2} \quad (3)$$

где k — номер отсчёта, $x_{\text{неискаж}}[k]$ — значение неискажённого сигнала в момент времени дискретизации k , $y_{\text{отфильтр}}[k]$ — значение отфильтрованного сигнала в момент времени дискретизации k , N — общее количество отсчётов искажённого сигнала.

Мера «средняя абсолютная ошибка» (САО, Mean Absolute Error, MAE), которая определяется выражением [4]:

$$M_{\text{САО}}(\{x_{\text{неискаж}}[k]\}, \{y_{\text{отфильтр}}[k]\}) = \frac{1}{N} \cdot \sum_{k=0}^{N-1} |x_{\text{неискаж}}[k] - y_{\text{отфильтр}}[k]| \quad (4)$$

где k — номер отсчёта, $x_{\text{неискаж}}[k]$ — значение неискажённого сигнала в момент времени дискретизации k , $y_{\text{отфильтр}}[k]$ — значение отфильтрованного сигнала в момент времени дискретизации k , N — общее количество отсчётов искажённого сигнала.

Постановка эксперимента

С помощью разработанной программной библиотеки, для оценки характеристик целевых функций сформированных из значений выбранных мер различия, был сгенерирован тестовый искажённый синусоидальный сигнал. В таблице 1 приведены характеристики искажённого сигнала.

Таблица 1

Характеристики тестового искажённого сигнала

Форма сигнала	синусоида
Амплитуда сигнала, В	1.0
Частота сигнала, Гц	1
Распределение шумовой составляющей	Гаусса
Среднеквадратичное отклонение, В	0.7
Частота дискретизации, Гц	100
Продолжительность регистрации сигнала, с	10

Чтобы оценить уровень шума было рассчитано значение отношения сигнал/шум (ОСШ), которое определяется выражением [6]:

$$\text{ОСШ}_{\text{отфильтр}} = 10 \cdot \log_{10} \left(\frac{\sum_{k=0}^{N-1} (x_{\text{неискаж}}[k])^2}{\sum_{k=0}^{N-1} (y_{\text{отфильтр}}[k] - x_{\text{неискаж}}[k])^2} \right) \quad (5)$$

где k — номер отсчёта, $x_{\text{неискаж}}[k]$ — значение неискажённого сигнала в момент времени дискретизации k , $y_{\text{отфильтр}}[k]$ — значение отфильтрованного сигнала в момент времени дискретизации k , N — общее количество отсчётов искажённого сигнала.

Значение ОСШ для сгенерированного сигнала равно 0.14 дБ. Сгенерированный сигнал был обработан при помощи фильтра ПСС со значениями параметра m в диапазоне от 1 до 250, где m — количество искажённых значений сигнала, участвующих в расчёте отфильтрованного значения для заданного момента времени k [см. выражение (1)]. Были проведены 250 сравнений и рассчитаны значения мер различия для двух временных рядов: временных рядов исходного неискажённого сигнала и 250 временных рядов отфильтрованных сигналов.

Поскольку значения мер различия могут находиться вне интервала $[0, 1]$, для приведения мер к одному диапазону значений, в их выражения было включено деление на максимальное значение (из возможных) для каждой меры различия.

Нормированные значения мер различия (значения целевых функций) изображены на рисунке 2. По оси абсцисс отложены значения параметра m , по оси ординат отложены нормированные значения мер различия, рассчитанные при сравнении значений исходного неискаженного и отфильтрованного дискретных квантованных сигналов.

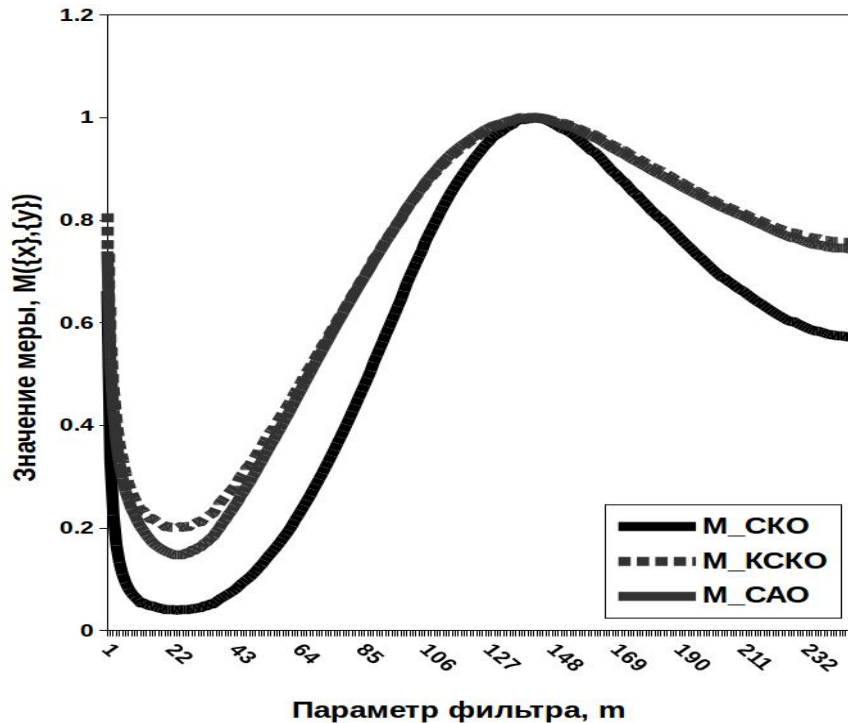


Рис. 2 Нормированные значения мер различия (ось ординат) для значений параметра m от 1 до 250 (ось абсцисс) при ОСШ=0.14 дБ

Из рисунка 2 видно, что нормированные значения мер различия имеют локальные минимальные значения (экстремумы), т. е. могут быть рассмотрены в качестве целевых функций в задаче поиска оптимальных значений параметра цифрового фильтра. Из рисунка 2 следует, что для значений меры СКО визуально можно выделить два локальных минимума при $m_1 = 28$ и $m_2 = 249$ (см. рисунок 3).

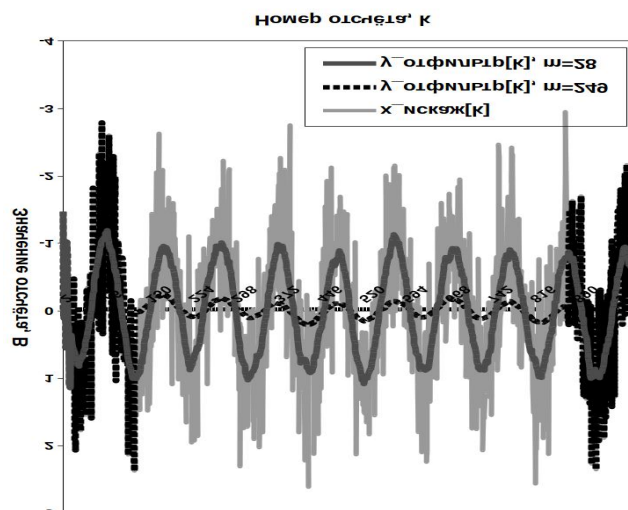


Рис. 3. Значения отфильтрованных сигналов для заданных значений параметра m цифрового фильтра

Из рисунка 3 следует, что применение цифровой фильтрации с оптимальными значениями параметра m позволяет выделить различные низкочастотные составляющие искажённого сигнала и проанализировать амплитудные и частотные характеристики исходного неискаженного сигнала.

Одинаковым значениям меры различия (целевой функции) соответствуют значения расположенные на прямолинейных участках целевой функции либо на интервалах значений с колебаниями целевой функции относительно прямой линии (см. рисунок 2). Количество локальных минимумов целевой функции

определяет количество значений параметра m с равными значениями целевой функции. Для подсчёта количества значений параметра m с одинаковыми значениями целевой функции все значения мер различия были округлены до сотых долей (см. рисунок 4).

На рисунке 5 изображены гистограммы, построенные для количеств значений параметра m с одинаковыми нормированными округлёнными значениями мер различия.

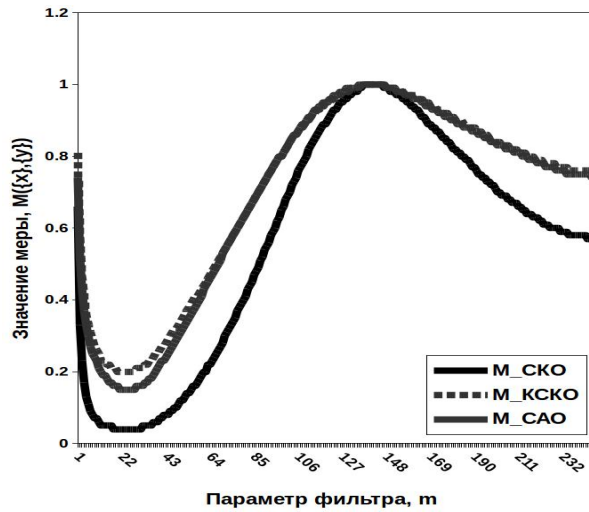


Рис. 4 Нормированные округлённые значения мер различия (ось ординат) для значений параметра m от 1 до 250 (ось абсцисс) при ОСШ=0.14 дБ

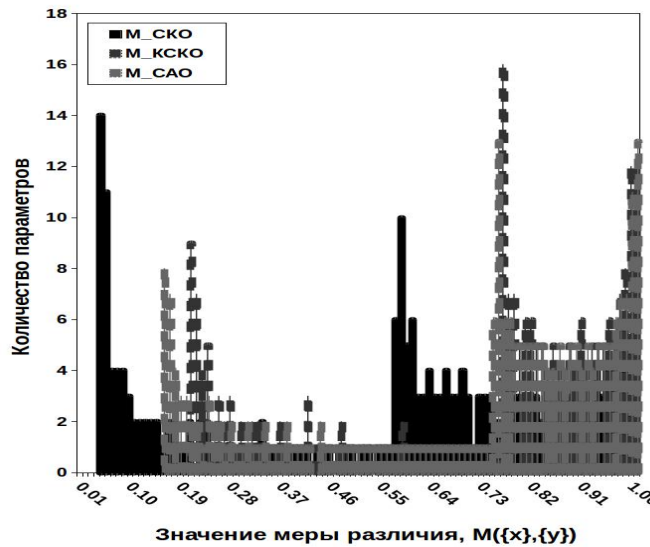


Рис. 5 Количество значений параметра m от 1 до 250 (ось ординат) с одинаковыми нормированными значениями мер различия (ось абсцисс)

Для оценки распределения последовательности значений в теории вероятностей применяют меру эксцесса, определяемую выражением (9) [7]:

$$\bar{M} = \frac{1}{N} \cdot \sum_{m=1}^N M(m) \tag{8}$$

$$\varepsilon = \frac{N \cdot \sum_{m=1}^N (M(m) - \bar{M})^4}{\left(\sum_{m=1}^N (M(m) - \bar{M})^2 \right)^2} \tag{9}$$

где m — параметр цифрового фильтра, $M(m)$ — нормированное округлённое значение меры различия для исходного и отфильтрованного (с параметром m) временных рядов, N — количество значений параметра m . Далее продемонстрированы значения эксцесса, рассчитанные для сигналов разной формы и различных уровней шумовой составляющей.

Сравнение мер различия

В результате серии экспериментов были сгенерированы искажённые сигналы с различными

значениями среднеквадратичного отклонения σ шумовой составляющей от 0.01 до 5.0 с шагом $\Delta\sigma = 0.01$, которым соответствуют значения ОСШ от 37.04 до -16.94 дБ (см. рисунок 6).

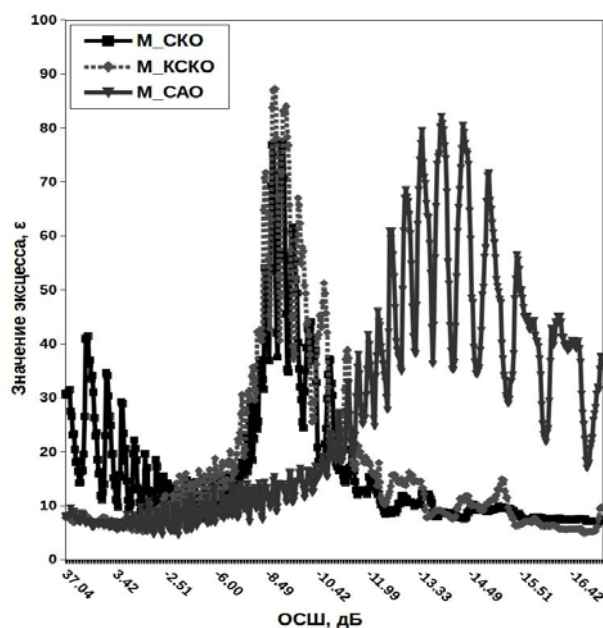


Рис. 6 - Значения эксцесса для мер различия при заданных ОСШ

Из рисунка 6 следует, что максимальные значения эксцесса для мер различия следующие: $\epsilon_{\text{CKO}} = 77.21$ дБ при ОСШ = -9.11 дБ; $\epsilon_{\text{KCKO}} = 87.23$ при ОСШ = -8.80 дБ; $\epsilon_{\text{CAO}} = 82.16$ при ОСШ = -13.84 дБ. Нормированные округлённые значения мер различия (значения целевых функций) при ОСШ = -9.11 дБ изображены на рисунке 7, при ОСШ = -13.84 дБ на рисунке 8.

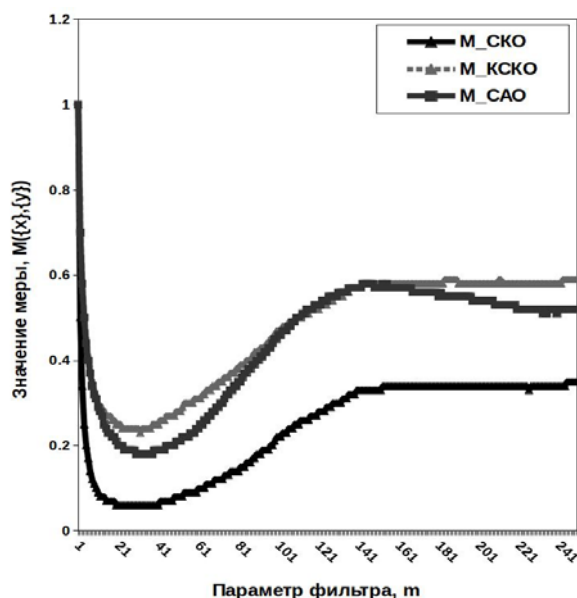


Рис. 7 - Нормированные значения мер различия (ось ординат) для значений параметра m от 1 до 250 (ось абсцисс) при ОСШ = -9.11 дБ

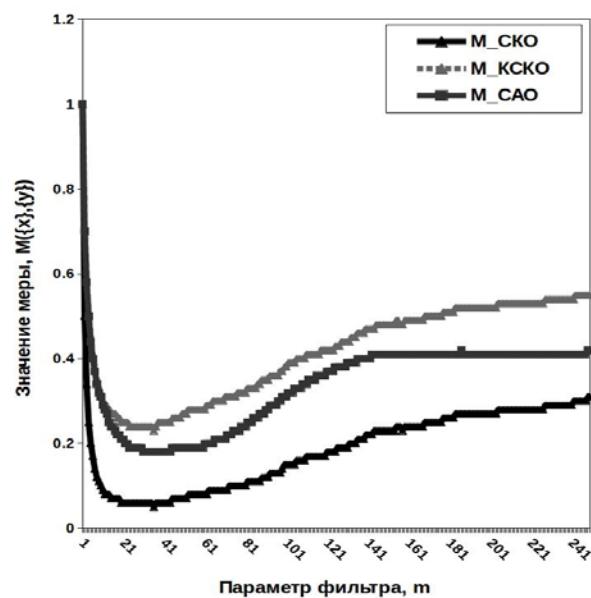


Рис. 8 - Нормированные значения мер различия (ось ординат) для значений параметра m от 1 до 250 (ось абсцисс) при ОСШ = -13.84 дБ

По оси абсцисс отложены значения параметра m , по оси ординат — значения мер различия, рассчитанные при сравнении неискаженного и отфильтрованного временных рядов. Из результатов проведённой серии экспериментов, изображённых на рисунках 7-8 следует, что максимальные значения эксцесса на рисунке 6 соответствуют участкам целевых функций, в которых локальные минимумы преобразуются в прямые линии, в результате чего увеличиваются количества значений параметра m с равными нормированными округлёнными значениями мер различия (см. рисунок 5).

Из рисунков 7-8 следует, что в отличие от мер различия СКО и КСКО, для которых различимость локальных минимумов целевой функции уменьшается при ОСШ ≥ -8.80 дБ, для меры различия CAO различимость уменьшается при ОСШ ≥ -13.84 дБ, следовательно мера CAO более устойчива к высоким уровням шума.

В целях полноты экспериментов количество форм тестовых сигналов было расширено до четырех: синусоидального, прямоугольного, треугольного и пилообразного. В таблице 2 для всех форм сигналов приведены значения ОСШ соответствующие максимальным значениям эксцесса распределений количеств параметра m с одинаковыми нормированными округлёнными значениями мер различия сигналов: исходного неискаженного и отфильтрованного (для значений σ , выбранных в пределах от 0.01 до 5.0 с шагом $\Delta\sigma = 0.01$).

Таблица 2

Значения ОСШ, дБ соответствующие максимальному значению эксцесса

Мера различия	Форма сигнала			
	Синусоидальная	Прямоугольная	Треугольная	Пилообразная
Среднеквадратичная ошибка	-9.11	-8.61	-8.69	-7.90
Корень среднеквадратичной ошибки	-8.80	-8.28	-8.80	-7.83
Средняя абсолютная ошибка	-13.84	-13.88	-13.60	-11.63

Проанализировав значения ОСШ, приведённые в таблице 2, можно сделать вывод, что мера «средняя абсолютная ошибка» более устойчива к высоким уровням шума, поскольку максимумам значений эксцесса соответствуют наименьшие значения ОСШ среди значений, рассчитанных для трёх анализируемых мер различия при различных уровнях шума, следовательно мера САО увеличивает различимость локальных экстремумов в задаче поиска оптимальных значений параметра цифрового фильтра при высоких уровнях шума (малых значениях отношения сигнал/шум).

В разделе «Выводы» обобщены результаты, полученные в разделах «Выбор мер различия», «Постановка эксперимента» и «Сравнение мер различия».

Выводы

Сравнение сигналов требуется при решении различных технических задач, в том числе при проектировании адаптивных цифровых фильтров. Формальной оценкой степени подобия сигналов может служить мера различия. От выбора меры различия зависит то, насколько эффективным будет поиск оптимального значения параметра цифрового фильтра.

Для того, чтобы исследовать данный вопрос, была создана тестовая среда (программная библиотека), включающая в себя модули, генерирующие сигналы различной формы, и модули анализирующие эти сигналы. Основными компонентами библиотеки являются: модуль генерирования искажённых дискретных квантованных периодических сигналов, модуль обработки искажённых сигналов, модуль оценки результатов фильтрации.

Для оценки варьирования значений мер различия для сигналов различной формы с помощью разработанной библиотеки были сгенерированы сигналы следующих форм: синусоидальный, прямоугольный, треугольный и пилообразный. К сгенерированным значениям была добавлена случайная составляющая (шум), после чего полученный сигнал был отфильтрован с помощью нерекурсивного фильтра. Было показано, что применяя к искажённому сигналу фильтрацию с различными значениями параметра m , можно найти оптимальные значения, при которых мера различия принимает локальные минимальные значения.

Для сравнения мер различия предложено использовать эксцесс. В результате оценки значений эксцесса при различных уровнях шума было установлено, что мера «средняя абсолютная ошибка» более устойчива к высоким уровням шума, поскольку глобальному максимальному значению эксцесса соответствует наименьшее значение отношения сигнал/шум среди значений, рассчитанных для трёх анализируемых мер. Из результатов экспериментов по обработке искажённых периодических сигналов различных форм следует, что выбор меры «средняя абсолютная ошибка» увеличивает различимость оптимальных значений параметра цифрового фильтра при высоких уровнях шума.

Литература

1. Сергиенко А. Цифровая обработка сигналов, 3-е издание / А. Сергиенко // Санкт-Петербург, 2011. – БХВ-Петербург, 2011. – С. 593-595.
2. Sergeev-Horchynskiy A. Periodic signal filtration using digital filtering system calculation optimized by approximation / A. Sergeev-Horchynskiy, V. Rogoza // Science-Based Technologies. – 2014. – Т. 23, № 3. – P. 284-288.

3. Willmott C.J., Matsuura K. Advantages of the mean absolute error (MAE) over the root mean square error (RMSE) in assessing average model performance / C.J. Willmott, K. Matsuura // *Climate Research*. – Inter-Research, 2005. – Vol. 30. – P. 79-82.

4. Chai T., Draxler R.R. Root mean square error (RMSE) or mean absolute error (MAE)? – Arguments against avoiding RMSE in the literature / T. Chai, R.R. Draxler // *Geoscientific Model Development*. – Copernicus Publications, 2014. – Vol. 7. – P. 1247-1250.

5. Cha S. Comprehensive Survey on Distance/Similarity Measures between Probability Density Functions / S. Cha // *International journal of mathematical models and methods in applied sciences*. – Singapore, 2007. – Issue 4, Vol. 1. – P. 300-307.

6. Oppenheim A. *Discrete-Time Signal Processing*, 3rd ed. / A. Oppenheim, R. Schaffer // London, 2010. – Pearson, 2010. – P. 223-227.

7. Смирнов Н.В., Дунин-Барковский И.В. Курс теории вероятностей и математической статистики / Н.В. Смирнов, И.В. Дунин-Барковский // Наука. – Москва, 1969. – С. 85-94.

References

1. Sergienko A. *Cifrovaja obrabotka signalov*, 3-e izdanie / A. Sergienko // Sankt-Peterburg, 2011. – BHV-Peterburg, 2011. – С. 593-595.

2. Sergeev-Horchynskiy A. Periodic signal filtration using digital filtering system calculation optimized by approximation / A. Sergeev-Horchynskiy, V. Rogozha // *Science-Based Technologies*. – 2014. – Т. 23, № 3. – P. 284-288.

3. Willmott C.J., Matsuura K. Advantages of the mean absolute error (MAE) over the root mean square error (RMSE) in assessing average model performance / C.J. Willmott, K. Matsuura // *Climate Research*. – Inter-Research, 2005. – Vol. 30. – P. 79-82.

4. Chai T., Draxler R.R. Root mean square error (RMSE) or mean absolute error (MAE)? – Arguments against avoiding RMSE in the literature / T. Chai, R.R. Draxler // *Geoscientific Model Development*. – Copernicus Publications, 2014. – Vol. 7. – P. 1247-1250.

5. Cha S. Comprehensive Survey on Distance/Similarity Measures between Probability Density Functions / S. Cha // *International journal of mathematical models and methods in applied sciences*. – Singapore, 2007. – Issue 4, Vol. 1. – P. 300-307.

6. Oppenheim A. *Discrete-Time Signal Processing*, 3rd ed. / A. Oppenheim, R. Schaffer // London, 2010. – Pearson, 2010. – P. 223-227.

7. Smirnov N.V., Dunin-Barkovskij I.V. *Kurs teorii veroyatnostej i matematicheskoj statistiki* / N.V. Smirnov, I.V. Dunin-Barkovskij // Nauka. – Moskva, 1969. – С. 85-94.

Рецензія/Peer review : 21.5.2017 р.

Надрукована/Printed :26.6.2017 р.

Стаття рецензована редакційною колегією