



...увеличивая
СКОРОСТЬ реакции



1967ВЦ2Ф

Микросхема процессора цифровой обработки сигнала предназначена для применения в системах радиолокации, эхолокации, системах связи и другой аппаратуре, где требуется высокая производительность.

- система команд - Ts201;
- внутренняя память - 24 Мбит (SRAM);
- частота ядра - 450 МГц;
- частота внешней шины - 100 МГц;
- разрядность внешней шины - 64 бит;
- пропускная способность внутренней памяти - 25,2 Гбайт/с;
- пропускная способность внешней шины - 800 Мбайт/с;
- количество внешних последовательных портов - 4;
- пропускная способность внешнего последовательного порта - 1 Гбайт/с.
- 576-выводной BGA-корпус МК 8303.576-1.

ЦИФРОВАЯ ОБРАБОТКА СИГНАЛОВ

Научно-технический журнал
№ 1/2015

Издается с 1999 года
Выходит четыре раза в год

ГЛАВНЫЙ РЕДАКТОР
д.т.н., чл. кор. РАН Ю.Б. ЗУБАРЕВ

ЗАМЕСТИТЕЛИ ГЛАВНОГО РЕДАКТОРА:

д.т.н., проф. В.В. ВИТЯЗЕВ,
д.т.н., проф. В.П. ДВОРКОВИЧ

РЕДАКЦИОННАЯ КОЛЛЕГИЯ:

д.т.н., проф. Ар.С. Аджемов, д.т.н., проф. Б.А. Алпатов,
д.т.н., проф. В.Г. Бартенов, д.т.н. Ю.И. Борисов,
д.т.н., проф. Ю.А. Брюханов, д.т.н., проф. В.В. Еремеев,
д.т.н. Г.В. Зайцев, Р.В. Зубарев, А.П. Кирпичников,
д.т.н., проф. М.И. Кривошеев, д.т.н., акад. РАН Н.А. Кузнецов,
В.Г. Мистюков, д.т.н., проф. С.Л. Мишенков,
д.т.н., проф. А.А. Петровский, д.т.н., проф. Ю.Н. Прохоров,
д.т.н., проф. В.Г. Санников, к.т.н., проф. В.С. Сперанский,
д.т.н., проф. Ю.С. Шинаков

Адрес редакции:

г. Москва, ул. Авиамоторная, д. 8
Научный центр МТУСИ
Тел.: 8 903 221 79 79
E-mail: rntores@mail.ru
vityazev.v.v@rsreu.ru
http://www.dspa.ru

Издатель:

Российское научно-техническое общество
радиотехники, электроники и связи им. А.С. Попова
Компьютерная верстка: И.А. Благодарова
Дизайн: М.В. Аверин

Подписной индекс по каталогу
ОАО «Роспечать» – **82185**

Подписано в печать 24.03.15 г.
Формат 60x90/8.

Гарнитура «Arial». Печать офсетная.
Бумага офсетная. Печ.л. 6,5. Тираж 700 экз.

Заказ № 1674. Отпечатано в
ООО НПЦ «Информационные технологии»
Рязань, ул. Островского, д. 21/1
тел.: (4912) 98-69-84

Издание зарегистрировано в Министерстве
Российской Федерации по делам печати,
телерадиовещания и средств
массовых коммуникаций.

Свидетельство о регистрации ПИ № 77-1488
от 14.01.2000 г.

Журнал «Цифровая обработка сигналов» включен
в Перечень ведущих рецензируемых научных изданий,
в которых по рекомендации Минобрнауки РФ, должны
быть опубликованы основные научные результаты дис-
сертаций на соискание ученой степени доктора
и кандидата наук.

УЧРЕДИТЕЛИ: ЗАО «Инструментальные системы» ФГУП «НИИ радио»

В НОМЕРЕ:

**Азаров И.С., Вашкевич М.И.,
Лихачев Д.С., Петровский А.А.**

*Повышение качества слухового восприятия
в шумовой обстановке на основе частотной декомпозиции
и адаптивной компрессии динамического диапазона сигнала 3*

Мандрикова О.В., Залаяев Т.Л.

*Моделирование вариаций космических лучей на основе
совмещения кратномасштабных вейвлет-разложений
и нейронных сетей переменной структуры 11*

Запорожцев И.Ф., Середа А.-В.И.

*Методика краткосрочного прогноза изменения
значений характеристик динамической системы
на основе численного анализа
многомерных временных рядов 17*

Плужников А.Д., Потапов Н.Н.

*Мультипроцессорная реализация алгоритмов
корреляционно-экстремальной обработки сигналов 25*

Чекушкин В.В., Михеев К.В.

*Быстродействующие алгоритмы преобразования
ортогональных составляющих сигналов
в амплитуду и фазу 32*

**Туровский Я.А., Кургалин С.Д., Вахтин А.А.,
Борзунов С.В., Белобродский В.А.**

*Обобщение метода цепочек локальных экстремумов
для анализа сигналов различной природы 35*

**Топников А.И., Нестеров М.С.,
Новоселов С.А., Приоров А.Л.**

*Неэталонная оценка разборчивости
зашумленных речевых сигналов 39*

Безруков В.Н.

*Принципы вычислительного анализа
характеристик квантования спектра блоков пикселей
сигналов изображений в системах цифрового телевидения 45*

Зубарев Ю.Б., Арзуманян Э.П., Миколайчук Д.В.

*Оценка качества работы телевизионных (тепловизионных)
координаторов в системах наблюдения и безопасности 52*

Нефедов В.И., Решетняк С.А., Третьяков Г.Н.

*Стохастическая фильтрация
периодического биполярного сигнала 56*

Ефимов И.Н.

*Локальные бинарные шаблоны медианного пикселя –
эффективные информативные признаки технологии
распознавания образов 61*

Давыдочкин В.М.

*Преобразование Фурье в задаче измерения расстояния
частотным дальномером в пространстве с дисперсией 66*

Ботвинёнок А.А., Кирпичников А.П.

Экранирование в системах цифровой обработки сигналов 71

Подписной индекс по каталогу
ОАО «Роспечать» – **82185**

Digital Signal Processing

Science & Technical Magazine
Issue 1, 2015 year

Is published quarterly since 1999

THE EDITOR-IN-CHIEF:

Dr.Sci. (Tech.), Professor, Corresponding Member
of Russian Academy of Sciences U.N. Zubarev

DEPUTY EDITORS-IN-CHIEF:

Dr.Sci.(Tech.), Professor V.V. Vityazev, Dr.Sci.(Tech.),
Professor V.P. Dvorkovich

EDITORIAL BOARD:

Adzhemov A.S., Dr.Sci.(Tech.), Professor

Alpatov B.A., Dr.Sci.(Tech.), Professor

Bartenev V.G., Dr.Sci.(Tech.), Professor

Borisov Y.I., Dr.Sci.(Tech.)

Bruchanov Y.A., Dr.Sci.(Tech.), Professor

Eremeyev V.V., Dr.Sci.(Tech.), Professor

Zaitsev G.V., Dr.Sci.(Tech.)

Zubarev R.V.,

Kirpichnikov A.P.,

Krivocheev M.I., Dr.Sci.(Tech.), Professor

Kuznetsov N.A., Dr.Sci.(Tech.), Academician

Mistyukov V.G.,

Mishenkov S.L., Dr.Sci.(Tech.), Professor

Petrovsky A.A., Dr.Sci.(Tech.), Professor

Prokhorov Y.N., Dr.Sci.(Tech.), Professor

Sannikov V.G., Dr.Sci.(Tech.), Professor

Speranskii V.S., Dr.Sci.(Tech.), Professor

Shinakov Y.S., Dr.Sci.(Tech.), Professor

Editorial office address:

Aviamotornaya, 8, Moscow, Russia

Research Center of MTUCI

Phone: +7 903 221 79 79

E-mail: rntores@mail.ru,

vityazev.v.v@rsreu.ru

Web: <http://www.dspa.ru>

Publisher:

Russian A.S. Popov Society for Radioengineering,
Electronics & Communications

Computer makeup:

I.A. Blagodarova

FOUNDERS:

InSys

Radio Research and Development Institute

CONTENTS:

E. Azarov, M. Vashkevich, A. Petrovsky

*Listening enhancement in noisy environment
based on spectral decomposition and adaptive
dynamic range compression of the signal*.....3

O.V. Mandrikova, T.L. Zalyaev

*Modeling of variations of cosmic rays on the basis
of combination of multiresolution wavelet
expansions and neural networks with variable structure*.....11

I.F. Zaporozhtsev, A-V.I. Sereda

*Numerical Method for Short-Term Forecasting
of Multivariate Time Series Evaluated in Nodes
of the Planar Uniform Grid*.....17

A.D. Pluzhniko., N.N. Potapov

*Multiprocessor realization
of correlational-extremal signal processing*.....25

V.V. Chekushkin, I.V. Panteleev, K.V. Mikheev

*High-speed algorithms of the signal orthogonal
components transformation in amplitude and phase*.....32

Y.A. Turovsky, S.D. Kurgalin, A.A. Vahtin,

S.V. Borzunov, V.A. Belobrodsky
*Generalization of the method of chains of local extrema
for analysis of signals of different origins abstract*.....35

A.Topnikov , M. Nesterov,

S. Novoselov, A. Priorov
Non-intrusive Estimation of Noisy Speech signal Intelligibility39

V.N. Bezrukov

*Principles for computational analysis
of image signal pixel block spectrum quantization
characteristics in digital television systems*.....45

E.P. Arzoumanian, D.V. Mikolaychuk,

Y.B. Zubarev
*Quality assessment of TV (IR) coordinator operation
in observing and security systems*.....52

V.I. Nefedov, S.A. Reshetnyak, G.N. Tretyakov

lc filtering of periodic bipolar signal56

I.N. Efimov .

*Local binary patterns median pixel's – effective
technology of pattern recognition* 61

V.M. Davydochkin

*Application of Fourier transform to the problem
of distance measurement by frequency-modulated
rangefinder in the presence of dispersion*66

A.A. Batvinionak, A.P.Kirpichnikov

The Shielding Effects in DSP Systems71

Subscription index: 82185 ("Rospechat")

ISSN: 1684-2634 9 771684 263005

УДК 004.4'277

ПОВЫШЕНИЕ КАЧЕСТВА СЛУХОВОГО ВОСПРИЯТИЯ В ШУМОВОЙ ОБСТАНОВКЕ НА ОСНОВЕ ЧАСТОТНОЙ ДЕКОМПОЗИЦИИ И АДАПТИВНОЙ КОМПРЕССИИ ДИНАМИЧЕСКОГО ДИАПАЗОНА СИГНАЛА

Азаров И.С., к.т.н., Белорусский государственный университет информатики и радиоэлектроники, e-mail: azarov@bsuir.by;

Вашкевич М.И., аспирант кафедры Электронных вычислительных средств Белорусского государственного университета информатики и радиоэлектроники, e-mail: vashkevich@bsuir.by;

Лихачев Д.С., Белорусский государственный университет информатики и радиоэлектроники;

Петровский А.А., д.т.н., профессор, зав. кафедрой Электронных вычислительных средств Белорусского государственного университета информатики и радиоэлектроники, e-mail: palex@bsuir.by.

Ключевые слова: качество слухового восприятия, звуковой сигнал, шумовая обстановка, нелинейное усиление, адаптивная компрессия.

Введение

При использовании мобильных звуковоспроизводящих устройств в шумовой обстановке пользователь воспринимает смесь чистого сигнала (от устройства) и внешнего акустического шума. Во многих практических ситуациях интенсивность шума столь высока, что приводит к существенному ухудшению слухового восприятия за счет эффекта маскирования полезного сигнала шумом. Данный эффект возникает как во время разговора с удаленным абонентом, так и при прослушивании различных звуковых записей (музыка, аудиокниги и проч.).

На рис. 1 показаны сигналы чистой речи и шума метрополитена для которых отношение сигнал/шум (ОСШ) составляет -5 дБ. В результате наложения сигналов спектральные компоненты речи полностью или частично маскируются шумом, что существенно ухудшает разбор-

Исследуется возможность автоматизированного повышения качества восприятия звуковых сигналов в шумовой обстановке. С целью повышения качества восприятия предлагается выполнять нелинейное усиление частотно-временных компонент воспроизводимого сигнала, зависящее от спектральных характеристик окружающего шума. Предлагаемый метод обработки является универсальным, поскольку одинаково хорошо применим для обработки различных звуковых сигналов, в том числе речевых и музыкальных. Выполнено исследование практической применимости предложенного метода для речевых сигналов с использованием объективных оценок.

чивость и общее качество восприятия.

В описываемой ситуации шум невозможно устранить при помощи классических способов шумоподавления (таких, например, как спектральное взвешивание), поскольку шум добавляется к полезному сигналу после его воспроизведения устройством. Таким образом, единственной практической возможностью улучшить восприятие полезного сигнала является его предварительная обработка, с учетом характеристик окружающего акустического шума. На практике пользователь воспроизводящего устройства вынужден самостоятельно регулиро-

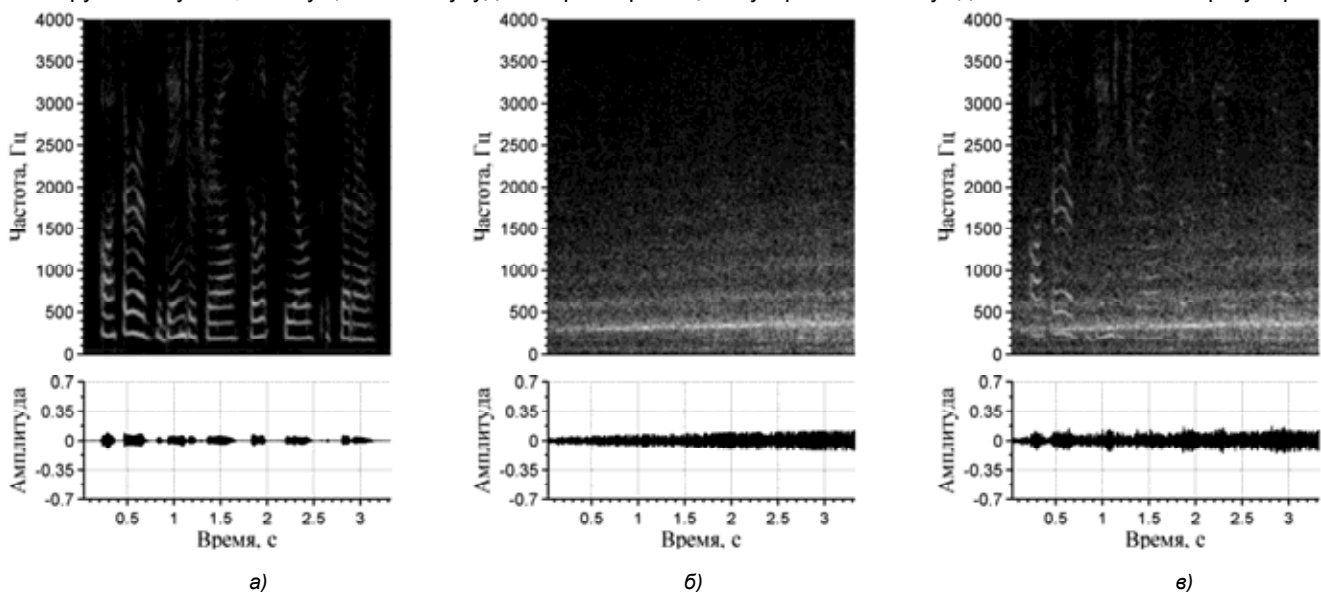


Рис. 1. Частотно-временное представление сигнала:
а) чистая речь; б) шум метрополитена; в) сумма речи и шума

вать громкость воспроизведения в зависимости от шумовой обстановки, однако при этом возникают следующие неудобства: 1) при наличии нестационарного источника шума требуется частая ручная регулировка; 2) регулировка громкости одинаково усиливает все частотные компоненты вне зависимости от того маскируются они шумом или нет, и в результате некоторые важные компоненты полезного сигнала остаются неслышными, в то время как другие получают чрезмерное усиление, вызывающее дискомфорт и быстрое утомление при прослушивании.

Современные мобильные телекоммуникационные устройства, часто используемые для воспроизведения звука, имеют внешний микрофон, который можно использовать для оценки параметров шума. Это делает возможным создание мобильной системы повышения качества воспроизведения звуковых сигналов с автоматическим учетом шума окружающей среды (рис. 2).



Рис. 2. Система для прослушивания сигнала в шумовой обстановке на основе мобильного телекоммуникационного устройства

В настоящее время предложено несколько подходов к решению указанной задачи. Например, в [1] предлагается способ выборочного усиления частотно-временных составляющих полезного сигнала. Целью обработки является обеспечение заданного минимального ОСШ между усредненным спектром полезного сигнала и измеренным спектром шума. Недостатком метода является то, что в результате обработки между всеми тихими звуками теряется относительная громкость – они становятся одинаково громкими на фоне шума. Также теряется относительная громкость и между спектральными составляющими, поскольку они усиливаются независимо друг о друга, что приводит к нежелательному эффекту «выбеления»¹ полезного сигнала. Похожий способ, предложенный в [2], состоит в избирательном усилении частотно-временных составляющих полезного сигнала для увеличения индекса разборчивости речи (ИРР) (SII –

speech intelligibility index). Здесь показано, что обработка сигнала с учетом ИРР позволяет получить более качественный результат, чем обработка с учетом ОСШ [1]. Однако ИРР применим только в случае, когда полезный сигнал является речевым, в то время как на практике, сигналы часто содержат неречевые фрагменты (музыка, звуки природы, аудио-эффекты и пр.). Общим недостатком обоих методов является то, что для анализа и обработки сигнала используется кратковременное дискретное преобразование Фурье (ДПФ), которое обеспечивает равномерное частотное разрешение, т.е. не учитывается нелинейное восприятие частот слуховой системой человека.

Недавно был предложен способ повышения разборчивости речи в шумах, сохраняющий энергию исходного сигнала [3] и состоящий из двух последовательных этапов: адаптивного изменения формы спектральной огибающей и расширения/компрессии динамического диапазона. Адаптивное изменение формы спектральной огибающей выполняется путем подчеркивания формантной структуры речи и уменьшения наклона спектральной огибающей корректирующим фильтром. Степень обработки зависит от уровня вокализованности речевого фрагмента – чем ближе он к гласному звуку, тем сильнее вносимый эффект. Способ предназначен для обработки исключительно речевых сигналов и предполагает наличие в сигнале специфических для речи характеристик, таких как формантные частоты и вокализованные фрагменты, а также подразумевает возможность произвести классификацию диктора (мужчина либо женщина). Поскольку сигнал сохраняет свою энергию, невозможно достичь удовлетворительной разборчивости при низких ОСШ. Недостатком способа является также и то, что даже при наличии идеальных условий (в полной тишине) в обрабатываемый сигнал вносятся искажения. Способ плохо применим для нестационарных шумов, поскольку в нем не учитываются изменяющиеся спектральные и динамические характеристики шума.

Целью данной работы является исследование возможности построения схемы обработки сигнала, обеспечивающей повышение разборчивости и натуральности звучания речи и музыки при прослушивании в шумной акустической обстановке за счет снижения эффекта маскирования полезного звукового сигнала нестационарными акустическими шумами путем использования частотно-зависимого адаптивного усиления. Важным условием является использование минимального необходимого усиления, достаточного для устранения эффекта маскирования, но при этом не создающего дискомфорта для восприятия и не приводящего к быстрой утомляемости слуха.

Предлагаемая схема обработки использует декомпозицию сигнала неравнополосным банком фильтров, позволяющим моделировать неравномерное частотное разрешение слуховой системы человека. Усиление каждой субполосы сигнала выполняется в зависимости от интенсивности шума в этой полосе при помощи адаптивного компрессора динамического диапазона с изменяющимися параметрами.

¹ Спектр сигнала становится плоским, как для белого шума, отсюда и название эффекта

Повышение качества слухового восприятия звуковых сигналов в шумовой обстановке
Схема обработки сигнала

Ухудшение восприятия звукового сигнала в шумовой обстановке вызвано тем, что шум маскирует [4] часть составляющих полезного сигнала. Эффект маскирования возможно уменьшить за счет компрессии (сужения) динамического диапазона (ДД) полезного сигнала [4, 5]. На рис. 3 показаны плотности распределения вероятности уровня частотных компонент речевого сигнала (рис. 3, а), шума метрополитена (рис.3, б) и речевого сигнала после компрессии (рис.3, в). Заметно, что частотные составляющие речи лежат в широком динамическом диапазоне, причем средний уровень компонент шумового сигнала для большого числа частотных полос выше среднего уровня компонент речевого сигнала. В результате компрессии громкие звуки получают малое усиление, а тихие – большое за счет чего эффект маскирования тихих компонент речевого сигнала становится менее выраженным.

Поскольку интенсивность шума зависит от частоты, то целесообразно использовать отдельные компрессоры для разных частотных диапазонов сигнала. Предлагаемая схема обработки заключается в выполнении следующих шагов: 1) декомпозиция полезного сигнала и шума на субполосы; 2) вычисление кратковременной энергии в каждой субполосе; 3) компрессия каждой субполосы полезного сигнала с учетом кратковременной энергии шума; 4) реконструкция выходного широкополосного сигнала из обработанных субполосных сигналов.

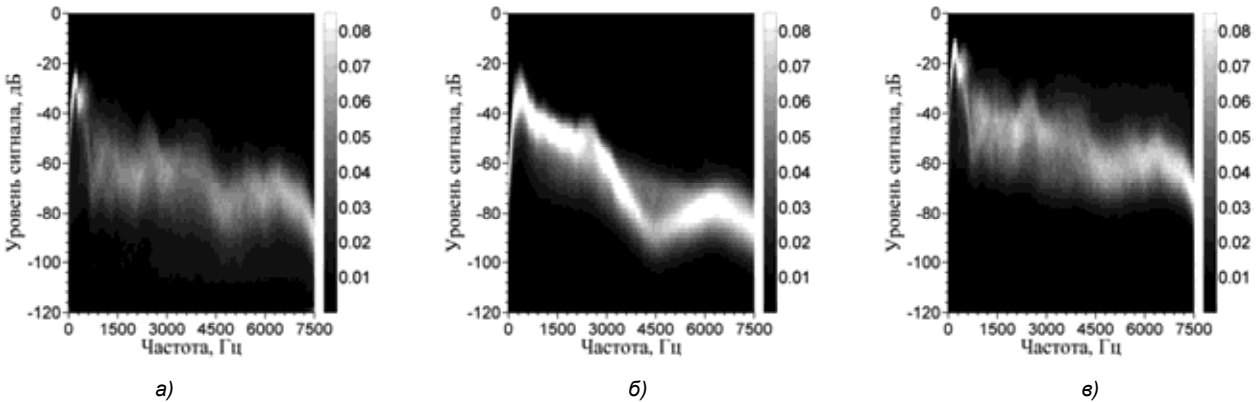


Рис. 3. Плотность распределения вероятности уровня частотных составляющих: а) речевого сигнала; б) шума метрополитена; в) речевого сигнала после компрессии



Рис. 4. Схема обработки сигнала для повышения качества его воспроизведения в шумовой обстановке

Декомпозиция и реконструкция полезного сигнала и шума выполняется при помощи неравнополосного косинусно-модулированного банка фильтров [6]. Для улучшения восприятия ширина полос выбирается в соответствии с разрешающей способностью слуховой системы человека. На рис. 4 показана предлагаемая схема обработки сигнала.

Схема имеет два входа, на которые поступают полезный и шумовой сигналы. На выходе формируется обработанный полезный сигнал. Оба входных сигнала раскладываются на субполосы при помощи банка фильтров анализа. Отсчеты субполосных сигналов набираются с прореживанием для того, чтобы снизить вычислительные затраты при практической реализации. Прореживание не приводит к потере информации, поскольку каждый субполосный сигнал имеет узкий частотный диапазон. Для восстановления полезного сигнала используется банк фильтров синтеза.

Адаптивный компрессор динамического диапазона

Уровень сигнала на выходе компрессора зависит от уровня входного сигнала и амплитудной характеристики компрессора. Амплитудная характеристика определяет вносимое усиление в зависимости от уровня входного сигнала (часто амплитудную характеристику отображают в виде графика, который показывает зависимость выходного уровня от входного).

На рис. 5, а показана амплитудная характеристика КДД, в соответствии с которой выполняется компрессия входного сигнала, если его уровень не превосходит по-

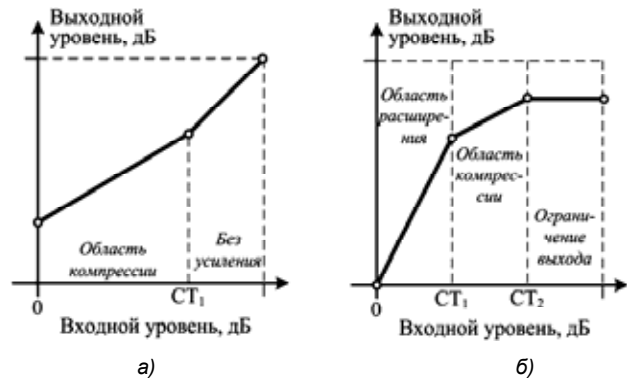
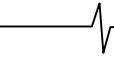


Рис. 5. Типовые амплитудные характеристики КДД



рога срабатывания CT_1 . В тоже время сигналы, чей уровень выше CT_1 , пропускаются компрессором без изменения. Более сложная амплитудная характеристика, сочетающая в себе три стратегии обработки, показана на рис. 5, б. Низкоуровневые сигналы (меньшие порога CT_1) попадают в область расширения, в которой усиление сигнала увеличивается по мере роста входного уровня. Применение операции расширения динамического диапазона делает более различимыми сигналы, обладающие малой энергией. Сигналы с уровнем, попадающим в диапазон от CT_1 до CT_2 , подвергаются компрессии, т.е. на этом участке коэффициент усиления падает с возрастанием уровня сигнала. Третий участок функции на рис. 5, б соответствует работе компрессора в режиме ограничителя, при этом выполняется ослабление, а не усиление сигнала.

Обычно амплитудная характеристика является фиксированной и не изменяется в процессе работы компрессора, что не подходит для решения поставленной задачи, поскольку в данном случае степень компрессии должна зависеть от уровня регистрируемого шума – при агрессивном шуме она должна быть высокой, а в тишине должна быть нулевой. Необходимо использовать перестраиваемый (адаптивный) компрессор, с изменяющейся амплитудной характеристикой.

Математически адаптивный компрессор описывается действием нелинейного оператора

$$y[n] = \psi \{x[n], w[n]\},$$

где $x[n]$ – полезный сигнал, $w[n]$ – шумовой сигнал, а $y[n]$ – выходной сигнал компрессора. Для обработки компрессором входной и шумовой сигналы разбиваются на фреймы размером R отсчетов:

$$x^{(m)}[r] = x[mR + r],$$

$$w^{(m)}[r] = w[mR + r],$$

где m – номер фрейма, а $r = 1, 2, \dots, R$ – индекс отсчета внутри фрейма. Выходной сигнал компрессора рассчитывается следующим образом

$$y^{(m)}[r] = x^{(m)}[r]g^{(m)}[r]. \quad (1)$$

Коэффициенты усиления удовлетворяют условию $g^{(m)}[r] \geq 1, \forall r$ и обеспечивают ослабление эффекта маскирования полезного сигнала шумовым. Для того, чтобы избежать появления слышимых артефактов при состыковке коэффициентов $g^{(m)}[r]$ необходимо, чтобы последовательность коэффициентов была без разрывов на границах соседних фреймов, что достигается путем интерполяции.

Адаптивный КДД, предлагаемый в данной работе, имеет изменяющуюся амплитудную характеристику (см. рис. 6), которая зависит от уровня шума W_{nos} (дБ) и определяется двумя параметрами:

1) ΔG (дБ) определяет динамический диапазон сигналов (в зависимости от уровня шума), к которым применяется усиление и компрессия;

2) ΔR (дБ) совместно с ΔG определяет коэффициент компрессии (угол наклона амплитудной характеристики).

Особенностью работы адаптивного КДД является то, что в случае отсутствия шума (в тихой обстановке), он не вносит искажений в полезный сигнал (характеристика имеет вид прямой с углом наклона 45°). Для слежения за мощностью шумового сигнала используется схема экспоненциального усреднения:

$$W = 10 \log_{10} \left(\frac{1}{R} \sum_{r=1}^R (w^{(m)}[r])^2 \right),$$

$$W_{nos}^{(m)} = \alpha W_{nos}^{(m-1)} + (1-\alpha)W, \quad 0 < \alpha < 1,$$

где $W_{nos}^{(m)}$ – мощность текущего фрейма, $W_{nos}^{(m-1)}$ – мощность предыдущего фрейма, α – параметр усреднения.

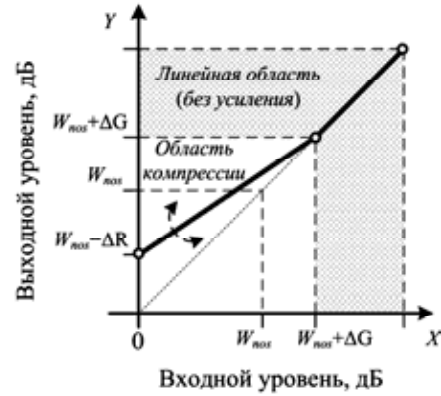


Рис.6. Амплитудная характеристика адаптивного КДД

Для текущего m -го фрейма полезного сигнала в адаптивном компрессоре рассчитывается мощность

$$X^{(m)} = 10 \log_{10} \left(\frac{1}{R} \sum_{r=1}^R (x^{(m)}[r])^2 \right),$$

которая вместе с текущей мощностью шума $W_{nos}^{(m)}$ используется для определения выходного уровня сигнала $Y^{(m)}$ (см. рис. 6). Коэффициент усиления для m -го фрейма рассчитывается следующим образом:

$$g_{avg}^{(m)} = 10^{(Y^{(m)} - X^{(m)})/20}.$$

Корректирующие коэффициенты $g^{(m)}[r]$, применяемые к входному сигналу (1) компрессора, рассчитываются путем интерполяции коэффициентов усиления текущего, последующего и предыдущего фреймов (т.е. $g_{avg}^{(m-1)}$, $g_{avg}^{(m)}$ и $g_{avg}^{(m+1)}$). При этом каждый коэффициент g_{avg} ассоциируется с центром соответствующего фрейма. Процедура интерполяции показана на рис. 7.

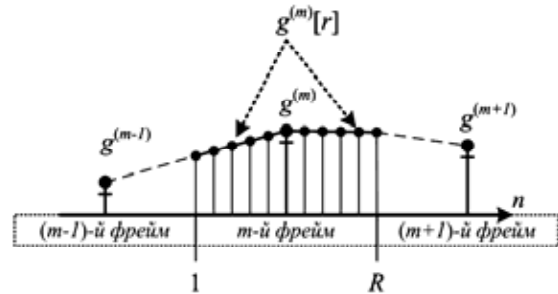


Рис. 7. Интерполяция коэффициентов усиления

Общая схема обработки сигнала в блоке адаптивного компрессора представлена на рис. 8.

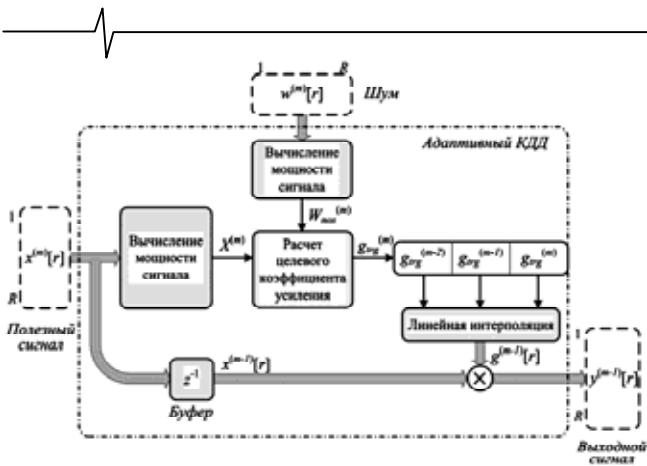


Рис. 8. Блок адаптивного КДД

Достоинством предлагаемого адаптивного КДД является то, что он позволяет ослабить эффект маскирования тихих звуков сигнала, не искажая громких звуков, которые не маскируются акустическим шумом.

Неравнополосный банк фильтров

Для реализации субполосного метода повышения качества звукового сигнала (рис. 4) используется неравнополосный косинусно-модулированный банк фильтров (НКМБФ), который аппроксимирует психоакустическую частотную шкалу барков [7]. В [8] показано, что повышение разборчивости речевого сигнала достигается путем разбиения входного сигнала на небольшое число субполос с последующей независимой компрессией каждой из них. Установлено [8], что максимизация разборчивости происходит уже при разбиении сигнала на 5 каналов, а дальнейшее увеличение их числа разборчивости практически не вызывает. Кроме того, небольшое количество полос позволяет избежать эффекта «выбеливания» и сохранить натуральное звучание полезного. Исходя из этих фактов и стремясь к уменьшению вычислительных затрат в практической реализации схемы обработки сигнала (рис. 4), предлагается использовать 5-канальный НКМБФ.

В банке фильтров используется принцип модуляции [9], позволяющий получить из одного фильтра-прототипа $h[n]$ импульсные характеристики полосовых фильтров $h_k[n]$, равномерно покрывающих весь частотный диапазон. Импульсные характеристики фильтров анализа $h_k[n]$ и синтеза $f_k[n]$ имеют следующий вид

$$h_k[n] = 2h[n] \cos\left((2k+1) \frac{\pi}{2M} \left(n - \frac{N-1}{2} \right) + (-1)^k \frac{\pi}{4} \right),$$

$$f_k[n] = 2h[n] \cos\left((2k+1) \frac{\pi}{2M} \left(n - \frac{N-1}{2} \right) - (-1)^k \frac{\pi}{4} \right),$$

где $k = 0, 1, \dots, M-1$ – номер каналов банка фильтров; M – число каналов банка фильтров, а N – длина фильтра-прототипа, которая должна удовлетворять уравнению $N = 2sM$, где s – произвольное натуральное число.

Преимуществом такого способа формирования импульсных характеристик является то, что для фильтрации сигнала существует эффективная схема (полифазная реализация). Схема основывается на следующих положениях:

- 1) вычисление выходных сигналов выполняется одновременно для всех каналов;
- 2) модулирующая функция является симметричной

$$c(k, n) = \cos\left((2k+1) \frac{\pi}{2M} \left(n - \frac{N-1}{2} \right) \pm (-1)^k \frac{\pi}{4} \right),$$

для которой из свойств периодичности косинуса имеем

$$c(k, l + 2pM) = (-1)^p c(k, l). \tag{2}$$

Для наглядности описания реализации необходимо выполнить запись в терминах z -преобразования:

$$\begin{bmatrix} H_0(z) \\ H_1(z) \\ \vdots \\ H_{M-1}(z) \end{bmatrix} = \begin{bmatrix} c(0,0) & c(0,1) & \dots & c(0,N-1) \\ c(1,0) & c(1,1) & \dots & c(1,N-1) \\ \vdots & \vdots & \ddots & \vdots \\ c(M-1,0) & c(M-1,1) & \dots & c(M-1,N-1) \end{bmatrix} \times \tag{3}$$

$$\times \text{diag}(1, z^{-1}, \dots, z^{-N}) \cdot \begin{bmatrix} h[0] \\ h[1] \\ \vdots \\ h[N-1] \end{bmatrix},$$

где $H_{0,M-1}(z)$ – z -преобразования импульсных характеристик полосовых фильтров.

В процессе фильтрации входной сигнал умножается на коэффициенты фильтра-прототипа и на матрицу косинусной модуляции M_C . Используя (2), можно показать, что данная матрица обладает регулярной структурой

$$M_C = \underbrace{\begin{bmatrix} C & -C & C & -C & \dots \end{bmatrix}}_{m \text{ повторов}},$$

где $C = [c(k, n)]_{0 \leq k < M, 0 \leq n < 2M}$, и может быть представлена в виде произведения

$$M_C = C \cdot \underbrace{\begin{bmatrix} I & -I & I & -I & \dots \end{bmatrix}}_{m \text{ повторов}} = C \cdot A, \tag{4}$$

где I – единичная матрица размера $2M \times 2M$. Факторизация (4), по сути, представляет запись быстрого алгоритма умножения на матрицу M_C . При этом количество требуемых операций умножения уменьшается в s раз, поскольку умножение на матрицу A требуют только операций сложения. Учитывая (4), выражение (3) можно переписать в следующем виде

$$\begin{bmatrix} H_0(z) \\ H_1(z) \\ \vdots \\ H_{M-1}(z) \end{bmatrix} = C \cdot \begin{bmatrix} h[0] - z^{-2M} h[2M] + \dots \\ z^{-1} h[1] - z^{-(2M+1)} h[2M+1] + \dots \\ \vdots \\ z^{-(2M-1)} h[2M-1] - z^{-(4M+1)} h[4M-1] + \dots \end{bmatrix} =$$

$$= \mathbf{C} \cdot \begin{bmatrix} \sum_{p=0}^{m-1} d[2Mp]z^{-2Mp} \\ \sum_{p=0}^{m-1} d[2Mp+1]z^{-(2Mp+1)} \\ \vdots \\ \sum_{p=0}^{m-1} d[2M(p+1)-1]z^{-(2M(p+1)-1)} \end{bmatrix} = \mathbf{C} \cdot \mathbf{P}, \quad (5)$$

где $d(q+2pM) = (-1)^p h[q+2pM]$, $0 \leq q < 2M$, $0 \leq p < m$. Элементами матрицы \mathbf{P} являются полифазные компоненты фильтра-прототипа. Дополнительное уменьшение вычислительной сложности косинусно-модулированного банка фильтров (КМБФ) заключается в поиске эффективного способа умножения на матрицу \mathbf{C} , которое может быть выполнено с использованием дискретного косинусного преобразования [7, 10].

Выражение (5) служит основой для построения полифазной реализации косинусно-модулированного банка фильтров, которая показана на рис. 9.

Выше описан способ получения эффективной реализации только анализирующей части банка фильтров. Используя аналогичные построения можно получить поли-

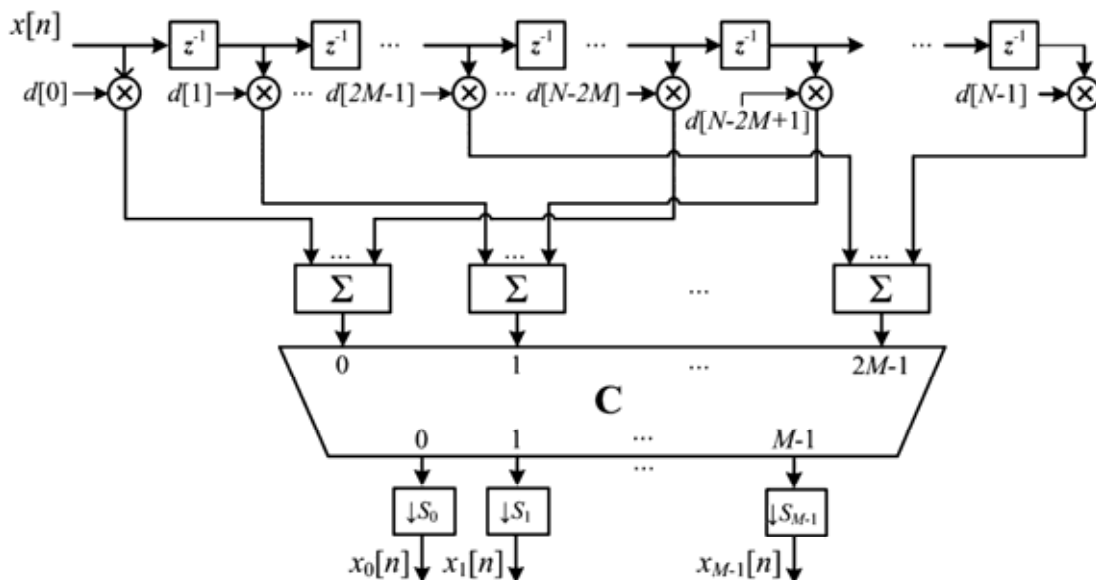


Рис. 9. Полифазная реализация косинусно-модулированного банка фильтров анализа

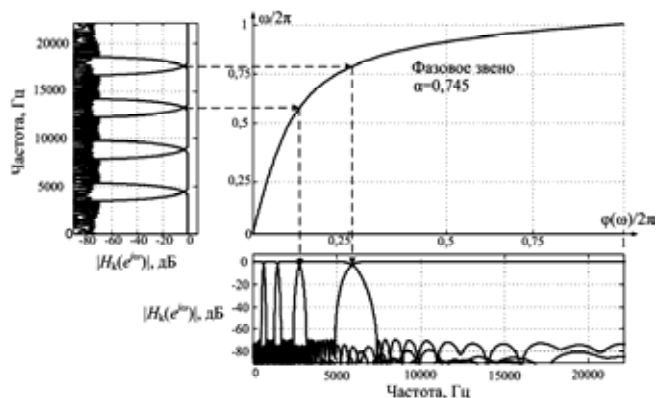


Рис. 10. Формирование неравнополосного банка фильтров (частота дискретизации 44,1 кГц)

фазную реализацию и для банка фильтров синтеза [7].

Для получения неравнополосной версии банка фильтров необходимо все элементы задержки структуры на рис. 9 заменить на фазовые звенья [7]:

$$A(z) = \frac{\alpha + z^{-1}}{1 + \alpha z^{-1}}, \quad |\alpha| \leq 1, \quad A(e^{j\omega}) = e^{j\varphi(\omega)},$$

где

$$\varphi(\omega) = -\omega + 2 \arctg(\alpha \sin \omega / (\alpha \cos \omega - 1))$$

Замена $z^{-1} \rightarrow A(z)$ вызывает нелинейное отображение частотной оси $\omega \mapsto \varphi(\omega)$, благодаря которому банк фильтров становится неравнополосным. Процесс синтеза 5-канального банка фильтров для практической реализации предлагаемого метода повышения качества звукового сигнала показан на рис. 10.

Коэффициент деформации частотной оси $\alpha = 0,745$ (при частоте дискретизации 44,1 кГц) позволяет получить банк фильтров, у которого частотные полосы имеют равную ширину в шкале барков. Для банка фильтров также определены коэффициенты децимации/интерполяции $S_k = \{21, 7, 3, 1, 1\}$. Подробности процедуры расчета НКМБФ можно найти в работах [6, 7].

Экспериментальные исследования

На рис. 11 показан результат обработки сигнала предлагаемой схемой (рис. 4), причем использован тот же сигнал, который был показан ранее на рис. 1. Сопоставляя спектрограммы речи до (рис. 1, а) и после (рис. 11, а) обработки со спектральными характеристиками шума (рис. 11, б), можно заметить, что усиление частотных компонент речи тем больше, чем больше интенсивность шума в соответствующей полосе. Также видно, что усиление полезного сигнала плавно возрастает с усилением интенсивности шума, что сохраняет для пользователя комфортное соотношение сигнала и шума при изменении уровня зашумленности внешней среды. Из рис. 11, в видно, что в отличие от (рис. 1, в), все компоненты речи легко различимы при прослушивании.

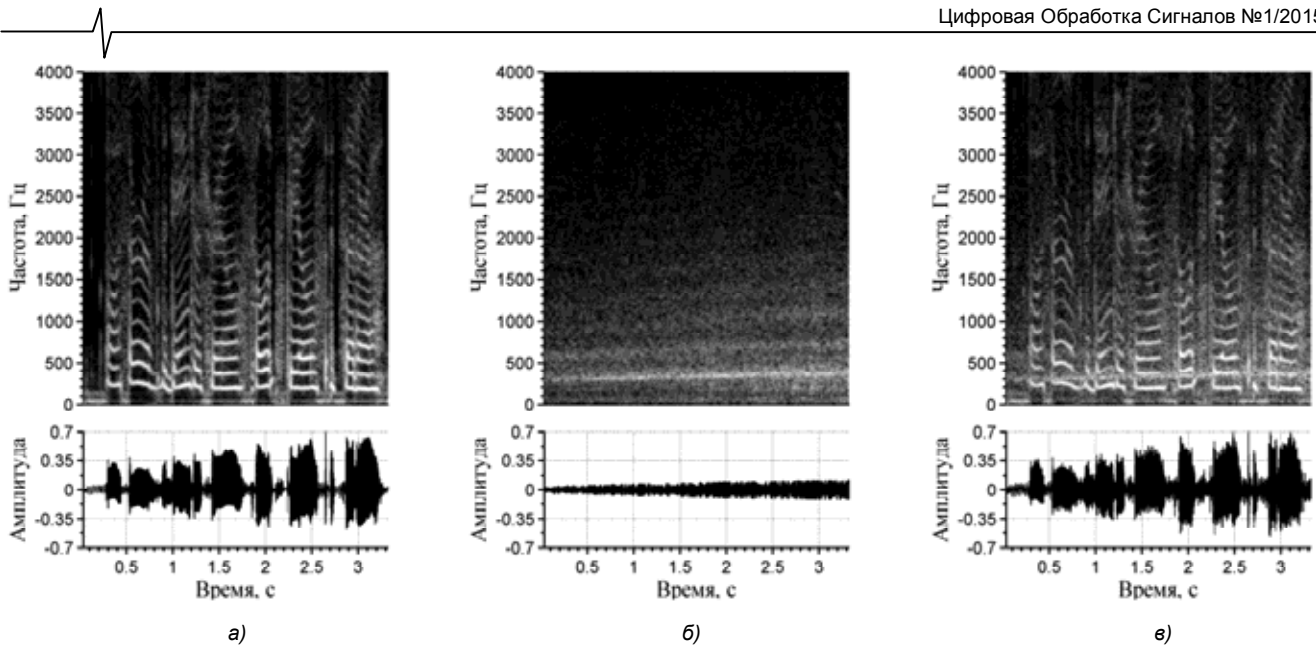


Рис. 11. Частотно-временное представление сигнала:

а) обработанная чистая речь; б) шум метрополитена; в) сумма обработанного сигнала и шума

Объективная оценка предлагаемого метода повышения качества восприятия звукового сигнала выполнена при помощи индекса разборчивости речи (ИРР) [11]. ИРР представляет собой объективную величину, которая коррелирует с разборчивостью речи в различных неблагоприятных условиях прослушивания. ИРР основан на вычислении энергии сигнала в критических частотных полосах. При расчете ИРР анализируется эквивалентный уровень спектра речевого сигнала и эквивалентный уровень спектра шумового сигнала в каждой из критических полос, вносящих вклад в разборчивость. Эквивалентный уровень рассчитывается, как усредненная по времени мощность сигнала в критической полосе, поделенная на ширину полосы. Поскольку в рассматриваемой задаче отдельно доступны и полезный сигнал и шумовой, то ИРР можно вычислить достаточно легко. Величина ИРР принимает значение от 0 до 1. Считается, что ИРР выше 0,75 является хорошим показателем, в то время как значение меньше 0,45 является неудовлетворительным.

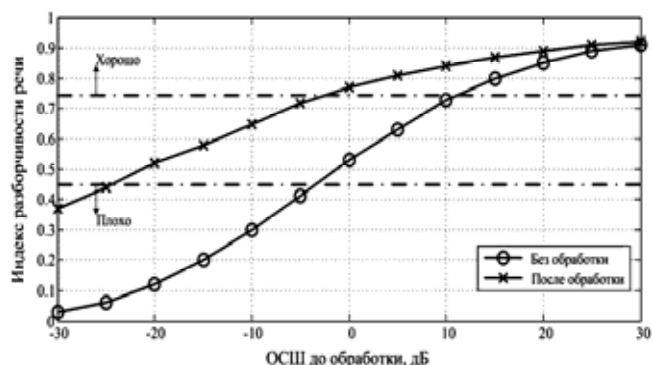


Рис. 12. Сравнение среднего значения ИРР для речевых сигналов

до и после обработки предложенным методом

Речевой материал для выполнения экспериментов взят из свободно распространяемой речевой базы CMU ARCTIC [12]. В качестве шумового сигнала использована запись акустической обстановки в метрополитене, запи-

санная при помощи смартфона. Работа предложенного способа проверялась для различных ОСШ, во всех случаях измерялась разборчивость речевого сигнала на фоне шума до обработки и после обработки. Полученные результаты экспериментов показаны на рис. 12.

В агрессивной шумовой обстановке (ОСШ < -20 дБ) метод обеспечивает удовлетворительную разборчивость, а при снижении интенсивности шума (ОСШ > 0 дБ) – хорошую. Необработанный сигнал имеет заметно меньшую разборчивость для ОСШ < 0 дБ и становится неразборчивым уже при ОСШ < -10 дБ.

Предлагаемая схема обработки сигнала реализована в виде приложения для iPhone².

Заключение

В работе предложена схема субполосной обработки звуковых сигналов для улучшения их восприятия в шумной акустической обстановке. Обработка выполняется при помощи неравнополосного банка фильтров, аппроксимирующего разрешающую способность слуховой системы, за счет чего достигается натуральное звучание выходного сигнала. Каждая субполоса полезного сигнала обрабатывается адаптивным компрессором динамического диапазона, который регулируется в зависимости от текущей энергии шума в этой субполосе. Эффективность применения предложенной схемы обработки сигнала подтверждается субъективными наблюдениями и объективными оценками на основе индекса разборчивости речи. Данная схема обработки может применяться в мобильных телекоммуникационных устройствах имеющих внешний микрофон для регистрации шума окружающей среды.

² Сайт приложения доступен по ссылке: <http://urbandenoiser.com/>

Литература

1. Sauert B., Vary P. Near and listening enhancement: Speech intelligibility improvement in noisy environments // Processing of International Conference on Acoustics, Speech, and Signal Processing (ICASSP), Toulouse, France, 14–19 May, 2006. – Toulouse, 2006. – P. 493–496.
2. Sauert B., Vary P. Near and listening enhancement optimized with respect to speech intelligibility index // Processing of 17th European Signal Processing Conference (EUSIPCO 2009), Glasgow, Scotland, 24–28 August, 2009. – Glasgow, 2009. – P. 1844–1848.
3. Zorila, T.-C. Speech-in-noise intelligibility improvement based on spectral shaping and dynamic range compression / T.-C. Zorila, V. Kandida, Y. Stylianou // In Proc. Interspeech, 2012. – Portland, Oregon, 2012. – P. 635–638
4. Фонлантен А. Слуховые аппараты / А. Фонлантен, А. Хорст.– Ростов н/Д.: Феникс, 2009. – 304 с.
5. Blesser B.A. Audio dynamic range compression for minimum perceived distortion / B.A. Blesser // IEEE Trans. Audio and Electroacoustics. – 1969. – vol. 17. – no. 1. – P. 22-32.
6. Вашкевич, М.И. Проектирование передискретизированного неравнополосного косинусно-модулированного банка фильтров / М.И. Вашкевич, А.А. Петровский // Информатика. – 2011. – № 2 (30). – С. 21–39.
7. Вашкевич М.И. Косинусно-модулированные банки фильтров с фазовым преобразованием: реализация и применение в слуховых аппаратах / М.И. Вашкевич, И.С. Азаров, А.А. Петровский – М.: Горячая линия-Телеком, 2014. – 210 с.
8. Trine T.D., Tasell D.V. Digital hearing aid design: facts vs. fantasy // The Hearing Journal. – 2002. – Vol. 55, №2. – P. 36–42.
9. Koilpillai D. Cosine-modulated FIR filter banks satisfying perfect reconstruction / D. Koilpillai, P.P. Vaidyanathan // IEEE transaction on signal processing. – 1992. – Vol. 40, № 4. – P. 770–783.
10. Вашкевич, М.И. Применение полиномиальных алгебр и теории Галуа для синтеза быстрых алгоритмов дискретных косинусных преобразований / М.И. Вашкевич, А.А. Петровский // Цифровая обработка сигналов. – 2011. – № 3. – С. 2–10.
11. American National Standard. Methods for the Calculation of the Speech Intelligibility Index. ANSI S3.5-1997, 1997.
12. Kominek J., Black A. The CMU ARCTIC speech databases for speech synthesis research / Language Technologies Institute, Carnegie Mellon University, Pittsburgh, PA, Tech. Rep. CMU-LTI-03-177, 2003.

LISTENING ENHANCEMENT IN NOISY ENVIRONMENT BASED ON SPECTRAL DECOMPOSITION AND ADAPTIVE DYNAMIC RANGE COMPRESSION OF THE SIGNAL

Azarov E.S., Vashkevich M.I., Lihachev D.S., Petrovsky A.A.

The paper investigates possibility of automatic listening enhancement in noisy environment. In order to improve subjective quality time-frequency components of the signal are nonlinearly amplified according to spectral components of surrounding noise. The method, proposed in the paper, is applicable to different signal including speech and audio. Practical value of the method is evaluated for speech signals using objective measures.

Уважаемые коллеги!

Приглашаем Вас принять участие в формировании тематических выпусков журнала «Цифровая обработка сигналов» и размещению рекламы продукции (услуг) Вашей организации на его страницах. В случае положительного решения просим представить в редакцию журнала Ваши предложения по плановому размещению информационных материалов и макет рекламы продукции (услуг) с указанием желаемого её месторасположения: обложка (2-я, 3-я или 4-я стр.), цветная внутренняя полоса (объем полосы).

Журнал «Цифровая обработка сигналов» издается с 1999 года. Выходит ежеквартально, тиражом – 700 экз. Распространяется по подписке через агентство «Роспечать» в России (индекс 82185), СНГ и странах Балтии (индекс 20630), а также на Конференции: «Цифровая обработка сигналов и ее применение – DSPA’».

Научно-технический журнал «Цифровая обработка сигналов» включен в Перечень изданий, рекомендуемый ВАК РФ для публикации результатов научных исследований соискателями ученой степени доктора и кандидата технических наук в области радиотехники, связи, вычислительной техники, электроники, приборостроения, информационных технологий, информационно-измерительных и управляющих систем. По предварительным итогам за 2014 год по рейтингу Science Index базы РИНЦ (3,394) журнал «Цифровая обработка сигналов» занимает 344-ю позицию из почти 3000 представленных изданий.

Планируемые сроки издания отдельных номеров журнала:

№ 1 – апрель 2015 г. Тематический выпуск: «ЦОС в информационно-измерительных системах»

№ 2 – июнь 2015 г. Тематический выпуск: «Теория и методы цифровой обработки сигналов» (по материалам международной научно-технической конференции «Цифровая обработка сигналов и ее применение – DSPA’ 2015».

№ 3 – сентябрь 2015 г. Тематический выпуск: «Цифровая обработка изображений».

№ 4 – декабрь 2015 г. Тематический выпуск: «ЦОС в радиотехнике и системах телекоммуникаций».

Ориентировочная стоимость рекламных услуг:

4-я (внешняя) страница цветной обложки – 25 тысяч рублей.

2-я и 3-я (внутренние) страницы цветной обложки – 15 тысяч рублей.

1/2 цветной внутренней полосы – 8 тысяч рублей.

Ждем Ваших предложений.

С наилучшими пожеланиями, зам. главного редактора
д.т.н., профессор Витязев Владимир Викторович, телефон 8-903-834-81-81.

Предложения прошу направлять по адресу: E-mail: vityazev.v.v@rsreu.ru или info@dspa.ru

УДК 519

МОДЕЛИРОВАНИЕ ВАРИАЦИЙ КОСМИЧЕСКИХ ЛУЧЕЙ НА ОСНОВЕ СОВМЕЩЕНИЯ КРАТНОМАСШТАБНЫХ ВЕЙВЛЕТ-РАЗЛОЖЕНИЙ И НЕЙРОННЫХ СЕТЕЙ ПЕРЕМЕННОЙ СТРУКТУРЫ

Мандрикова О.В., д.т.н., ведущий научный сотрудник Института космофизических исследований и распространения радиоволн ДВО РАН, профессор кафедры систем управления Камчатского государственного технического университета, e-mail: oksanam1@mail.ru;

Заляев Т.Л., младший научный сотрудник Института космофизических исследований и распространения радиоволн ДВО РАН, e-mail: tim.aka.gerald@mail.ru.

Ключевые слова: вейвлет-разложение, нейронные сети, переменная структура, вариации космических лучей, аномальные особенности.

Введение

Наблюдения вариаций интенсивности космических лучей (КЛ) используются при проведении ряда фундаментальных и прикладных исследований, связанных с мониторингом и прогнозом космической погоды [1]. В периоды сильных возрастаний интенсивности (GLE-события [2]) обширные потоки высокоэнергичных частиц создают главную проблему для космического оборудования и других технологий, для радиосвязи в полярных регионах, а также для космонавтов [2]. Поэтому выделение и предсказание подобных событий является весьма важным.

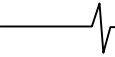
Сложности решения задачи моделирования и анализа вариаций космических лучей связаны с необходимостью понимания внутренней структуры исходных массивов данных, исследованием носителей информации, выделением информации, связанной непосредственно с процессом исследования. Наблюдаемые на поверхности Земли вариации космических лучей являются интегральным результатом различных солнечных, гелиосферных и атмосферных явлений и имеют сложную внутреннюю структуру [1, 3]. Их интенсивность зависит от метеорологических параметров, в частности от температуры и давления воздуха, состояния геомагнитного поля, электромагнитной обстановки в Солнечной системе и физических условий в Галактике [1]. В вариациях космических лучей находит отражение 11-летний цикл и 27-дневный солнечный период вращения [3]. Помимо перечисленного в вариациях присутствует суточный ход, обусловленный асимметрией формы магнитосферы, которая изменяется во времени при изменении параметров солнечного ветра [5].

В виду сложной структуры данных традиционные методы моделирования временных рядов не позволяют адекватно описывать вариации космических лучей и приводят к потере информации [5]. Исследования данной работы основаны на совместном применении методов вейвлет-преобразования и нейронных сетей переменной структуры [8]. Вейвлет-преобразование является

Предложен метод моделирования данных космических лучей, основанный на совместном применении вейвлет-преобразования и нейронных сетей, позволяющий выполнить детальный анализ и прогноз данных. Полученные на основе вейвлет-преобразования информативные составляющие вариаций космических лучей аппроксимируются нейронными сетями переменной структуры. Путём анализа ошибок сетей выявляются аномальные особенности, возникающие в периоды повышенной солнечной активности. На основе предложенного метода исследованы данные нейтронных мониторов станций «Мыс Шмидта» и «Апатиты» и выявлены аномальные особенности, возникающие в вариациях космических лучей во время магнитных бурь и в периоды событий сильного повышения интенсивности космических лучей. Проведена оценка точности метода и границ его применимости.

эффективным средством изучения сложных нестационарных процессов и дает возможность выполнять детальный анализ локальных структур [6, 7, 15, 16]. Преимущество нейросетевого представления аппроксимируемой функции заключается в большой гибкости базовых функций и их способности к адаптации [4, 10]. Аппарат нейронных сетей также позволяет путем анализа ошибок решения обученной сети выявлять скрытые закономерности в данных, связанные с изменением их структуры [10, 11]. В данной работе на основе кратномасштабных вейвлет-разложений [6, 7] выполняется выделение информативных составляющих вариаций космических лучей. Полученные составляющие далее аппроксимируются нейронной сетью. Для возможности построения адаптивных аппроксимаций предлагается использовать архитектуру прямонаправленной сети переменной структуры [8]. Для выбора вейвлетов используются критерии, предложенные в работе [12]. Алгоритм выбора уровня вейвлет-разложения и выделения информативных составляющих основан на минимизации погрешности получаемой аппроксимации.

На основе разработанного средства авторами проведено исследование данных нейтронных мониторов станций «Мыс Шмидта» и «Апатиты». В процессе моделирования данных в вариациях космических лучей выявлены аномальные особенности, возникающие в периоды сильных магнитных бурь, а также наблюдающиеся накануне и после GLE-событий.



Теоретическая часть
Кратномасштабные разложения
временного ряда на компоненты

Не нарушая общности, будем считать, что исходные данные принадлежат пространству с разрешением $j = 0$ (до кратномасштабного разложения на компоненты). На основе кратномасштабных вейвлет-разложений до уровня m получаем представление временного ряда в виде [13, 14]:

$$f_0(t) = \sum_{j=-1}^{-m} f^d [2^j t] + f^a [2^{-m} t], \quad (1)$$

где $f^d [2^j t] \in W_j$, $f^a [2^{-m} t] \in V_{-m}$, $W_j = \text{clos}_{L^2(\mathbb{R})} \times (2^{j/2} \Psi(2^j t - n)) : n \in \mathbb{Z}$, Ψ – базисный вейвлет; $V_j = \text{clos}_{L^2(\mathbb{R})} (2^{j/2} \varphi(2^j t - n)) : n \in \mathbb{Z}$, φ – скейлинг-функция, j – разрешение. Компонента ряда $f^a [2^{-m} t] = \sum_k c_{-m,n} \varphi_{-m,n}(t)$, где $c_{-m,n} = \langle f, \varphi_{-m,n} \rangle$, является сглаженной компонентой, компоненты $f^d [2^j t] = \sum_n d_{j,n} \Psi_{j,n}(t)$, где $d_{j,n} = \langle f, \Psi_{j,n} \rangle$, являются разномасштабными детализирующими компонентами.

С целью восстановления исходного разрешения $j = 0$, для полученных после преобразования компонент (см. (1)) выполняем операцию вейвлет-восстановления (отдельно для каждой компоненты). Получаем следующее представление ряда:

$$f_0(t) = f_0^{a,(-m)}(t) + \sum_{j=-1}^{-m} f_0^{d,j}(t), \quad (2)$$

где $f_0^{a,(-m)}(t) = \sum_n c_{0,n}^{(-m)} \varphi_{0,n}(t)$, $f_0^{d,j}(t) = \sum_n d_{0,n}^j \Psi_{0,n}(t)$, верхние индексы $(-m)$, j соответствуют разрешению компоненты до выполнения операции вейвлет-восстановления.

Путем изменения уровня разложения m , мы можем получить различные кратномасштабные представления ряда вида (2). Для выбора уровня вейвлет-разложения, обеспечивающего наилучшее представление ряда, определяемое его структурой, будем использовать алгоритм, описанный далее.

Следуя критериям выбора аппроксимирующих вейвлетов, предложенным в работе [12], для кратномасштабных вейвлет-разложений будем использовать вейвлеты семейства Койфлеты. Койфлеты – единственные из ортогональных вейвлетов, которые имеют носитель наименьшего размера при достаточном числе нулевых моментов в скейлинг-функции φ [7]:

$$\int_{-\infty}^{+\infty} t^k \Psi(t) dt = 0, k = \overline{0, s}; \quad \int_{-\infty}^{+\infty} \varphi(t) dt = 1, \int_{-\infty}^{+\infty} t^k \varphi(t) dt = 0$$

при $k = \overline{1, s}$.

В работе [9] показано, что данное свойство обеспечивает наилучшее приближение функции в сглаженных компонентах кратномасштабных вейвлет-разложений:

если функция f принадлежит C^k (C^k – пространство функций k раз непрерывно дифференцируемых) в окрестности $2^{-m} n$ с $k \leq s$, то

$$2^{-m/2} \langle f, \varphi_{-m,n} \rangle \approx f(2^{-m} n) + O(2^{-m(k+1)}).$$

Порядок приближения возрастает с ростом k , результирующий Койфлет имеет носитель размера $3s - 1$.

В работе [13] показано, что при совместном применении вейвлет-преобразования и нейронных сетей наименьшую погрешность аппроксимации вариаций космических лучей позволяют получить Койфлеты порядка $s = 3$.

Аппроксимация компонент временного ряда
нейронной сетью

На основе нейронной сети для сглаженной компоненты получаем отображение

$$y: f_0^{a,(-m)} \rightarrow f_0^{*a,(-m)}$$

Если $\hat{f}_0^{*a,(-m)}$ действительный выход сети, а $f_0^{*a,(-m)}$ – желаемый, то $f_0^{*a,(-m)} = y(f_0^{a,(-m)})$ – неизвестная функция, а $\hat{f}_0^{*a,(-m)}$ – ее аппроксимация, которую воспроизводит нейронная сеть.

При подаче на вход обученной нейронной сети значений функции $f_0^{a,(-m)}$ из интервала $[l - Q + 1, l]$, сеть становится способной вычислить упрежденные ее значения на временном интервале $[l + 1, l + I]$, где l – текущий дискретный момент времени; I – длина интервала упреждения. Ошибка сети в момент времени t определяется как разность между желаемым $f_0^{*a,(-m)}(t)$ и действительным $\hat{f}_0^{*a,(-m)}(t)$ выходными значениями сети.

Алгоритм построения сети
и выбора уровня разложения

Шаг 1. На основе соотношения (2) получаем сглаженную компоненту ряда в виде

$$f_0^{a,(-m)}(t) = \sum_n c_{0,n}^{(-m)} \varphi_{0,n}(t),$$

где $m = 1$.

Шаг 2. Массив данных $\{c_{0,n}^{(-m)}\}_{n=1}^N$, где N – число элементов массива, делим на блоки:

$$\{c_{0,n}^{(-m)}\}_{n=1}^Q,$$

$\{c_{0,n}^{(-m)}\}_{n=2}^{Q+1}, \dots, \{c_{0,n}^{(-m)}\}_{n=N-Q+1}^N$. Длина блока Q определяется путем минимизации погрешности получаемой аппроксимации данных.

Шаг 3. Из полученных блоков данных формируем обучающую матрицу размерности $Q \times V$, где Q – длина входного вектора сети, V – количество обучающих векторов.

Шаг 4. Строим сеть переменной структуры (сеть переменной структуры – это многослойная прямона-

правленная сеть, архитектура которой определяется путем минимизации ошибки решения на множестве обучающих векторов [8]).

Шаг 5. Используя тестовые данные $\{c_{0,l}^{(-m)}\}_{l=l_0}^L$, оцениваем ошибку сети:

$$E_m = \sum_{l=1}^L \sum_{i=1}^I |\varepsilon_{i,m}(l)|,$$

где $\varepsilon_{i,m}(l) = \hat{c}_{0,l}^{i,(-m)} - c_{0,l}^{i,(-m)}$ – ошибка сети в дискретный момент времени l с шагом упреждения i , $c_{0,l}^{i,(-m)}$ – желаемое, $\hat{c}_{0,l}^{i,(-m)}$ – действительное выходное значение сети, I – длина выходного вектора сети, L – длина массива тестовых данных.

Шаг 6. Оцениваем разницу: $\Delta = E_m - E_{m-1}$. Если $\Delta \leq 0$, то переходим к шагу 7. Если $\Delta > 0$ то искомый уровень разложения $m^* = m - 1$, шаг 7 не выполняем.

Шаг 7. Если $m \leq M$, где $M = \log_2 N$ (максимальный допустимый уровень разложения M определяется числом элементов массива данных $N : M \leq \log_2 N$), увеличиваем на 1 уровень разложения ($m = m + 1$) и выполняем шаги 2-5. Если $m > \log_2 N$, искомый уровень разложения $m^* = m$.

Результаты экспериментов

В экспериментах использовались минутные данные нейтронных мониторов станций «Апатиты» и «Мыс Шмидта» за 2008г. Этот период времени является периодом минимума солнечной активности, что обеспечивает наилучшие условия для построения аппроксимации характерного хода КЛ. Поскольку динамика наблюдаемых на поверхности Земли вариаций КЛ существенно зави-

сит от состояния геомагнитного поля и электромагнитной обстановки в солнечной системе [3], обучающие множества нейронных сетей формировались из данных, регистрируемых в периоды спокойного геомагнитного поля. В качестве базисных функций использовались ортонормированные вейвлеты семейства Койфлеты порядка $s = 3$. На основе описанного выше алгоритма для уровней разложения $m = \overline{1,6}$ были построены нейронные сети. Наилучшие результаты при тестировании полученных нейронных сетей показала сеть, воспроизводящая сглаженную компоненту $f^a [2^{-5}t]$ (см. соотношение 1). Данная компонента и исходные данные КЛ для станции «Апатиты» представлены на рис. 1. Архитектура построенных нейронных сетей для станций «Апатиты» и «Мыс Шмидта» представлена на рис. 2, 3, соответственно. Результаты работы данных нейронных сетей показаны на рис. 4, 5. При анализе результатов использовались Н-компоненты геомагнитного поля (данные станции «Паратунка», Камчатский край), характеризующие интенсивность геомагнитных возмущений в анализируемые периоды времени. Анализ рис. 4, 5 показывает, что в периоды спокойного геомагнитного поля абсолютные значения ошибок сетей не превышают значения 0.05, что говорит о хорошем качестве их аппроксимирующих свойств. В периоды повышенной геомагнитной активности на станции «Мыс Шмидта» (3.05.2008 – 5.05.2005, рис. 5, д) наблюдается существенное увеличение ошибок сети, которые характеризуют возникновение аномалий, связанных с изменением хода процесса. На рис. 6 представлены результаты обработки данных КЛ в период GLE-события и сильных магнитных бурь. Анализ рис. 6 показывает, что в периоды магнитных бурь наблюдается существенное увеличение ошибок сети, обусловленных понижением уровня КЛ (Форбуш-эффект). Увеличение ошибок сети также наблюдается накануне и после GLE-события.

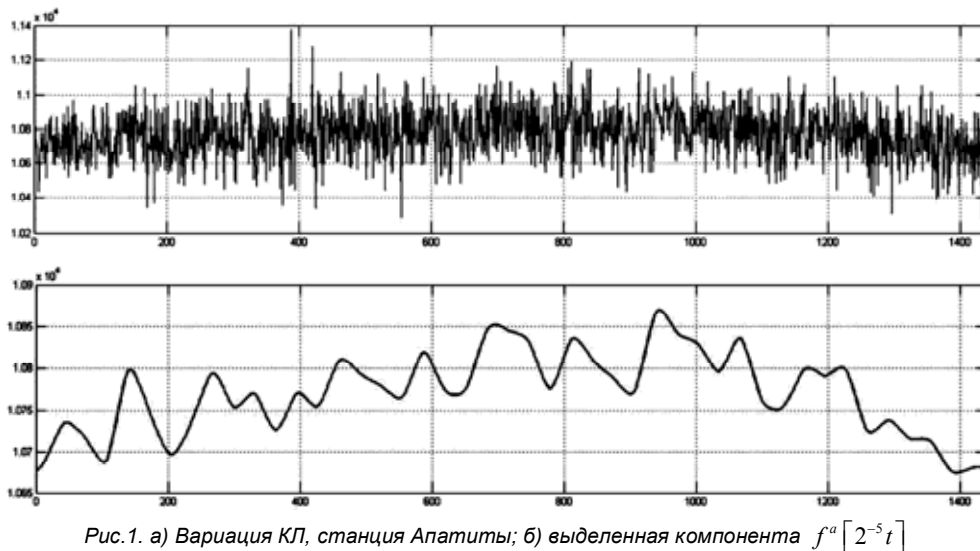


Рис.1. а) Вариация КЛ, станция Апатиты; б) выделенная компонента $f^a [2^{-5}t]$



Рис 2. Архитектура нейронной сети по аппроксимации данных КЛ станции «Апатиты»

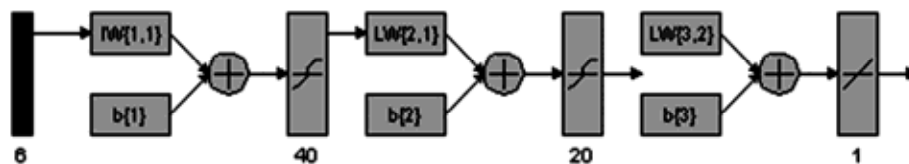


Рис 3. Архитектура нейронной сети по аппроксимации данных КЛ станции «Мыс Шмидта»

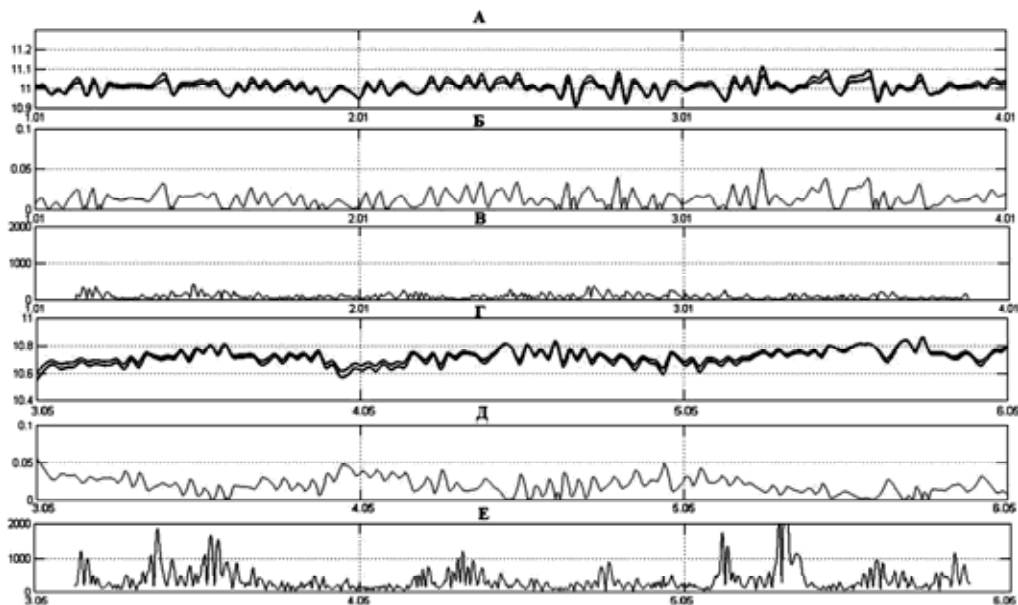


Рис. 4. Результаты работы нейронной сети, станция «Апатиты».

- а) сглаженная компонента вариации КЛ за период 1-3 января 2008 года (серый цвет) и её аппроксимация нейронной сетью (черный цвет) (период спокойного геомагнитного поля);
 б) абсолютные значения ошибок нейронной сети;

- в) интенсивность геомагнитных возмущений; г) сглаженная компонента вариации КЛ за период 3-5 мая 2008 года (серый цвет) и её аппроксимация нейронной сетью (черный цвет) (период возмущенного геомагнитного поля);
 д) абсолютные значения ошибок нейронной сети; е) интенсивность геомагнитных возмущений

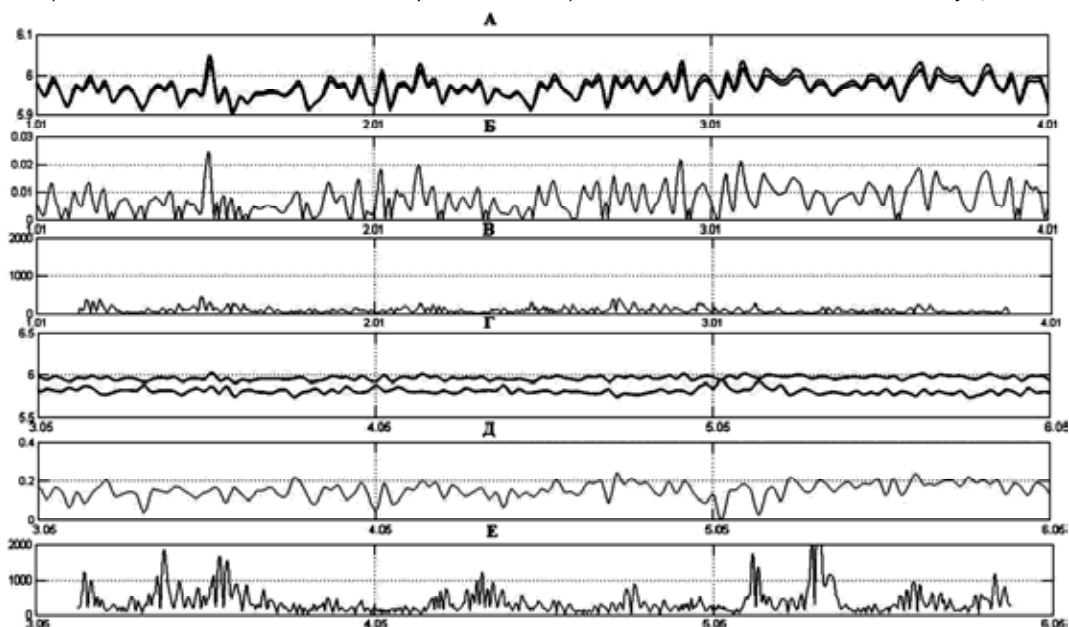


Рис. 5. Результаты работы нейронной сети, станция «Мыс Шмидта».

- а) сглаженная компонента вариации КЛ за период 1-3 января 2008 года (серый цвет) и её аппроксимация нейронной сетью (черный цвет) (период спокойного геомагнитного поля);
 б) абсолютные значения ошибок нейронной сети;

- в) интенсивность геомагнитных возмущений; г) сглаженная компонента вариации КЛ за период 3-5 мая 2008 года (серый цвет) и её аппроксимация нейронной сетью (черный цвет) (период возмущенного геомагнитного поля);
 д) абсолютные значения ошибок нейронной сети; е) интенсивность геомагнитных возмущений

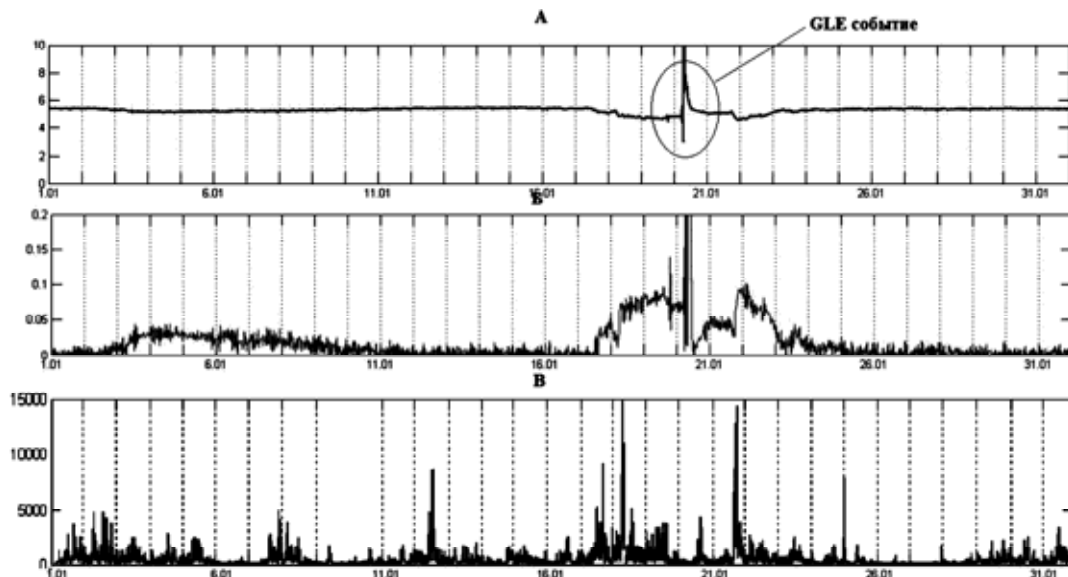


Рис. 6. Результаты работы нейронной сети, станция «Мыс Шмидта».

а) сглаженная компонента вариации КЛ за период 01.01.2005-31.01.2005 года (период GLE-события и сильных геомагнитных бурь); б) абсолютные значения ошибок нейронной сети; в) интенсивность геомагнитных возмущений

Заключение

Результаты работы нейронных сетей, реализующих предложенный метод аппроксимации КЛ для станций «Мыс Шмидта» и «Апатиты», показали, что в периоды спокойного геомагнитного поля, характеризующего низкую активность Солнца, абсолютные значения ошибок не превышают значения 0.05, что говорит о хорошем качестве аппроксимирующих свойств. В периоды повышенной геомагнитной активности наблюдается существенное увеличение ошибок сетей, обусловленных изменением состояния околоземного космического пространства и понижением уровня космических лучей (Форбуш-эффекты). Увеличение ошибок сетей также наблюдается накануне и после анализируемого GLE-события. Полученные результаты подтверждают эффективность предлагаемого метода и возможность его использования в задачах детального анализа вариаций космических лучей и выявления аномалий, возникающих в периоды повышенной активности Солнца.

Благодарности

Работа поддержана грантом РФФИ № 14-11-00194 и грантом «УМНИК» – № 862ГУ1/2013 от 09.12.2013 г. Авторы благодарят институты, поддерживающие станции регистрации космических лучей и магнитные обсерватории, данные которых были использованы в исследовании, а также выражают признательность сотрудникам Института земного магнетизма, ионосферы и распространения радиоволн РАН, разработавшим интерактивную среду и программное обеспечение баз данных, обеспечивающих оперативное получение информации о параметрах состояния космического пространства.

Литература

1. Топтыгин И.Н. Космические лучи в межпланетных магнитных полях. – М.: Наука, 1983. – 301 с.
2. Eroshenko E.A., Belov A.V., Kryakunova O.N.,

Kurt V.G., Yanke V.G. The alert signal of GLE of cosmic rays // Proceedings of the 31st ICRC, 2009.

3. Тясто М.И., Данилова О.А, Дворников В.М, Сдобнов В.Е. Большие снижения геомагнитных порогов космических лучей в период возмущений магнитосферы. Известия РАН, серия физическая, Т. 73, № 3, стр. 385-388. 2009.

4. P. Paschalis, C. Sarlanis, H. Mavromichalaki – Artificial Neural Network Approach of Cosmic Ray Primary Data Processing. Solar Physics, 2013; 182(1):303-318.

5. J. Kóta, A. Somogyi – Some problems of investigating periodicities of cosmic rays – Acta Physica Academiae Scientiarum Hungaricae, Tomus 27, pp, 523-548 (1969).

6. Mallat S. A Wavelet tour of signal processing [пер. с англ.]. – М.: Мир, 2005. – 671 с.

7. Daubechies I. Ten Lectures on Wavelets. – SIAM, 1992.

8. Нейроматематика: Учеб. пособие для вузов / А.Д. Агеев и др.; общ. ред. А.И. Галушкина. – М.: ИПРЖР, 2002. – 448 с.

9. Мандрикова О.В. Многокомпонентная модель сигнала со сложной структурой // Проблемы эволюции открытых систем. – 2008. – Вып. 10. – Т. 2. – С. 161–172.

10. Полозов Ю.А. Метод формирования обучающего множества для нейронной сети на основе вейвлет-фильтрации // Известия вузов, Северо-Кавказский регион. Ростов-на-Дону. – 2010. – № 3. – С. 12-16.

11. Мандрикова О.В. Оптимизация процесса обучения нейронной сети на основе применения конструкции вейвлет-преобразования (на примере модельного представления ионосферного сигнала) // Автоматизация и современные технологии. – 2009. – № 3. – С. 14.

12. Мандрикова О.В., Полозов Ю.А. Критерии выбора вейвлет-функции в задачах аппроксимации природных временных рядов сложной структуры // Информационные технологии. – Москва. 2012 г. -№1. – С. 31–36

13. Мандрикова О.В., Залаяев Т.Л., Белов А.В.,

Янке В.Г. Метод выявления аномалий в вариациях космических лучей на основе совмещения вейвлет-преобразования с нейронными сетями. Сборник докладов VI международной конференции «Солнечно-земные связи и физика предвестников землетрясений», – 2013 – С. 304-310.

14. Мандрикова О.В., Заляев Т.Л. Моделирование вариаций космических лучей на основе совмещения кратномасштабного анализа и сетей переменной структуры // Сборник тезисов докладов VI-й Международной научно-технической конференции по мягким вычислениям и измерениям (SCM'2013). Т.2. – СПб, 2013, стр. 111-117.

15. Мандрикова О.В., Глушкова Н.В., Полозов Ю.А. Алгоритмы выделения и анализа аномалий в параметрах критической частоты ионосферы fOF2 на основе совмещения вейвлет-преобразования и авторегрессионных моделей // Цифровая обработка сигналов. – Москва: РНТОРЭС. 2013 г. №1. С. 47-53.

16. Мандрикова О.В., Соловьев И.С. Вейвлет-технология обработки и анализа геомагнитных данных // Цифровая обработка сигналов – Москва: РНТОРЭС. 2012, – №2, – С. 24-29.

MODELING OF VARIATIONS OF COSMIC RAYS ON THE BASIS OF COMBINATION OF MULTIREOLUTION WAVELET EXPANSIONS AND NEURAL NETWORKS WITH VARIABLE STRUCTURE

Mandrikova O.V., Zalyaev T.L.

In this paper, we propose a method of cosmic rays data modeling, which is based on the joint application of the wavelet transform and neural networks, and allows to perform data analysis and forecasting. Allocated via wavelet decomposition, informative components of cosmic ray are further approximated by neural networks. Abnormal features that appear during periods of high solar activity, are revealed by analyzing the network error vector.

Based on the developed method, authors have investigated the data of neutron monitors station «Cape Schmidt» and «Apatity» and successfully detected abnormal features that appeared in the cosmic ray variations during magnetic storms and on the eve of events of a major increase of the ground level intensity of cosmic rays.

Н О В Ы Е К Н И Г И

Горбунов Ю.Н., Лобанов Б.С., Куликов Г.В.

Введение в стохастическую радиолокацию: учеб. пособие. – М.: Горячая линия – Телеком, 2015. – 376 с.

Рассмотрены основы радиолокации, вопросы реализации когерентно-импульсных систем радиолокации и их компонент – цифровых систем обнаружения и измерения параметров радиолокационных сигналов систем селекции движущихся целей, обеспечивающих выделение полезных сигналов от целей на фоне коррелированных помех по доплеровским и угловым признакам (пространственным частотам). В отличие от традиционного изложения материала, упор сделан на стохастические методы обработки сигналов и информации.

Для студентов, обучающихся по специальности 210601 – «Радиоэлектронные системы и комплексы», а также других радиотехнических и инфокоммуникационных специальностей.

Рембовский А.М., Ашихмин А.В., Козьмин В.А.

Радиомониторинг: задачи, методы, средства: научн. издание под ред. Рембовского А.М. – М.: Горячая линия – Телеком, 2015. – 640 с., 4-е изд., с испр.

Рассмотрены типовые задачи, решаемые с помощью средств радиомониторинга в самых различных областях – от управления использованием радиочастотного спектра до контроля радиообстановки при проведении контртеррористических операций. Изложены теоретические и практические вопросы построения и функционирования радиотехнических систем радиомониторинга и технической защиты информации, в том числе автоматизированных стационарных и мобильных систем пеленгования, определения местоположения источников радиоизлучения, измерения параметров сигналов, видов радиоизлучений и утечек информации. Рассмотрены математические методы и программное обеспечение для выявления и анализа радиоизлучений, измерения параметров сигналов и помех, определения положения источников радиосигналов и помех. Представлен анализ особенностей построения современных одноканальных и многоканальных широкополосных цифровых панорамных измерительных радиоприемников и радиопеленгаторов, в том числе и CDMA.

Для специалистов в области радиомониторинга, операторов и руководителей радиоконтрольных служб, а также студентов и аспирантов.



УДК 519.688

МЕТОДИКА КРАТКОСРОЧНОГО ПРОГНОЗА ИЗМЕНЕНИЯ ЗНАЧЕНИЙ ХАРАКТЕРИСТИК ДИНАМИЧЕСКОЙ СИСТЕМЫ НА ОСНОВЕ ЧИСЛЕННОГО АНАЛИЗА МНОГОМЕРНЫХ ВРЕМЕННЫХ РЯДОВ

*Запорожцев И.Ф., аспирант Мурманского государственного технического университета,
e-mail: zaporozhtsev.if@gmail.com;*

Серёда А.-В.И., д.т.н., e-mail: avis_14@mail.ru.

Ключевые слова: краткосрочный прогноз, пространственное распределение физической характеристики, многомерный временной ряд, сингулярный спектральный анализ, декомпозиция на эмпирические моды, модель TARX.

Введение

Задача разработки и совершенствования математических моделей и численных методов для анализа многомерных временных рядов, отражающих изменение во времени значений целевой характеристики для различных точек некоторого физического пространства, является на данный момент перспективным направлением исследований для специалистов из различных отраслей. Разработано большое количество теоретических и узко специализированных эмпирических моделей, допущений и эвристик, которые справедливы для конкретной прикладной задачи и предметной области. Тем не менее, ни одно из существующих решений не позволяет выстроить системный адаптивный подход к прогнозированию временной изменчивости пространственного распределения значений произвольной характеристики без учёта специфики анализируемых данных, обусловленной их природой. В данной статье делается определённый шаг в этом направлении: предлагается как общая концепция, так и детальная реализация авторской методики прогноза при работе с гридированными океанографическими данными. Для определения места этой методики среди имеющихся работ в первом разделе статьи приведена краткая характеристика некоторых групп моделей, положенных в основу современных концепций прогнозирования значений временных рядов. Затем (раздел 2) следует краткое описание многомерного ряда, с которым идёт работа. В разделе 3 представлены сведения о начальном внутрикластерном прогнозе, который является первым этапом предлагаемой методики. В разделе 4 вводятся так называемые кластерные характеристики, а в разделе 5 рассмотрена процедура исправления значений внутрикластерного прогноза. Общая схема методики прогноза приведена в разделе 6. Завершается статья анализом результатов вычислительных экспериментов, имевших целью апробацию методики на модельных данных (раздел 7).

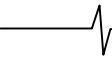
Представлена методика краткосрочного прогноза временной изменчивости пространственно распределённых характеристик природных динамических систем, основанная на анализе многомерных временных рядов. Предполагается, что значения характеристик заданы в узлах планарной равномерной сетки. Сформулирована как общая концепция этой методики, так и алгоритмическая детализация некоторых её этапов. На данный момент предлагаемая версия методики полностью сформирована и реализована в виде компьютерной технологии. Приводятся результаты численных экспериментов для модельных данных, а также их анализ.

1. Особенности проблемы классификации моделей прогноза

Стоит отметить, что классификация моделей и методов прогнозирования, как и классификация математических моделей, является в настоящее время не вполне устоявшейся. Некоторые критерии выделения и названия классификационных групп, предложенные отдельными авторами в последнее десятилетие и используемые нами в данной статье, не следует рассматривать как общепринятые.

В основу теоретических моделей анализа временной изменчивости характеристик природных динамических систем положены теории и методы решения задач математической физики. С развитием этого подхода стали появляться предметно-ориентированные модели, представляющие пространственную и временную изменчивость дифференциальными уравнениями, отражающими фундаментальные законы. Примеры подобных моделей, например, для динамики уровня поверхности моря можно найти в работах [1-3]. Эти модели могут быть достаточно легко адаптированы для применения в самых разных предметных областях. Одна из существенных объективных проблем их практической реализации вызвана отсутствием некоторых необходимых данных для расчётов или их недостаточной точностью. Также известной проблемой являются большие затраты временных ресурсов численной реализации этих моделей, которые не находят оправдание с точки зрения точности получаемых значений.

В основе другого подхода к прогнозированию лежат так называемые физико-статистические методы, по сути реализующие модели множественной линейной регрес-



сии по времени [2]. Для физико-статистических методов ставят целью получение статистических закономерностей изменения целевого параметра в зависимости от значений модельных параметров, имеющих чёткий физический смысл.

Аппроксимация результатов пассивных наблюдений и измерений, которая осуществляется подбором подходящего уравнения, включающего то или иное число независимых переменных, не несущих в себе никакой информации с точки зрения физики конкретных протекающих в реальной системе процессов, стала основой эмпирических моделей. Они являются итогом обработки результатов наблюдения изменения количественных значений наблюдаемых характеристик с целью обобщения результатов экспериментальных исследований для получения некоторой качественной информации и выработки управляющих решений. Наиболее распространённые эмпирические модели являются статистическими. Классическим примером для временных рядов является модель стационарных рядов ARIMA.

В настоящее время среди множества эмпирических регрессионных моделей выделяется классификационная единица, а именно так называемые структурные модели [4], не требующие выполнения условия стационарности. Они обычно используются в задачах декомпозиции временного ряда на аддитивные составляющие и прогноза его значений. Эти модели стали результатом усложнения процесса построения аналитических зависимостей, но по-прежнему отображают случайный (стохастический) характер некоторых параметров в представлении анализируемого временного ряда. Некоторые из этих моделей можно считать статистическими, например, модель эволюции выборочной плотности функции распределения [5]. Другие модели ориентированы на исключение случайных составляющих посредством задания количественных значений некоторых параметров без необходимости содержательной интерпретации этих значений, при условии, что модельные данные практически соответствуют реальным. Суть этого задания сводится к априорной предустановке значения отношения «сигнал-шум», как, например, в рамках сингулярного спектрального анализа [9-10]. Наконец, к структурным моделям можно отнести и нейронные сети, в частности, многослойный перцептрон, который вовсе не различает в обработке зашумлённые данные и данные, содержащие только полезный сигнал. Общей чертой всех структурных моделей является их алгоритмичность: необходимость выполнения ряда шагов, не все из которых хорошо формализованы. Неотъемлемой частью представления моделей подобного типа являются блок-схемы, структурные схемы, диаграммы состояний и т.д.

2. Многомерный временной ряд

Методика прогноза, предлагаемая нами, относится к классу структурных эмпирических моделей и основана на анализе многомерных временных рядов, каждый одномерный компонент которого поставлен в соответствие некоторому узлу регулярной сетки UG (uniform grid), заданной в прямоугольной области Ω . Предполагается, что одномерные временные ряды для каждого из узлов

сформированы единообразно: совпадают стартовый временной отсчёт, шаг по времени между измерениями и анализируемая длина временного ряда. Такой подход к формированию рядов позволяет говорить, что в области Ω задан многомерный временной ряд $F_N^{UG}(t) = (F_N^1(t), F_N^2(t), \dots, F_N^s(t))$ длины N , где s – число одномерных рядов. Каждый из одномерных рядов может быть также проиндексирован двумя координатами соответствующего ему узла сетки UG. На рис. 1 изображения узлов дублируются для всех моментов времени, что означает необходимость задания значений характеристики в узлах в эти моменты времени.

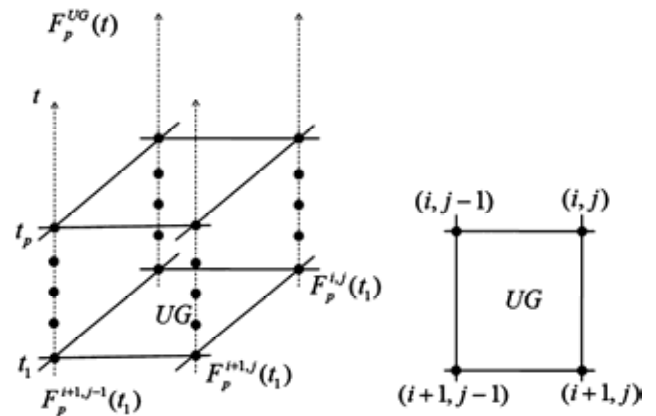


Рис. 1. Многомерный временной ряд

Основным ограничением в рамках формируемой методики является доступность для анализа только одного многомерного временного ряда (данных о значениях одной характеристики), что позволит использовать такую методику для прогноза различных характеристик, не учитывая природу данных.

3. Начальный внутрикластерный прогноз

Первым этапом анализа многомерного ряда в задаче прогноза является выделение кластеров узлов сетки UG; в дальнейшем временные ряды (кластерные ряды), которые им соответствуют, будут обрабатываться в отдельности. Разработанный авторами алгоритм кластеризации, представленный в статье [6], учитывает как взаимное расположение узлов, так и специфику временной изменчивости соответствующих им временных рядов.

Основные идеи, а также алгоритм начального внутрикластерного прогноза описаны детально в статьях [7, 8]. Здесь следует лишь упомянуть, что в основе алгоритма лежит модель и метод К-продолжения многомерного временного ряда, использующий результаты сингулярного спектрального анализа (МССА) [9, 10], а также (вариативно) многомерный вариант декомпозиции на эмпирические моды (МДЭМ) [11, 12], авторский опыт использования которого обобщён в статье [8]. Начальный внутрикластерный прогноз был протестирован, как на некоторых модельных данных, так и апробирован на экспериментальных данных геофизической природы: аномалиях уровня поверхности моря [7, 8]. В качестве новизны подхода следует отметить совместное использование математических моделей, которые лежат в основе современных технологий обработки временных рядов: МССА, МДЭМ и методики подбора оптимального

с точки зрения предустановленной границы ошибки прогноза длины фрагмента временного ряда, значения которого непосредственно участвуют в прогнозе (эта методика, основанная на выделении квазистационарных фрагментов ряда, изложена в книге [5]). Авторами разработан и реализован программно адаптированный алгоритм, который позволяет выстроить взаимодействие процедур, соответствующим трём упомянутым технологиям.

Предположим, что горизонт прогноза (число последовательных временных отсчётов, для которых необходимо построить прогноз значений характеристики) известен и равен τ . О точности прогноза значений характеристики по фрагменту кластерного многомерного ряда $F_N^C = (F_N^1, F_N^2, \dots, F_N^{s_c})$ длины T , для которого последним временным отсчётом является $(t-1)$, на горизонт τ будем судить по нормированной среднеквадратичной ошибке $\delta_{i,T,\tau}^C$ (в %):

$$\delta_{i,T,\tau}^C = \frac{100}{d(t)} \cdot \max_{i=1, \dots, s_c} \sqrt{\frac{\sum_{j=0}^{\tau-1} (F_N^i(t+j) - \tilde{F}_\tau^i(t+j))^2}{\tau}}, \quad (1)$$

где $\tilde{F}_\tau^C = (\tilde{F}_\tau^1, \tilde{F}_\tau^2, \dots, \tilde{F}_\tau^{s_c})$ – соответствующий многомерный ряд прогнозных значений длины τ (его значения соответствуют временным отсчётам $(t+j)$, $j = 0, \dots, \tau-1$), вычисленных по фрагменту ряда F_N^C длины T , а нормирующий множитель определяется разностью $d(t)$:

$$d(t) = \max_{i=1, \dots, s_c} \left\{ \max_{t' < t} F_N^i(t') - \min_{t' < t} F_N^i(t') \right\}. \quad (2)$$

В следующем разделе вводятся так называемые кластерные характеристики, которые будут участвовать в процедуре коррекции начального прогноза, формируя таким образом финальный внутрикластерный прогноз.

4. Кластерные характеристики

На множестве кластеров, а также на множестве упорядоченных пар кластеров зададим несколько функционалов, которые будем называть кластерными характеристиками. Их значения будут *использованы* на этапе коррекции результатов начального внутрикластерного прогноза.

Коэффициент кросс-корреляции. Коэффициент кросс-корреляции можно рассматривать как кластерную характеристику двумя способами: выбрать наименьшее его значение, вычисленное для всех *возможных* пар одномерных временных рядов одного кластера при нулевом лаге, или ввести для пары кластеров, определяя корреляционную связь для временных рядов различных кластеров.

Отношение сингулярных чисел. Множество сингулярных чисел может быть определено для каждого временного ряда, если использовать одномерный вариант сингулярного спектрального разложения траекторной матрицы в рамках метода МССА. Составим отношения первого сингулярного числа ко второму, второго к третьему и т.д. Затем выбираем для каждой пары номеров сингулярных чисел наибольшее и наименьшее значение

отношения среди всех рядов кластера. Представляет интерес рассмотрение таких отношений небольшого количества первых сингулярных чисел. Как критерий качества, эти отношения можно использовать следующим образом: чем меньше отличие наибольшего и наименьшего отношений для фиксированных номеров сингулярных чисел, тем в большей степени кластерный ряд является однородным. Также заметим, что значительная близость значений пары соседних сингулярных чисел показывает наличие гармонической составляющей в модели временного ряда [9]. Если одно значение в паре отношений сингулярных чисел (k_1, k_2) указывает на присутствие гармоники, а другое – нет, то кластер можно признать неоднородным. Подводя итог сказанному выше, будем считать кластер однородным, если имеет место один из следующих случаев:

- 1) $\min\{k_1, k_2\} \leq 1,1$ и $|k_1 - k_2| \leq 0,5$;
- 2) $\min\{k_1, k_2\} > 1,1$ и $\frac{\max\{k_1, k_2\} - 1}{\min\{k_1, k_2\} - 1} \leq 2$.

Очевидно, что представленные выше константы необходимо определять *эмпирически*. В приведённых условиях они зафиксированы определённым образом лишь для примера.

Дисперсия кластера (дисперсия пространственной изменчивости приращений характеристики). Проанализируем пространственную и временную изменчивость приращений исследуемой характеристики в пределах кластера C за небольшой временной промежуток τ . Для этого найдём разности

$$\Delta F^i(t) = F^i(t) - F^i(t-1), \quad (3)$$

где i – номер узла кластера C . Вычислим среднее значение для каждого временного ряда отдельно

$$\Delta \bar{F}^i(t|\tau) = \frac{\sum_{j=0}^{\tau-2} \Delta F^i(t-j)}{\tau}, \quad (4)$$

и нормируем найденные разности

$$\Delta f^i(t-j) = \Delta F^i(t-j) / \bar{F}^i(t|\tau). \quad (5)$$

При фиксированных значениях t_0 и τ эту процедуру можно распространить на все кластеры сетки UG, сформировав временные ряды и соответствующие карты $\Delta f(t)$ (множества единовременных значений в узлах сетки UG).

Вычислим для каждого момента времени $t \in [t_0 - \tau - 2; t_0]$ оценку дисперсии значений карты в области, ограниченной кластером. Наконец, найдём среднеквадратичную дисперсию:

$$D(C, t, \tau) = \sqrt{\frac{\sum_{j=0}^{\tau-2} D^2(\Delta f(t-j))}{\tau-1}}, \quad (6)$$

которую и будем считать дисперсией пространственной изменчивости приращений характеристики в рамках кластера C на момент t для временного масштаба τ (промежутка времени длины τ).

Градиент кластера. Для каждого узла кластера C в заданный момент времени находим градиент как век-

торную сумму конечных разностей значений характеристики для пары ближайших узлов вдоль горизонтальной и вертикальной осей. На рис. 2 приведён пример расчёта градиента при заданных значениях в соседних узлах: $(3-4; 2-(-1)) = (-1; 3)$. Векторную сумму градиентов для всех узлов одного кластера назовём градиентом этого кластера. Для простоты направление градиента кластера будем приводить к одному из четырёх вариантов, аналогичных расположению ближайших соседей отдельного узла сетки UG. Если предположить, что градиент в выбранном узле на рис. 2 совпадает с градиентом кластера, то результатом применения к нему такой аппроксимации будет направление на север.

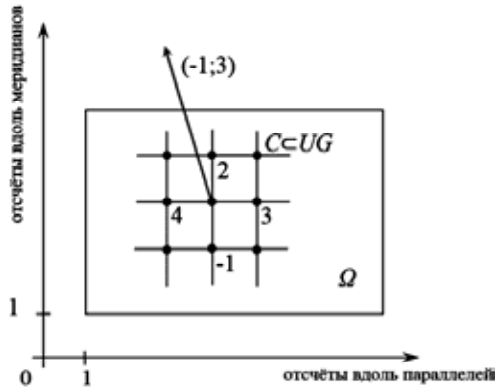


Рис. 2. Градиент характеристики в узле при заданных значениях в 4-х соседних узлах

Эту характеристику можно интерпретировать как критерий качества кластеризации по значению модуля градиента: чем больше его значение (при условии малости дисперсии $D(C, \tau, t)$), тем более однородным является кластер C в момент времени t для временного масштаба τ .

Отметим, что введённые характеристики можно разделить на две группы: меры однородности кластера как меры согласованности в динамике одномерных временных рядов представлены показателями локального (дисперсия и градиент) и глобального (корреляция и отношения сингулярных чисел) временного масштаба.

5. Коррекция результатов начального внутрикластерного прогноза

Метод прогноза значений ряда \tilde{F}_τ^C для кластера C , упомянутый в разделе 3, выполняется исключительно по значениям кластерного ряда F_τ^C , вследствие чего данный этап называется начальным внутрикластерным прогнозом. Методика коррекции его результатов должна использовать информацию о динамике других кластерных рядов, составленных для кластеров узлов сетки UG. Выполнение этого требования позволяет предположить повышение оправдываемости результирующего прогноза.

Принятая в данной работе коррекция результатов начального внутрикластерного прогноза, выполненного с помощью МССА, применяемого к исходному кластерному ряду или его модам МДЭМ, основана на модели одномерного временного ряда TARX (*Threshold AutoRegressive model with eXternal input*). Эта модель представляет собой авторегрессионную модель с дополнительной авторегрессионной частью относительно *внешних*

переменных, причём значения констант в линейной комбинации зависят в том числе от так называемого режима, определяемого моментом времени. Последний факт стал причиной другого названия модели TARX – модели с переключением режимов (*regime switching model*). На данный момент литература по этой модели были представлена исключительно зарубежными работами, например [13].

Общий вид модели TARX в приложении к решаемой задаче имеет следующий вид:

$$e_i(C) = \beta_0^r + \sum_{j=1}^{\tau} \beta_j^r e_{i-j}(C) + \sum_{i=1}^p \sum_{j=1}^{\tau} \beta_{ij}^r e_{i-j}(C_i) \quad (7)$$

где C – *корректируемый кластер*, то есть кластер, прогнозные значения \tilde{F}_τ^C которого необходимо корректировать; $e_i(C)$ – средняя ошибка начального внутрикластерного прогноза \tilde{F}_τ^C для момента t и $\tau=1$ значений ряда кластера C ; C_i – i -ый корректирующий кластер, то есть кластер, ошибка прогноза в котором выступает внешней переменной TARX; p – число корректирующих кластеров; r – номер режима модели TARX.

Отметим, что в модели TARX в данном случае используется один тип регрессора – ошибка начального внутрикластерного прогноза как среднее арифметическое разностей фактического и прогнозного значений, вычисленных для всех одномерных рядов кластера.

Для практического использования модели необходимо провести её структурную и параметрическую идентификацию, а именно выполнить следующие действия:

- 1) задать количество режимов и способы их переключения (смены);
- 2) определить число корректирующих кластеров и провести отбор этих кластеров из множества всех имеющихся;
- 3) найти значения параметров β .

Для пункта (3) из списка выше при условии выполнения двух предыдущих действий есть очевидное решение – метод МНК. Количество режимов зададим равным четырём – по количеству ближайших соседей неограниченных узлов сетки UG. Режим в момент времени t определим как направление градиента корректируемого кластера (см. раздел 4) в момент времени $(t-1)$. Число уравнений, реализующих модель TARX, и число корректирующих кластеров будем выбирать из условия получения квадратной матрицы коэффициентов системы линейных алгебраических уравнений, которую необходимо решать в рамках МНК. Для выбора корректирующих кластеров будем использовать правила, основанные на некоторых свойствах пары (корректируемый кластер C , потенциальный корректирующий кластер K) и значениях кластерных характеристик, введённых в разделе 4.

Случаи запрета использования кластера K :

- 1) непопадание кластера K в заданную окрестность кластера C в рамках сетки UG;
- 2) отставание временных рядов кластера K , то есть наличие пары одномерных рядов (F^i, F^j) , $F^i \in F^K$, $F^j \in F^C$, в которой первый отстаёт от второго в том

смысле, что наибольший коэффициент кросс-корреляции приходится не на нулевой лаг, причём фрагмент ряда F^j использует временные отсчёты, которые не представлены во фрагменте F^i .

Далее введём критерии для отбора p наилучших кластеров среди тех, которые не были опущены по запрещающим правилам. Правила отбора следует рассматривать в соответствии со следующим списком убывания их приоритета:

- 1) задача максимизации: число совпадений направления градиента кластеров K и C для промежутка времени длины τ , непосредственно предшествующего первой прогнозной точке t ;
- 2) задача минимизации: значение дисперсии кластера;
- 3) задача максимизации: значение коэффициента корреляции для пары кластеров (K, C), определяемое как минимум по всем возможным парам одномерных временных рядов из разных кластеров;

Таким образом, выделение корректирующих кластеров есть многокритериальная задача дискретной оптимизации. По каждому из значений критериев можно установить порог, превышение которого является необходимым условием для результата решения этой задачи. Ситуацию, при которой в TARX используется только авторегрессионная часть без внешних переменных, то есть множество корректирующих кластеров является пустым, будем именовать автокоррекцией.

Следует также упомянуть меру однородности кластера по динамике глобального временного масштаба – отношения сингулярных чисел. Наличие свойства однородности, что определяется по численному критерию, сформулированному в разделе 4, можно считать как необходимым условием в определении корректирующего кластера, так и дополнительной информацией при выборе из нескольких альтернатив.

6. Общая схема методики прогноза

Сформируем итоговый вариант внутрикластерного прогноза. Для удобства ссылок введём некоторые обозначения, используя рис. 3. Выполним разметку временной оси, предполагая, что t^* – первая прогнозная точка, T_{opt} – оптимальная длина фрагмента для прогноза, вычисленная по значениям, соответствующим $1, \dots, t^*$. Момент t_0 есть первая прогнозная точка первого начального внутрикластерного прогноза, который повторяется последовательно u раз для формирования данных, на которых настраивается модель TARX. Длина T_{opt} фрагмента временного ряда, который непосредственно участвует в каждом начальном внутрикластерном прогнозе, не изменяется и определяется из условия квазистационарности [5]. Моменту t_0 предшествует фрагмент длины τ , на котором определяется оптимальное количество значимых сингулярных чисел n_{opt} первого прогноза по фрагменту длины T_{opt} , далее для каждого из оставшихся $(u-1)$ прогнозов значение n_{opt} обновляется. В рамках начального внутрикластерного прогноза для обучающего временного ряда [7] используется заданное

значение горизонта τ , а для тестового будем указывать моменты времени, в которые определяются значения, например, «МССА: t_0 » означает вычисления для одного шага по времени (отсчёта t_0), то есть горизонт равен единице. Общая схема методики приведена на рис. 4.

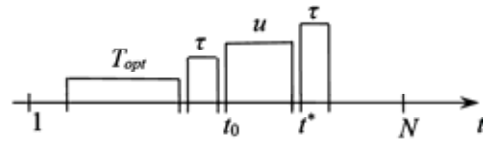


Рис. 3. Разметка временной оси

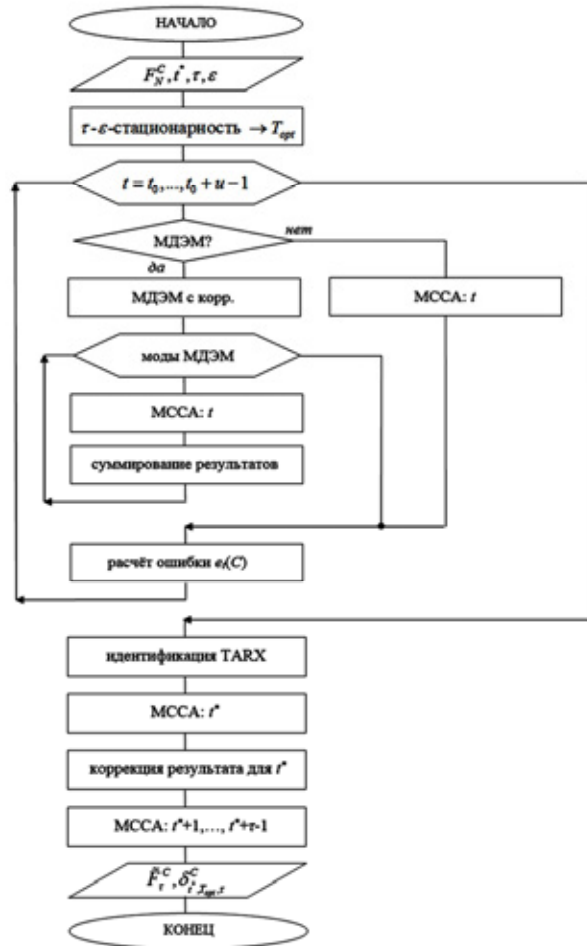


Рис. 4. Блок-схема методики прогноза

7. Вычислительный эксперимент

Апробация отдельных элементов авторской методики прогноза выполнялась для данных альтиметрии (масивов значений аномалий уровня поверхности моря) в работах [6-8]. Полный вариант алгоритма, представленного в предыдущем разделе, был полноценно апробирован на модельных данных. Результаты экспериментов и их анализ представлены в данном разделе.

Рассмотрим сетку UG размером (30×20) и построим для неё временной ряд $F_N^{UG}(t)$ длиной $N = 350$ отсчётов. Априори будем считать, что временные ряды в значительной степени согласованы и образуют несколько кластеров (в соответствии с авторским алгоритмом кластеризации). Каждый из рядов так или иначе строится по следующей модели полезного сигнала:

$$y(t) = P_4(\tilde{t}) + \sin \omega \tilde{t}, \quad (8)$$

для определённости примем, что

$$P_4(\tilde{t}) = 2,7356\tilde{t}^4 - 0,0714\tilde{t}^3 - 10,106\tilde{t}^2 - 4,0196\tilde{t} + 19,0154, \quad (9)$$

$$\tilde{t} = t/90 - 2, \quad t = 0, \dots, 349, \quad \omega = 2,5\pi.$$

К полезному сигналу будет добавляться шум $z(t)$ – случайная составляющая, распределённая по нормальному закону:

$$x(t) = y(t) + z(t), \quad z(t) \sim N(0, \sigma^2), \quad (10)$$

где σ^2 определяется из условия для оценки математического ожидания отношения шума к сигналу:

$$M\left(\frac{z(t)}{y(t)}\right) = 0,05. \quad \text{В данном случае } \sigma = 0,7.$$

Априорные кластеры сформированы так:

- кластер 0 (15×10): модель $x(t)$;
- кластер 1 (15×5): линейная аппроксимация модели $x(t)$: $f(t) = 25 - 0,0667t$;
- кластер 2 (15×5): модель $x(t)$;
- кластер 3 (7×5): модель $x(t-10)$;
- кластер 4 (8×5): модель $y(t)$;
- кластер 5 (7×5): модель $x(t-5)$;
- кластер 6 (8×5): модель $x(t+3)$;
- кластер 7 (15×10): модель $x(t)$, в которой установлен более высокий уровень шума: $\sigma = 0,9$.



Рис. 5. Профили временных рядов, являющихся представителями кластеров: 0 (зашумлённый), 1 (нелинейный), 4 (линейный)

Таким образом, одномерные временные ряды, составляющие кластерный ряд для кластера 1 (как и для кластера 4), в точности совпадают; для остальных кластеров они отличаются лишь значениями шумовой составляющей (рис. 5).

При выполнении кластеризации получено несколько отличающееся разбиение: кластер с наибольшими шу-

мовыми эффектами разделился на пять независимых кластеров (№№ 7-11, рис. 6). Эти результаты не противоречат закономерностям в модельной генерации рядов, поэтому именно это апостериорное множество кластеров далее используется для вычисления прогноза. Изображение сгенерировано созданной автором утилитой кластеризации (номера добавлены в результате постобработки вручную); прямоугольник наименьшей площади соответствует одному узлу.

Прогноз выполнялся независимо для целевого временного интервала I , содержащего 20 последовательных временных отсчётов (с 220 по 239, рис. 5) на горизонт, равный 5. Для данного примера оценкой математического ожидания оптимальной длины фрагмента временного ряда для прогноза с ошибкой не более 10% по условию квазистационарности из монографии [5] стало значение, близкое к 20 ($T = T_{opt} = 20$). Для экспериментального сравнения результатов прогноза взято ещё одно значение: $T = 30$.

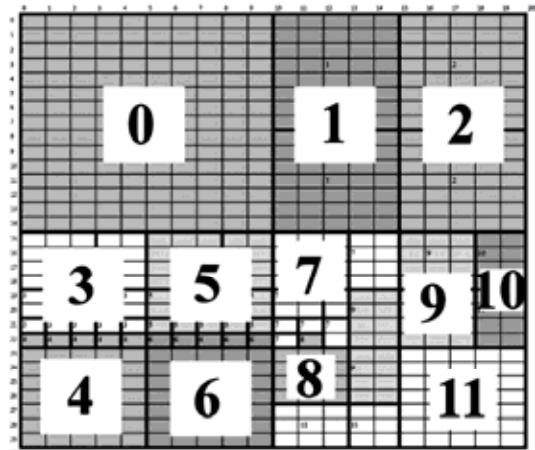


Рис. 6. Результаты кластеризации временных рядов и узлов сетки UG

Была проведена серия опытов в предположении, что потенциально возможным корректирующим кластером является только кластер 4, как содержащий только идентичные ряды с полезным сигналом (без шума). Результаты показаны на рис. 7 и рис. 8. Изображения сгруппированы парами и строятся следующим образом: изображение слева отражает результаты коррекции, при которой используются сторонние (корректирующие) кластеры, а справа – коррекции по ошибкам прогноза в самом кластере (автокоррекция). На последующих рисунках тонкая линия показывает ошибку в процентах начального внут-

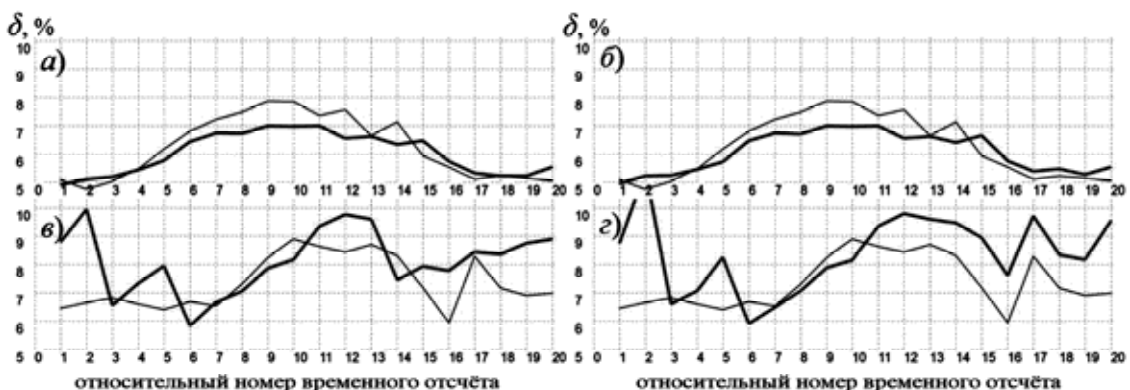


Рис. 7. Средняя ошибка прогноза при $T = 20$ (а), б) и $T = 30$ (в), г), автокоррекция – б) и г)

рикластерного прогноза, утолщённая – ошибку в процентах прогноза после коррекции. Для каждого кластера ошибки вычисляются по формуле (2), итоговое значение для графика выбирается из анализа всех кластеров: либо как средняя ошибка, либо как наибольшая ошибка.

Далее были введены такие ограничения на кластерные характеристики, которые сделали возможным использование большего количества кластеров для коррекции (но не более трёх, не считая сам корректируемый кластер). Результаты второй серии экспериментов приведены на рис. 9 и рис. 10.

Из анализа полученных результатов (рис. 7-10) можно сделать вывод, что начальный внутрикластерный прогноз даёт наибольшую ошибку в центральной части целевого интервала I – эта область соответствует существенному увеличению по модулю величины производной полезного сигнала (8), что видно на рис. 4.

Коррекция прогноза показала наилучшие результаты

при длине фрагмента в 20 отсчётов, что соответствует оценке, полученной из условия квази-стационарности. Стоит отметить, что при использовании более одного корректирующего кластера (рис. 9 а, 10 а), и средняя, и наибольшая ошибки на границах целевого интервала I уменьшаются по сравнению с результатами коррекции с помощью одного 4-го кластера (рис. 7 а, 8 а), а также меньше значений при автокоррекции (рис. 7 б, 8 б). Примечательно, что временные отсчёты, соответствующие границам целевого интервала I , характеризуются инерционностью – смены монотонности профиля временного ряда полезного сигнала (8) не происходит. На рис. 10 а наихудший результат имеет место для отсчёта 16 (отсчёт 235 на рис. 5), что соответствует точке перегиба для полезного сигнала (8). При $T = 30$ начальный внутрикластерный прогноз корректируется плохо во всех случаях, а значения ошибки прогноза без коррекции больше на 2-4% аналогичных значений при $T = 20$.

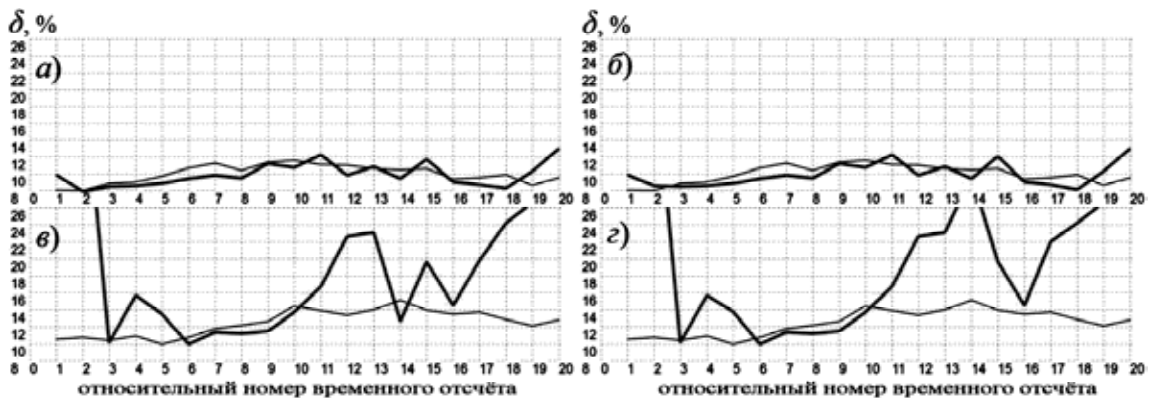


Рис. 8. Наибольшая ошибка прогноза при $T = 20$ (а, б)) и $T = 30$ (в, г)), автокоррекция – б) и г)

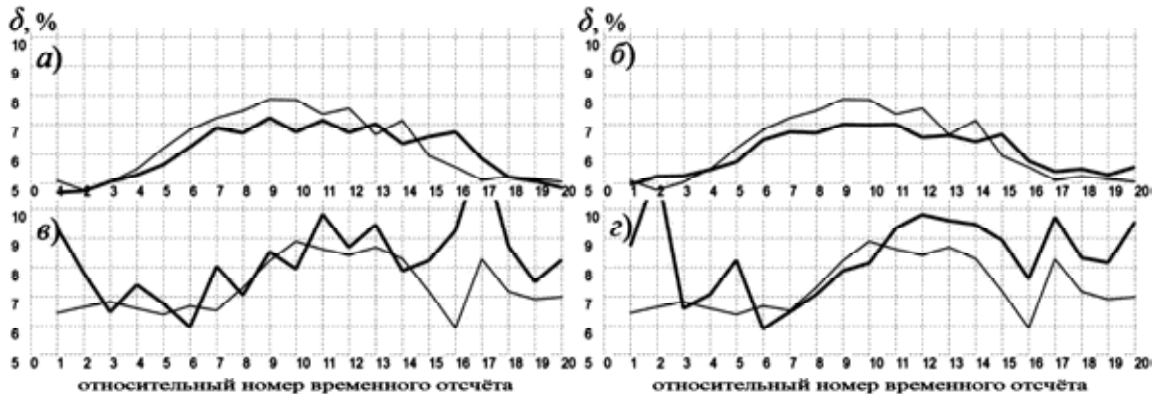


Рис. 9. Средняя ошибка прогноза при $T = 20$ (а, б)) и $T = 30$ (в, г)), автокоррекция – б) и г)

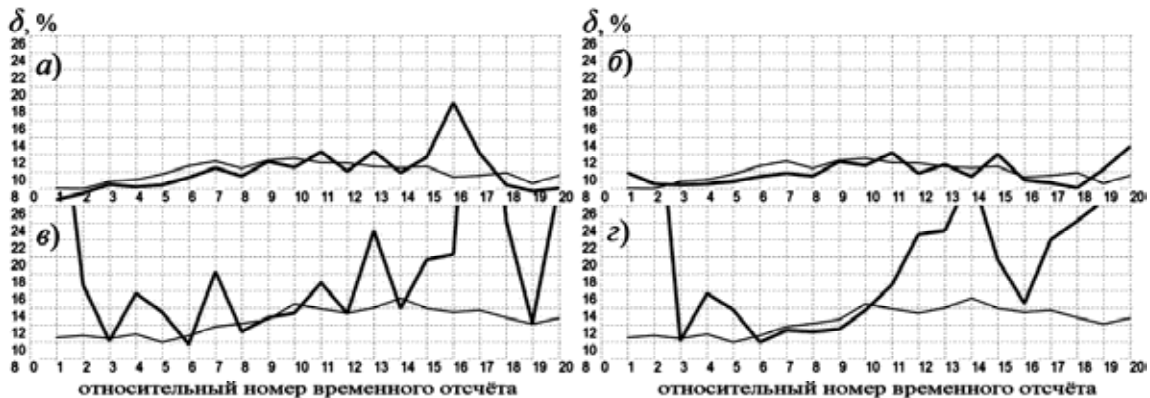


Рис. 10. Наибольшая ошибка прогноза при $T = 20$ (а, б)) и $T = 30$ (в, г)), автокоррекция – б) и г)



Таким образом, коррекция с участием нескольких корректирующих кластеров наиболее эффективна в случае продолжения монотонного участка профиля временного ряда и сохранения знака второй производной полезного сигнала (8). В окрестности точки перегиба начальный внутрикластерный прогноз оказался оптимальным. Наихудшие результаты в случае $T = 30$ как раз приходится на случай коррекции с участием более одного корректирующего кластера (рис 9 в, 10 в).

Заключение

Изложенная в статье методика прогноза является первым шагом авторского коллектива к построению оригинальной технологии краткосрочного прогноза временной изменчивости характеристик естественных процессов, которая удовлетворяла бы требованиям точности, надёжности, адаптивности к различным предметным областям исследований. Важно отметить, что в настоящей статье предложена как общая концепция методики, так и детализация её некоторых шагов. В настоящее время продолжается тестирование методики на модельных данных и начата её апробация в рамках решения задачи прогноза динамики уровня поверхности океана в акваториях Арктики с использованием реальных данных. Эти данные уже участвовали в расчётах по отдельным элементам методики. Стремление к высокой степени автоматизации обработки временных рядов с целью прогноза не исключает существенной роли экспертов в процессе использования данной методики. Предполагается, что её конечная версия будет содержать ряд параметров, для которых потребуется априорное задание рекомендуемых значений, а также предоставление возможности задания других конкретных значений в результате ручной настройки. Таким образом, методика будет реализована в виде интерактивной (на каждом шаге обработки данных) компьютерной технологии, применимой для краткосрочного прогнозирования пространственно-временной изменчивости различных физических характеристик при исследовании природных динамических процессов.

Литература

1. Ашик И.М. Численный гидродинамический метод прогноза колебаний уровня в юго-восточной части Баренцева и юго-западной части Карского морей [Электронный ресурс] / Режим доступа: <http://method.hydro-met.ru/publ/sb/sb31/sb31.html>, свободный. – 2005.
2. Вербицкая О.Н. Гидродинамический метод прогноза синоптических колебаний уровня и течений Каспийского моря: дис. канд. физ.-мат. наук [Электронный ресурс] // МГОУНБ. Диссертации. – 2009. – 175 с.
3. Попов С.К. Прилив в оперативной модели краткосрочного прогноза скоростей течений и уровня моря в Баренцевом и Белом морях / С.К. Попов [и др.] // Метеорология и гидрология. – 2013. – №6. – С. 68-82.
4. Чучуева И.А. Модель прогнозирования временных рядов по выборке максимального подобия: дис. канд. техн. наук [Электронный ресурс] // МГОУНБ. Диссертации. – 2012. – 153 с.
5. Орлов Ю.Н., Осминин К.П. Нестационарные вре-

менные ряды: Методы прогнозирования с примерами анализа финансовых и сырьевых рынков. – М.: Либроком, 2011. – 384 с.

6. Запорожцев И.Ф. Кластеризация узлов планарной регулярной сетки в задаче прогноза временной изменчивости пространственного распределения физических характеристик [Электронный ресурс] / И.Ф. Запорожцев, А.-В.И. Серeda // Наука и образование – 2014: материалы междунар. науч.-техн. конф., Мурманск, 24–28 марта 2014 г. / Мурман. гос. техн. ун-т. – Мурманск : МГТУ, 2014. – С. 382-385.

7. Zaporozhtsev I.F. Analysis and Forecast of Sea Level Anomalies Spatio-Temporal Variability in the Barents Sea / I.F. Zaporozhtsev, A.-V.I. Sereda // Computer Modeling and Simulation : труды междунар. науч.-техн. конф., Санкт-Петербург, 2-4 июля 2014 года. – СПб.: Изд-во Политехн. ун-та, 2014. – С. 81-85.

8. Запорожцев И.Ф. Декомпозиция на эмпирические моды в задаче краткосрочного прогнозирования многомерных временных рядов геофизической природы / И.Ф. Запорожцев, А.-В.И. Серeda // Цифровая обработка сигналов. – 2014. – №2. – С. 34-40.

9. Голяндина Н.Э. Метод «Гусеница»-SSA: прогноз временных рядов: Учеб. пособие / Н.Э. Голяндина. – СПб.: С.-Петербургский государственный университет, 2004. – 52 с.

10. Степанов Д.А. Варианты метода «Гусеница»-SSA для прогноза многомерных временных рядов / Д.А. Степанов, Н.Э. Голяндина // Труды IV Международной конференции «Идентификация систем и задачи управления» SICPRO'05. – Москва, 2005.

11. Huang N.E. «The Empirical Mode Decomposition and the Hilbert Spectrum for Nonlinear and Nonstationary Time Series Analysis» [Электронный ресурс] / Режим доступа: http://keck.ucsf.edu/~schenk/Huang_etal98.pdf, свободный. – 1998.

12. Fleureau J. et al. Multivariate Empirical Mode Decomposition and Application to Multichannel Filtering // Signal Processing, vol. 91, pp. 2783-2792, 2011.

13. Tastu J. et al. Spatio-Temporal Analysis and Modeling of Short-Term Wind Power Forecast Errors // Wind Energy, vol. 14, pp. 43-60, 2011.

NUMERICAL METHOD FOR SHORT-TERM FORECASTING OF MULTIVARIATE TIME SERIES EVALUATED IN NODES OF THE PLANAR UNIFORM GRID

Zaporozhtsev I.F., Sereda A.-V.I.

Short-term temporal variability forecasting methodology for spatially distributed characteristics of natural processes based on multivariate time series analysis is presented in the article. Both its general concept and some stages algorithmic specification are designed. Methodology version introduced currently is formed completely and is implemented as a computer technology. The results of numerical experiments for model data and corresponding analysis are given.

УДК 629.705:681.3.06

МУЛЬТИПРОЦЕССОРНАЯ РЕАЛИЗАЦИЯ АЛГОРИТМОВ КОРРЕЛЯЦИОННО-ЭКСТРЕМАЛЬНОЙ ОБРАБОТКИ СИГНАЛОВ³

Плужников А.Д., доктор тех. наук, профессор кафедры информационных радиосистем Нижегородского государственного технического университета им. Р.Е. Алексеева, e-mail: pluzhnikov@nntu.nnov.ru;

Потапов Н.Н., научный сотрудник кафедры информационных радиосистем Нижегородского государственного технического университета им. Р.Е. Алексеева, e-mail: nick_potapov@nntu.nnov.ru.

Ключевые слова: корреляционно-экстремальная обработка, навигация, алгоритм, сигнальный процессор, мультипроцессорные системы.

Введение

Для решения задач навигации летательных аппаратов в настоящее время широко применяются автономные системы определения собственных координат [1]. Наиболее точное определение собственных координат обеспечивается обзорно-сравнительными радионавигационными системами, в частности, теми, которые строятся на основе использования двумерных карт местности. В таких системах часто реализуются алгоритмы корреляционно-экстремальной обработки информации о разнообразных физических полях. Цифровые данные об этих полях, представленные в форме двумерного массива, называют двумерными картами местности. В дальнейшем для определенности рассматривается тот характерный случай, когда каждый элемент двумерного массива представляет собой значение высоты рельефа в некоторой точке с заданной парой ее декартовых координат на местности. Корреляционно-экстремальная обработка должна представлять собой сравнение двух карт. Одна из них является эталонной картой. Она составляется заранее в соответствии с полетным заданием и заносится в память системы. Причем определенному значению времени полета должен соответствовать свой вариант эталонной карты. Другую карту называют текущей или измеренной картой местности. Она создается в процессе полета путем обработки данных, получаемых разнообразными измерительными средствами. Так, для получения данных о высотах рельефа (о подстилающей поверхности) могут использоваться, например, радиолокационные средства. При этом данные о рельефе вычисляются в процессе решения геометрических задач на основе использования результатов радиолокационных измерений – измерений наклонных дальностей до участков подстилающей поверхности.

Рассмотрены корреляционные и разностные алгоритмы корреляционно-экстремальной обработки навигационной информации. Определены требуемое время вычислений и надежность оценки координат. Даны рекомендации по выбору алгоритма для решения задач навигации летательного аппарата с многолучевой радиолокационной системой определения высот рельефа местности. Оценены возможности практической реализации такой системы на базе интегрированного модуля с высокопроизводительными сигнальными процессорами.

Сравнение двух упомянутых карт осуществляется посредством вычисления некоторого показателя их близости. В результате должна быть определена зависимость показателя близости от взаимного смещения карт по двум координатам. Тогда координаты глобального экстремума вычисленной зависимости будут определять отклонение реального положения летательного аппарата от расчетного.

Постановка задачи

В качестве основной решаемой задачи нами ставится задача поиска надежного (в смысле обеспечения достаточно высокой надежности определения координат) и практически реализуемого алгоритма корреляционно-экстремальной обработки навигационной информации, извлекаемой бортовым радиолокатором с многолучевой антенной системой при таком обзоре подстилающей поверхности, который осуществляется за счет движения летательного аппарата. Такая задача распадается на следующие две частные задачи.

1. Поиск надежного алгоритма корреляционно-экстремальной обработки с минимальной вычислительной сложностью.

2. Оценка возможностей использования современных отечественных мультипроцессорных аппаратных средств для решения задач корреляционно-экстремальной обработки навигационных данных.

Решение первой частной задачи позволит минимизировать требования к аппаратным средствам, реализующим обработку навигационных данных, сокращая, в частности, массо-габаритные показатели, что весьма важно для применения в бортовой аппаратуре беспилотных летательных аппаратов. Упомянутую вычислительную сложность обычно характеризуют количеством и характером необходимых арифметических операций, что определяет время вычислений. Поэтому необходимое время вычислений будем использовать в качестве

³ Работа выполнена при поддержке Министерства образования и науки РФ в рамках договора № 02.G25.31.0061 от 12 февраля 2013 года (в соответствии с Постановлением Правительства Российской Федерации от 9 апреля 2010 г. № 218).



количественной характеристики вычислительной сложности. Отметим также, что минимизация вычислительной сложности делает разумным использование целочисленной арифметики (с фиксированной запятой).

Современные отечественные мультипроцессорные аппаратные средства, упомянутые при формулировке второй частной задачи, считаем целесообразным конкретизировать так, чтобы проанализировать возможности разрабатываемых с нашим участием интегрированных модулей цифровой обработки, обладающих относительной универсальностью.

Наряду со взаимосвязанностью (согласно постановке основной задачи) двух сформулированных частных задач, на наш взгляд, целесообразно отметить и то, что каждая из них имеет также самостоятельное значение.

Математическое описание алгоритмов корреляционно-экстремальной обработки

При вышеперечисленных условиях алгоритмы корреляционно-экстремальной обработки различаются выбором показателя близости. Ниже осуществляется сравнительная оценка следующих алгоритмов [2, 3]:

- классический корреляционный алгоритм (ККА);
- алгоритм минимума абсолютной разности (МАР);
- алгоритм минимума абсолютной разности с вычтенным средним (МАВ);
- алгоритм минимума квадрата разности с вычтенным средним (МКВ).

Последние три алгоритма объединяют общим названием «разностные алгоритмы».

Для формализации перечисленных алгоритмов обозначим через

$$\mathbf{H}_R = \begin{pmatrix} h_R(0, 0) & h_R(0, 1) & \dots & h_R(0, N_{yR} - 1) \\ h_R(1, 0) & h_R(1, 1) & \dots & h_R(1, N_{yR} - 1) \\ \vdots & \vdots & \ddots & \vdots \\ h_R(n_x, 0) & h_R(n_x, 1) & \dots & h_R(n_x, N_{yR} - 1) \\ \vdots & \vdots & \ddots & \vdots \\ h_R(N_{xR} - 1, 0) & h_R(N_{xR} - 1, 1) & \dots & h_R(N_{xR} - 1, N_{yR} - 1) \end{pmatrix} \quad (1)$$

измеренную карту местности, т.е. двумерный массив высот $h_R(n_x, n_y)$ рельефа, найденный, например, с использованием радиолокационных средств. При этом натуральные числа n_x и n_y являются номерами значений двух координат (например, декартовых координат x и y) на местности и могут быть названы дискретными координатами. Согласно (1) массив \mathbf{H}_R имеет размер $N_{xR} \times N_{yR}$, т.е. $n_x = 0, 1, \dots, N_{xR} - 1$ и $n_y = 0, 1, \dots, N_{yR} - 1$.

Аналогичным массивом

$$\mathbf{H}_0 = \begin{pmatrix} h_0(0, 0) & h_0(0, 1) & \dots & h_0(0, N_{y0} - 1) \\ h_0(1, 0) & h_0(1, 1) & \dots & h_0(1, N_{y0} - 1) \\ \vdots & \vdots & \ddots & \vdots \\ h_0(n_x, 0) & h_0(n_x, 1) & \dots & h_0(n_x, N_{y0} - 1) \\ \vdots & \vdots & \ddots & \vdots \\ h_0(N_{x0} - 1, 0) & h_0(N_{x0} - 1, 1) & \dots & h_0(N_{x0} - 1, N_{y0} - 1) \end{pmatrix} \quad (2)$$

размера $N_{x0} \times N_{y0}$ представим эталонную карту местности.

На практике размеры измеренной карты существенно меньше размеров эталонной карты. При этом взаимное смещение карт в процессе корреляционно-экстремальной обработки осуществляется так, что измеренная карта не выходит за пределы эталонной карты. Кроме того, для определенности и упрощения анализа сделаем следующее допущение: в том случае, когда отсутствует отклонение реального положения летательного аппарата от расчетного, элемент $h_R(0, 0)$ измеренной карты и элемент $h_0(0, 0)$ эталонной карты являются значениями высоты рельефа одной и той же точки реальной местности. Последнее допущение не всегда обеспечивается на практике, однако не способно привести к ошибочным результатам решения задачи сравнительной оценки рассматриваемых алгоритмов.

Определяя показатель близости $C(n_{sx}, n_{sy})$ для случая ККА в соответствии с [4], при сделанных допущениях сможем записать:

$$C(n_{sx}, n_{sy}) = \frac{1}{N_{xR} N_{yR}} \sum_{n_x=0}^{N_{xR}-1} \sum_{n_y=0}^{N_{yR}-1} h_R(n_x, n_y) h_0(n_x + n_{sx}, n_y + n_{sy}), \quad (3)$$

где n_{sx} определяет взаимное смещение измеренной и эталонной карт по координате x , а n_{sy} – по координате y , причем n_{sx} и n_{sy} являются натуральными числами и удовлетворяют неравенствам $0 \leq n_{sx} \leq N_{x0} - N_{xR}$, $0 \leq n_{sy} \leq N_{y0} - N_{yR}$.

Пара значений $n_{sx} = n_{sxe}$ и $n_{sy} = n_{sye}$, соответствующих глобальному максимуму функции $C(n_{sx}, n_{sy})$ дискретных аргументов n_{sx} и n_{sy} , будет определять отклонение реального положения летательного аппарата от расчетного. При этом точку с координатами (n_{sxe}, n_{sye}) часто называют точкой совмещения эталонной и измеренной карт.

Чтобы повысить надежность оценки реального положения требуется уменьшить количество локальных максимумов функции $C(n_{sx}, n_{sy})$ и увеличить главный максимум относительно остальных [5]. Для этого ККА реализуют при определении показателя близости с использованием нормирования правой части формулы (3) и центрирования сомножителей под знаком суммы в указанной формуле [2]:

$$C^2(n_{sx}, n_{sy}) = \left[\sum_{n_x=0}^{N_{xR}-1} \sum_{n_y=0}^{N_{yR}-1} (h_R(n_x, n_y) - M_{HR}) \times (h_0(n_x + n_{sx}, n_y + n_{sy}) - M_{H0}(n_{sx}, n_{sy})) \right]^2 / \left\{ \left[\sum_{n_x=0}^{N_{xR}-1} \sum_{n_y=0}^{N_{yR}-1} (h_R(n_x, n_y) - M_{HR})^2 \right] \times \left[\sum_{n_x=0}^{N_{xR}-1} \sum_{n_y=0}^{N_{yR}-1} (h_0(n_x + n_{sx}, n_y + n_{sy}) - M_{H0}(n_{sx}, n_{sy}))^2 \right] \right\}, \quad (4)$$

где M_{HR} – среднее значение элементов массива (1); $M_{H0}(n_{sx}, n_{sy})$ – среднее значение тех элементов массива

(2), которые используются при вычислениях по формуле (3) при заданных значениях n_{sx} и n_{sy} .

Разностные корреляционные алгоритмы основаны на вычислении разностей элементов эталонной и измеренной карт, а не их произведений, как в формуле (3). Благодаря этому сокращается время вычислений. По быстрдействию разностные алгоритмы эффективнее, поскольку в них отсутствует операция умножения. Для алгоритма МАР показатель близости определяется как частный случай величины

$$C(n_{sx}, n_{sy}) = \sum_{n_x=0}^{N_{xR}-1} \sum_{n_y=0}^{N_{yR}-1} |h_R(n_x, n_y) - h_0(n_x + n_{sx}, n_y + n_{sy})|^p,$$

а именно, как частный случай при $p = 1$. При $p = 2$ реализуется так называемый алгоритм минимума квадрата разности.

Разностные алгоритмы с вычтенным средним предполагают центрирование величин при вычислении их разности:

$$C(n_{sx}, n_{sy}) = \sum_{n_x=0}^{N_{xR}-1} \sum_{n_y=0}^{N_{yR}-1} |h_R(n_x, n_y) - M_{HR} - h_R(n_x + n_{sx}, n_y + n_{sy}) + M_{H0}(n_{sx}, n_{sy})|^p. \quad (5)$$

При $p = 1$ соотношение (5) соответствует алгоритму МАВ, а при $p = 2$ – алгоритму МКВ.

При использовании разностных алгоритмов отклонение реального положения летательного аппарата от расчетного оценивается путем поиска глобального минимума показателя близости. При этом упомянутое центрирование позволяет сократить количество локальных минимумов и уменьшить их глубину [3], благодаря чему повышается надежность оценки.

Сравнение алгоритмов корреляционно-экстремальной обработки

Проведем сравнение алгоритмов в условиях, характерных для того случая, когда измеренная карта является результатом обработки информации, извлекаемой в процессе обзора подстилающей поверхности бортового радиолокатором с многолучевой антенной системой. Лучи в таких системах, как правило, неподвижны относительно летательного аппарата, а обзор подстилающей поверхности реализуется благодаря движению аппарата. Количество лучей при этом сравнительно невелико (1–5), что ограничивает один из размеров реальной измеренной карты, например, $1 \leq N_{yR} \leq 5$ при таком выборе системы декартовых координат x и y , когда направление отсчета координаты x параллельно направлению движения летательного аппарата.

Для сравнения алгоритмов необходимо определить и сопоставить их характеристики. В качестве сопоставляемых характеристик выберем требуемое время вычислений и надежность оценки координат.

Указанная сравнительная оценка проведена методом математического моделирования. При этом эталонная карта задавалась как массив взаимно независимых случайных величин (элементов массива), имеющих некоторое математическое ожидание m_0 и среднеквадратичное

отклонение σ_0 . Плотность вероятности элементов массива задавалась в двух вариантах: как плотность вероятности равномерно распределенной случайной величины и как плотность вероятности нормального распределения. Измеренная карта моделировалась следующим образом. Из эталонной карты, т.е. из массива размером $N_{x0} \times N_{y0}$ выделялся участок \mathbf{H}_{0R} меньшего размера $N_{xR} \times N_{yR}$, образованный элементами $h_0(n_x, n_y)$ эталонной карты с номерами $n_{sx0} \leq n_x \leq n_{sx0} + N_{xR} - 1$ и $n_{sy0} \leq n_y \leq n_{sy0} + N_{yR} - 1$. Причем величины n_{sx0} и n_{sy0} выбирались случайно (равновероятно) из интервалов $0 \leq n_{sx0} \leq N_{x0} - N_{xR}$ и $0 \leq n_{sy0} \leq N_{y0} - N_{yR}$ соответственно. После этого измеренная карта вычислялась как результат суммирования выделенного участка с дополнительно формируемым массивом \mathbf{H}_a размера $N_{xR} \times N_{yR}$, моделирующим ошибки измерений:

$$\mathbf{H}_R = \mathbf{H}_{0R} + \mathbf{H}_a,$$

где массив \mathbf{H}_a задавался как массив взаимно независимых нормально распределенных случайных величин, имеющих математическое ожидание m_a и среднеквадратичное отклонение σ_a .

Величину m_a можно интерпретировать как систематическую ошибку измерений. В случае радиолокационных измерений она может быть довольно значительной. Действительно, радиолокационные средства в рассматриваемых системах обеспечивают измерение лишь разностей высот определенной точки рельефа и летательного аппарата. В то же время высота летательного аппарата (например, относительно уровня моря) может быть не известна с достаточной точностью. Тогда данные эталонной карты (составленной достаточно точно) и измеренной карты будут отличаться на некоторую константу, которую можно назвать систематической ошибкой измерений.

В процессе статистического эксперимента моделировалось Q_0 реализаций эталонной карты \mathbf{H}_0 . Для каждой из них моделировалось Q_a реализаций массива \mathbf{H}_a со взаимно независимыми парами значений параметров n_{sx0} и n_{sy0} . Если моделирование какого-либо из исследуемых алгоритмов при заданной паре реализаций массивов \mathbf{H}_0 и \mathbf{H}_a приводит к результату $n_{sxe} = n_{sx0}$ и $n_{sye} = n_{sy0}$, то это означает правильное решение задачи корреляционно-экстремальной обработки. Обозначив через Q_p количество таких результатов, полученных за время эксперимента, надежность оценки координат сможем приближенно определять по формуле

$$P = Q_p / (Q_0 Q_a). \quad (6)$$

На рис. 1 – 4 представлены результаты расчета величины (6), проведенного при описанном математическом моделировании и следующих значениях параметров: $N_{x0} \times N_{y0} = 400 \times 80$, $N_{xR} \times N_{yR} = 100 \times 3$, $Q_0 = 20$, $Q_a = 10$, $\sigma_0/m_0 = 0,577$ для равномерного распределения и $\sigma_0/m_0 = 1/3$ для нормального распределения. При этом ККА реализован согласно формуле (4). Рис. 1 и 2 со-

ответствуют равномерному распределению элементов массива H_0 . Рис. 3 и 4 соответствуют нормальному распределению этих элементов. Для рис. 1 и 3 $m_a = 0$, а для рисунков 2 и 4 $m_a/m_0 = 1$.

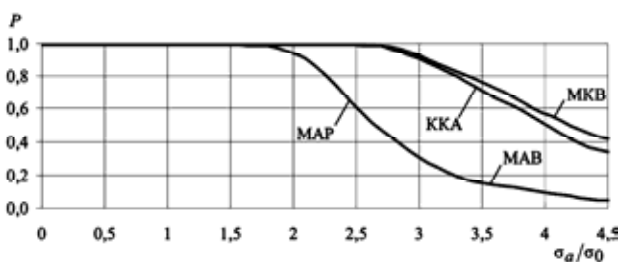


Рис. 1. Зависимость вероятности правильного решения от дисперсии равномерно распределенной ошибки измерений при $m_a = 0$

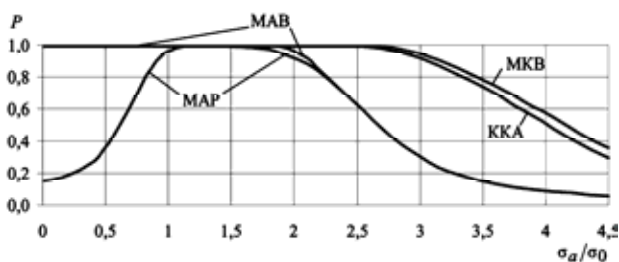


Рис. 2. Зависимость вероятности правильного решения от дисперсии равномерно распределенной ошибки измерений при $m_a/m_0 = 1$

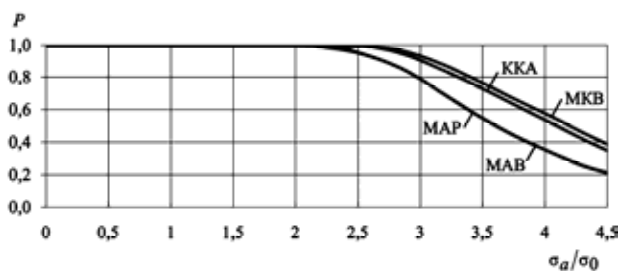


Рис. 3. Зависимость вероятности правильного решения от дисперсии нормально распределенной ошибки измерений при $m_a = 0$

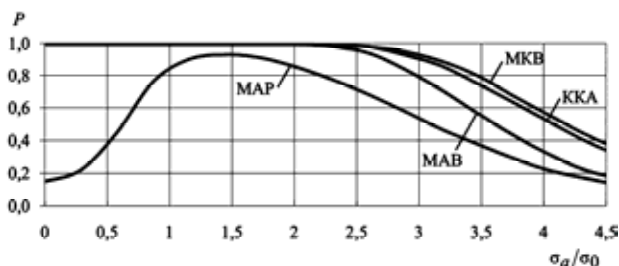


Рис. 4. Зависимость вероятности правильного решения от дисперсии нормально распределенной ошибки измерений при $m_a/m_0 = 1$

Из анализа численных результатов (рис. 1 – 4) следует, что алгоритм MAP во всех случаях дает наихудшую надежность оценки координат. Это можно объяснить влиянием упоминавшейся систематической ошибки измерений. Использование операций центрирования в процессе вычислений по остальным трем из сравниваемых алгоритмов позволяет скомпенсировать систематическую ошибку и повысить надежность измерений.

Показатели надежности, достигаемые при использовании ККА, не являются самыми высокими. Последнее связано с накоплением ошибок при достаточно сложных вычислениях с использованием целочисленной арифметики (с фиксированной запятой), которой отдается предпочтение с целью понижения требований к аппаратно-программным средствам. Действительно, в случае вычислений с плавающей запятой надежность оценки координат при использовании ККА оказывается примерно на 5% выше и является наилучшей. Алгоритм МКВ, согласно представленным численным результатам при целочисленной арифметике, является предпочтительным с точки зрения обеспечиваемой надежности оценки координат.

Табл. 1 содержит результаты оценки времени вычислений, осуществляемых по различным алгоритмам. Согласно этой таблице за единицу принято время вычислений, осуществляемых в соответствии с формулой (4) для ККА. Заметим, что алгоритмы ККА, МAB и МКВ включают в себя вычисления значений $M_{H0}(n_{sx}, n_{sy})$. Относительное время этих вычислений составляет 0,279. Причем для заданной (составленной при подготовке полетного задания) эталонной карты и всевозможных комбинаций значений n_{sx}, n_{sy} величина $M_{H0}(n_{sx}, n_{sy})$ может быть рассчитана заранее. Тогда время вычислений, осуществляемых по различным алгоритмам в полете, сократится и будет соответствовать данным правого столбца табл. 1. При этом за единицу принято новое (сокращенное) время вычислений, осуществляемых в соответствии с формулой (4) для ККА.

Таблица 1. Относительное время вычислений

Алгоритм	Относительное время вычислений	
	с учетом вычисления величины $M_{H0}(n_{sx}, n_{sy})$	без учета вычисления величины $M_{H0}(n_{sx}, n_{sy})$
ККА, формула (4)	1	1
MAP	0,499	0,638
MAB	0,787	0,727
МКВ	0,806	0,752

Из табл. 1 следует, что ККА имеет наихудшие показатели по времени выполнения вычислений, что объясняется большим количеством операций умножения. Алгоритмы MAB и МКВ являются предпочтительными и с точки зрения необходимого времени вычислений. Таким образом, из сопоставляемых алгоритмов наилучшим соотношением надежности оценки координат и вычислительной сложности характеризуется алгоритм МКВ.

Реализация корреляционно-экстремальной обработки на базе интегрированного многопроцессорного модуля

В ряде ситуаций реализация рассмотренных алгоритмов корреляционно-экстремальной обработки требует выполнения сравнительно большого количества арифметических операций за ограниченное время. Поэтому анализ возможностей аппаратно-программной реализации таких алгоритмов представляет особый интерес. В частности, целесообразно проанализировать

упомянутые возможности для случая использования интегрированных модулей цифровой обработки [6], разрабатываемых в настоящее время согласно Постановлению Правительства Российской Федерации от 9 апреля 2010 г. № 218 в качестве основы построения мультипроцессорных систем, обладающих относительной универсальностью.

Упомянутый интегрированный модуль разрабатывается для использования на неподвижных и мобильных объектах. Он содержит высокопроизводительные микропроцессоры, способные взаимодействовать с памятью большого объема. В качестве высокопроизводительных микропроцессоров в модуле применяются отечественные аналоги микропроцессора ADSP-TS201 (TigerSHARC), тоже находящиеся в стадии разработки.

Структурная схема системы корреляционно-экстремальной обработки на основе интегрированного модуля приведена на рис. 5.



Рис. 5. Структурная схема системы корреляционно-экстремальной обработки

Система (рис. 5) содержит интегрированный модуль цифровой обработки сигналов и промышленную ЭВМ, к которой через шину PCI Express подключен радиолокационный высотомер. Радиолокационный высотомер представляет собой набор аппаратно-программных средств, обеспечивающих получение данных о высотах рельефа (радиолокационные измерения), их представление в необходимом формате, а также средства реализации обмена данными с промышленной ЭВМ и интегрированным модулем.

Для осуществления связи радиолокационного высотомера с интегрированным модулем целесообразно использовать высокоскоростной коммуникационный порт, наличие которого предусмотрено в разрабатываемом интегрированном модуле.

В интегрированном модуле должно выполняться сравнение эталонной карты и измеренной в соответствии с алгоритмом МКВ, выбираемым согласно результатам проведенного сравнения алгоритмов (см. выше).

Промышленная ЭВМ используется для управления средствами радиолокационного высотомера и интегрированным модулем. Так, ЭВМ передает в интегрированный модуль данные эталонной карты H_0 (как уже упоминалось, они загружаются в ЭВМ при подготовке полетного задания), а также принимает результаты вычислений от интегрированного модуля. Кроме того, промышленная ЭВМ решает задачи анализа и статистической обработки результатов вычислений (с оценкой точности и надежности указанных результатов), а также их протоколирования.

Один из вариантов структурной схемы интегриро-

ванного модуля, а также его связи, соответствующие рис. 5, показаны на рис. 6. В других вариантах количество процессорных кластеров в модуле может быть иным (до 5). При этом количество процессоров (ЦСП) в кластере может принимать значения от 2 до 4.

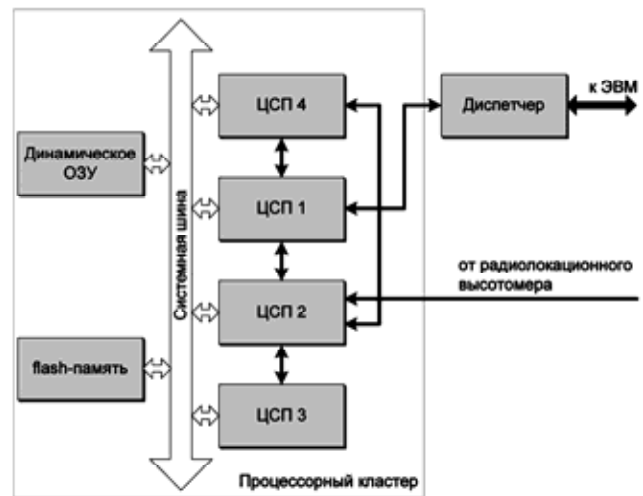


Рис. 6. Структурная схема интегрированного модуля

На рис. 6 интегрированный модуль имеет в своем составе 1 кластер, содержащий 4 высокопроизводительных процессора цифровой обработки сигналов – цифровых сигнальных процессора (ЦСП1 – ЦСП4), микросхему программируемой логики (диспетчер), оперативное запоминающее устройство, а также постоянное запоминающее устройство. Анализ возможных параметров тактовой синхронизации, определяющих быстродействие компонентов интегрированного модуля, а также пропускную способность линий обмена данными между компонентами, приведен в [6]. Диспетчер на рис. 6 является так называемой конечной точкой шины PCI Express и транслирует команды от ЭВМ к процессорам, а также ответы на команды, передаваемые в обратном направлении. Обмен данными между диспетчером и процессором ЦСП1, а также межпроцессорный обмен организуется при использовании высокоскоростных коммуникационных портов. Доступ процессоров к постоянному и оперативному запоминающим устройствам (к flash-памяти и динамическому ОЗУ на рис. 6) осуществляется за счет объединения упомянутых компонентов системной шиной.

Задачи распределяются между процессорами следующим образом. Процессор ЦСП 2 принимает данные от радиолокационного высотомера по коммуникационному порту и формирует из них измеренную карту. Эти данные передаются пакетами, каждый из которых соответствует одному столбцу в матричном представлении измеренной карты H_R . Обработку измеренной карты после ее формирования процессором ЦСП 2 выполняют три из четырех процессоров, а именно ЦСП 2 – 4. Принятые процессором ЦСП 2 пакеты транслируются без изменений процессорам ЦСП 3 и 4. Каждый из трех процессоров выполняет вычисления по алгоритму МАВ для набора возможных гипотез о точке совмещения эталонной и измеренной карты, формируя парциальный (соответствующий процессору) показатель близости. Парциальные показатели близости передаются процессору

ЦСП 1, который составляет объединенный показатель близости эталонной и измеренной карт, а также выполняет поиск экстремума этого показателя для определения точки совмещения карт.

Результаты оценки положения экстремума передаются диспетчеру, откуда могут быть считаны управляющей ЭВМ. Таким образом, управляющая ЭВМ имеет возможность отслеживать текущую оценку местоположения летательного аппарата в реальном времени.

Для памяти, которая содержится в интегрированном модуле и системе в целом, функции распределяются следующим образом. Программное обеспечение процессоров хранится во flash-памяти и загружается во внутреннюю память процессоров при включении питания модуля. При подготовке полетного задания эталонная карта местности после ее записи в память ЭВМ сразу же передается из ЭВМ в динамическое ОЗУ. Для получения значений элементов эталонной карты процессоры ЦСП 2 – 4 обращаются к динамическому ОЗУ по системной шине. Для хранения промежуточных результатов вычислений сигнальные процессоры используют собственные внутренние оперативные запоминающие устройства.

Оценка возможностей интегрированного модуля по решению задач корреляционно-экстремальной обработки

Для оценки возможностей решения задач корреляционно-экстремальной обработки наиболее важными являются такие параметры интегрированного модуля как производительность, объем оперативной памяти и пропускная способность линий передачи данных. Значения упомянутых параметров приведены в табл. 2.

Таблица 2. Параметры интегрированного модуля

Параметр	Значение
Теоретическая пиковая производительность одного процессора (P_t), млн. операций в секунду	600
Объем динамического ОЗУ (V_{RAM}), МБайт	256
Скорость обмена данными с внешними устройствами при помощи высокоскоростного коммуникационного порта (R_{text}), Мбит/с	600
Скорость обмена данными между процессорами (R_{int}), Мбит/с	1280

Определим максимальное количество столбцов и строк матрицы измеренной карты, т.е. такое их количество, при котором возможно выполнение расчета координат летательного аппарата в реальном масштабе времени. Поскольку измеренная карта формируется по столбцам, то количество строк измеренной карты определяется количеством операций для обработки одного элемента и производительностью процессоров. Пусть формирование одного столбца измеренной карты осуществляется за время t_m . Как показывают оценки, обработка одного элемента измеренной карты в соответствии с алгоритмом минимума абсолютной разности с вычтенным средним требует $O_{elm} \approx 16500$ целочисленных операций при условии, что $N_{x0} - N_{xR} = N_{y0} - N_{yR} = \Delta_{0R} = 40$. Максимальное количество N_{yRmax} строк из-

меренной карты определяется как $N_{yRmax} = 3P_t t_m / O_{elm}$. Здесь коэффициент 3 соответствует количеству процессоров в модуле, выполняющих вычисления (см. выше). Потребуем выполнения равенства $t_m = 1$ мс, чего, согласно нашим расчетам, вполне достаточно даже для высокоскоростных летательных аппаратов – объектов навигации. Тогда получим $N_{yRmax} = 109$.

Максимальное количество столбцов измеренной карты можно вычислить, зная количество столбцов соответствующей эталонной карты. В свою очередь максимальный размер эталонной карты определяется доступным объемом динамического ОЗУ. Элементы матрицы эталонной карты H_0 представляются 32-разрядными числами. Следовательно, объем памяти для хранения каждого такого элемента равен $V_{el0} = 4$ байт. Пусть вся информационная емкость динамического ОЗУ (см. табл. 2) использована для хранения эталонной карты. При этом $N_{x0} N_{y0} = V_{RAM} / V_{el0}$. Тогда $(N_{xR} + \Delta_{0R})(N_{yR} + \Delta_{0R}) = V_{RAM} / V_{el0}$, а значит, $N_{xR} = V_{RAM} / [(N_{yR} + \Delta_{0R})V_{el0}] - \Delta_{0R}$. Подставляя в последнее выражение ранее найденное значение $N_{yR} = N_{yRmax}$, найдем максимальное количество столбцов измеренной карты $N_{xRmax} = V_{RAM} / [(N_{yRmax} + \Delta_{0R})V_{el0}] - \Delta_{0R} = 453615$.

Таким образом, рассматриваемый модуль способен выполнять необходимую корреляционно-экстремальную обработку в реальном времени, если размер измеренной карты не превышает 109×453615 . Указанное условие в реальных корреляционно-экстремальных системах навигации выполняется с запасом.

Определим теперь требования к скорости передачи данных через интерфейс связи модуля с радиовысотометром. За время t_m обработки одного столбца измеренной карты передается объем данных, равный $N_{yRmax} V_{elR}$, где $V_{elR} = V_{el0}$ – объем памяти для хранения одного элемента измеренной карты. Тогда средняя скорость передачи данных через интерфейс связи, требуемая для выполнения обработки информации модулем, определяется как $R_{IR} = N_{yRmax} V_{elR} / t_m$, и получается равной $R_{IR} \approx 0,45$ Мбит/с. Найденное значение требуемой скорости передачи данных много меньше скоростей, обеспечиваемых линиями передачи интегрированного модуля (см. табл. 2).

Таким образом, возможности интегрированного модуля перекрывают требования, предъявляемые к вычислительному устройству предлагаемой корреляционно-экстремальной обработки, и обеспечивают запас для модернизации алгоритмов обработки при их усложнении с целью повышения показателей качества (точности и надежности), а также для использования упомянутых возможностей при решении задач формирования данных измеренной карты: обработки отсчетов комплексной огибающей радиолокационных сигналов, определения высот в узлах измеренной карты. Сформулированный результат с одной стороны положительно характеризуют

разработку рассматриваемых модулей как относительно универсальной основы построения мультипроцессорных систем, а с другой стороны показывает, что выбранный алгоритм обработки может быть реализован существенно более простыми аппаратными средствами, что обеспечивает возможность значительного ужесточения массо-габаритных и иных (например, по надежности) требований к аппаратуре, весьма важную для реализации многих типов беспилотных летательных аппаратов.

Заключение

Сопоставление различных алгоритмов корреляционно-экстремальной обработки навигационной информации показывает, что алгоритм минимума абсолютной разности с вычтенным средним обеспечивает необходимую надежность оценки координат летательного аппарата, имея при этом значительные преимущества по сравнению с классическим корреляционным алгоритмом в отношении требуемых аппаратных ресурсов и времени вычислений. Указанные преимущества алгоритма делают возможной его аппаратно-программную реализацию на современной элементной базе. В частности, для реализации алгоритма может быть применена отечественная разработка – один из вариантов интегрированного модуля, выполняемого на базе высокопроизводительных отечественных процессоров.

Предлагаемая структура корреляционно-экстремальной обработки информации позволит решить задачу определения поправок к местоположению летательного аппарата при использовании отечественной элементной базы, а именно, простейшей модификации разрабатываемых универсальных мультипроцессорных систем – одного интегрированного модуля с одним кластером, содержащим четыре отечественных сигнальных процессора.

При этом раздельная оценка разрабатываемых аппаратных средств и предлагаемого алгоритма позволяет сделать два вывода:

- разрабатываемые отечественные мультипроцессорные аппаратные средства обеспечивают решение практически важных задач, в том числе задач корреляционно-экстремальной обработки навигационных данных;

- предлагаемый алгоритм корреляционно-экстремальной обработки навигационной информации не только обеспечивает необходимые точность и надежность определения координат, но и реализуется аппаратными средствами со сравнительно невысокими требованиями к ним, что весьма важно для применения в бортовой аппаратуре беспилотных летательных аппаратов при характерных для них жестких массо-габаритных требованиях и требованиях по надежности аппаратуры.

Литература

1. Бакулев П.А., Сосновский А.А. Радиолокационные и радионавигационные системы: Учеб. пособие для вузов.– М.: Радио и связь, 1994.– 296 с.
2. Андреев Г.А., Потапов А.А. Алгоритмы обработки навигационной пространственно-временной информации (часть 1) // Зарубежная радиоэлектроника.– 1989.– № 3.– с. 3–18.
3. Андреев Г.А., Потапов А.А. Алгоритмы обработки на-

вигационной пространственно-временной информации (часть 2) // Зарубежная радиоэлектроника.– 1989.– № 4.– с. 3–21.

4. Бочкарев А.М. Корреляционно-экстремальные системы навигации // Зарубежная радиоэлектроника.– 1981.– № 9.– с. 28–53

5. Куликов Е.И., Трифонов А.П. Оценка параметров сигналов на фоне помех. – М.: Сов. радио, 1978. – 296 с.

6. Кузин А.А., Плужников А.Д. и др. Анализ временных соотношений для сигналов в проектируемых цифровых модулях и оценка работоспособности. // Цифровая обработка сигналов. – 2014.- № 2. – С. 70-77.

MULTIPROCESSOR REALIZATION OF CORRELATIONAL-EXTREMAL SIGNAL PROCESSING

Pluzhnikov A.D., Potapov N.N.,

Correlational and differential algorithms of correlational-extremal processing are discussed. Required calculation time and coordinates estimation reliability are defined. Recommendations concerning choice of algorithm to solve problem of navigation of aircraft with multiray altimeter radar are given. Possibility of hardware realization of such system based on the integrated module which contains high-performance signal processors is estimated.

At present the most accurate standalone solution of aircraft navigation problem is achieved by map-matching radionavigation systems, especially by systems built on the base of two-dimensional topography maps usage. Comparison of prepared reference map to created during flight measured map performed in such systems by calculating so called proximity index. Global extremum of proximity index defines deviation between real aircraft coordinates and expected coordinates.

Algorithms of correlational-extremal processing differ in the choice of proximity index type. Type of proximity index specifies calculation complexity of particular algorithm. For classic correlational algorithm (CCA) proximity index is represented by mutual correlation function of reference and measured maps. For the family of differential algorithms multiplication is replaced by subtraction, allowing to significantly reduce requirements of calculation unit performance. This paper deals with analysis of the following differential algorithms: algorithm of absolute difference minimum (ADM), algorithm of squared difference minimum with average subtracting (SSM), algorithm of absolute difference minimum with average subtracting (ASM).

Measured maps obtained by processing of information extracted during the area relief scanning is commonly used in map-matching navigation systems. Relief scanning performed by onboard multiray radar. Radar rays are stationary relative to aircraft. Relief scanning is realized due to aircraft motion. Different characteristics of algorithms were compared using mathematical modeling and showed ADM algorithm to provide the worst coordinate estimation reliability. Three other algorithms calculations contain operation of centering which allows to compensate systematic error and to increase the reliability of coordinate estimation. CCA provides worst calculation time, which caused by large amount of multiplication operations. SSM algorithm is preferred with respect to higher coordinate estimation reliability and lower calculation time requirement.

Calculation time required to perform calculation according to SSM algorithm allows to realize this algorithm using modern hardware. In particular, recently developed integrated module with high-performance signal processors can be used for algorithm implementation.

БЫСТРОДЕЙСТВУЮЩИЕ АЛГОРИТМЫ ПРЕОБРАЗОВАНИЯ ОРТОГОНАЛЬНЫХ СОСТАВЛЯЮЩИХ СИГНАЛОВ В АМПЛИТУДУ И ФАЗУ

Чекушкин В.В., д.т.н., профессор кафедры САПР Муромского института (филиала) Владимирского государственного университета имени Александра Григорьевича и Николая Григорьевича Столетовых, e-mail: chekvv@gmail.com;

Михеев К.В., аспирант кафедры САПР Муромского института (филиала) Владимирского государственного университета имени Александра Григорьевича и Николая Григорьевича Столетовых, e-mail: kiri-mikheev@yandex.ru.

Ключевые слова: измерение сигналов, преобразование координат, погрешность измерительной системы, полином, специализированный вычислитель.

Введение

При вычислении амплитуд сигналов со случайной начальной фазой по квадратурным составляющим A и B прямая форма реализации преобразования комплексного сигнала из алгебраического представления в экспоненциальное использует выражение $Z = \sqrt{A^2 + B^2}$ [1]. В радиолокационных станциях вычисление амплитуды такого сигнала должно производиться в реальном масштабе времени в каждом текущем элементе разрешения по дальности, например, с линейным размером в 1 м, что соответствует временному интервалу вычислений 6,6 нс. При этом необходимо обеспечить относительную приведенную погрешность результата измерений порядка (0,4...2,5)%. Сравнительный анализ и улучшение методов вычисления стандартных функций, преобразования координат, совершенствования специализированных вычислителей проведено в [2-7]. В тоже время имеются потенциальные возможности упрощения реализации функциональных зависимостей путём предварительного моделирования вычислительного процесса. Обеспечение предельных, оптимальных соотношений показателей точности, быстродействия и программно-аппаратных затрат на основе применения и совершенствования методов компьютерной математики повышает эффективность вычислительных структур систем цифровой обработки сигналов.

Цель данной статьи – совершенствование быстродействующих методов и алгоритмов совместного вычисления функций $Z = \sqrt{A^2 + B^2}$ и $\beta = \arctg(A/B)$. В специализированных вычислителях систем цифровой обработки сигналов при исключении не востребуемой избыточной точности с формированием порядка от 3 до 64-х значащих двоичных разрядов в формате с фиксированной запятой перед старшим разрядом исследуются методы обеспечения максимального приращения количества значащих цифр представления результата (M) при соответствующем ему минимальном увеличении числа вычислительных операций N и обращений к памяти P .

Вычисление амплитуды сигнала

Для ускоренного приближенного нахождения значения амплитуды сигнала Z по его квадратурным состав-

Исследованы быстродействующие алгоритмы воспроизведения функций $Z = \sqrt{A^2 + B^2}$, $\beta = \arctg(A/B)$ с оптимизацией вычислительного процесса по точностным характеристикам, быстродействию и программно-аппаратным затратам на примерах измерений амплитуд сигналов с неизвестной начальной фазой, напряжений переменного тока и фазовых углов.

ляющим A и B был предложен алгоритм с проверкой одного условия соотношения значений A и B [8]:

$$Z_1 = 1,04\sqrt{A^2 + B^2} \approx A + 0,4143B \quad \text{при } A \geq B. \quad (1)$$

$$Z_1 = 1,04\sqrt{A^2 + B^2} \approx B + 0,4143A \quad \text{при } B > A.$$

Этот алгоритм обеспечивает максимальную относительную погрешность $\delta_{mo} = 4,08\%$ при 4-х операциях $N + P$: сравнение, умножение, сложение и извлечение константы. Текущее значение относительной погрешности определялось выражением

$$\delta_M = |1 - M1/M2| \cdot 100\%, \quad (2)$$

где $M1$ – эталонное значение $\sqrt{A^2 + B^2}$, а $M2$ – аппроксимируемое в соответствии с (1). При исключении масштабирующего коэффициента 1,04 в (1) число операций $N + P$ возрастает до 6. В то же время для вычисления на интервале $x \in [0; 1]$ с погрешностью 4,6% только для полинома наилучшего приближения $\sqrt{x} = 0,0459 + x(2,866 - x(4,172 - 2,305x))$ требуется реализовать 10 операций: 6 N и 4 P . При вычислениях по алгоритму (1) не происходит как переполнения разрядной сетки, так и исчезновения порядка чисел A и B [7].

Для уменьшения значения погрешности введем большее число секторов (подинтервалов) аппроксимации, определяемых дополнительными условиями проверки числа соотношений A и B . При двух условиях сравнения соотношений получены полиномы, аппроксимирующие функцию $Z = \sqrt{A^2 + B^2}$ (рис. 1 а) с значением погрешности δ_{mo} менее 1,4% (рис. 1 б). График строился для диапазона значений A и B от 0 до 99. В соответствии с (2) введение других интервалов переменных не изменяет график значений погрешности. При реализации алгоритма преобразования, представленного на рис. 1а, необходимо выполнить 9 $N+P$ операций и хранить в памяти 5 констант. По сравнению с алгоритмом (1) с одним условием при увеличении числа операций на 3 погрешность δ_{mo} уменьшается в 2,9 раза.

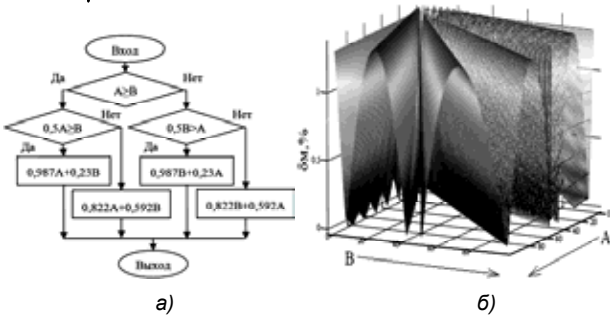


Рис. 1. Структурная схема (а)

и график погрешности (б) алгоритма вычисления функции $Z = \sqrt{A^2 + B^2}$ с двумя условиями

Дальнейшее увеличение количества условий усложняет реализацию алгоритма. В связи с этим для цифровых систем с более высокими классами точности предпочтительнее использовать непосредственные методы вычисления квадратного корня [4, 5].

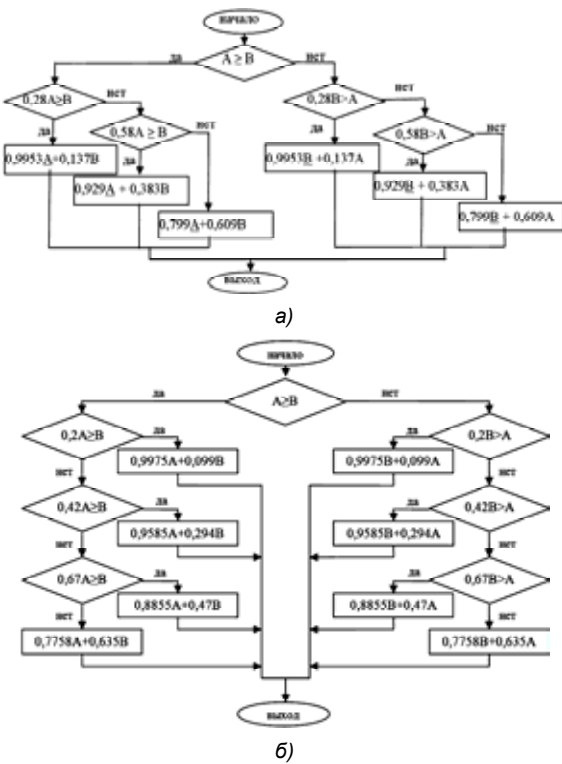


Рис. 2. Структурные схемы алгоритмов с 3-мя (а) и 4-мя (б) условиями

При введении 3-х условий (рис. 2 а) погрешность $\delta_{\text{лю}}$ уменьшилась до 0,5 %, что в 2,8 раза меньше чем для алгоритма с двумя условиями. Число операций $N + P$ на самой длинной ветви алгоритма соответствует 12, а в памяти необходимо хранить 8 констант. Для алгоритма с 4-мя условиями (рис. 2 б) погрешность не превышает 0,26 % при реализации процесса за 15 операций. В памяти необходимо хранить 11 констант. При увеличении числа операций $N + P$ на 3 погрешность уменьшилась почти в 2 раза, что соответствует увеличению числа двоичных цифр результата M только на 1.

Преобразование ортогональных составляющих сигнала в фазу

Для определения фазы используем аппроксимацию функции $\beta = \arctg(A/B)$ на интервале $\beta \in [0^\circ; 90^\circ]$ с пос-

ледующим нахождении фазы в диапазоне $\beta \in [0^\circ; 360^\circ]$ по знаковым разрядам составляющих A и B . Для упрощения вычислений еще раз сократим интервалы изменения аргумента и функции соответственно до $x \in [0; 1]$, $\beta \in [0^\circ; 45^\circ]$ (здесь $x = A/B$ при $A \geq B$). При $A \geq B$ вычисления производятся по алгоритму $\beta = \arctg(A/B)$. Если $A < B$, то вычисляется значение угла $\beta' = \arctg(B/A)$. Для получения действительного значения используется формула $\beta = 90^\circ - \beta'$. Путем моделирования на интервале $x \in [0; 1]$ (при $A \geq B$) для трех подинтервалов аппроксимации, определенных 3-мя условиями отношения ортогональных составляющих (рис. 2 а) и трех подинтервалов с равными максимальными абсолютными значениями погрешностей $\delta = |L_n(x) - \arctg(x)|$ приближения функции полиномами $L_n(x)$ получен набор схем аппроксимации 1.1, 1.2 и 2.1, 2.2 соответственно (табл. 1). Схемы 1.1 и 1.2 не обеспечивают равные максимальные погрешности на каждом из подинтервалов всего интервала от 0 до 1 (рис. 3), но исключаются 4 дополнительные операции (2 сравнения и 2 извлечения констант). Это обеспечивает в целом выигрыш по сравнению со схемами 2.1 и 2.2. Так, для схемы 1.2 по сравнению со схемой 2.1 получаем при меньшем количестве операций вычисления $\beta = \arctg(A/B)$ выигрыш по точности в $3,53 \cdot 10^{-3} / 3,066 \cdot 10^{-4} = 11,5$ раз. Для схемы 1.2 также имеем выигрыш по точности в $1,3 \cdot 10^{-3} / 3,066 \cdot 10^{-4} = 4,24$ раза по сравнению с использованием полинома наилучшего приближения на интервале аппроксимации от 0 до 1 по схеме 3 (табл. 1).

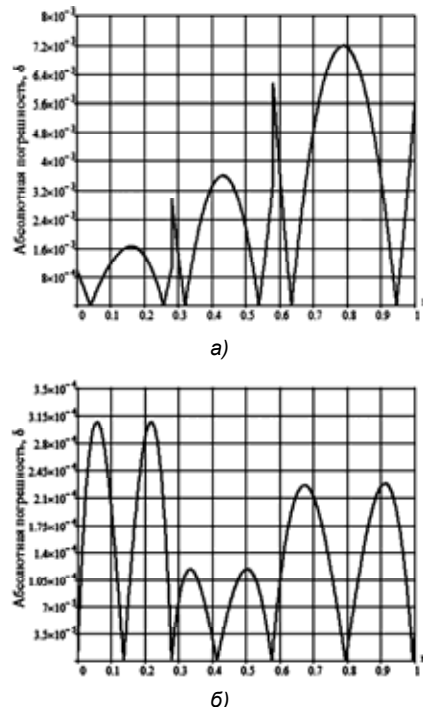


Рис. 3. Погрешности аппроксимации полиномами первой (а) и второй степени (б), приведенными в соответствии с алгоритмом для рис. 2 а

Операция деления A/B заменяется операцией умножения числителя на обратную величину знаменателя: $A/B = A \times (1/B)$. Функция $F(B) = 1/B$ аппроксимируется в

Таблица 1. Полиномы для аппроксимации функции арктангенса при $x \in [0; 1]$.

Схема	Полином наилучшего приближения	Подинтервал (интервал)	Максимальная погрешность	Число операций
1.1	$0,975x+0,001$	0; 0,28	$7,185 \cdot 10^{-3}$	4
	$0,843x+0,04$	0,28; 0,58		
	$0,617x+0,174$	0,58; 1		
1.2	$x(1,011773-0,131118x)$	0; 0,28	$3,066 \cdot 10^{-4}$	7
	$-0,011526 + x(1,100332 - 0,300443x)$	0,28; 0,58		
	$-0,005932 + x(1,089098 - 0,297799x)$	0,58; 1		
2.1	$0,9528x+0,0034$	0; 0,393	$3,53 \cdot 10^{-3}$	8
	$0,7731x+0,0741$	0,393; 0,689		
	$0,5855x+0,2033$	0,689; 1		
2.2	$-0,0001308 + x(1,01004 - 0,114636x)$	0; 0,24	$1,4 \cdot 10^{-4}$	11
	$-0,010573 + x(1,0957497 - 0,2951596x)$	0,24; 0,59		
	$-0,004363 + x(1,085874 - 0,2962549x)$	0,59; 1		
3	$x(1,02687+x(-0,1649-0,0779x))$	0; 1	$1,3 \cdot 10^{-3}$	8

Таблица 2. Аппроксимация функции $F(B) = 1/B$

Подинтервалы аппроксимации	полином наилучшего приближения	δ_{MA}
[0,03;1]	$40+B(-404+B(1379+B(-1830+819B)))$	4
[0,004;0,03]	$506+B(-89604+B(6948525+B(-240965372+3051998356B)))$	5,1
$[2^{-10};0,004]$	$2671+B(-2674061+B(1255458564+B(-277781760978+23324283858392B)))$	5,3

диапазоне $B \in [2^{-10}; 1]$ (табл. 2) тремя полиномами наилучшего приближения на трех подинтервалах с примерно равными абсолютными максимальными значениями погрешностей, причем $\delta_{MA} = 5.3$. При реализации алгоритма число операций равно 17: проверка двух условий, выполнение 8-ми алгебраических операций, извлечение из памяти одновременно 7-ми констант. Для интервала $B \in [2^{-10}; 1]$ получено значение приведенной относительной погрешности результата 0,53%. В соответствии с табл. 2 это значение получено по отношению к максимальной величине функции $1/B = 1/2^{-10} = 1024$. Этот метод вычислений при значении погрешности более 0,5% будет более эффективным чем итерационный [9, 10].

Заключение

Алгоритмы преобразования ортогональных составляющих в амплитуду для значений относительной погрешности результата от 4% до 0,26% по сравнению с алгоритмами с непосредственным вычислением квадратного корня по схемам аппроксимации $1L_1 \dots 2L_3$ обеспечивают выигрыш по быстродействию в 2...2,75 раза [5]. При этом не происходит как переполнения разрядной сетки, так и исчезновения порядка чисел A и B [7]. Измерение фазы с использованием предварительно определенных подинтервалов аппроксимации вычисления амплитуды сигнала по ортогональным составляющим обеспечивает уменьшение погрешности примерно в 10 раз при фиксированных остальных критериях вычислительного процесса.

Работа выполнена при поддержке гранта РФФИ 14-07-00293.

Литература

1. Акчурич, Э.А. Программная реализация взаимных преобразований алгебраического и экспоненциального представления комплексного сигнала на цифровых сигнальных процессорах / Э.А. Акчурич // Радиотехника. – 1995. – №1-2. – С. 21-23.
2. Ercegovas Milos Dand. Others. Reciprocation square root, inverse square root, and some elementary functions Using small multipliers // IEEE Trans Comput. 2000 / 49, №7, – 628-637 с.
3. Schulte Michall J., Swarzlander Earl E. (Jr). A family of variable – precision interval arithmetic processors // IEEE Trans Comput. 2000/ 49, № 5, – 387-397 с.

4. Чекушкин В.В., Аверьянов А.М., Богатов А.Д. Способ и устройство вычисления квадратного корня – патент РФ № 2438160 Бюл. №36, – 2011г.

5. Аверьянов А.М., Пантелеев И.В., Чекушкин В.В. Методы повышения быстродействия и точностных характеристик преобразователей ортогональных составляющих сигнала в амплитуду // Измерительная техника. – 2012г. – № 8. – с.9-14. Averyanov A.M., Chekushkin V.V., Panteleev I.V. Methods of increasing the speed and accuracy characteristics of converters of orthogonal components of a signal into amplitude // Measurement Techniques. November 2012, Volume 55, Issue 8, pp 858-866.

6. Авт. свид. СССР. Устройство для извлечения квадратного корня из суммы квадратов / И.Я. Мионов, Ю.В. Малинин, Т.Г. Лазебник, Л.И. Новикова // 1983. Бюл. № 8.

7. Каханер Д., Моулер К., Нэш С. Численные методы и программное обеспечение: пер. с англ. – М.: Мир, 2001. – 575 с.

8. Авт. свид. СССР. Устройство для вычисления функции $Z = \sqrt{A^2 + B^2}$ / Чекушкин В.В./1982. Бюл.№ 37.

9. Чекушкин В.В., Гришин В.Ю., Костров В.В. Совершенствование алгоритмов деления чисел в информационно-измерительных системах // Метрология. – 2013. – № 11. – С. 3-14.

10. Гришин В.Ю., Пантелеев И.В., Чекушкин В.В. Совершенствование методов, математических моделей реализации вычислительных процессов в радиолокационных системах // Вестник воздушно-космической обороны. – 2014. – № 3. – С. 31-34.

HIGH-SPEED ALGORITHMS OF THE SIGNAL ORTHOGONAL COMPONENTS TRANSFORMATION IN AMPLITUDE AND PHASE

Chekushkin V.V., Panteleev I.V., Mikheev K.V.

A study of fast-speed algorithms of the joint computation of the functions $R = \sqrt{A^2 + B^2}$, $\beta = \arctg(A/B)$ from the viewpoint of optimization of the computational process criteria: precision characteristics, high-speed operation and hardware and software costs on the base of the examples of measurement of amplitudes of signals with unknown initial phase, AC voltage and phase angles, shaft vibration parameters, implementation of transformation of a complex signal from the algebraic representation to the exponential one has been conducted.

УДК 004.5

ОБОБЩЕНИЕ МЕТОДА ЦЕПОЧЕК ЛОКАЛЬНЫХ ЭКСТРЕМУМОВ ДЛЯ АНАЛИЗА СИГНАЛОВ РАЗЛИЧНОЙ ПРИРОДЫ

Туровский Я.А., к.м.н., зав. лабораторией медицинской кибернетики Воронежского государственного университета (ВГУ), e-mail: yaroslav_turovsk@mail.ru;

Кургалин С.Д., д.ф.-м.н., зав. кафедрой цифровых технологий ВГУ, e-mail: kurgalin@bk.ru;

Вахтин А.А., к.ф.-м.н., доцент кафедры программирования и информационных технологий ВГУ, e-mail: alvahtin@gmail.com;

Борзунов С.В., к.ф.-м.н., доцент кафедры цифровых технологий ВГУ, e-mail: kurgalin@bk.ru;

Белобродский В.А., аспирант кафедры цифровых технологий ВГУ, e-mail: belobrodsky@yandex.ru.

Ключевые слова: биомедицинские сигналы, спектральный анализ, вейвлет-преобразование, цепочки локальных максимумов и минимумов, рибонуклеиновая кислота, метеорология, биржевые котировки.

Введение

Методы спектрального анализа одномерных биомедицинских сигналов достаточно давно и прочно вошли не только в научные исследования, но и в клиническую практику [1, 2]. Они широко используются в анализе электроэнцефалограмм (ЭЭГ), являющихся результатом регистрации электрической активности мозга через неповреждённые покровы головы [3], электромиограмм – регистрации электрической активности мышц [4], вариабельности сердечного ритма – изменения частоты пульса во времени [5] и т.д. Одним из важных направлений совершенствования теории и практики спектрального оценивания применительно к одномерным сигналам, регистрируемым с человека, является создание новых методов выделения научно-значимой информации.

Большое распространение получили в настоящее время подходы, осуществляющие развитие уже хорошо зарекомендовавших себя преобразования Фурье или вейвлет-преобразования для исследования разнообразных сигналов. При этом совершенствование методов анализа одномерных сигналов на основе указанных выше преобразований во многом не затрагивает методов расчёта их коэффициентов. Основной упор разработчики новых методов делают на улучшение подходов к обработке коэффициентов, полученных в ходе преобразования, и на извлечение из них интересующей исследователей информации.

В работах [6-9] был разработан «метод цепочек локальных экстремумов» – метод построения цепочек локальных максимумов (ЦЛМ) и минимумов (ЦЛМин) на основе матрицы квадратов коэффициентов вейвлет-преобразования для исследования сигналов ЭЭГ. Полученные ЦЛМ в рамках принятой интерпретации могут рассматриваться как частотно-временные области с максимальной энергией сигнала в заданном диапазоне

Обобщен метод построения цепочек локальных экстремумов матрицы квадратов коэффициентов вейвлет-преобразования, разработанный и использованный ранее для анализа сигналов ЭЭГ («метод цепочек локальных экстремумов»). Существенным моментом развития метода является формулировка общего правила для присоединения имеющегося экстремума локального спектра к уже сформированной цепочке экстремумов. Предлагаемый подход позволяет избежать чрезмерного объединения экстремумов в цепочки с большими разрывами во временной области. Обобщенный метод цепочек локальных экстремумов был проверен на данных разной природы: перекодированные РНК-последовательности генома вируса гепатита С, метеорологические данные – среднемесячная температура, сведения о биржевых котировках акций.

времени и частоты. Для ЦЛМин верно обратное – это частотно-временные области с минимальной энергией сигнала в заданном диапазоне времени и частоты. Данный метод был апробирован на сигналах ЭЭГ и рассчитанных на их основе т.н. «вызванных потенциалах головного мозга» [10], а также на ряде других сигналов, зарегистрированных с человека [11]. Однако, для эффективного внедрения этого метода в разные области исследований необходимо провести и обосновать его обобщение на сигналы различной природы.

Целью данной работы является разработка наиболее общего, не зависящего от природы одномерного сигнала метода построения цепочек локальных экстремумов матрицы квадратов коэффициентов вейвлет-преобразования для извлечения из данных дополнительной значимой информации.

Метод построения ЦЛМ И ЦЛМин на основе матриц вейлет-преобразования

В предлагаемом методе используется тот факт, что после непрерывного вейвлет-преобразования квадраты элементов полученной матрицы $W^2(a, b)$ отражают распределение энергии сигнала по масштабам вейвлет-преобразования. Величины $W^2(a, b)$ формируют трёхмерную область, содержащую как локальные максимумы энергии сигнала на определённом масштабе a и сдвиге вейвлет-функции b , отражающем сдвиг вейвлета по оси времени t , так и локальные минимумы. Очевидно, что локальные максимумы и минимумы величин $W^2(a, b)$

несут информацию о свойствах исследуемых сигналов. Эти свойства могут быть содержательно интерпретированы, исходя из современных представлений о причинах появления в анализируемых сигналах тех или иных частотно-временных особенностей и их характеристиках. Для значительной части сигналов различной природы свойственно формирование достаточно близко расположенных в частотно-временном пространстве последовательностей локальных максимумов и минимумов.

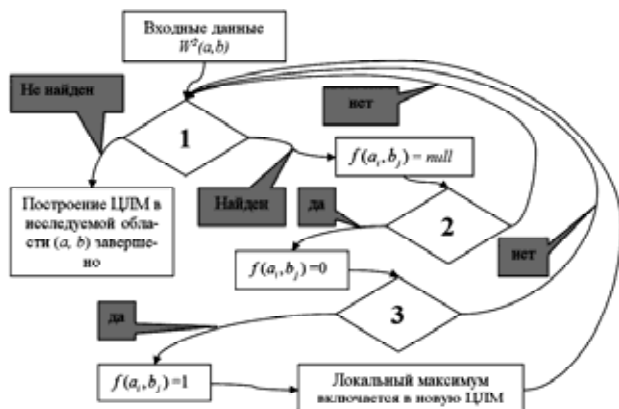


Рис. 1. Схема формирования цепочек локальных максимумов. На схеме: 1. Нахождение локального максимума с координатой b_j , не включенного ни в одну из уже построенных ЦПМ и не проверенного для текущей конфигурации построенных и строящихся ЦПМ

2. $(a_i \in [a_{i-1} - (\nu a_s + u), a_{i-1} + (\nu a_s + u)]) \vee (b_j \in [b_{j-1}, b_{j-1} + (\nu a_s + u)])$
3. $(a_i \in [a_{i-1} - (\nu a_s + u), a_{i-1} + (\nu a_s + u)]) \wedge (b_j \in [b_{j-1}, b_{j-1} + (\nu a_s + u)])$

Схема формирования таких цепочек на примере локальных максимумов представлена на рис. 1, где использовано: a_{i-1} – масштаб вейвлет-преобразования для локального максимума концевой точки ЦПМ, ближайшей по величине масштаба к предполагаемому для включения в ЦПМ локальному максимуму; b_{j-1} – номер отсчёта (номер скейлграммы) [6-10] концевой точки ЦПМ, ближайшей к аналогичному значению предполагаемого для включения в ЦПМ локального максимума; a_s – значение масштаба первой точки формирующейся ЦПМ; индексы $i = 1, 2, \dots$ и $j = 1, 2, \dots$ характеризуют текущие номера локальных экстремумов, потенциально принадлежащих той или иной ЦПМ (или ЦПМин); $f(a_i, b_j)$ – логическая функция, описывающая решение о включении или не включении новой точки локального экстремума в уже существующие ЦПМ или ЦПМин (для локальных максимумов и минимумов, соответственно).

Константы u и ν получены эмпирически [6-10] для данного типа сигналов. Параметр b в нашем методе расчета является безразмерным, он представляет собой номер отсчёта АЦП, а в размерность времени его значения переводятся путём умножения на частоту дискретизации сигнала, нетрудно получить размерность частоты (в Гц) тех или иных его компонентов.

Из представленной на рис. 1 схемы видно, что формирование ЦПМ и ЦПМин определяется константами u и ν . Для расчёта констант используются два параметра: k – число точек в ЦПМ (или ЦПМин) и Δb – разность первого и последнего порядковых номеров локальных спектров матрицы $W^2(a, b)$, в которых находились локальные максимумы (минимумы). В случае если локальный спектр строился при каждом временном отсчёте, то Δb будет равным разности порядковых номеров первого и последнего отсчёта ЦПМ (или ЦПМин). Величина $k/\Delta b$ принадлежит диапазону значений $(0, 1]$ и, по сути, выражает долю насыщения ЦПМ определённой длины локальными максимумами. Очевидно, что в случае, если число локальных максимумов (или минимумов) равно длине цепочки Δb , то $k/\Delta b = 1$. Такую цепочку следует признать удачно построенной, т.к. отсутствуют локальные спектры, принадлежащие отрезку Δb , но не содержащие локальных максимумов (минимумов) для включения в цепочку. Поскольку в каждый момент времени (или для конкретного порядкового номера отсчёта) b в ЦПМ (ЦПМин) включается только один локальный экстремум, отношение $k/\Delta b$ принципиально не может быть большим единицы. В случае же, если происходит объединение двух далеко отстоящих по времени ЦПМ (ЦПМин) в одну, то отношение $k/\Delta b \ll 1$. Подобное объединение следует признать «искусственным», не отражающим реальное распределение энергии сигнала по элементам матрицы $W^2(a, b)$. Таким образом, задача поиска оптимальных значений u и ν для корректного построения ЦПМ (ЦПМин) сводится к максимизации отношения $k/\Delta b$ (т.е. $k/\Delta b \rightarrow \max$) в пространстве изменения параметров (u, ν) .



Рис. 2. Схема выбора параметров u и ν для расчета ЦПМ и ЦПМин

В настоящее время не существует теоретических подходов, позволяющих корректно рассчитать значения u и ν для получения максимальных величин $k/\Delta b$ для всего пула ЦПМ (или ЦПМин) в пространстве параметров (a, b) . Простейшим решением этой проблемы является перебор с определённым шагом всех возможных значений u и ν для поиска максимума отношения $k/\Delta b$ (см. рис. 2).

Апробация метода на сигналах различной природы

Подход к выбору параметров формирования ЦПМ, схематически представленный на рис. 2, ранее был использован [6-10] для анализа ЭЭГ вызванных потенциа-

лов. Для этого класса сигналов были получены значения $u=3$ и $v=0.05$.

Рассмотрим применение данного подхода для исследования различных одномерных данных. Возьмём, во-первых, данные о последовательности нуклеотидов в нуклеиновых кислотах вируса гепатита С одного из изолятов [12].

Произведём перекодировку нуклеотидов:

$$T(d) = \begin{cases} T(d-1)-1, P_i \\ T(d-1)+1, P_u \end{cases} \quad (1)$$

где $T(d)$ – текущие значения перекодированной функции последовательности РНК; $d=1,2,3,\dots$ – порядковый номер (позиция) нуклеотида в исследуемом ряду; присутствие P_i – означает наличие в позиции d нуклеотида с пиримидиновым основанием, а присутствие P_u – наличие в позиции d нуклеотида с пуриновым основанием. При $d=1$ $T(d)=0$.

На рис. 3 представлены результаты оценки структур ЦЛМ для перекодированных согласно (1) РНК-последовательностей вируса гепатита С. Глобальный максимум величины $k/\Delta b$ при $v \in [0, 0.1]$ и $u \in [3, 7]$ имеет координаты (0,3), что дает возможность считать координаты этой области пространства (u, v) как параметры для формул, изложенных на рис.1. Однако при $v \in [0, 0.1]$ и $u \in [6, 7]$ выполняется условие $k/\Delta b > 0.97$, что также позволяет рассматривать и эту область пространства (u, v) в качестве приемлемого варианта для использования констант u и v при построении ЦЛМ (ЦЛМин), согласно схеме формирования цепочек, представленной на рис. 1.

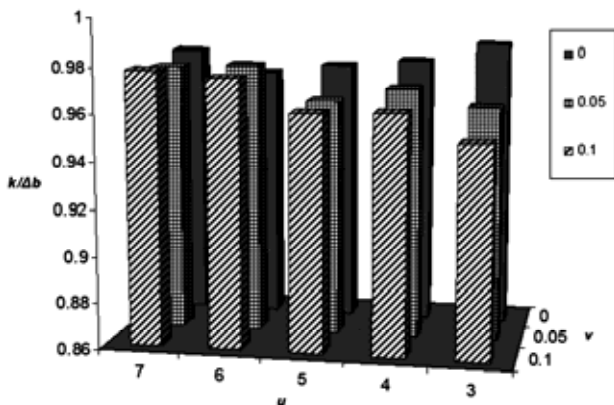


Рис. 3. Результаты оценки величин $k/\Delta b$ в пространстве (u, v)

для РНК- последовательностей вируса гепатита С

Во-вторых, исследуем информацию о среднемесячной температуре по Москве [13], начиная с 1989 г. В этом случае выявляется глобальный максимум функции $k/\Delta b$ с координатами (0,3). Несколько уступают ему значения $k/\Delta b$ в координатах (0,4) и (0.05,4). Остальные значения $k/\Delta b$, рассчитанные для других координат v и u , существенно ниже (рис. 4). Следовательно, для этого случая выделяется только одна область параметров (u, v) , в отличие от примера с РНК вируса гепатита С, где таких областей было выделено две.

В-третьих, отличную от предыдущих двух случаев структуру распределения значений $k/\Delta b$ в пространстве (u, v) продемонстрировали данные биржевых котировок

обыкновенных акций Внешторгбанка [14] за период с 1 по 4 декабря 2014 г. Оценивалось время открытия каждого из одноминутных интервалов торгов на ММВБ. Полученные результаты (см. рис. 5), продемонстрировали наличие только одного максимума величины $k/\Delta b$. При этом, в отличие от приведённых выше примеров, остальные значения функции $k/\Delta b$ были существенно ниже глобального максимума. Следовательно, соответствующие этим значениям $k/\Delta b$ координаты в пространстве (u, v) не могут быть использованы в качестве параметров формул, указанных на рис. 1, т.к. для них не выполняется указанное ранее условие $k/\Delta b \rightarrow \max$.

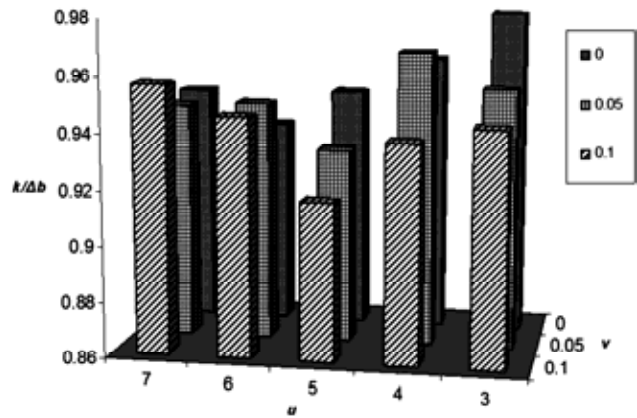


Рис. 4. Результаты оценки величин $k/\Delta b$ в пространстве (u, v)

для среднемесячной температуры по г. Москве

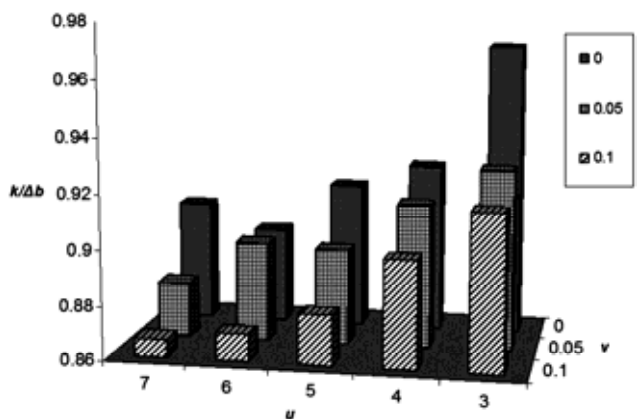
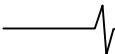


Рис. 5. Результаты оценки величин $k/\Delta b$ в пространстве параметров (u, v)

для биржевых котировок обыкновенных акций ВТБ за период с 1 по 3 декабря 2014 г.

Заключение

Метод построения цепочек локальных экстремумов матрицы квадратов коэффициентов вейвлет-преобразования, разработанный и использованный ранее для анализа сигналов ЭЭГ («метод цепочек локальных экстремумов»), обобщен в настоящей работе для исследования любых одномерных сигналов или данных, вне зависимости от их природы. Существенным моментом развития метода является формулировка общего правила для присоединения имеющегося экстремума локального спектра к уже сформированной цепочке экстремумов путем определения величин параметров u и v , от которых зависит характеристики построенной ЦЛМ или



ЦЛМин. Выбор параметров должен обеспечить максимизацию отношения числа экстремумов (только минимумов или только максимумов) ко времени существования цепочки. Данный подход позволяет избежать чрезмерного объединения экстремумов в цепочки с большими разрывами во временной области.

Предложенный универсальный метод был проверен на данных разной природы: на перекодированных РНК-последовательностях генома вируса гепатита С, на метеорологических данных – среднемесячных температурах, на сведениях о биржевых котировках акций. Продемонстрировано, что для каждого из этих типов данных величина $k/\Delta b$ имеет свои максимумы разной амплитуды в различных областях пространства параметров (u, v). Можно сделать вывод, что величина $k/\Delta b$ является достаточно чувствительным критерием для выбора значений u и v , необходимых для корректного построения ЦЛМ и ЦЛМин. Следовательно, данный подход позволяет получать более детальную информацию об исследуемых объектах на основе улучшения спектральных оценок анализируемых сигналов и анализа динамики их поведения во времени.

Литература

1. Биоэлектрическая активность мозга у больных с невротическими и неврозоподобными расстройствами (по данным спектрального анализа) Шульц Е.В., Бабулин И.Н., Караваева Т.А., Карвасарский Б.Д., Слезин В.Б. Обзорение психиатрии и медицинской психологии им. В.М. Бехтерева. 2010. № 3. С. 26-31.
2. Взаимосвязь быстрых изменений индивидуального поддиапазона волн электроэнцефалограммы и кардиоритма во время сна Васильев Е.Н., Урываев Ю.В. Физиология человека. 2006. Т. 32. № 4. С. 18-23.
3. Зенков Л.Р. Клиническая электроэнцефалография (с элементами эпилептологии). – М.: МЕДПресс-информ, 2011.– 356 с.
4. Анализ мышечных биопотенциалов методом спектрографии Дегтярев Ю.Г., Осипов А.Н., Ковалев В.В., Кульчицкий В.А. Военная медицина. 2013, – № 3 (28). – С. 94-97.
5. Состояние вегетативной нервной регуляции у больных атопическим дерматитом Аксенова О.И., Марченко В.Н., Монахов К.Н. Вестник современной клинической медицины. 2014. Т. 7. № 4. – С. 15-17.
6. Анализ электроэнцефалограмм на основе исследования изменяющейся во времени структуры локальных максимумов матрицы вейвлет-коэффициентов Туровский Я.А., Кургалин С.Д., Максимов А.В., Семёнов А.Г. Вестник Воронежского государственного университета. Серия: Системный анализ и информационные технологии. 2012. № 2. – С. 69-73.
7. Анализ энцефалограмм на основе исследования цепочек локальных максимумов скейлограмм Туровский Я.А., Кургалин С.Д., Семёнов А.Г. Цифровая обработка сигналов. 2013. № 2. – С. 20-23.
8. Динамика цепочек локальные максимумов спектров электроэнцефалограмм человека Туровский Я.А., Кургалин С.Д., Семёнов А.Г. Биофизика. 2014. Т. 59. № 1. – С. 185-190.
9. Фактор времени при реализации непрерывного вейвлет-преобразования для анализа сигналов ЭЭГ Туровский Я.А., Кургалин С.Д., Вахтин А.А., Максимов А.В. Информационные технологии в проектировании и производстве. 2012. № 2. С. 61-66.
10. Исследование динамики максимумов локальных вейвлет-спектров вызванных зрительных потенциалов головного мозга Туровский Я.А., Кургалин С.Д., Семенов А.Г. Информационные технологии. 2013. № 10. С. 46-50.
11. Вегетативная регуляция сердечно-сосудистой системы плодов и новорожденных детей, перенесших хроническую внутриутробную гипоксию Туровский Я.А. автореферат диссертации на соискание ученой степени кандидата медицинских наук / Воронежская государственная медицинская академия им. Н.Н. Бурденко. Воронеж, 2005
12. <http://www.ncbi.nlm.nih.gov/nuccore/EF032883.1?report=genbank&log>
13. http://www.atlas-yakutia.ru/weather/climate_russia-III.html
14. <http://www.finam.ru/>

GENERALIZATION OF THE METHOD OF CHAINS OF LOCAL EXTREMA FOR ANALYSIS OF SIGNALS OF DIFFERENT ORIGINS ABSTRACT

Turovsky Y.A., Kurgalin S.D., Vahtin A.A., Borzunov S.V., Belobrodsky V.A.

The method of construction of chains of local extrema for the matrix of squared coefficients of the wavelet transformation, developed and applied previously to analyze EEG signals («the method of chains of local extrema»), is summarized in this paper for the research of all one-dimensional signals or data, regardless of their nature. An essential feature of the development of the method is the formulation of a general rule for adding an existing extremum of the local spectrum to the already formed chain by identifying values of the parameters u and v , which determine the characteristics of the formed Chain of Local Maxima or Chain of Local Minima. Selection of parameters must ensure the maximization of the ratio of extrema numbers (only minimums or only maximums) at the time of the existence of the chain. This approach avoids an excessive union of extrema into the chains with large gaps in the time domain. The proposed universal method has been tested on data of different nature: on the recoded RNA genome sequences of hepatitis C virus, on meteorological data - the average monthly temperatures, on information about stock exchange quotations. It may be concluded that the value of $k / \Delta b$ is sufficiently sensitive criterion for selecting values of u and v necessary for the correct construction of CLM and CLMin signals of different nature.

УДК. 004.934

НЕЭТАЛОННАЯ ОЦЕНКА РАЗБОРЧИВОСТИ ЗАШУМЛЕННЫХ РЕЧЕВЫХ СИГНАЛОВ

Топников А.И., к.т.н., ассистент кафедры динамики электронных систем Ярославского государственного университета им. П.Г. Демидова, e-mail: toptnikov@gmail.com;

Нестеров М.С., аспирант кафедры динамики электронных систем Ярославского государственного университета им. П.Г. Демидова, e-mail: taximator89@mail.ru;

Новоселов С.А., к.т.н., научный сотрудник отдела верификации и идентификации диктора, «Центр Речевых Технологий», e-mail: novoselov@speechpro.com;

Приоров А.Л., д.т.н., доцент кафедры динамики электронных систем Ярославского государственного университета им. П.Г. Демидова, e-mail: andcat@yandex.ru.

Ключевые слова: речевой сигнал, разборчивость, неэталонная оценка, шумоподавление, методика, парная регрессия.

Введение

Одной из важнейших составляющих акустической экспертизы каналов речевой коммуникации является измерение разборчивости речи [1]. Под разборчивостью понимается степень, с которой определенные речевые единицы (фонемы, слоги, слова, фразы) могут быть верно восприняты слушателем. Различные искажающие факторы (шумы, реверберация, нелинейные искажения) могут привести к значительному снижению разборчивости и потери части информации. Вместе с разборчивостью снижается и качество речевого сигнала [2], однако эти понятия стоит разграничивать.

Наиболее общим подходом при классификации методов оценки разборчивости является разделение их на субъективные и объективные. Отличительная особенность субъективных методов – непосредственное участие дикторов и аудиторов (слушателей) в процессе оценки. Эти методы обладают высокой достоверностью, но ресурсоемки и непригодны для проведения экспертизы каналов коммуникации в режиме реального времени. Поэтому примерно с середины 50-х годов прошлого века активно развиваются объективные методы измерения разборчивости.

В настоящее время выделяются две наиболее крупные группы этих методов: формантные и модуляционные. К первой группе относятся зарубежные методы: AI (Articulation Index) и SII (Speech Intelligibility Index), их предшественники и модификации [1, 3-8], а также отечественные методы, предложенные Ю.С. Быковым, Н.Б. Покровским, М.А. Сапожковым [1, 9-12]. Модуляционные методы представлены различными модификациями метода STI (Speech Transmission Index), включая методы RASTI, STIPA и STITEL [1, 13-15]. Предпринимаются попытки и по объединению двух подходов [16, 17]. Общей чертой указанных выше мер разборчивости является их эталонность – потребность в «чистом» (незашумленном, неискаженном) сигнале, без которого процедура оценки невозможна. Отсутствие таких сигналов в реальных условиях приводит к необходимости разработки неэталонных методов [18–21]. Это

Предлагается методика неэталонной оценки значений показателя разборчивости зашумленной речи SNR loss, а также ее модификация, отличающаяся меньшими вычислительными затратами. Достоверность полученных неэталонных оценок сравнивается путем сопоставления со значениями, полученными с использованием эталонного метода.

одна из наиболее сложных и актуальных задач в рассматриваемой области. До сих пор ни один неэталонный метод не включен в известные стандарты, регламентирующие оценку разборчивости в помещениях или телекоммуникационных системах.

Целью данной работы является исследование возможности получения неэталонных оценок разборчивости зашумленных речевых сигналов с использованием эталонного метода и подсистемы, восстанавливающей эталонный (квазиэталонный) сигнал из искаженного. Для случая, когда основным искажающим фактором является аддитивный шум, роль такой системы может выполнять шумоподавитель, реализуемый по методу [22]. В данной работе рассматривается техника, основанная на применении меры разборчивости речи SNR loss [7] и метода шумоподавления Скалара (Scalart) [23]. Выбранный подход, заключающийся в неэталонной оценке значений уже существующего эталонного метода, позволяет существенно упростить процесс разработки. Так как исходный эталонный метод SNR loss верифицирован (сравнен по достоверности с субъективными оценками), то достоверность предлагаемого неэталонного подхода можно оценивать по близости полученных оценок к истинным значениям SNR loss.

Описание оригинального метода SNR loss

Метод SNR loss является развитием формантного метода AI и в отличие от него учитывает влияние на разборчивость речи работы нелинейных методов, применяемых для подавления шума в речи. Кратко рассмотрим метод SNR loss, подробно описанный в [7].

Вычисления осуществляются для отдельных временных интервалов (кадров). Для каждой j -й критической полосы каждого m -го кадра вычисляется следующая величина:

$$L(j, m) = SNR_x(j, m) - SNR_{\hat{x}}(j, m) =$$

$$= 10 \lg \frac{X^2(j, m)}{D^2(j, m)} - 10 \lg \frac{\hat{X}^2(j, m)}{D^2(j, m)} = 10 \lg \frac{X^2(j, m)}{\hat{X}^2(j, m)},$$

где $SNR_X(j, m)$ – входное отношение сигнал/шум (ОСШ) в полосе j и интервале m , $SNR_{\hat{X}}(j, m)$ – ОСШ после обработки в полосе j и интервале m , $X(j, m)$ – j -я составляющая спектра незашумленного (эталонного) сигнала, вычисленного с учетом критических полос для m -го временного интервала, $\hat{X}(j, m)$ – j -я составляющая спектра обработанного сигнала, вычисленного с учетом критических полос для m -го временного интервала. Очевидно, что при $\hat{X}(j, m) = X(j, m)$ величина $L(j, m) = 0$. В целом значение $L(j, m)$ может быть как положительным, так и отрицательным.

Значения $L(j, m)$ ограничиваются в пределах определенного диапазона $[-SNR_{lim}, SNR_{lim}]$:

$$\hat{L}(j, m) = \min(\max(L(j, m), -SNR_{lim}), SNR_{lim}).$$

Полученные на предыдущем этапе значения $\hat{L}(j, m)$ масштабируются на интервал $[0, 1]$:

$$SNR_{LOSS}(j, m) = \begin{cases} -\frac{C_-}{SNR_{lim}} \hat{L}(j, m), & \text{если } \hat{L}(j, m) < 0 \\ \frac{C_+}{SNR_{lim}} \hat{L}(j, m), & \text{если } \hat{L}(j, m) \geq 0, \end{cases}$$

где C_- и C_+ – параметры масштабирующей функции.

Для каждого временного интервала осуществляется усреднение $SNR_{LOSS}(j, m)$ по всем критическим полосам:

$$fSNR_{LOSS}(m) = \frac{\sum_{j=1}^K W(j) SNR_{LOSS}(j, m)}{\sum_{j=1}^K W(j)},$$

где $W(j)$ – весовая функция, учитывающая психоакустические закономерности восприятия речевых сигналов.

Вычисляется среднее значение $\overline{SNR_{LOSS}}$ путем усреднения $fSNR_{LOSS}(m)$ по всем окнам:

$$\overline{SNR_{LOSS}} = \frac{1}{M} \sum_{m=0}^{M-1} fSNR_{LOSS}(m),$$

где M – число окон, на которые разделен сигнал. Получившаяся величина, для удобства обозначенная как SNR_{loss} , изменяется в интервале от 0 до 1. Нулевое значение соответствует идеальной разборчивости, единичное – ее полному отсутствию. Графики, характеризующие взаимосвязь разборчивости в процентах со значениями SNR_{loss} , представлены в [7].

Стоит пояснить, что в тексте для удобства используется два варианта написания: SNR_{loss} – для названия метода и SNR_{loss} – для обозначения значений, полученных при помощи этого метода.

Методика неэталонной оценки разборчивости

Идея, лежащая в основе предлагаемой неэталонной версии критерия SNR_{loss} , состоит в том [22], чтобы вычислять значение SNR_{loss} для оцениваемого (зашумленного) сигнала и сигнала, полученного шумоподавлением с использованием метода, предложенного Скаларом и Филхо в работе [23]. При этом сигнал на выходе шумоподавителя считается чистым (эталонным). В данном случае это допустимо, так как сигнал на выходе шумоподавителя является оценкой чистого (незашумленного) сигнала.

Проанализируем взаимосвязь значений SNR_{loss} , вычисленных таким образом, (обозначим их как SNR_{loss}') с истинными значениями SNR_{loss} , полученными при использовании чистого сигнала в качестве эталона. Для моделирования использовалось шесть незашумленных речевых фрагментов, каждый из которых зашумлялся аддитивным белым гауссовским шумом (АБГШ). Для каждого фрагмента формировались 230 зашумленных версий: 5 реализаций шума для каждого ОСШ из диапазона от -15 до 30 дБ [22]. В ходе моделирования значения SNR_{loss} измерялись для 1380 сигналов (6 фраз по 230 вариантов зашумления каждая). Диаграмма рассеяния между истинными значениями SNR_{loss} и значениями SNR_{loss}' представлена на рис. 1. Видно, что при изменении ОСШ в пределах от -15 до 30 дБ (рис. 1) истинное (т.е. измеренное эталонно – с использованием чистого сигнала) значение критерия SNR_{loss} изменяется в пределах примерно от 0,4 до 1. Множество точек на рис. 1 распределено таким образом, что позволяет сделать предположение о возможности линейной аппроксимации зависимости SNR_{loss} от SNR_{loss}' . На основе имеющихся данных и метода наименьших квадратов взаимосвязь может быть описана уравнением парной линейной регрессии:

$$SNR_{loss} = b_1 * SNR_{loss}' + b_0. \quad (1)$$

Установлено, что для данного типа шума коэффициенты уравнения (1) принимают следующие значения: $b_1 = 0,8909$; $b_0 = 0,043$. Даже при использовании линейной регрессии наблюдается высокое значение коэффициента детерминации (выше 0,99), характеризующего связь между истинными значениями SNR_{loss} и неэталонными оценками SNR_{loss} , полученными путем подстановки значений SNR_{loss}' в (1).

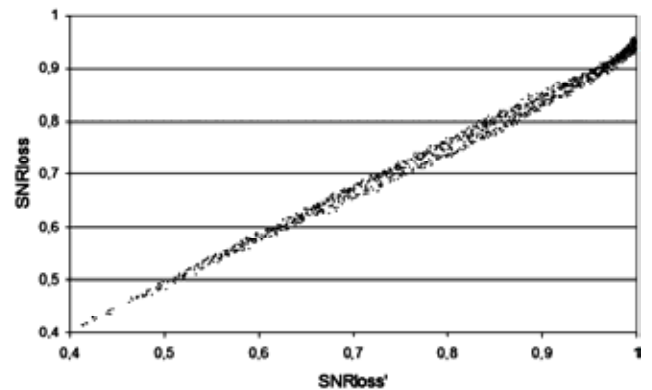


Рис. 1. Диаграмма рассеяния значений SNR_{loss}' и SNR_{loss} для АБГШ

Применение степенной или полиномиальной функций для описания исследуемой зависимости приводит к дополнительному незначительному повышению точности аппроксимации. Выбор в пользу линейной функции сделан с учетом возможности достижения высокой точности и сохранения простоты.

Для проверки предложенной неэталонной методики оценки значений SNRloss проведено дополнительное моделирование. Для обеспечения достоверности выбран набор речевых фрагментов, отличный от того, который использовался в первой части исследований. Для 24 речевых фрагментов было измерено в общей сложности 5520 пар значений SNRloss и SNRloss' (для разных реализаций шума и ОСШ). Полученные значения SNRloss' подставлялись в выражение (1). Таким образом было найдено 5520 неэталонных оценок значений SNRloss. Сравнение неэталонных оценок с истинными значениями критерия, полученными эталонным методом, позволяет судить о достаточно высокой точности предложенного метода неэталонной оценки (рис. 2).

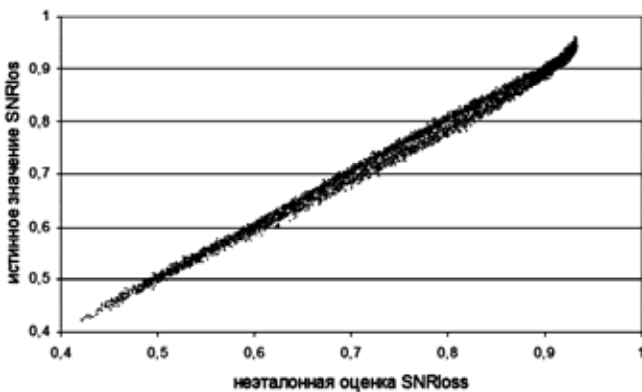


Рис. 2. Диаграмма рассеяния истинных значений SNRloss и их неэталонных оценок для АБГШ

При использовании линейной функции средняя величина абсолютной ошибки составляет примерно 0,008, а ее максимальная величина равна 0,036. Средняя величина относительной ошибки составляет примерно 1,05 %, а ее максимальное значение равно 4,72 %. Стоит отметить, что величины ошибок могут быть дополнительно снижены благодаря применению полиномиальной или кусочно-линейной аппроксимаций. Полученные результаты свидетельствуют о перспективности применения предложенного подхода к неэталонной оценке значений показателя SNR loss.

Следующим шагом исследования является проверка применимости данного подхода для других видов шумов, наиболее часто встречающихся при решении практических задач. С этой целью использовалась специализированная база шумов Noisex-92, состоящая из записей реальных шумов различных типов. Так как частота дискретизации тестовых речевых сигналов составляла 8 кГц, то частота дискретизации записей шумов также была понижена до 8 кГц. В работе приводятся оригинальные названия записей шумов на английском языке.

Для осуществления неэталонной оценки значений SNRloss в случае воздействия на сигнал шума произвольного типа возможно три подхода: использование

регрессионного выражения (1) с параметрами, полученными для АБГШ (подход 1); использование этого регрессионного выражения с параметрами, полученными для смешанной выборки, состоящей из речевых сигналов, зашумленных шумами разных типов (подход 2); нахождение новых параметров зависимости (1), индивидуальных для каждого конкретного типа шума (подход 3). Первые два подхода более универсальны и удобны на практике, однако третий подход потенциально способен обеспечить большую точность.

Для реализации второго и третьего подхода необходима оценка новых коэффициентов уравнения (1). При нахождении коэффициентов для реализации второго подхода использовалась смешанная выборка, состоящая из речевых сигналов, искаженных шумами различных типов. Установлено, что в этом случае коэффициенты уравнения (1) принимают следующие значения: $b_1=0,8879$; $b_0=0,0479$. При этом наблюдается высокое значение коэффициента детерминации $R^2=0,9927$.

Для реализации третьего подхода необходимо найти коэффициенты уравнения регрессии, индивидуальные для каждого типа шума. Как и в случае с АБГШ, линейная регрессия хорошо подходит для описания взаимосвязи между значениями SNRloss и SNRloss'. Для нахождения коэффициентов проведено моделирование. Для каждого типа шума формировалось 1380 сигналов (6 фраз по 230 вариантов зашумления каждой фразы). На основе этих выборок для каждого типа шума найдены коэффициенты линейной регрессии, связывающей значения SNRloss и SNRloss', а также оценено среднеквадратическое отклонение (СКО) и коэффициент детерминации R^2 (табл. 1). Значения коэффициента детерминации, близкие к единице, свидетельствуют о применимости линейной модели и тесной взаимосвязи между значениями SNRloss и SNRloss'. Установлено, что наибольшее СКО и наиболее низкое значение коэффициента детерминации наблюдается в случае нахождения параметров линейной регрессии для шума «Speech babble».

Далее проводилось сравнение указанных подходов по точности. С этой целью для каждого типа шума, как и в случае с АБГШ, для 24 фраз измерялось по 5520 пар значений SNRloss и SNRloss' (для разных реализаций шума и ОСШ). Полученные значения SNRloss' подставлялись в выражение (1) с соответствующими коэффициентами, зависящими от применяемого подхода. После этого полученные оценки значений SNRloss сравнивались с истинными значениями, полученными при помощи эталонного метода. Вычислялись значения средней абсолютной ошибки, корня из среднеквадратической ошибки, относительной ошибки, выраженной в процентах:

$$MAE = \frac{1}{N_{\text{ш}}} \sum_{l=1}^{N_{\text{ш}}} \left| \text{SNRloss}_l^{\text{оцен.}} - \text{SNRloss}_l^{\text{ум.}} \right|,$$

$$RMSE = \sqrt{\frac{1}{N_{\text{ш}}} \sum_{l=1}^{N_{\text{ш}}} (\text{SNRloss}_l^{\text{оцен.}} - \text{SNRloss}_l^{\text{ум.}})^2},$$

$$MAPE = \frac{100\%}{N_{\text{ш}}} \sum_{l=1}^{N_{\text{ш}}} \frac{\left| \text{SNRloss}_l^{\text{оцен.}} - \text{SNRloss}_l^{\text{ум.}} \right|}{\text{SNRloss}_l^{\text{ум.}}},$$

Таблица 1. Параметры линейной регрессии для взаимосвязи SNR_{loss} и SNR_{loss}^*

Тип шума	b_1	b_0	CKO	R^2
АБГШ (AWGN)	0,891	0,043	0,010	0,995
Car interior noise	0,894	0,039	0,010	0,997
Destroyer engine room noise	0,892	0,042	0,015	0,991
Destroyer operations room noise	0,879	0,056	0,011	0,995
F-16 cockpit noise	0,891	0,040	0,011	0,995
Factory floor noise 1	0,877	0,068	0,015	0,990
Factory floor noise 2	0,872	0,057	0,012	0,995
Jet cockpit noise 1	0,891	0,042	0,010	0,995
Jet cockpit noise 2	0,889	0,046	0,010	0,995
Military vehicle noise	0,870	0,057	0,011	0,996
Speech babble	0,889	0,070	0,022	0,982
Tank noise	0,869	0,055	0,010	0,997

Таблица 2. Точность неэталонной оценки значений SNR_{loss} при использовании разных подходов

Тип шума	Средняя относительная ошибка, %		
	Подход 1	Подход 2	Подход 3
АБГШ (AWGN)	1,05	1,07	1,05
Car interior noise	1,87	2,43	1,60
Destroyer engine room noise	1,58	1,65	1,57
Destroyer operations room noise	1,27	1,17	1,14
F-16 cockpit noise	1,18	1,34	1,13
Factory floor noise 1	2,11	1,87	1,56
Factory floor noise 2	1,42	1,40	1,28
Jet cockpit noise 1	1,06	1,12	1,06
Jet cockpit noise 2	1,09	1,09	1,08
Military vehicle noise	1,35	1,36	1,17
Speech babble	3,55	3,22	2,20
Tank noise	1,43	1,48	1,26
В среднем	1,58	1,60	1,34

где $SNR_{loss}_i^{оцен}$ – неэталонная оценка значения SNR_{loss} для i -го сигнала, $SNR_{loss}_i^{ист}$ – истинное значение SNR_{loss} для i -го сигнала, $N_{зн.}$ – общее число сигналов, для которых измеряются значения SNR_{loss} .

Результаты, полученные с использованием моделирования, позволяют сравнить точность предложенных подходов (табл. 2). Как и ожидалось, третий подход обеспечивает наилучшую точность. Первый и второй подходы обладают сопоставимой точностью. Версию метода, использующую индивидуальные коэффициенты уравнения регрессии, стоит использовать в случаях, когда есть возможность оценить коэффициенты уравнения (1), соответствующие конкретным шумовым условиям. В случаях, когда шумовые условия заранее не известны или тип шума может существенно меняться с течением времени, применение третьего подхода невозможно и рекомендуется применять первый или второй подход.

Дальнейшее совершенствование предложенной методики

После того как моделирование с различными типами шумов подтвердило жизнеспособность предлагаемой неэталонной методики, встал вопрос ее дальнейшего совершенствования. Анализ количества операций сложения и умножения, необходимых для неэталонной оценки разборчивости речевого сигнала, показал, что значительная их часть приходится на быстрое преобразование Фурье (БПФ) и обратное БПФ (ОБПФ) [24]. Од-

нако используемая схема вычислений (рис. 3) содержит операции, которые можно исключить. При этом крайне важным является приведение параметров спектрального преобразования в блоке шумоподавления к тем, которые используются в методе SNR loss, так как они влияют на достоверность оценки разборчивости. Схема предлагаемой модификации, основанной на исключении повторных и взаимоисключающих операций, имеет более простую структуру и содержит один блок вычисления БПФ (рис. 4).

Предложенная модификация проанализирована и сравнена с исходной методикой по числу требуемых вычислительных операций. Выявлено, что предложенная модификация позволяет уменьшить количество операций примерно на 30,5 % по сравнению с показателями исходной методики [24].

Также проводилось сравнение скорости работы двух версий методик на основе моделирования в среде Matlab. Для этого использовался персональный компьютер следующей конфигурации: Intel(R) Pentium(R) D 930 CPU 3.00GHz, RAM 3,00GB 400MHz, 64-bit operating system Windows 8. В результате выигрыш предлагаемой модификации по времени составил 29,9 % [24].

Как показало моделирование, предложенная модификация обладает несколько меньшей точностью. Результаты сравнения по показателям MAE, RMSE и MAPE представлены на рис. 5. На графиках приводятся усредненные по всем типам шумов характеристики.

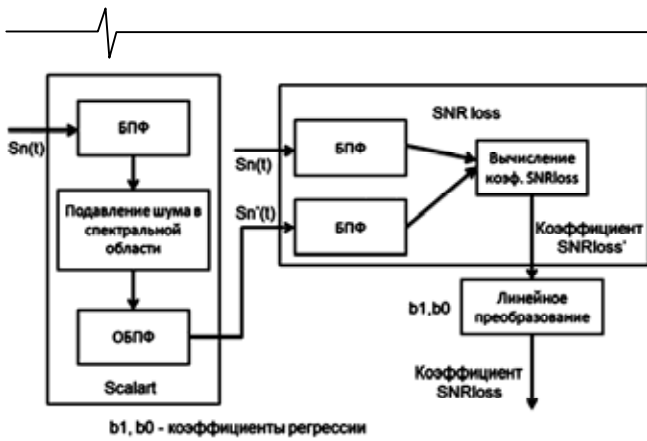


Рис. 3. Блок-схема исходной нестандартной методики оценки разборчивости

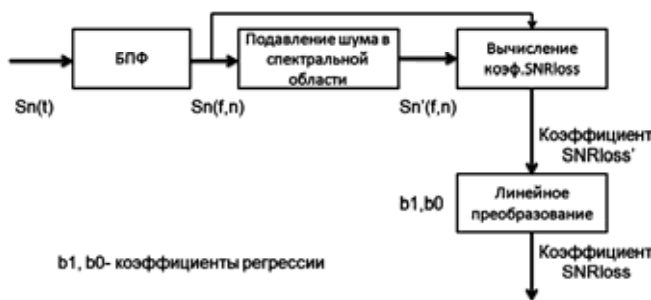


Рис. 4. Блок-схема модифицированной методики

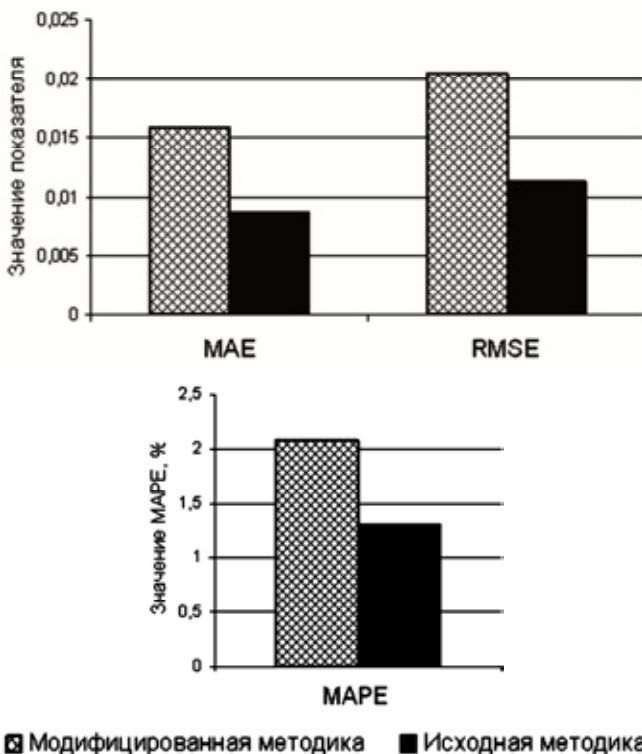


Рис. 5. Сравнение точности методик на основе показателей MAE, RMSE и MAPE

Представленные диаграммы демонстрируют, что точность оценок предложенной модификации несколько уступает точности исходной методики. Отношение показателей точности для исходной методики и ее модификации не столь информативны, как абсолютное изменение этих показателей. Например, средняя ошибка, выраженная в процентах (MAPE), выросла примерно с

1,3 % до 2,1 %. Относительное изменение в 1,6 раза кажется значительным, однако на практике указанное ухудшение характеристик вполне приемлемо. Кроме того, не исключено, что дальнейшая доработка предлагаемой модификации позволит ещё сильнее приблизиться по точности к исходной методике.

Заключение

Таким образом, предложена нестандартная версия меры разборчивости зашумленных речевых сигналов SNR loss, базирующаяся на применении исходной (эталонной) версии метода SNR loss, метода шумоподавления и парной регрессии. Исследована точность предложенной нестандартной меры разборчивости в случае воздействия на речевые сигналы шумов различных типов. Полученные результаты свидетельствуют об относительно высокой точности предложенной методики нестандартной оценки (средняя относительная ошибка составляет 1,05–3,55%).

Также представлена более быстрая версия нестандартной методики. Из результатов исследований следует, что модификация имеет достаточно высокую точность, которая незначительно уступает оригиналу, при этом скорость работы возросла за счет значительного сокращения числа осуществляемых вычислительных операций. Предложенная модифицированная методика может применяться для автоматического управления системами шумоподавления, а так же для выбора режима передачи, обеспечивающего допустимый уровень разборчивости, в системах связи.

Часть результатов, представленных в работе, получена при поддержке гранта РФФИ № 14-07-31335.

Литература

1. Дидковский В.С., Дидковская М.В., Продеус А.Н. Акустическая экспертиза каналов речевой коммуникации. Монография. – Киев: Имэкс-ЛТД, 2008. 420 с.
2. Новоселов С.А., Топников А.И., Савватин А.И., Приоров А.Л. Подавление шума в речевых сигналах на основе метода нелокального усреднения // Цифровая обработка сигналов. 2011. №4. – С. 23-28.
3. Collard J. A theoretical study of the articulation and intelligibility of a telephone circuit // Electrical Communication. 1929. V. 7. – P. 168.
4. French N.R., Steinberg J.C. Factors governing the intelligibility of speech sounds // The journal of the Acoustical Society of America. 1947. V. 19. Is. 1. – P. 90-119.
5. Kryter K.D. Methods for the calculation and use of the articulation index // The Journal of the Acoustical Society of America. 1962. V. 34. Is. 11. – P. 1689-1697.
6. Kryter K.D. Validation of the articulation index // The Journal of the Acoustical Society of America. 1962. V. 34. Is. 11. – P. 1698-1702.
7. Ma J., Loizou P. SNR loss: a new objective measure for predicting the intelligibility of noise-suppressed speech // Speech Communication. 2011. V. 53. Is. 3. – P. 340-354.
8. Loizou P., Ma J. Extending the articulation index to account for non-linear distortions introduced by noise-suppression algorithms // Journal of the Acoustical Society of America. 2011. V. 130. Is. 2. P. 986–995.

9. Быков Ю.С. Теория разборчивости речи и повышение эффективности радиотелефонной связи. – М.-Л.: Госэнергоиздат, 1959. 350 с.
10. Покровский Н.Б. Расчет и измерение разборчивости речи. – М.: Связьиздат, 1962. 390 с.
11. Сапожков М.А. Речевой сигнал в кибернетике и связи. – М.: Связьиздат, 1963. 452 с.
12. Сапожков М.А., Михайлов В.Г. Вокодерная связь. – М.: Радио связь, 1983. – 248 с.
13. Houtgast T., Steeneken H.J.M. Evaluation of speech transmission channels by using artificial signals // Acta Acustica united with Acustica. 1971. V. 25. Is. 6. P. 355-367.
14. Steeneken H.J.M., Houtgast T. A physical method for measuring speech transmission quality // The Journal of the Acoustical Society of America. 1980. V. 67. Is. 1. P. 318-326.
15. Steeneken H.J.M., Houtgast T. Validation of the revised STIr method // Speech Communication. 2002. V. 38. Is. 3. – P. 413-425.
16. Prodeus A. On possibility of advantages join of formant and modulation methods of speech intelligibility evaluation // Proceedings of VII International Conference on Perspective Technologies and Methods in MEMS Design (MEMSTECH). 2010. P. 254-259.
17. Prodeus A. Assessment of Speech Intelligibility by Formant-Modulation Method // Journal of Basic and Applied Physics. 2013. V. 2 Is. 5. P. 10-18.
18. Falk T.H., Zheng C., Chan W.Y. A non-intrusive quality and intelligibility measure of reverberant and dereverberated speech // IEEE Transactions on Audio, Speech, and Language Processing. 2010. V. 18. Is. 7. – P. 1766-1774.
19. Santos J.F., Senoussaoui M., Falk T.H. An improved non-intrusive intelligibility metric for noisy and reverberant speech // 14th International Workshop on Acoustic Signal Enhancement (IWAENC). 2014. – P. 55-59.
20. Li F.F. Speech intelligibility of VoIP to PSTN interworking – a key index for the QoS. // Proceedings of Interna-

tional Conference «Telecommunications Quality of Service: The Business of Success». 2004. – P. 104-108.

21. Chen F., Hazrati O., Loizou P. C. Predicting the intelligibility of reverberant speech for cochlear implant listeners with a non-intrusive intelligibility measure // Biomedical signal processing and control. 2013. V. 8. Is. 3. – P. 311-314.

22. Веселов И.А., Новоселов С.А., Топников А.И. Методика неэталонной оценки разборчивости зашумленных речевых сигналов // 15-я международная конференция «Цифровая обработка сигналов и её применение – DSPA-2013». М., 2013. Т. 1. – С. 256-259.

23. Scalart P., Filho J. Speech enhancement based on a priori signal to noise estimation // IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP-96). 1996. V. 2. – P. 629-632.

24. Нестеров М.С., Топников А.И. Модификация методики неэталонной оценки разборчивости речевых сигналов // 16-я международная конференция «Цифровая обработка сигналов и её применение – DSPA-2014». М., 2014. Т. 1. – С. 208-212.

NON-INTRUSIVE ESTIMATION OF NOISY SPEECH SIGNAL INTELLIBILITY

Topnikov A., Nesterov M., Novoselov S., Priorov A.

It is proposed non-intrusive technique for estimation of noisy speech signals intelligibility. Technique is based on the modern intelligibility measure SNR loss and Scalart speech denoising method. It is provided a comparison of the accuracy of the proposed technique with accuracy of the original reference SNR loss method for speech signals which are noised by different types of noise. The modification of proposed technique is described also that spent allows to reduce the number of computational operations.

Н О В Ы Е К Н И Г И

Вашкевич М.И., Азаров И.С., Петровский А.А.

Косинусно-модулированные банки фильтров с фазовым преобразованием: реализация и применение в слуховых аппаратах: научн. изд. – М.: Горячая линия – Телеком, 2014. – 212 с.

Рассмотрены вопросы построения цифровых слуховых аппаратов. В общем виде описаны технические аспекты коррекции нарушений. Большая часть материала посвящена разработке и эффективной реализации неравнополосных банков фильтров, играющих ключевую роль при проектировании цифровых слуховых аппаратов. Рассмотрены вопросы построения алгоритмов адаптивной компенсации акустической обратной связи, компрессии и шумоподавления, необходимых для реализации основных функций слухового аппарата.

Для научных и инженерно-технических работников, специализирующихся в области проектирования устройств цифровой обработки сигналов, биомедицинской техники, а также студентов и аспирантов соответствующих специальностей.



УДК 621.397.6

ПРИНЦИПЫ ВЫЧИСЛИТЕЛЬНОГО АНАЛИЗА ХАРАКТЕРИСТИК КВАНТОВАНИЯ СПЕКТРА БЛОКОВ ПИКСЕЛЕЙ СИГНАЛОВ ИЗОБРАЖЕНИЙ В СИСТЕМАХ ЦИФРОВОГО ТЕЛЕВИДЕНИЯ

Безруков В.Н., д.т.н., профессор, заведующий кафедрой телевидения и звукового вещания Московского технического университета связи и информатики, e-mail: ldty@mail.ru.

Ключевые слова: цифровое телевидение, сигнал изображения, сжатие, многомерный спектр, квантование, блок пикселей, вычислительный анализ.

Введение

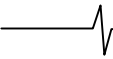
Проблема адаптивного сжатия многомерного спектра сигналов телевизионных (ТВ) изображений представляет большой теоретический и практический интерес для разработчиков и производителей оборудования цифровых систем передачи видеoinформации. Устранение избыточности в данном случае тесно связано с визуальным качеством передачи текущих изменений пространственной (внутрикадровой), межкадровой и цветовой структур сигналов изображений [1, 2, 3, 8, 9]. Увеличение степени её устранения или огрубления существенно влияет на точность передачи указанных изменений, что вызывает появление заметных для визуального контроля искажений пространственно-временной и цветовой структур ТВ изображений в условиях, когда фиксирована и является относительно низкой скорость выходного цифрового потока.

Метод сжатия спектра с фрагментацией пространственно-временной структуры сигнала ТВ изображения независимыми блоками пикселей (отсчётов) в реальном пространстве (в пределах одного и последовательности кадров) определяет раздельное сжатие спектра каждого из независимых блоков, с допустимыми при декодировании и воспроизведении погрешностями, отражающими в своей совокупности текущую пространственно-временную структуру исходных сигналов ТВ изображений.

В стандарте JPEG, например, последовательностями независимых блоков представлена структура общего ансамбля пикселей (отсчётов) изображений в пределах каждого отдельного кадра. Далее отсчёты каждого блока накапливают и подвергают ДКП с последующим, учитывающим особенности характеристик зрительной системы, частотно-зависимым взвешиванием полученных коэффициентов спектра в пределах соответствующей матрицы квантования. Дальнейшая селекция последних, по абсолютному значению уровня, обеспечивает в ходе работы систем цифрового телевидения адаптивную регулировку степени сжатия и сохранение (с учетом энтропийного кодирования и обратного частотного преобразования) наиболее заметных и интенсивных составляющих спектра сигналов изображений в условиях фиксированной скорости сформированного цифрового потока. Од-

Матричное представление таблиц частотного взвешивания при квантовании коэффициентов спектра в стандартах JPEG, MPEG затрудняет выявление специфики данного преобразования и его связь с характеристиками зрения. Реализована разработка принципов вычислительного анализа таких связей с использованием варианта таблицы частотного взвешивания блоков сигналов яркости, предложенного группой JPEG. Осуществлён вычислительный анализ двумерной импульсной характеристики, соответствующей данному варианту таблицы, и её сечений. Предложено использовать дискретизацию при преобразовании указанной импульсной характеристики или её сечений в частотное пространство для вычислительного анализа соответствующих одномерных (двумерных) функций коэффициентов передачи.

нако увеличение степени сжатия приводит к снижению точности передачи и к появлению нарушений на границах сопряжения смежных блоков текущей структуры сигнала изображения. Причиной является соответствующее снижение числа передаваемых уровней (огрубление) шкалы квантования коэффициентов ДКП. С учетом того, что наиболее высокочастотные коэффициенты имеют относительно большие коэффициенты матрицы квантования, то, по мере увеличения степени сжатия, полученные в результате взвешивания низкоинтенсивные высокочастотные составляющие спектра всё в большей мере обнуляются, обнуляются и другие, изначально относительно малого уровня, составляющие пространственного спектра. В связи с этим в пределах смежных блоков, чаще всего, имеют место индивидуальные трансформации формы функции устранения составляющих частот пространственного спектра сигнала изображения, которые определяются спецификой распределения их уровней. В соответствии с регулируемыми величинами шага шкалы квантования коэффициентов различной частоты, при текущем увеличении степени сжатия в структуре пространственного спектра сигнала ТВ изображений возрастает степень деградации всех частотных составляющих малой интенсивности. По мере дальнейшего увеличения шага шкалы квантования полностью исчезают низкочастотные и всё более высокие пространственные частотные составляющие относительно большой амплитуды. Такое нелинейное стробирование коэффициентов пространственного спектра сигнала изображения и определяет специфику искажений, возникающих в пределах каждого отдельного блока. В структуре распределения спектральных составляющих соответствующей матрицы в результате этого появляются ступенчатые перепады, ортогональные плоскости $\omega_x \ominus \omega_y$ (при строби-



ровании до нулевого значения). Появление последних отражается в реальном пространстве возникновением сопутствующих (колебательного типа) изменений среднего уровня и выбросов в структуре совокупности пикселей декодированного блока сигналов изображений, которые и провоцируют указанную выше заметность последовательностей блоков и их границ в локальных участках внутрикадрового пространства.

Известно, что анизотропия усреднённой пространственной спектральной интенсивности реальных изображений и пространственной частотной характеристики зрения имеют ромбовидную конфигурацию области граничных частот двумерного пространственного спектра [4, 5]. При этом максимальная протяжённость указанного пространственного спектра ориентирована в горизонтальном и вертикальном направлениях внутрикадрового пространства. Интегральные искажения (относительно резких ограничений) таких высокочастотных пространственных составляющих сигналов изображений, возникающих при сжатии, существенно влияют на субъективное качество восприятия ТВ изображений. Соответственно падает чёткость декодированных ТВ изображений, появляются выбросы на резких границах [3, 9]. Меньшее влияние на визуальное качество изображения, в соответствии с указанной выше спецификой пространственной спектральной интенсивности реальных изображений и пространственной частотной характеристики зрения, оказывает устранение малоинтенсивных высокочастотных пространственных составляющих в диагональном направлении.

Вычислительный анализ специфики импульсной характеристики частотного взвешивания коэффициентов пространственного спектра блоков сигналов яркости по стандарту JPEG

В процессе квантования, чаще всего, имеет место известное [8, 9] преобразование коэффициентов ДКП отдельных блоков в соответствии с соотношением:

$$C_{q(m,n)} = \text{Round} \left(\frac{C(m,n)}{\rho Q(m,n)} \right),$$

где $Q(m,n)$ – распределение весовых коэффициентов квантования, задаваемых в виде функции или таблицы целых чисел; ρ – параметр, определяющий степень сжатия изображения; Round – операция округления до ближайшего целого значения; $C(m,n)$ – исходные коэффициенты ДКП, соответствующие данному блоку; $C_q(m,n)$ – полученные в результате выполнения квантования коэффициенты ДКП, которые могут быть как положительными, так и отрицательными. Как показали проведённые группой JPEG исследования, более высокую эффективность сжатия обеспечивает применение в качестве распределения $Q(m,n)$ эмперически полученных таблиц в матричном представлении. Информация об используемой таблице квантования передаётся на приёмную сторону в общем потоке данных. Матричное представление данных таблиц затрудняет выявление специфики частотного взвешивания, реализуемого с помощью этих таблиц, и её связь с характеристиками зрения. Разра-

ботка методов вычислительного анализа таких связей является целью данной статьи.

В общем виде строки таблицы частотного взвешивания представим в виде матрицы с числом элементов $\bar{\alpha}_{\pm\mu, \pm\chi}$, равным $(2n_x + 1) \times (2n_y + 1)$, следующим образом:

$$\begin{matrix} \bar{\alpha}_{-n_x, n_y} & \bar{\alpha}_{-n_x+1, n_y} & \bar{\alpha}_{-n_x+2, n_y} & \dots & \bar{\alpha}_{0, n_y} & \dots & \bar{\alpha}_{n_x-2, n_y} & \bar{\alpha}_{n_x-1, n_y} & \bar{\alpha}_{n_x, n_y} & \dots \\ \bar{\alpha}_{-n_x, n_y-1} & \bar{\alpha}_{-n_x+1, n_y-1} & \bar{\alpha}_{-n_x+2, n_y-1} & \dots & \bar{\alpha}_{0, n_y-1} & \dots & \bar{\alpha}_{n_x-2, n_y-1} & \bar{\alpha}_{n_x-1, n_y-1} & \bar{\alpha}_{n_x, n_y-1} & \dots \\ \bar{\alpha}_{-n_x, n_y-2} & \bar{\alpha}_{-n_x+1, n_y-2} & \bar{\alpha}_{-n_x+2, n_y-2} & \dots & \bar{\alpha}_{0, n_y-2} & \dots & \bar{\alpha}_{n_x-2, n_y-2} & \bar{\alpha}_{n_x-1, n_y-2} & \bar{\alpha}_{n_x, n_y-2} & \dots \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ \bar{\alpha}_{-n_x, n_y-2} & \bar{\alpha}_{-n_x+1, 0} & \bar{\alpha}_{-n_x+2, 0} & \dots & \bar{\alpha}_{0, 0} & \dots & \bar{\alpha}_{n_x-2, 0} & \bar{\alpha}_{n_x-1, 0} & \bar{\alpha}_{n_x, 0} & \dots \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ \bar{\alpha}_{-n_x, -n_y+2} & \bar{\alpha}_{-n_x+1, -n_y+2} & \bar{\alpha}_{-n_x+2, -n_y+2} & \dots & \bar{\alpha}_{0, -n_y+2} & \dots & \bar{\alpha}_{n_x-2, -n_y+2} & \bar{\alpha}_{n_x-1, -n_y+2} & \bar{\alpha}_{n_x, -n_y+2} & \dots \\ \bar{\alpha}_{-n_x-1, -n_y+2} & \bar{\alpha}_{-n_x, -n_y+2} & \bar{\alpha}_{-n_x+1, -n_y+2} & \dots & \bar{\alpha}_{-n_x+2, -n_y+2} & \dots & \bar{\alpha}_{-n_x+3, -n_y+2} & \bar{\alpha}_{-n_x+4, -n_y+2} & \bar{\alpha}_{-n_x+5, -n_y+2} & \dots \\ \dots & \bar{\alpha}_{0, -n_y+1} & \bar{\alpha}_{n_x-2, -n_y+1} & \bar{\alpha}_{n_x-1, -n_y+1} & \bar{\alpha}_{n_x, -n_y+1} & \bar{\alpha}_{n_x+1, -n_y+1} & \bar{\alpha}_{n_x+2, -n_y+1} & \bar{\alpha}_{n_x+3, -n_y+1} & \bar{\alpha}_{n_x+4, -n_y+1} & \dots \\ \bar{\alpha}_{-n_x+2, -n_y} & \bar{\alpha}_{-n_x+1, -n_y} & \bar{\alpha}_{-n_x, -n_y} & \bar{\alpha}_{-n_x-1, -n_y} & \bar{\alpha}_{-n_x-2, -n_y} & \bar{\alpha}_{-n_x-3, -n_y} & \bar{\alpha}_{-n_x-4, -n_y} & \bar{\alpha}_{-n_x-5, -n_y} & \bar{\alpha}_{-n_x-6, -n_y} & \dots \end{matrix} \quad (1)$$

Чётный характер табл. 1 частотного взвешивания определяет равенство численных значений уровня её элементов $\bar{\alpha}_{\pm\mu, \pm\chi}$, симметрично расположенных по отношению к позиции элементов $\bar{\alpha}_{0,0}$ и центральных строки и столбца. Относительные значения элементов табл. 1 получим делением каждого элемента на значение $\bar{\alpha}_{0,0}$ и затем найдём преобразованные значения всех элементов $\alpha_{\pm\mu, \pm\chi}$ таблицы, в которой каждый из них определяется соответствующим частным от деления единицы на полученное относительное значение.

Таблица 1

16	11	10	16	24	40	51	61
12	12	14	19	26	58	60	55
14	13	16	24	40	57	69	56
14	17	22	29	51	87	80	62
18	22	37	56	68	109	103	77
24	35	55	64	81	104	113	92
49	64	78	87	103	121	120	101
72	92	95	98	112	100	103	99

Таблица 2

1.000	1.454	1.600	1.000	0.667	0.400	0.314	0.262
1.333	1.333	1.143	0.842	0.616	0.276	0.267	0.291
1.143	1.231	1.000	0.667	0.400	0.281	0.232	0.286
1.143	0.941	0.727	0.551	0.314	0.184	0.200	0.258
0.889	0.727	0.432	0.286	0.235	0.147	0.155	0.208
0.667	0.457	0.286	0.250	0.197	0.154	0.142	0.174
0.326	0.250	0.205	0.184	0.155	0.132	0.133	0.158
0.222	0.174	0.168	0.163	0.143	0.16	0.155	0.162

Нижний правый квадрант (оси координат делят общую плоскость на четверти или квадранты) варианта таблицы частотного взвешивания коэффициентов пространственного спектра блоков сигналов яркости с включением соответствующих элементов участков центральной строки и центрального столбца, предложенный

группой JPEG, представлен в табл. 1. Преобразованные значения $\alpha_{\mu, -\chi}$ элементов правого нижнего квадранта табл. 1, показанные в табл. 2, иллюстрируют степень изменения значений коэффициентов с увеличением значений пространственных частот по горизонтальному, вертикальному и другим направлениям. По существу, все четыре квадранта, соответствующие табл. 2, определяют совокупностью своих элементов $\alpha_{\pm j, \pm i}$ дискретные отсчеты характеристики коэффициента передачи фильтра низких пространственных частот, которая (в идеализированном случае) может быть отражена соответствующей совокупностью взвешенных дельта-функций

$$S_d(\omega_x, \omega_y) = \sum_{\mu=-n_x}^{\mu=n_x} \sum_{\chi=-n_y}^{\chi=n_y} \alpha_{\pm\mu, \pm\chi} \delta(\omega_x - \mu\omega_{x0}) \delta(\omega_y - \chi\omega_{y0}). \quad (2)$$

Соответствующая импульсная характеристика определяется соотношением

$$g_0(t_x, t_y) = \sum_{\mu=-n_x}^{\mu=n_x} \sum_{\chi=-n_y}^{\chi=n_y} \alpha_{\pm\mu, \pm\chi} \exp(-jt_x \mu\omega_{x0} - jt_y \chi\omega_{y0}). \quad (3)$$

С учётом значения частоты дискретизации и числа отсчётов в блоке ТВ сигнала, определим в частотном пространстве $f_{x0} = f_{y0} = 0.9 \cdot 10^6$ условной единицы. Общая импульсная характеристика (ИХ) может быть, с некоторым приближением, вычислена в данном случае как сумма характеристик, соответствующих строкам преобразованной матрицы (1). При этом общая ИХ такого фильтра может быть представлена в виде, например, суммы составляющих, соответствующих строкам всех четырёх квадрантов. С учётом чётности преобразованной матрицы и значений $\alpha_{\pm j, \pm i}$, входящих в выражение (3), каждая из таких составляющих выражается суммой соответствующих косинусов. Следовательно, имеют место следующие составляющие $g_k(t_x, t_y)$ для общего (интегрального) выражения:

$$\begin{aligned} g_{00}(0, 0) &= 1 \\ g_{10}(t_x, 0) &= 2 \cdot 1.454 \cos(2\pi f_{0x} t_x) + 2 \cdot 1.6 \cos(2\pi 2 f_{0x} t_x) + \dots + 2 \cdot 0.314 \cdot \cos(2\pi 6 f_{0x} t_x) + 2 \cdot 0.262 \cos(2\pi 7 f_{0x} t_x) \\ g_{01}(0, t_y) &= 2 \cdot 1.333 \cos(2\pi f_{0y} t_y) + 2 \cdot 1.143 \cos(2\pi 2 f_{0y} t_y) + \dots + 2 \cdot 0.326 \cos(2\pi 6 f_{0y} t_y) + 2 \cdot 0.222 \cos(2\pi 7 f_{0y} t_y) \\ g_2(t_x, t_y) &= 2 \cdot 1.333 (\cos(2\pi f_{0x} t_x + 2\pi f_{0y} t_y) + \cos(2\pi f_{0x} t_x - 2\pi f_{0y} t_y)) + 2 \cdot 1.143 (\cos(2\pi 2 f_{0x} t_x + 2\pi f_{0y} t_y) + \cos(2\pi 2 f_{0x} t_x - 2\pi f_{0y} t_y)) + \dots + 2 \cdot 0.291 (\cos(2\pi 7 f_{0x} t_x + 2\pi f_{0y} t_y) + \cos(2\pi 7 f_{0x} t_x - 2\pi f_{0y} t_y)) \\ g_3(t_x, t_y) &= 4 \cdot 1.231 (\cos(2\pi f_{0x} t_x + 2\pi 2 f_{0y} t_y) + \cos(2\pi f_{0x} t_x - 2\pi 2 f_{0y} t_y)) + 2 \cdot 1.0 (\cos(2\pi 2 f_{0x} t_x + 2\pi 2 f_{0y} t_y) + \cos(2\pi 2 f_{0x} t_x - 2\pi 2 f_{0y} t_y)) + \dots + 2 \cdot 0.286 (\cos(2\pi 7 f_{0x} t_x + 2\pi 2 f_{0y} t_y) + \end{aligned}$$

$$\begin{aligned} &+ \cos(2\pi 7 f_{0x} t_x - 2\pi 2 f_{0y} t_y)) \\ &\quad \vdots \\ g_8(t_x, t_y) &= 2 \cdot 0.174 (\cos(2\pi f_{0x} t_x + 2\pi 7 f_{0y} t_y) + \cos(2\pi f_{0x} t_x - 2\pi 7 f_{0y} t_y)) + 2 \cdot 0.168 (\cos(2\pi 2 f_{0x} t_x + 2\pi 7 f_{0y} t_y) + \cos(2\pi 2 f_{0x} t_x - 2\pi 7 f_{0y} t_y)) + \dots + \\ &+ 2 \cdot 0.162 (\cos(2\pi 7 f_{0x} t_x + 2\pi 7 f_{0y} t_y) + \cos(2\pi 7 f_{0x} t_x - 2\pi 7 f_{0y} t_y)) + \end{aligned} \quad (4)$$

Как уже отмечалось, результирующее выражение для ИХ в данном частном случае определяется суммой составляющих (4). Относительное значение данной расчётной функции запишем в следующем виде:

$$g_0(t_x, t_y) = \sum_{k=1}^{k=8} g_k(t_x, t_y) / \sum_{k=1}^{k=8} g_k(0, 0). \quad (5)$$

На рис.1, а) представлены результаты вычислений относительной результирующей ИХ частотного взвешивания коэффициентов пространственного спектра блоков сигналов яркости по стандарту JPEG, осуществлённых в соответствии с табл. 2 и соотношениями (4, 5). Вид результирующей ИХ в данном случае, с некоторым приближением, определяется первым лепестком вычисленной функции. Следует учитывать возможность реализации различной конфигурации сечений расчётной результирующей ИХ при анализе частных случаев. Можно выбирать, например, различные зависимости между координатами по осям $0t_x, 0t_y$. Другим вариантом сечений является выбор различных, например, последовательностей (распределений коэффициентов спектра) пространственных частот в соотношениях (4, 5) и, наконец, можно получать комбинации распределений по частотам и по осям $0t_x, 0t_y$. Простейшими реализациями первого варианта является фиксация $t_y = 0$; $t_x = 0$; $t_x = t_y$. Подстановка указанных значений $t_y = 0$; $t_x = 0$; $t_x = t_y$ в соотношения (4, 5) определяют соответствующие частные совокупности данных соотношений. Простейшими реализациями второго варианта является фиксация $\mu = 0$; $\chi = 0$; $\mu = \chi$ ($-\mu = -\chi$, $\mu = -\chi$, $-\mu = \chi$). Подстановка указанных, например, значений $\chi = 0$; $\mu = 0$; $\mu = \chi$ в соотношения (2, 3, 4, 5) позволит получить соответствующие другие частные совокупности данных соотношений. При выполнении условия $t_y = 0$ имеем в реальном пространстве зависимости преобразования общей совокупности соотношений. Данный частный случай с некоторым приближением определяет изменение, после частотного взвешивания (коэффициентов спектра) в пределах блока, формы протяжённых, вертикально ориентированных границ (линий).

Зависимость 1 на рис. 1, б) отражает специфику изменения формы ИХ для такого условия. Зависимость 2 на рис. 1, б) демонстрирует вариант выполнения условия $t_x = 0$, а зависимость 3 $t_x = t_y$. Присутствие достаточно протяжённых, трапецеидального вида, компонент в форме главного лепестка зависимостей 1 и 2 (рис. 1, б)), рассматриваемых реализаций ИХ, свидетельствует о наличии соответствующего подъёма перед квантованием относительного уровня составляющих

спектра воздействия, ограниченного блоком, в области низких пространственных частот. И в то же время, наличие относительно короткой, импульсного типа, компоненты главного лепестка в сочетании с колебательной отражает относительно высокую степень сохранения и даже наличие подъёма частотных составляющих в области верхних пространственных частот. Реализация ИХ, изображённая зависимостью 3 (рис. 1. б), свидетельствует об относительно плавном подавлении, в данном случае ($t_x = t_y$) частотного взвешивания, уровня составляющих пространственного спектра воздействия с увеличением пространственных частот.

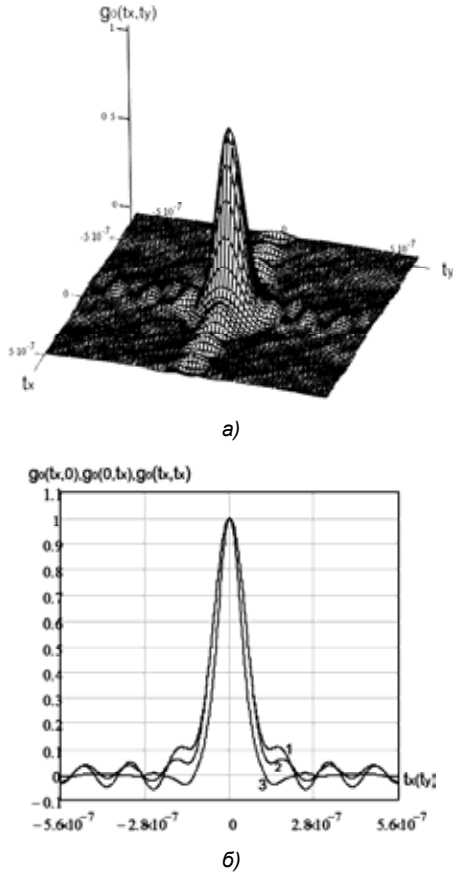


Рис. 1. Результаты вычисления относительной ИХ частотного взвешивания коэффициентов пространственно-го спектра блоков сигналов яркости по стандарту JPEG:

а) результирующая расчётная ИХ;

б) результаты расчётов сечений ИХ:

1 - при $t_y = 0$; 2 - при $t_x = 0$; 3 - при $t_x = t_y$.

Подстановка значений $\chi = 0$; $\mu = 0$; $\mu = \chi$ в (4, 5) определяется частными относительными соотношениями:

$$g_{0\mu}(t_x, 0) = \frac{g_{00}(0, 0) + g_{10}(t_x, 0)}{g_{00}(0, 0) + g_{10}(0, 0)};$$

$$g_{0\chi}(0, t_y) = \frac{g_0(0, 0) + g_{01}(0, t_y)}{g_0(0, 0) + g_{01}(0, 0)}; \quad g_{0\mu\mu}(t_x, t_y) = \frac{g_{\mu\mu}(t_x, t_y)}{g_{\mu\mu}(0, 0)},$$

где

$$g_{\mu\mu}(t_x, t_y) = g_{00}(0, 0) + 2 \cdot 1.333 \cos(2\pi f_{0x} t_x + 2\pi f_{0y} t_y) + 2 \cdot 1.0 \cos(2\pi 2 f_{0x} t_x + 2\pi 2 f_{0y} t_y) + 2 \cdot 0.552 \times \cos(2\pi 3 f_{0x} t_x + 2\pi 3 f_{0y} t_y) + 2 \cdot 0.235 \cos(2\pi 4 f_{0x} t_x + 2\pi 4 f_{0y} t_y) + 2 \cdot 0.154 \cos(2\pi 5 f_{0x} t_x + 2\pi 5 f_{0y} t_y) +$$

$$+ 2 \cdot 0.133 \cos(2\pi 6 f_{0x} t_x + 2\pi 6 f_{0y} t_y) + 2 \cdot 0.162 \times \cos(2\pi 7 f_{0x} t_x + 2\pi 7 f_{0y} t_y) \quad (6)$$

В первом случае предполагается наличие постоянно-го уровня составляющих спектра воздействия только по оси $0f_x$ и обнуление по всем другим координатам, во втором – только по оси $0f_y$, в третьем – только под со-рок пять градусов (по сечению через первый и третий квадранты).

Наличие соотношений для частных реализаций ИХ позволяет осуществить приблизительный анализ специфики соответствующих характеристик в частотном пространстве. Одним из относительно простых вариантов выполнения такого анализа является представление (в интервале главного лепестка) полученных сечений ИХ в виде последовательности дискретных отсчётов, следующих с достаточно высокой (в реальном пространстве) частотой, обеспечивающей низкую степень вносимых нелинейных искажений. При этом, за счёт ограничения интервала анализа в реальном пространстве имеет место выявление огибающей в частотной области. Задаваясь, например, интервалом дискретизации, равным в пространстве $T_x = 0,1 \cdot 10^{-7}$ условной единицы реального пространства, получим с использованием соотношений (4, 5, 6) совокупность отсчётов, соответствующих, например дискретизированным с заданным периодом, различным вариантам указанных выше функций сечений $g_0(t_x, 0)$, $g_{0\mu}(t_x, 0)$, $g_{0\mu\mu}(0,5t_x, 0,5t_y)$, представленных соответствующими зависимостями на рис. 1, б и 2, а. В частности, последовательность (в количестве 56) отсчётов, соответствующая дискретному варианту $g_0^d(t_x, 0)$ – зависимость 1 на рис. 1, б, может быть вычислена и представлена в следующем виде:

$$g_0^d(\pm 0, 0) = 1.0, \quad g_0^d(\pm 0.1 \cdot 10^{-7}, 0) = 0.983, \\ g_0^d(\pm 0.2 \cdot 10^{-7}, 0) = 0.932, \quad g_0^d(\pm 0.3 \cdot 10^{-7}, 0) = 0.853, \dots \\ g_0^d(\pm 1.1 \cdot 10^{-7}, 0) = 0.135, \dots, \quad g_0^d(\pm 1.4 \cdot 10^{-7}, 0) = 0.108, \\ g_0^d(\pm 1.5 \cdot 10^{-7}, 0) = 0.112, \quad g_0^d(\pm 1.6 \cdot 10^{-7}, 0) = 0.114, \\ g_0^d(\pm 1.7 \cdot 10^{-7}, 0) = 0.11, \dots, \quad g_0^d(\pm 2.2 \cdot 10^{-7}, 0) = -4,115 \cdot 10^{-3}, \\ \dots, \quad g_0^d(\pm 2.6 \cdot 10^{-7}, 0) = -0.05, \dots \\ g_0^d(\pm 2.9 \cdot 10^{-7}, 0) = 2,124 \cdot 10^{-3}, \dots, \quad g_0^d(\pm 3.3 \cdot 10^{-7}, 0) = 0.056, \\ \dots, \quad g_0^d(\pm 3.7 \cdot 10^{-7}, 0) = 1,12 \cdot 10^{-3}, \dots \\ g_0^d(\pm 4.1 \cdot 10^{-7}, 0) = -0.043, \dots \\ g_0^d(\pm 4.4 \cdot 10^{-7}, 0) = -8,451 \cdot 10^{-3}, \dots, \quad g_0^d(\pm 4.8 \cdot 10^{-7}, 0) = 0.041, \\ \dots, \quad g_0^d(\pm 5.3 \cdot 10^{-7}, 0) = -8,927 \cdot 10^{-3}, \dots \\ g_0^d(\pm 5.6 \cdot 10^{-7}, 0) = -0.027. \quad (7)$$

В частотном пространстве ограниченной по протяженности последовательности отсчётов (7) соответствует характеристика передачи:

$$S_0(f_x) = 1.0 + 2 \cdot 0.983 \cos(2\pi T_x f_x) + 2 \cdot 0.932 \cos(2\pi 2T_x f_x) + 2 \cdot 0.853 \cos(2\pi 3T_x f_x) + \dots \\ + 2 \cdot 0.11 \cos(2\pi 17T_x f_x) + \dots - 2 \cdot 0.004115 \cos(2\pi 22T_x f_x) - \\ - 2 \cdot 0.05 \cos(2\pi 26T_x f_x) - \dots + 2 \cdot 0.056 \cos(2\pi \cdot 33T_x f_x) + \dots \\ - 2 \cdot 0.043 \cos(2\pi 41T_x f_x) - \dots + 2 \cdot 0.041 \cos(2\pi 48T_x f_x) + \dots$$

$$-2 \cdot 0.027 \cos(2\pi 56 T_x f_x). \quad (8)$$

Аналогичным образом могут быть вычислены, согласно полученным выше аналитическим выражениям, последовательности отсчётов и характеристики передачи $S_{0\mu}(f_x)$ и $S_{0\mu\mu}(f_x, f_x)$, соответствующие функциям сечений $g_{0\mu}(t_x, 0)$ и $g_{\mu\mu}(t_x, t_y)$.

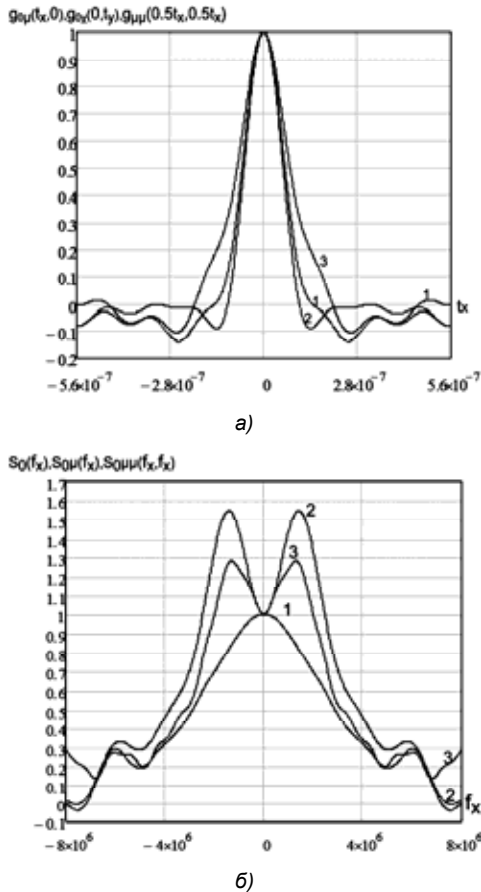


Рис. 2. Результаты вычисления сечений ИХ и характеристик передачи составляющих спектра частных воздействий:

а) расчётные сечения ИХ при $\chi = 0$, $\mu = 0$, $\mu = \chi$ - соответственно зависимости 1, 2, 3;

б) результаты расчётов характеристик передачи, полученных для сечений $g_0(t_x, 0)$, $g_{0\mu}(t_x, 0)$, $g_{0\mu\mu}(0.5t_x, 0.5t_y)$ - соответственно зависимости 1, 2, 3

Зависимостями на рис. 2, а) продемонстрированы расчётные варианты частных сечений $g_{0\mu}(t_x, 0)$; $g_{0\chi}(0, t_y)$; $g_{0\mu\mu}(t_x, t_y)$ при $\chi = 0$, $\mu = 0$ и $\mu = \chi$, зависимостями – рис. 2, б) – вычисленные характеристики передачи, соответствующие в частотном пространстве частным сечениям $g_0(t_x, 0)$, $g_{0\mu}(t_x, 0)$, $g_{0\mu\mu}(0.5t_x, 0.5t_y)$ соотношения (4) для ИХ.

Согласно результатам расчёта, представленным на рис. 1 и рис. 2, сечения ИХ, полученные для различных вариантов распределения (последовательностей) отсчётов спектра воздействия фиксированной амплитуды в пределах блока, существенно отличаются по виду и свидетельствуют о том, что при частотном взвешивании имеют место отличающиеся по типу преобразования исходного вида таких последовательностей. Указанные преобразования можно подразделить на варианты интегрального и варианты интегрально-дифференциального

типов. К первому типу можно отнести трансформации вида (сечения) ИХ, приведённые на рис. 1, б, а ко второму, с некоторым приближением, сечения ИХ, показанные на рис. 2, а. При этом, в первом случае, согласно зависимости 1 рис. 2, б, имеет место, при частотном взвешивании, относительный подъём составляющих спектра воздействия в области низких частот, который во втором случае (зависимости 2, 3 на рис. 2, б) сочетается с более высокой степенью подъёма в локальной зоне ($\approx (0.3 \div 1.5) \cdot 10^6$) низкочастотного участка спектра исходного воздействия.

Вычислительный анализ специфики двумерной характеристики частотного взвешивания коэффициентов пространственного спектра блоков сигналов яркости по стандарту JPEG

Совокупность дискретных элементов (табл. 2) частотного взвешивания определяет специфику распределённого в двумерном частотном пространстве амплитудного преобразования конкретной совокупности отсчётов (коэффициентов) спектра воздействия в пределах блока при квантовании. С учётом чётности, таблица частотного взвешивания, в идеализированном случае, может быть отражена соответствующей двумерной матрицей взвешенных дельта-функций (соотношение 2).

Выявление непрерывного вида дискретной характеристики частотного взвешивания $S_d(\omega_x, \omega_y)$ удобно осуществить на основе интегрального преобразования её дискретного варианта. Такое интегральное преобразование должно обеспечить эффективное выделение составляющих двумерной огибающей дискретного спектра и подавление гармоник дискретизации более высокого порядка. Для уменьшения протяжённости переходных процессов осуществим выделение упомянутой огибающей в соответствии с соотношением следующего вида:

$$S_n(\omega_x, \omega_y) = S_d(\omega_x, \omega_y) \otimes P(\omega_x, \omega_y) = \sum_{\mu=-n_x}^{\mu=n_x} \sum_{\chi=-n_y}^{\chi=n_y} \alpha_{\pm\mu, \pm\chi} P(\omega_x - \mu\omega_{x0}) P(\omega_y - \chi\omega_{y0}),$$

где

$$P(\omega_x, \omega_y) = \left[\frac{\sin(2\pi f_x T_x \cdot 7.5 + 10^{-16})^3}{(2\pi f_x T_x \cdot 7.5 + 10^{-16})} \right]^3 \times \left[\frac{\sin(2\pi f_y T_y \cdot 7.5 + 10^{-16})}{(2\pi f_y T_y \cdot 7.5 + 10^{-16})} \right]^3 \text{ и } T_x = T_y = 0.0465 \cdot 10^{-6}. \quad (9)$$

По существу вычислительное выявление непрерывного вида дискретной характеристики частотного взвешивания в данном случае, как следует из соотношения (9), выполняется за счёт позиционирования функции $P(\omega_x, \omega_y)$ с заданным (для конкретной точки наличия дискретных элементов табл. 2) весом в каждой точке четырёх квадрантов плоскости $\omega_x 0 \omega_y$. При этом, имеет место интегральное преобразование дискретной характеристики частотного взвешивания $S_d(\omega_x, \omega_y)$ в непрерывный вид за счёт заполнения соответствующей матрицы в плоскости $\omega_x 0 \omega_y$ совокупностью взвешенных функций $P(\omega_x, \omega_y)$.

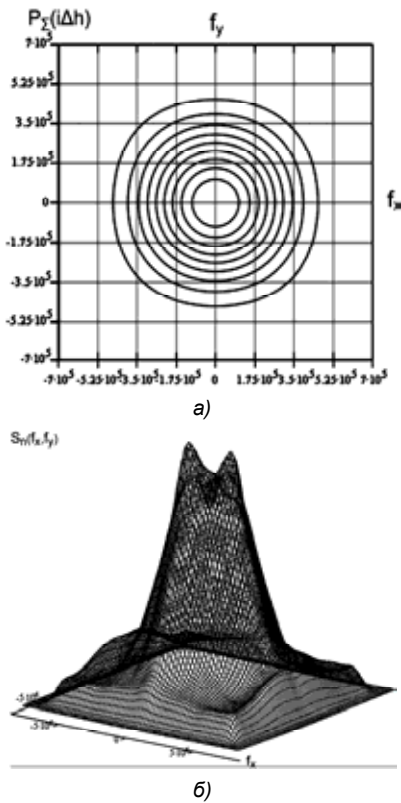


Рис. 3. Проекция совокупности на плоскость $f_x 0 f_y$

сечений фиксированных изменений уровня функции интегрального преобразования дискретной характеристики частотного взвешивания (а) и общего вида вычисленной интегрированием двумерной функции её огибающей (б)

Выбор вида функции $P(\omega_x, \omega_y)$ должен обеспечивать в данном случае минимальную степень интегральных искажений огибающей и в достаточной мере подавить остаточные составляющие дискретизации. Соответственно, анизотропия в форме функции $P(\omega_x, \omega_y)$ должна быть сопряжена со спецификой распределения дискретных элементов табл. 2 в квадрантах плоскости $\omega_x 0 \omega_y$, т.е. учитывать увеличение периода следования дискретных элементов таблицы в направлении (плоскость $\omega_x 0 \omega_y$) под сорок пять градусов. На рис. 3, а, б отображена специфика изменений $P(i\Delta h)$ проекций совокупности сечений (квантование по уровню) фиксированных (Δh) изменений уровня h выбранной функции $P(2\pi f_x, 2\pi f_y)$ на плоскость $f_x 0 f_y$ и общего вида двумерной функции огибающей дискретной характеристики частотного взвешивания $S_n(2\pi f_x, 2\pi f_y)$, вычисленных в соответствии с соотношениями 9.

По виду совокупности сечений выбранной (соотношение 9), рис. 3, а функции интегрального преобразования типа $P(f_x, f_y) = [\sin c(\omega_x T_{x0}) \sin c(\omega_y T_{y0})]^3$ можно установить, что для данной функции по сравнению, например, с вариантом $P(f_x, f_y) = [\sin c(\omega_x T_{x1}) \sin c(\omega_y T_{y1})]$ характерен достаточно высокий и относительно равномерный градиент падения уровня при увеличении аргументов, очень низкий (менее процента) уровень сопутствующих

разнополярных выбросов по осям $0 f_x$ и $0 f_y$ и, согласно последнему, незначительная эквивалентная длительность переходных процессов при реализации интегрирования. К тому же, радиус показанных на рис. 3 сечений несколько возрастает под углом сорок пять градусов в плоскости $f_x 0 f_y$. Последнее позволяет снизить искажения, сопутствующие реализуемому, при выделении огибающей, интегрированию, которые обусловлены увеличением расстояния между отсчётами (табл. 2) под углом, в плоскости $f_x 0 f_y$, сорок пять градусов.

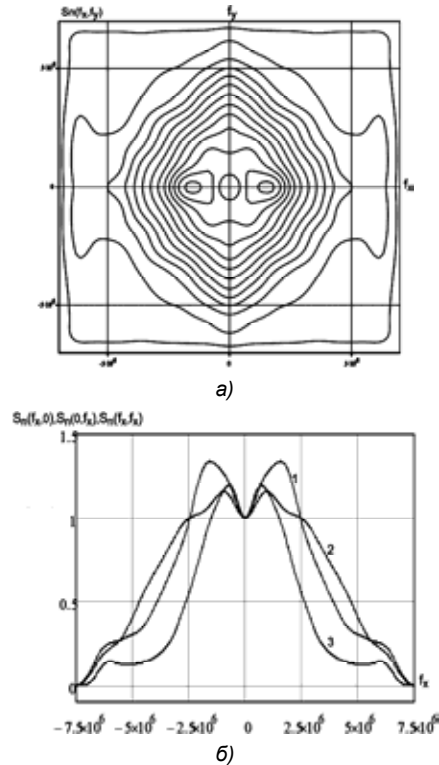


Рис.4. Проекция на плоскость $f_x 0 f_y$ совокупности сечений фиксированных изменений уровня функции интегральной огибающей характеристики частотного взвешивания – (а) и её сечений (б): 1 – $S_n(f_x, 0)$, $S_n(0, f_x)$, $S_n(f_x, f_x)$

На рис. 4, а показан результат вычисления проекций совокупности сечений (результат квантования по уровню) фиксированных изменений уровня расчётной функции огибающей дискретной характеристики частотного взвешивания $S_n(2\pi f_x, 2\pi f_y)$ на плоскость $f_x 0 f_y$; на рис. 4, б показаны зависимости 1, 2, 3 частных сечений данной огибающей, соответственно, при $f_y = 0$, $f_x = 0$ и $f_x = f_y$.

Результаты анализа. Заключение

Сопоставление расчётных данных и зависимостей, полученных выше при проведении вычислительного анализа, выявляет существенное влияние на результат реализации частотного взвешивания особенностей распределения коэффициентов спектра каждой частной реализации пространственной структуры изображения в пределах блока. Полученные, с использованием разработанных принципов вычислительного анализа соотношения и зависимости, представленные на рис. 1-4, отра-

зили жёсткую связь эмпирически полученных таблиц в матричном представлении с характеристиками зрительного анализатора человека [6, 7, 9]. Прежде всего следует отметить, что относительный подъём в области низких пространственных частот и соответствующее увеличение точности их передачи при сжатии обеспечивается при обратном частотном преобразовании весовую трансформацию пространственных составляющих спектра искажений сжатия и помех в область относительно высоких пространственных частот, т.е. в область пониженной заметности. Соответственно расчётным результатам, представленных зависимостями 1 рис. 1, б и рис. 2, б), наиболее высокая точность в части передачи низких пространственных частот обеспечивается при кодировании низкочастотных вертикально ориентированных составляющих значительной протяжённости. Именно такие составляющие существенно влияют на визуальную идентификацию объектов. Трансформация возникающих искажений сжатия (блочность) в область верхних пространственных частот, в данном случае, позволяет (из-за сокращения протяжённости переходов) эффективно использовать адаптивную сглаживающую фильтрацию для снижения заметности возникающих, в смежных блоках, составляющих блочной структуры и сопряжения пространственной структуры изображений на их границах. Следует также отметить наличие локальных корректирующих преобразований (относительный подъём составляющих спектра в диапазоне $\approx (0.3 \div 1.5) \cdot 10^6$ условной частоты) в области их максимальной заметности [6, 7, 9], что иллюстрируется зависимостью 2 на рис. 2, б, двумерной функцией огибающей на рис. 3, б и зависимостью 1 на рис. 4, б. При обратном частотном преобразовании имеет место сопутствующее подавление соответствующих составляющих спектра, возникших при реализации его сжатия.

По существу результат вычисления проекций на плоскость $f_x 0 f_y$ совокупности сечений фиксированных изменений уровня функции интегральной огибающей характеристики частотного взвешивания, представленный на рис. 4, а), отражает трансформации конфигурации границ области воспроизводимых (пропускания) пространственных частот при воздействии с фиксированным уровнем спектра и последовательном огрублении (возрастании степени сжатия) шкалы квантования. При выборе низкой степени сжатия имеет место существенное преимущество по сохранению (существенно влияющих на визуальное качество ТВ изображений) высокочастотных в горизонтальном направлении, но относительно низкочастотных в вертикальном, составляющих спектра реализаций в пределах блока. Следует отметить в данном случае и дополнительное увеличение точности передачи в вертикальном направлении наиболее высокочастотных составляющих горизонтального направления. В условиях последующего возрастания степени сжатия конфигурация границ области воспроизводимых пространственных частот, при воздействии с фиксированным уровнем пространственного спектра, приобретает, как и ожидалось, ромбовидный вид. При максимальной степени сжатия более высокая точность передачи обеспечивается, согласно рис. 4, а, в горизонтальном направлении, а конфигурации границ области воспроизводимых пространственных частот при воздей-

ствии с фиксированным уровнем спектра теряют ромбовидный характер по оси $0 f_y$.

По существу учтённые вариантом, подвергнутой данному вычислительному анализу эмпирической таблицы, особенности характеристик и функционирования зрительной системы человека не исчерпаны. Выполненные работы [7] показывают наличие здесь дополнительных резервов, использование которых может существенно увеличить степень сжатия спектров ТВ изображений в системах цифрового телевидения.

Литература

1. А.с. 748903 СССР Способ передачи и воспроизведения сигнала изображения / Безруков В.Н. – Заявл. 1976. – Опубл. 1980, Бюлл. № 26.
2. А.с.1211892 СССР Способ формирования телевизионного сигнала цветного изображения и устройство для реализации/Безруков В.Н. – Заявл. 1983. – Опубл. – 1986, Бюлл. № 6.
3. Безруков В.Н., Романов С.Г., Седов М.О. Адаптивность при сжатии спектра сигналов изображений в система цифрового телевидения // Электросвязь. – 2013. – № 3. – С.22-26.
4. Красильников Н.Н. Теория передачи и восприятия изображений. – М.: Радио и связь, 1986. – 248 с.
5. Безруков В.Н. О некоторых особенностях характеристик зрительной системы наблюдателя телевизионных изображений / Труды учебных институтов связи. – 1976. – Вып.74. – С.28 – 36.
6. Безруков В.Н., Королёв А.В., Ляпунов В.Н., Новаковская О.С. Выбор параметров систем телевидения высокой визуальной чёткости и качества // Техника кино и телевидения. – 1985, – №10, – С.22-28.
7. Безруков В.Н. Специфика видеоконтроля изображений вещательного телевидения, Материалы международного конгресса НАТ. – Москва, 2002, – С. 215-216.
8. Дворкович В.П., Дворкович А.В. Цифровые видеоинформационные системы (теория и практика). М.: Техносфера, 2012. – 1007 с
9. Безруков В.Н., Балобанов В.Г. Системы цифрового вещательного и прикладного телевидения. – М.: Горячая линия-Телеком, 2015. – 607 с.

PRINCIPLES FOR COMPUTATIONAL ANALYSIS OF IMAGE SIGNAL PIXEL BLOCK SPECTRUM QUANTIZATION CHARACTERISTICS IN DIGITAL TELEVISION SYSTEMS

Bezrukov V.N.

Matrix representation of frequency weights tables in spectrum coefficients quantation in JPEG and MPEG makes it difficult to find out specificity of that transform and its connection to the vision characteristics. The principles of computational analysis of these connections using the variant of frequency weights table for luma signal blocks proposed by JPEG are designed. The computational analysis of two-dimensional impulse response, corresponding to this table and its slices was held. It was proposed to use discretisation in transform of the mentioned impulse response or its slices into the frequency domain for computational analysis of corresponding one-dimensional (two-dimensional) functions of transform coefficients. Based on this integral transform the function of two-dimensional envelope is picked out and the analysis of considered frequency weights matrix table subtle features is held out. The rigid connection of this empirical variant of table with spatial frequency characteristics of human vision system is shown.

ОЦЕНКА КАЧЕСТВА РАБОТЫ ТЕЛЕВИЗИОННЫХ (ТЕПЛОВИЗИОННЫХ) КООРДИНАТОРОВ В СИСТЕМАХ НАБЛЮДЕНИЯ И БЕЗОПАСНОСТИ

Зубарев Ю.Б., д.т.н., чл.-кор. РАН, советник генерального директора, ЗАО «МНИТИ», e-mail: osa@mniti.ru;

Арзуманян Э.П., д.т.н., главный научный сотрудник, ЗАО «МНИТИ», e-mail: arzedouard@gmail.com;

Миколайчук Д.В., начальник сектора, ЗАО «МНИТИ», e-mail: amid_mik@mail.ru.

Ключевые слова: цифровые координаты, видеосигнал, чересстрочная развертка телевизионного изображения.

Введение

В настоящее время широкое распространение получили системы наблюдения и безопасности, использующие телевизионные или тепловизионные датчики для съема сигналов изображения. Со временем требования к таким системам повышаются, в частности, в сторону расширения области применения и увеличения функциональных возможностей. Одним из интересных направлений их применения являются мобильные системы, предназначенные для оперативной съемки локальных событий. В этом случае критерием качества работы комплекса является обеспечение минимальной динамической ошибки при максимальной угловой скорости оптической оси оптико-электронной системы (ОЭС), применяемой для широкого обзора местности. В данной работе рассматривается контур слежения с астатизмом первого порядка. Следящие контуры с более высокими порядками астатизма весьма чувствительны к флуктуационной составляющей сигналов, управляющих электроприводом ОЭС. Эта чувствительность особенно свойственна системам, использующим принцип управления следящим приводом по растровому изображению.

Постановка задачи

В цифровых координаторах из-за дискретизации видеосигнала, а также дискретности ТВ раstra возникает регулярная ошибка вычисления координат, составляющая ± 1 т.э. (телевизионный элемент, вертикальный размер которого соответствует шагу строк ТВ раstra). Данная ошибка имеет случайный характер и зависит от положения текущего изображения относительно структуры раstra. Таким образом, первыми источниками флуктуационной составляющей выходного сигнала координатора являются строчная структура ТВ раstra по вертикали и шаг дискретизации по другой оси. Следующим источником флуктуационной составляющей является чересстрочная структура ТВ развертки, широко используемая в вещательном стандарте телевидения. В этом случае координатная ошибка может возрасти до ± 2 т.э. (особенно при низких контрастах изображения).

В мгновенном поле зрения ОЭС одновременно могут находиться несколько подвижных объектов (человек,

Описывается один из подходов к оценке характеристик телевизионных (тепловизионных) систем автоматического слежения за объектами, обеспечивающий наблюдение пространства, заключенного в пределах угла зрения мобильной оптико-электронной системы. Приводятся результаты математического моделирования, в котором учтены внешние влияющие факторы, и различные режимы внутренних параметров контура слежения.

автомобиль и др.), из которых следует выделить только один и обеспечить за ним автосопровождение. Специфика ТВ принципа локализации объекта состоит в поиске его в ближайшей окрестности текущей координаты. Поиск объекта в более широкой области наблюдаемого изображения может привести к ложным идентификациям. Поэтому в каждом цикле Δt работы координатора вычисляются приращения координат Δx , Δy , а сами координаты формируются методом накопления дифференциальной ошибки следующим образом:

$$X(t) = X(t - \Delta t) + \Delta x;$$

$$Y(t) = Y(t - \Delta t) + \Delta y. \quad (1)$$

Обеспечение минимально допустимой динамической ошибки необходимо для предотвращения выхода изображения объекта за пределы мгновенного угла зрения при использовании длиннофокусного оптического средства. В контуре с астатизмом 1-го порядка ошибка определяется допустимой угловой скоростью оптической оси.

Для обеспечения быстрого поиска в окрестности текущей координаты объекта, целесообразно строить систему по двухконтурной схеме. При этом внутренний, т.е. электронный контур, осуществляет слежение в пределах мгновенного поля зрения оптического средства. Он вырабатывает сигналы координат относительно центра и в пределах ТВ раstra. Во внешнем контуре эти сигналы используются в качестве управляющих пространственным положением оптической оси ТВ камеры.

На рис. 1 приведен вариант структурной схемы для одного канала двухконтурной следящей системы ТВ координатора.

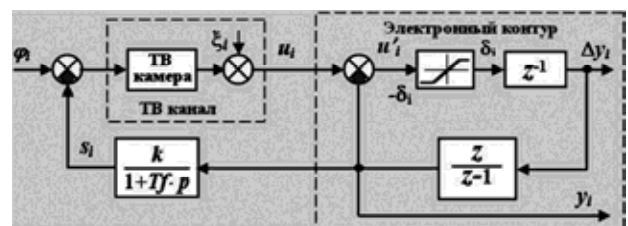


Рис. 1. Структурная схема ТВ координатора (ось y)

Входное значение φ_i мобильной ОЭС характеризует угловое положение линии визирования ТВ камеры. Сигнал изображения u_i подвергается входным возмущениям (изменениям φ_i) и поступает на вход электронного контура слежения. Далее выполняются следующие процедуры обработки (рис. 1):

- ограничение зоны поиска $\pm\delta$ в ближайшей окрестности объекта слежения, необходимое для достоверной идентификации отслеживаемого объекта;

- идентификация объекта с вычислением приращений координат Δy_i вследствие перемещения изображения;

- формирование координат объекта в пределах мгновенного поля зрения ТВ камеры.

Совокупность процедур обработки должна занимать минимальное время, поскольку от быстродействия внутреннего (электронного) контура полностью зависит быстродействие внешнего (электромеханического) контура.

Это минимальное время не может быть меньше времени формирования изображения u_i – основного источника информации. При использовании вещательного стандарта эта величина для полного ТВ кадра составляет $\Delta t = 0,04$ с, что соответствует 25 кадрам в секунду. В некоторых разработках в качестве исходного изображения используется не полный кадр, а каждое из двух полей, получаемых при чересстрочной развертке ТВ раstra. В этом случае время цикла обработки уменьшается до $\Delta t = 0,02$ с, а быстродействие электронного контура увеличивается вдвое, но вдвое ухудшается вертикальное разрешение изображения, поступающего на вход селектора объекта.

Для достоверной оценки качества работы ТВ координатора в модели вводится обобщенный аргумент ξ_i , характеризующий состав от двух источников ошибки координат – дискретности и чересстрочности раstra.

Зона поиска $\pm\delta$ ограничивает значения приращений координат, которые, в свою очередь, ограничивают допустимые динамические возмущения мобильной ОЭС. Поэтому для выявления пределов возможностей работы системы в модели используется гармонический сигнал возмущения с линейной частотной модуляцией (ЛЧМ) [1].

$$\varphi_i = A \sin \left[2\pi f_0 i \Delta t + \pi m \frac{(i \Delta t)^2}{T^2} \right], \quad (2)$$

где T – длительность реализации; A – амплитуда сигнала; f_0 – начальная частота ЛЧМ сигнала; i – переменная времени: $i = 1, \dots, \frac{T}{\Delta t} - 1$; m – база ЛЧМ сигнала; Δt – шаг дискретности цикла обработки.

На рис. 2 приведён график временной зависимости сигнала с ЛЧМ, подаваемого на вход модели.

Параметры ограничителя задаются следующим образом:

$$\delta = n \cdot \frac{\theta}{z}, \quad (3)$$

где n – число, задаваемое на основе алгоритма селекции объекта; θ – угол зрения оптического средства (в граду-

сах); Z – число ТВ элементов, заключённых в угле θ .

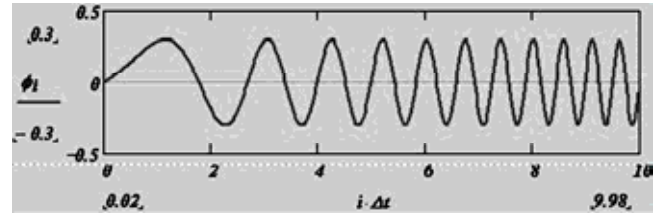


Рис. 2. Входной сигнал с линейно-частотной модуляцией

В электронном контуре осуществляется идентификация объекта и формирование координат. Минимальное время обработки изображения соответствует Δt – времени одного цикла работы системы, поэтому в структурную схему вводится компонент задержки $Z-1$ общий для всего электронного контура.

Окончательное формирование координаты y по (1)

осуществляется на накопителе $\frac{z}{z-1}$. При этом прохождение сигнала u_i от входа до выхода селектора можно рассматривать как

$$y_i = u_{i-1}. \quad (4)$$

Внешний контур координатора обладает инерционностью, поэтому его передаточная функция представляется аperiодическим звеном вида

$$w(p) = \frac{k}{T_f p + 1}, \quad (5)$$

где k – коэффициент усиления в тракте привода; T_f – постоянная времени звена; p – оператор Лапласа.

Известны другие подходы к моделированию передаточной функции внешнего контура, например [2, 3]. Любая модель стремится приблизиться к свойствам реального электромеханического привода системы. Тем не менее, в каждом варианте главными исходными данными являются: постоянная времени звена обратной связи; коэффициент усиления тракта и период дискретизации сигнала управления.

Применяя явный метод Эйлера представления производной [4] для сигнала s_i (см. рис. 1) на выходе передаточного звена $w(p)$, получаем следующее уравнение в дискретной форме:

$$s_{i+1} = s_i + \left[\frac{(y_i \cdot k) - s_i}{T_f} \right] \cdot \Delta t. \quad (6)$$

В выражении (6) значение y_i не определено, но, как следует из структурной схемы,

$$u_i = \varphi_i - s_i. \quad (7)$$

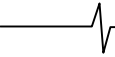
Поэтому, подставляя (7) в (6), с учётом (4), получаем окончательное разностное уравнение

$$s_{i+1} = s_i + \left[\frac{(\varphi_{i-1} - s_{i-1})k - s_i}{T_f} \right] \cdot \Delta t. \quad (8)$$

Сигнал, поступающий на вход ограничителя, можно определить, используя значение u_i из (7) и y_i из (4),

$$u'_i = u_i - y_i. \quad (9)$$

Условия прохождения сигнала u'_i через ограничитель и осуществление останова по критерию достижения



пределов $\pm\delta$ выполняются оператором «until», который в системе MathCAD представляется следующим образом:

$$lu'_i = \text{until} \left[- \left[(u'_i < -\delta) \vee (u'_i > \delta) \right], u'_i \right] \quad (10)$$

Диаграмма сигнала lu'_i (рис. 3) иллюстрирует пределы области динамических возмущений, допустимых для траектории мобильной ОЭС.

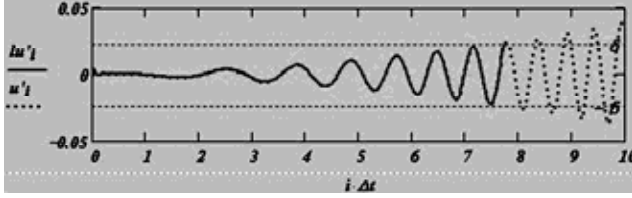


Рис. 3. Иллюстрация срыва автосопровождения (граница с пунктирной линией) при достижении границ ограничителя $\pm\delta$

Дальнейшая процедура заключается в определении граничной частоты f_m , выше которой слежение срывается. Частоту f_m найдем, используя полную фазу входного сигнала (2), которая для ЛЧМ сигнала выражается как

$$\psi(t) = 2\pi f_0 t + \pi m \frac{t^2}{T^2}, \quad (11)$$

а ее производная записывается в виде [5]:

$$\omega(t) = \frac{d\psi(t)}{dt} = 2\pi f_0 + \frac{\pi \cdot m}{T^2} \cdot 2t. \quad (12)$$

В дискретной форме выражение (12) запишется как

$$\omega_i = 2\pi f_0 + \frac{\pi m}{T^2} \cdot 2i\Delta t. \quad (13)$$

Искомая частота f_m определяется следующим образом:

$$f_m = \frac{\omega_i}{2\pi} = f_0 + \frac{m}{T^2} \cdot i\Delta t. \quad (14)$$

Рассматриваемая математическая модель выполнена в системе MathCAD. Ниже приводятся результаты моделирования, полученные для разных внешних влияющих факторов и режимов внутренних параметров контура слежения.

Для задания параметров модели следящей системы необходимо установить критерий устойчивости, в соответствии с которым добротность внешнего контура должна быть меньше добротности внутреннего. Добротность электронного контура соответствует частоте циклов f_c или

$$f_c = \frac{1}{\Delta t} \quad [1/c]. \quad (15)$$

Аналогичный параметр для внешнего контура определяется его передаточной функцией (5). Для этого производится обратное преобразование Лапласа:

$$L^{-1} \left\{ \frac{k}{Tf \cdot p + 1} \right\} = \frac{k}{Tf} \cdot e^{-\frac{\Delta t}{Tf}} = Q \quad [1/c], \quad (16)$$

где $\frac{k}{Tf}$ – добротность по [6], а множитель $e^{-\frac{\Delta t}{Tf}}$ – учёт периода дискретизации управляющего сигнала в замкнутом контуре.

Критерием устойчивости контура является выполнение неравенства

$$Q < f_c \quad \text{или} \quad C = \frac{f_c}{Q} > 1. \quad (17)$$

Как было показано, квантование и дискретизация видеосигнала приводят к возникновению координатной ошибки $\xi_{1i} = \pm 1$ т.э. С учётом чересстрочной структуры раstra координатная ошибка может возрасти до $\xi_{2i} = \pm 2$ т.э. (особенно при низких контрастах изображения). На каждом интервале Δt возникает одна из этих ошибок так, что помеха в целом ξ_i выражается суммой $\xi_i = \xi_{1i} + \xi_{2i}$ (рис. 4).

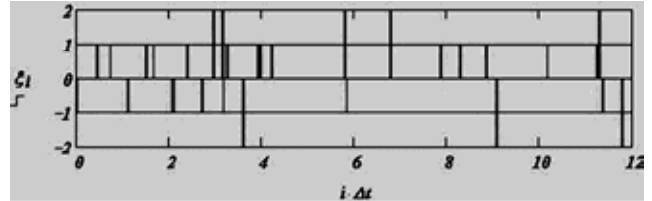


Рис. 4. Временная диаграмма обобщённой помехи

В модели задавались вероятностные характеристики помех, сопоставимые с существующими в реальных системах. В данной реализации длительностью $Tr = 12$ с, вероятность возникновения одиночной помехи за время Δt составляет

$$P_1 = \frac{1}{Tr} \cdot \sum_i |B_{1i}| \cdot \Delta t \quad (P_1 = 0,035), \quad (18)$$

где B_{1i} – вектор помехи в ± 1 т.э., а вероятность возникновения двойной помехи за тот же период времени составляет

$$P_2 = \frac{1}{Tr} \cdot \sum_i |B_{2i}| \cdot \Delta t \quad (P_2 = 0,013), \quad (19)$$

где B_{2i} – вектор помехи в ± 2 т.э.

Более достоверные результаты вероятностей P_1 и P_2 получены при выборе длительностей реализаций, равных 200 секундам, но и в этом случае отношение P_1/P_2 сохраняется в ранее полученных пределах.

Данная помеха является составной частью сигнала u'_i , поэтому сигнал lu'_i в (10) следует переписать как $lu'_i = lu'_i + \xi_i$.

Влияние помех при определении максимально допустимой частоты f_m возмущения мобильной ОЭС по (14) видно из диаграмм, приведённых на рис. 5 и 6, реализованных при равных прочих условиях:

$$Tf = 0,4 \text{ с}; \quad k = 7; \quad \Delta t = 0,02 \text{ с}; \quad A = 0,2^\circ; \quad f_0 = 0,1 \text{ Гц}.$$

В зависимости от момента воздействия одинарной или двойной помехи величина f_m будет определять допустимую угловую скорость

$$\Omega m = A 2\pi f_m. \quad (21)$$

На рис. 6 помеха от чересстрочности вывела сигнал u'_i за пределы границы $-\delta$ раньше, чем на рис. 5, сократив тем самым допустимую угловую скорость линии визирования. Следует учитывать ещё один важный фактор. Значение $f_m = 1,434$ Гц достигнуто благодаря монотонному нарастанию входной частоты f , начиная от значения $f_0 = 0,1$ Гц. При повышении параметра f_0 , например, до значения 0,4 Гц с большой вероятностью происходит срыв слежения в момент включения системы из-за

наличия больших угловых ускорений в переходных процессах. Этот факт устанавливает связь между величинами уровней $\pm \delta$ и допустимым угловым ускорением линии визирования.

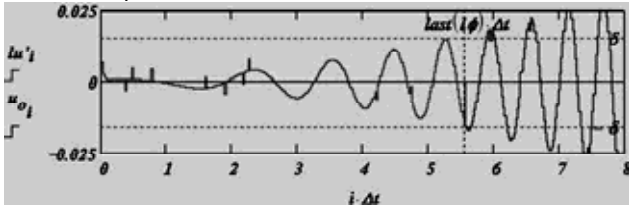


Рис. 5 Сигнал u' с аддитивной помехой и параметры-1 в момент срыва слежения ($f_m = 1,434$ Гц, $\Omega_m = 1,803^\circ/\text{с}$)

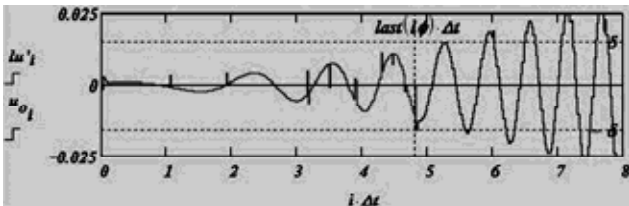


Рис. 6 Сигнал u' с аддитивной помехой и параметры-2 в момент срыва слежения ($f_m = 1,257$ Гц, $\Omega_m = 1,579^\circ/\text{с}$)

Подставив в (15) и (16) значения параметров, установленных при получении f_m и Ω_m на рис. 5 и 6, можно определить запас устойчивости контура по (17): $C = \frac{f_c}{Q} = 3,0$.

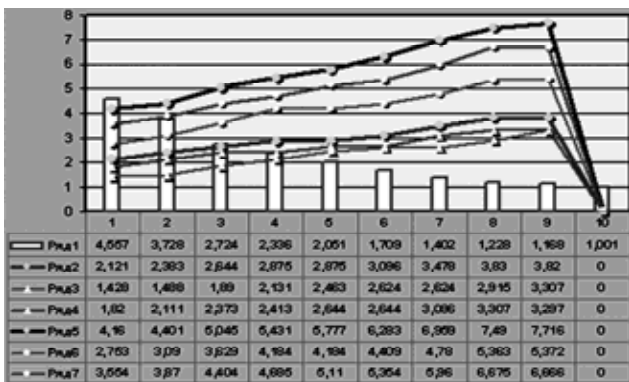


Рис. 7. Диаграммы допустимых угловых скоростей

Далее приводятся результаты, полученные при разных сочетаниях параметров системы и неизменной вероятности возникновения помех. При проведении эксперимента (рис. 7) установлены следующие параметры: $A = 0,4$; $\Delta t = 0,02$ с; $f_0 = 0,1$ Гц. Для получения достоверных результатов при определении минимальных и максимальных угловых скоростей в зависимости от величины ξ , реализации модели воспроизводились 20 раз для каждого отсчёта каждой диаграммы и двух режимов параметра Δt . Параметры T_f и k подбирали таким образом, чтобы было можно построить ряд по параметру C из (17), убывающий до уровня неустойчивости системы (на рис. 7 – гистограмма – «Ряд 1»). При этом ряды 2 и 5 иллюстрируют допустимые угловые скорости линии визирования в зависимости от запаса устойчивости C и в отсутствии импульса помехи в моменты определения f_m для $\Delta t = 0,02$ с и $\Delta t = 0,01$ с соответственно:

– ряды 3 и 6 – вероятностные угловые скорости в моменты воздействия импульсной помехи от чересстрочной развёртки;

– ряды 4 и 7 – вероятностные угловые скорости в моменты воздействия импульсной помехи от прогрессивной развёртки.

Заключение

Анализ результатов проведённых исследований показывает, что использование чересстрочной развёртки телевизионного изображения препятствует достижению максимальной угловой скорости из-за помех, обуславливаемых не только дискретностью раstra, но и смещением ТВ полей, удваивающих величину координатной ошибки. В то же время, двукратное увеличение частоты ТВ кадров (до $\Delta t = 0,01$ с) позволяет вдвое увеличить достижимую угловую скорость линии визирования. Таким образом, применение телевизионного вещательного стандарта не позволяет в ТВ координаторе обеспечить достижения лучших параметров. ТВ развёртка должна быть прогрессивного типа, а с повышением частоты кадров повышаются динамические характеристики системы автосопровождения объектов.

В режиме автосопровождения при заданных малых значениях ограничителя зоны поиска δ в контуре слежения можно обеспечить лишь невысокие значения угловых ускорений. Это следует учитывать при определении возможностей траекторного пути мобильной ОЭС. Значение δ может быть увеличено в зависимости от типа объектов, а также путем создания новых эффективных алгоритмов идентификации объектов.

Литература

- Каганов В.И. Радиотехнические цепи и сигналы. Компьютеризованный курс: Учебное пособие. – М.: ФОРУМ: ИНФРА-М, 2005. – 432 с. (Высшее образование).
- Алпатов Б.А., Бабаян П.В., Балашов О.Е, Степашкин А.И. Методы автоматического обнаружения и сопровождения объектов. Обработка изображений и управление / М.: Радиотехника, 2008. – 176 с.
- Воронин С.Г. Электропривод летательных аппаратов / Учебно-методический комплекс. Версия 1.0. – Челябинск, 1995–2011. – файлов 489, ил. См. http://epla.susu.ac.ru/vsg_udk.htm.
- Бабенко К.И. Основы численного анализа. – 2-е изд., испр. и доп. – Москва-Ижевск: НИЦ «Регулярная и хаотическая динамика», 2002. – 848 с.
- Баскаков С.И. Радиотехнические цепи и сигналы. М.: Высшая школа, 1988 – 448 с.
- Бесекерский В.А. Динамический синтез систем автоматического регулирования. – изд. «Наука», Москва, 1970 г.

QUALITY ASSESSMENT OF TV (IR) COORDINATOR OPERATION IN OBSERVING AND SECURITY SYSTEMS

Arzoumanian E.P., Mikolaychuk D.V., Zubarev Y.B.

This article describes one of methods for estimating parameters of television (or infrared) autotracking systems and which provides area observation within visual angle of mobile optoelectronic system. Results of mathematic modeling with considered external influencing factors and different modes of tracking circuit are presented.

СТОХАСТИЧЕСКАЯ ФИЛЬТРАЦИЯ ПЕРИОДИЧЕСКОГО БИПОЛЯРНОГО СИГНАЛА

Нефедов В.И., д.т.н., профессор, заведующий кафедрой телекоммуникационных систем (ТС) Московского государственного технического университета радиотехники, электроники и автоматики (МИРЭА), e-mail: nefedov@mirea.ru;

Решетняк С.А., д.ф.-м.н., профессор кафедры ТС МИРЭА, e-mail: reshets@bk.ru;

Третьяков Г.Н., к.т.н., доцент кафедры ТС МИРЭА, e-mail: gennady.tretyakov@rambler.ru.

Ключевые слова: биполярный сигнал, метод Эйлера-Коши, фильтрация, нелинейный фильтр, цифровое моделирование.

Введение

При прохождении сигнала и внешнего шума через нелинейную систему возникают процессы их взаимодействия, которые описываются уравнением, аналогичным уравнению движения броуновской частицы в определённом потенциальном поле при действии на неё гармонического сигнала и практически белого шума в виде стационарного гауссовского случайного процесса со временем корреляции, существенно меньшим других характерных времён задачи. В нелинейных системах с потенциальной функцией, имеющей два устойчивых положения равновесия и потенциальный барьер между ними, наблюдается рост выходного отношения сигнал/шум (S/N) с увеличением дисперсии шума в определённой области её изменения. Процесс сопровождается передачей энергии шума в слабый сигнал. Такое усиление сигнала вызвано совпадением его частоты со средней частотой Крамерса переходов через потенциальный барьер в результате действия шума. Этот эффект получил название стохастического резонанса и в настоящее время детально исследован [1-3].

В случае, когда мощность сигнала соизмерима с дисперсией шума в нелинейных системах возникает эффект стохастической фильтрации (СФ) сигнала, в результате которого отношение S/N на выходе нелинейной системы превышает аналогичное отношение на её входе. Данный эффект впервые был обнаружен в результате численного анализа [4] следующего уравнения:

$$\tau \frac{du}{dt} + W'(u) = a(t) + \xi(t), \quad (1)$$

где u – реакция на выходе системы, $W(u) = bu^4/4$ – потенциальная функция критического вида. Внешнее воздействие на систему определялось суммой гармонического сигнала $a(t) = a_0 \cos(\omega_0 t)$ и практически белого шума $\xi(t)$, $W'(u)$ – производная потенциальной функции, τ – постоянная времени системы. Теоретический анализ этого эффекта представлен в [5] и показывает, что СФ возникает в результате подавления шума сигналом.

Поскольку входной сигнал имеет в спектре всего одну гармонику, то отношение S/N определялось как отношение

Проведено цифровое моделирование процесса прохождения слабого сигнала и шума через нелинейный активный фильтр низкой частоты. Сигнал представлял собой периодическую последовательность биполярных импульсов с одинаковыми амплитудами. Были найдены области параметров нелинейного фильтра и сигнала, в которых коэффициент передачи по отношению сигнал/шум существенно выше, чем у линейного фильтра. Показано, что наилучшие условия фильтрации реализуются для сигналов типа меандра, амплитуда которых соизмерима со среднеквадратическим значением мощности шума.

квадрата амплитуды сигнала (мощности гармоники) к спектральной плотности шума на частоте ω_0 сигнала. Фильтрующие свойства системы оценивались на основе коэффициента передачи $q = (S/N)_{\text{вых}} / (S/N)_{\text{вх}}$. Очевидно, что для линейной системы $q = 1$. В [4] были найдены области параметров системы и сигнала, где $q = 1.3 \dots 1.4$. Эффект СФ наблюдался [6-8] также в результате аналогового моделирования (1) для акустических сигналов прямоугольной формы.

Результаты численного моделирования [4] были экспериментально подтверждены в [9] для нелинейного фильтра низкой частоты первого порядка, построенного на основе двух одинаковых туннельных диодов с постоянными напряжениями смещения. Отметим, что порядок фильтра совпадает с порядком стохастического дифференциального уравнения, описывающего прохождение сигнала и шума через систему. В [9] было установлено, что уравнению (1) подчиняется напряжение u на ёмкости фильтра с потенциальной функцией

$$W(u) = (1 - R_n A)u^2 / 2 + R_n B u^4 / 4, \quad (2)$$

где R_n – сопротивление нагрузки, A и B – параметры результирующей вольт-амперной характеристики системы диодов, $\tau = R_n C$ – постоянная времени, C – величина ёмкости в фильтре. Исследования [9] показали, что эффект СФ носит пороговый характер и реализуется, если мощность сигнала соизмерима с дисперсией шума. В [10] СФ была обнаружена также в нелинейном фильтре второго порядка. Так как аналитические методы исследования указанных эффектов достаточно трудоёмки, то большое значение в их анализе имеют численные методы. В последнее время в [11] разработан метод, показавший, что СФ имеет место не только в фильтрах низких, но и высоких частот сигнала, схемы которых представлены в [11].

Цель данной работы – на основе разработанного в [11] подхода проанализировать эффект СФ в случае сигнала, который в спектре имеет не одну, а большое число гармоник. Локальное по спектру определение отношения S/N , используемое в случае сигнала с одной гармоникой, здесь неприемлемо. Фильтрующие свойства системы, на наш взгляд, удобнее оценивать на основе интегрального по спектру отношения S/N :

$$S/N = \sum_n |A_n|^2 / \sigma^2,$$

где $\sigma^2 = \sum_n |\xi_n|^2$ – дисперсия шума с комплексными амплитудами его гармоник ξ_n , A_n – амплитуды гармоник сигнала. Данное определение S/N , применяемое как на входе, так и на выходе фильтра, есть отношение мощностей сигнала и шума. Сразу отметим, что при данном определении S/N коэффициент передачи q может принимать значения, существенно превышающие единицу даже для линейной системы, поэтому полученные результаты сравнивались с данными для линейного фильтра.

Параметры нелинейного фильтра и сигнала

Для исследования в настоящей работе был рассмотрен нелинейный фильтр низких частот первого порядка с потенциалом (2) критического вида, что и в эксперименте [9] со следующими параметрами:

$R_n = A^{-1} = 65 \text{ Ом}$ и $C = 15 \text{ нФ}$. После введения масштабов по напряжению $u_M = (BR_n)^{-1/2} = 0.5 \text{ В}$ и времени $t_M = \tau = 10^{-6} \text{ с}$ безразмерное уравнение (1) для напряжения на ёмкости принимает вид

$$\frac{du}{dt} + u^3 = h(t) + \zeta(t), \quad (3)$$

где $h(t) = a(t)/u_M$ и $\zeta(t) = \xi(t)/u_M$ – соответственно безразмерные сигнал и шум. В уравнении (3) напряжение u также является безразмерным.

В качестве входного сигнала рассматривалась периодическая последовательность биполярных прямоугольных импульсов одинаковой амплитуды h , длительности τ_u и периодом T_c . Вид входного сигнала представлен на рис. 1. На каждом периоде сигнал имеет нулевое среднее значение. На первом периоде

$$h(t) = \begin{cases} h, & 0 < t < \tau_u, \\ -h, & T_c/2 < t < T_c/2 + \tau_u. \end{cases} \quad (4)$$

Шаг решения уравнения (3) по времени Δt выбирался равным времени корреляции шума $\tau_k = 10^{-7} \text{ с}$ [9]. Поскольку τ_k есть величина на порядок меньшая постоянной τ времени фильтра, то шум можно считать практически белым. При этом использовалась модель шума, в которой не детализируется его поведение на интервале времени τ_k . В каждой точке отсчёта времени $t_n = n\Delta t$ включался датчик случайных чисел, распределённых по нормальному закону с нулевым средним значением и дисперсией $\sigma^2 = \xi^2$, где черта обозначает среднее зна-

чение по времени. Очевидно, что после усреднения по большому числу реализаций шума, его корреляционная функция имеет вид $\langle \zeta(t_n)\zeta(t_k) \rangle = \sigma^2 \delta_{nk}$. Полное число N отсчётов по времени определялось длительностью T исследуемых процессов: $N = T / \Delta t$.

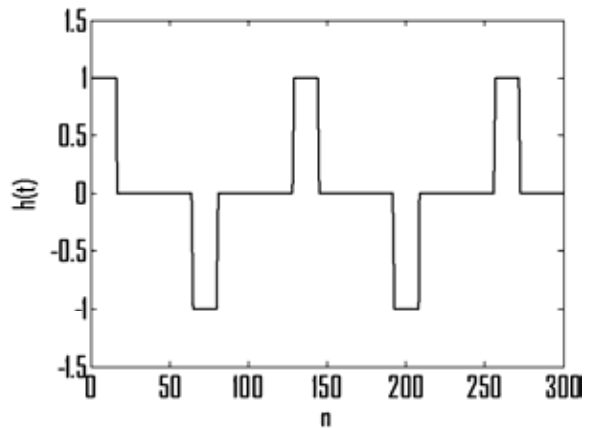


Рис. 1. Форма сигнала на входе фильтра в зависимости от числа n отсчётов моментов времени $t_n = \Delta t n$.

Параметры сигнала: $h = 1$, $\eta = 0.125$ и $f_c = s\Delta f = 78 \text{ кГц}$

В случае детерминированной правой части уравнения (3) его решение находится с помощью общепринятого метода Рунге-Кутты. Для исследуемого стохастического дифференциального уравнения и выбранной модели шума, с нашей точки зрения, наиболее корректным методом решения (3) является численный метод Эйлера-Коши [11].

Предварительный анализ спектров сигналов. Отношение сигнал/шум

Спектры входных и выходных процессов находились на основе дискретного преобразования Фурье (ДПФ), которое содержит полное число N отсчетов времени наблюдения. При этом амплитудный спектр является симметричным относительно гармоники с номером $N/2$, частота которой соответствует теореме отсчетов: $f_{\max} = \Delta f N / 2 = 1 / (2\tau_k)$, где $\Delta f = 1/T$ – разность между соседними частотами. Отметим, что при $N = 2^m$ ДПФ требует меньшее число вычислений, поэтому в работе полагалось $N = 2^{14}$, что соответствует экспериментальному значению $T = 1.6 \text{ мс}$, которое несколько превышает длительность сигналов в [9].

Простые вычисления показывают, что отличные от нуля амплитуды входного сигнала (4) имеют вид

$$|h_{(2k-1)s}| = \frac{h}{\pi(2k-1)} \left[2(1 - \cos(2\pi(2k-1)\eta)) \right]^{1/2}, \quad (5)$$

где $\eta = \tau_u / T_c$ – безразмерная длительность импульса, $s = f_c / \Delta f$ – число, определяющее частоту сигнала f_c или число периодов сигнала на интервале времени рассмотрения T .

Сопоставление точных значений (5) с расчетными величинами амплитуд на основе ДПФ показало, что метод ДПФ приводит к правильным результатам практически во всей области спектра кроме малой области вбли-

зи его границы, где погрешность вычислений не превышала 50 %. Анализ спектра выходного сигнала в отсутствие шума показал, что кроме частот основных гармоник с номерами $n = (2k - 1)s = s, 3s, 5s, \dots$ он содержит гармоники с комбинационными частотами, амплитуды которых малы. На рис. 2 изображены два выходных сигнала, один из которых является точным, так как построен с учётом всех гармоник и удовлетворяет начальному условию $u(0) = 0$, а второй построен только с учётом основных гармоник. Видно, что эти два решения незначительно отличаются между собой только на начальной стадии первого периода сигнала. На остальных периодах эти решения совпадают.

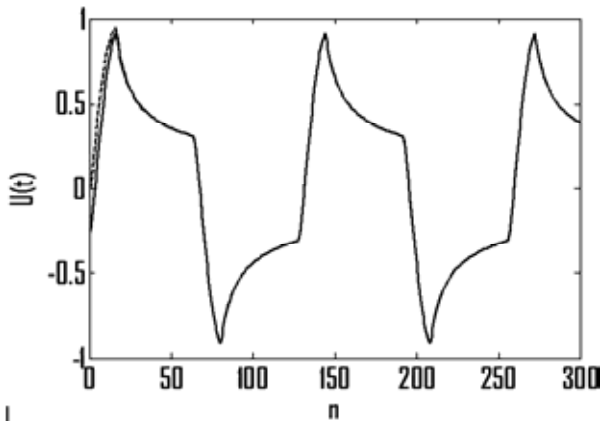


Рис. 2. Сигнал на выходе фильтра в зависимости от числа n отсчётов моментов времени $t_n = \Delta t n$,

соответствующий входному сигналу на рис. 1. Пунктирная кривая построена с учётом всех частот выходного сигнала, а сплошная кривая – с учётом только его основных гармоник

После установления номеров $n = (2k - 1)s$ сигнальных гармоник было использовано следующее интегральное по спектру отношение S/N :

$$S/N = \sum_{k=1}^{(2k-1)s \leq N} |A_{(2k-1)s}|^2 / \sigma^2, \quad (6)$$

где $|A_n|$ – амплитуды гармоник сигнала с частотами $\omega_n = 2\pi\Delta f n$, которые соответственно для сигналов на входе и выходе фильтра имеют вид

$$A_n^{(вх)} = \sum_n [h(t) + \zeta(t)] \exp(j\omega_n t),$$

$$A_n^{(вых)} = \sum_n u(t) \exp(j\omega_n t).$$

В расчёте дисперсии шума σ^2 сигнальные гармоники во внимание не принимались.

Результаты вычислений для различных реализаций входного шума незначительно отличались друг от друга, поэтому полученные ниже данные усреднялись по количеству пусков K . Для определения числа K проверялось свойство эргодичности шума, т.е. равенство его следующих средних значений по времени и по ансамблю реализаций:

$$\overline{\zeta^2} = \frac{1}{T} \int_0^T \zeta^2(t) dt = \sum_{n=0}^{N-1} |\zeta_n|^2 \approx \langle \zeta^2 \rangle = \frac{1}{K} \sum_{i=1}^K |\zeta^{(i)}|^2,$$

где ζ_n – амплитуды гармоник шума, $\zeta^{(i)}$ – значение

шума в фиксированный момент времени на его i -ой реализации. Как показали вычисления, это равенство при $K = 200$ выполняется с относительной погрешностью не более 1 %. Отсюда все полученные результаты усреднялись по 200 реализациям входного случайного процесса.

Основные результаты цифрового моделирования

Исследуемая нелинейная система является низкочастотным фильтром и в силу определения (8) выходное интегральное отношение S/N может существенно превышать входное отношение даже для линейной системы. Поэтому в работе проводилось сравнение коэффициента $q = (S/N)_{вых} / (S/N)_{вх}$ передачи отношения S/N с аналогичным коэффициентом для линейной системы или RC-фильтра, который реализуется при замене в уравнении (3) u^3 на u .

На рис. 3 представлена зависимость коэффициента передачи q от безразмерной длительности импульса η . Из неё следует, что при $\eta > 0.3$ нелинейная фильтрация преобладает над линейной, а для сигналов типа меандра ($\eta \approx 0.5$) коэффициент передачи q существенно превышает значение, соответствующее линейной системе. Эта зависимость объясняется тем, что с ростом длительности импульса большое число сигнальных гармоник расположено в области низких частот и не подавляются при фильтрации.

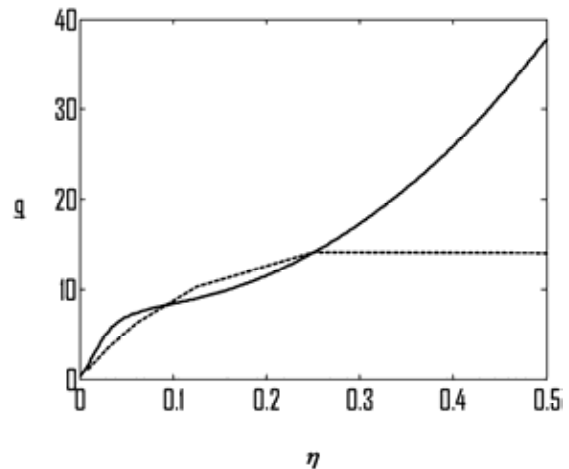


Рис. 3. Зависимость коэффициента передачи q от нормированной длительности импульса $\eta = \tau_u / T_c$ для частоты $f_c = 78 \text{ кГц}$ сигнала. Сплошная кривая – результаты в случае нелинейного фильтра, пунктирная кривая – данные для линейного RC-фильтра

Ниже основные результаты анализа приводятся для сигнала – меандра. На рис. 4 для двух значений параметра τ , определяемого величиной ёмкости C , представлены зависимости q от частоты сигнала f_c . Из них следует, что в максимуме частотной зависимости, которая наблюдается в области 5...10 кГц, коэффициент передачи q для нелинейного фильтра в 4 раза выше, чем у линейного фильтра. Падение q с ростом f_c является стандартной зависимостью для низкочастотного фильтра. С уменьшением f_c сигнальный спектр сдвига-

ется в низкочастотную область, в которой мощность выходного шума возрастает, в результате этого происходит падение q . С ростом параметра τ диапазон подавления высоких частот увеличивается, что приводит к увеличению выходного отношения S/N и q . Однако дальнейшее увеличение τ может привести к подавлению амплитуд гармоник сигнала и уменьшению коэффициента передачи q .

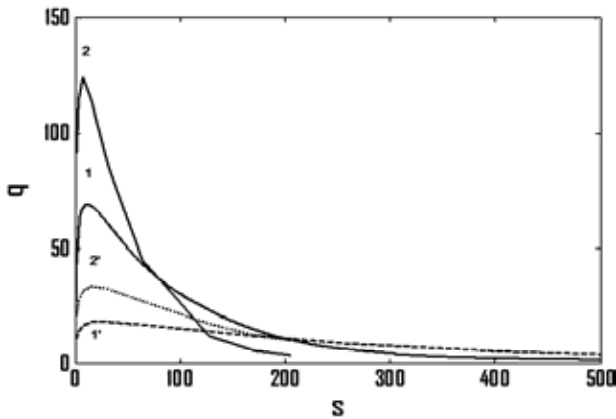


Рис. 4. Зависимость q от числа s , определяющего частоту $f_c = \Delta f s$ Гц сигнала. Сплошные кривые – результаты для нелинейного фильтра, пунктирные кривые – данные для линейного RC-фильтра при $h = 0.4$. Кривые 1 и 1' соответствуют $C = 15$ нФ, а кривые 2 и 2' – ёмкости $C = 30$ нФ

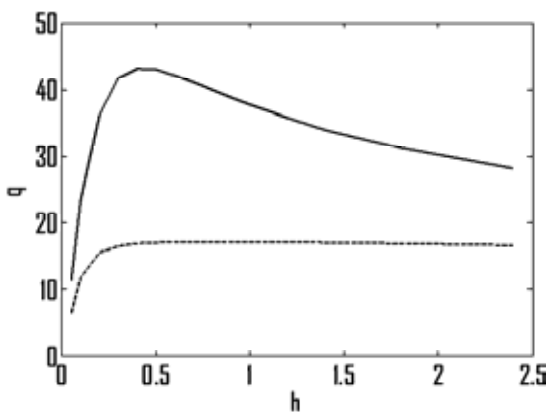


Рис. 5. Зависимость коэффициента передачи q от амплитуды h меандра для его частоты $f_c = 39$ кГц. Сплошная кривая – результаты в случае нелинейного фильтра, пунктирная кривая – данные для линейного RC-фильтра

Для фиксированных значений мощности шума, частоты и длительности импульса получена зависимость q от амплитуды h импульса, которая изображена на рис. 5. Как и в случае [9] гармонического сигнала, эта кривая указывает на пороговый характер возникновения эффекта СФ. Увеличение q с ростом h объясняется подавлением шума сигналом. При больших амплитудах h , как и в отсутствие шума, коэффициент передачи падает. На рис. 6 представлена зависимость q от среднеквадратичной интенсивности шума σ при фиксированных остальных параметрах фильтра и сигнала.

Отметим, что с ростом σ выходное отношение S/N монотонно убывает. Однако скорости убывания входного и выходного отношений S/N с ростом σ отличаются, поэтому кривая $q(\sigma)$ имеет максимум. Приведённые результаты численного моделирования практически не изменялись при варьировании начального условия и уменьшении в два раза времени корреляции шума.

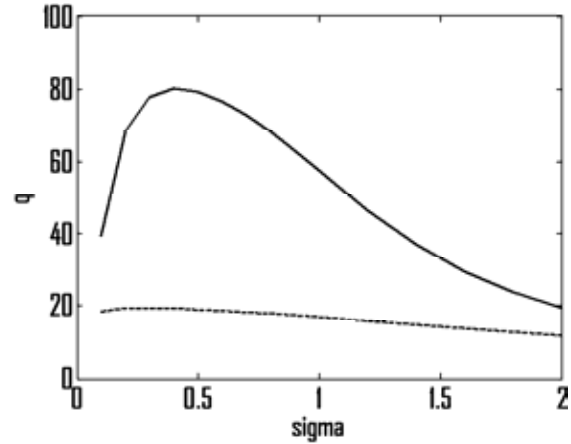


Рис. 6. Коэффициент передачи q в зависимости от среднеквадратичной интенсивности шума $\sigma = \sigma / u_m$ для меандра с высотой импульсов $h = 0.4$ и частотой $f_c \cong 10$ кГц

Заключение

Проведённое исследование позволяет сделать следующие выводы. Основные закономерности, присущие эффекту СФ в случае гармонического сигнала, сохраняются и в случае сигналов типа меандра. Численный анализ низкочастотного нелинейного активного фильтра позволил установить область параметров сигнала и фильтра, в которой его фильтрующие способности существенно выше, чем у аналогичного линейного фильтра. Зависимости коэффициента передачи q по отношению сигнал/шум указывают на наличие оптимальных частот сигнала, его амплитуды и мощности шума для достижения наибольших значений q . Оптимальные условия нелинейной фильтрации реализуются при мощностях указанных сигналов соизмеримых с дисперсией шума. Вполне возможно, что существуют другие формы сигналов, в случае которых их стохастическая фильтрация будет наблюдаться более ярко. Однако, на наш взгляд, оптимальные формы сигналов в основном определяются типом нелинейного фильтра.

Литература

1. Gammaitoni L., Hanggi P., Jung P., Marchesoni F. Stochastic resonance. Rev. of Mod. Phys., 1998, v.70, №1, p. 223-287.
2. Анищенко В.С., Нейман А.Б., Мосс Ф., Шиманский-Гайер Л. Стохастический резонанс как индуцированный шумом эффект увеличения степени порядка. УФН, 1999, т.169, №1, с. 7-38.
3. Климонтович Ю.Л. Что такое стохастическая фильтрация и стохастический резонанс? УФН, 1999, т.169, №1, с. 39-47.

4. Hanggi P., Inghiosa M.E., Fogliatti D., Bulsara A.R. Non-linear stochastic resonance: The saga of anomalous output-input gain. Phys. Rev. E., 2000. v.62, №5, p. 6155-6163.

5. Решетняк С.А., Третьяков Г.Н. Теоретическое исследование эффекта стохастической фильтрации. Радиотехника и электроника, 2013, т. 58, №4, с. 360-366.

6. Gingl Z., Makra P., Vajtai R. High signal-to-noise ratio gain by stochastic resonance in a double well. Fluctuation and Noise Lett., 2001, v.1, No.3, p. L181-L188.

7. Makra P., Gingl Z., Kish L.B. Signal-to-noise ratio gain in non-dynamical and dynamical bistable stochastic resonator. Fluctuation and Noise Lett., 2002, v.2, No.3, – p. L147-L155.

8. Makra P., Gingl Z., Fulei T. Signal-to-noise ratio gain in stochastic resonators driven by coloured noises. Phys. Lett., 2003, v.317, No.3-4, p. L228-L232.

9. Домбровский А.Н., Решетняк С.А. Стохастический резонанс и фильтрация сигналов в нелинейной электрической системе второго порядка. Радиотехника, 2007, №9, с. 19-25.

10. Домбровский А.Н., Решетняк С.А. О стохастической фильтрации сигналов в нелинейных электрических

системах. Радиотехника и электроника, 2009, т.54, №11, с. 1369-1371.

11. Абоелазм М.А., Мельчаков В.Н., Решетняк С.А., Третьяков Г.Н. Исследование цифровой модели нелинейного активного фильтра первого порядка. Цифровая обработка сигналов, 2014, №4, с. 62-64.

IC FILTERING OF PERIODIC BIPOLAR SIGNAL

Nefedov V.I., Reshetnyak S.A., Tret'yakov G.N.

The process of passing a weak signal and noise through non-linear active low frequency filter was numerically investigated. The signal was a periodic sequence of bipolar pulses with the same amplitudes. Optimal areas of nonlinear filter and signal were found in which the signal/noise ratio is substantially higher than that of a linear filter. Shows that the best conditions are for signals like meandering, the amplitude of which is comparable with the value of root mean-square of the noise power.



«РАДИОЭЛЕКТРОННЫЕ СРЕДСТВА ПОЛУЧЕНИЯ, ОБРАБОТКИ И ВИЗУАЛИЗАЦИИ ИНФОРМАЦИИ» (РСПОВИ-2015)

28 - 30 октября 2015 г. Москва

All-Russian conference (with the international participation)

«RADIO-ELECTRONIC MEANS OF RECEIVING, PROCESSINGS AND INFORMATION VISUALIZATION» (RMRPIV-2015)



Уважаемые коллеги!

ПРИГЛАШАЕМ ВАС ПРИНЯТЬ УЧАСТИЕ В РАБОТЕ КОНФЕРЕНЦИИ

ОРГАНИЗАТОРЫ:

- Институт радиотехники и электроники имени В.А. Котельникова РАН
- Российское научно-техническое общество радиотехники, электроники и связи им. А.С. Попова
- Министерство образования и науки РФ
- ОАО «Конструкторское бюро – 1»
- Вневедомственный экспертный совет по проблемам воздушно-космической обороны

ПРИ УЧАСТИИ:

- Федеральное агентство по промышленности РФ
- ОАО «Концерн радиостроения «Вега»
- ОАО «Корпорация «Комета»
- ФГУП «НИИР»
- Академия инженерных наук имени академика А.М. Прохорова
- Владимирский государственный университет
- Московский авиационный институт
- Московский государственный технический университет им. Н.Э. Баумана
- Московский институт радиотехники, электроники и автоматики (технический университет)
- Московский технический университет связи и информатики
- Московский институт электронной техники
- Московский энергетический институт
- Рязанский государственный радиотехнический университет
- Санкт-Петербургский государственный университет телекоммуникаций им. проф. А.М. Бонч-Бруевича
- Санкт-Петербургский государственный электротехнический университет
- Ульяновский государственный технический университет
- Ярославский государственный университет

Подробная информация на сайте: www.rntores.ru

УДК 004.93'11

ЛОКАЛЬНЫЕ БИНАРНЫЕ ШАБЛОНЫ МЕДИАННОГО ПИКСЕЛЯ – ЭФФЕКТИВНЫЕ ИНФОРМАТИВНЫЕ ПРИЗНАКИ ТЕХНОЛОГИИ РАСПОЗНАВАНИЯ ОБРАЗОВ

Ефимов И.Н., аспирант Самарский государственный университет путей сообщения,
e-mail: Mr.Efimov.IN@gmail.com

Ключевые слова: компьютерное зрение, локальные бинарные шаблоны, распознавание образов, информативные технологии.

Введение

Технологии компьютерного зрения находят применения в самых разнообразных областях жизни человека, будь то распознавание дорожных знаков, распознавание номеров транспортных средств, распознавание черт лица человека, эмоций или жестов. Создание вышеперечисленных технологий остаётся сложной технической и теоретической задачей.

Для успешного функционирования технологии распознавания образов необходимо обеспечить: высокую вероятность правильного распознавания, возможность работы в реальном времени и успешное определение подмены распознаваемого объекта. Основной задачей при создании технологии распознавания образов является корректный выбор наиболее эффективного информативного признака.

Информативный признак, представленный в данной статье, может эффективно использоваться как для распознавания образов, так и для определения подмены распознаваемого объекта [1]. Скорость вычисления информативного признака позволяет использовать его в реальном времени.

Вычисление кодовых комбинаций ЛБШ

По данным ряда исследований [3, 4], распознавание образов с помощью локальных бинарных шаблонов (далее ЛБШ) обеспечивает требуемые результаты с точки зрения скорости работы и качества распознавания. ЛБШ доказали свою состоятельность в особенности из-за частичной инвариантности к условиям освещения и вычислительной простоте формирования кодов, что позволяет использовать ЛБШ в реальном времени. ЛБШ являются простыми, но очень эффективными текстурными признаками.

В 2002 году группа под руководством Тимо Оджала [7] предложила метод вычисления кодов ЛБШ. Для представления ЛБШ используется обозначение $LBP_{P,R}$, где P – количество рассматриваемых точек окрестности, R – радиус окрестности. Коды ЛБШ вычисляются путем суммирования пороговых различий интенсивности с центром окрестности:

$$LBP_{P,R}(x,y) = \sum_{i=0}^{P-1} s(g_i - g_c) 2^i, \quad (1)$$

Приводятся описание, достоинства и недостатки оригинального информативного признака, основой которого являются локальные бинарные шаблоны. Автором делается попытка изменения формы и алгоритма построения кодовых комбинаций, что является новизной публикации, а также предлагается сравнительный анализ результатов работы оригинального информативного признака. Показаны преимущества признака применительно к распознаванию образов по сравнению с аналогами.

где x, y – координаты центрального пикселя, i – номер рассматриваемого пикселя, g_c – значение яркости центрального пикселя, g_i – яркость рассматриваемого пикселя, а $s(z)$ равна

$$s(z) = \begin{cases} 1, & z \geq 0; \\ 0, & z < 0. \end{cases}$$

В 2008 году были предложены два новых усовершенствованных ЛБШ [4]: Three-PatchЛБШ и Four-PatchЛБШ. Основное отличие предложенных ЛБШ от существующих в том, что сравниваются фрагменты изображения, а не пиксели, по какой-либо метрике (например, Хи-квадрат). Как следует из названия, шаблон Three-PatchЛБШ (TP) вычисляется путем сравнения значений трех фрагментов (рис. 1 а).

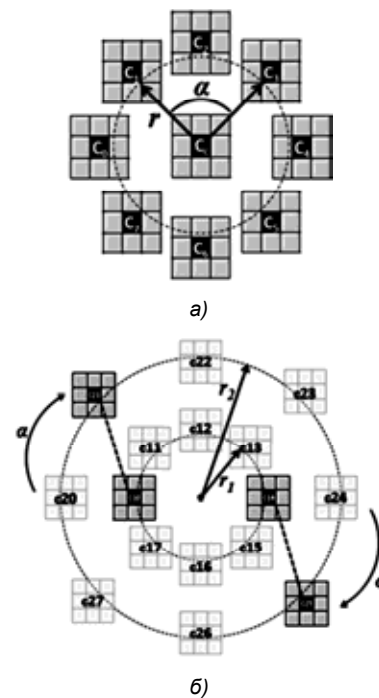


Рис. 1. а) $TP_{8,1,2}$ для 3-х фрагментов для $P=8, r=1$ и $\infty=2$. б) $FP_{8,1,2,1}$ для $P=8, r_1=1, r_2=2$ и $\infty=1$

Алгоритм Three-PatchЛБШ:

Шаг 1. Фрагменты C_i и $C_{i+\alpha}$ сравниваются по некоторой метрике d со значениями центрального фрагмента C_c , где α – число, определяющее второй фрагмент для сравнения.

Шаг 2. Если окрестность C_i ближе по значениям к C_c , чем $C_{i+\alpha}$ к C_c , то $s(z) = 1$, иначе $s(z) = 0$.

Шаг 3. Расчет производится для каждого пикселя по следующему выражению:

$$TP_{P,R,\alpha}(x,y) = \sum_{i=0}^{P-1} s(d(C_i, C_c) - d(C_{i+\alpha}, C_c)) 2^i, \quad (2)$$

где

$$s(z) = \begin{cases} 1, & z \geq \tau; \\ 0, & z < \tau. \end{cases}$$

На практике используется значение большее нуля (например, $\tau = 0,01$), для обеспечения стабильности в расчетах.

Шаблон Four-PatchЛБШ (FP) вычисляется путем сравнения значений четырех фрагментов (рис. 1 б).

Алгоритм Four-PatchЛБШ:

Шаг 1. Определяются 2 окружности фрагментов с радиусами R_1 и R_2 .

Шаг 2. По некоторой метрике d попарно сравниваются фрагменты $C_{1,i} - C_{2,i+\alpha}$ и $C_{1,i+p/2} - C_{2,i+p/2+\alpha}$, где α – число, определяющее следующий фрагмент для сравнения.

Шаг 3. Если первая пара фрагментов ближе по значениям друг к другу, чем вторая, то $s(z) = 1$, иначе $s(z) = 0$.

Шаг 4. Расчет производится для каждого пикселя по следующему выражению:

$$FP_{P,r_1,r_2,\alpha}(x,y) = \sum_{i=0}^{P-1} s(d(C_{1,i}, C_{2,i+\alpha}) - d(C_{1,i+p/2}, C_{2,i+p/2+\alpha})) 2^i, \quad (3)$$

где

$$s(z) = \begin{cases} 1, & z \geq \tau; \\ 0, & z < \tau. \end{cases}$$

На практике используется значение большее нуля (например, $\tau = 0,01$), для обеспечения стабильности в расчетах. На рис.6представлены иллюстрации к Three-PatchЛБШ и Four-PatchЛБШ.

Вычисление гистограмм кодовых комбинаций ЛБШМП

На данный момент, помимо Three-PatchЛБШ и Four-PatchЛБШ, существуют альтернативные ЛБШ [6, 9]. Одним общим недостатком является вычислительная сложность алгоритмов. Автором разработан новый вид ЛБШ. Предложенный ЛБШ медианного пикселя (далее ЛБШМП) (рис. 2), существенно ускоряет процесс вычисления шаблона, при сохранении свойств, присущих Three-PatchЛБШ и Four-PatchЛБШ.

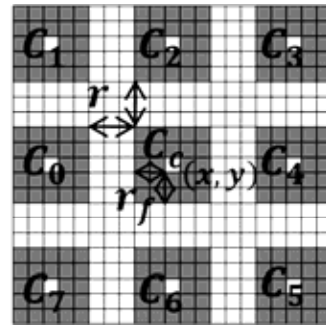


Рис. 2. Схема построения ЛБШМП для $P = 8$, $r = 2$, $P_f = 25$ и $r_f = 2$

Математическая модель ЛБШМП имеет вид:

$$M_{P,r,P_f,r_f}(x,y) = \sum_{i=0}^{P-1} 2^i \cdot S(z_i), \quad (4)$$

где P – количество рассматриваемых фрагментов, r – радиус окрестности, P_f – количество пикселей во фрагменте, r_f – радиус фрагмента, x, y – координаты центрального пикселя, i – номер сравниваемого фрагмента, $i = \overline{0, P-1}$, а выражение s :

$$s(z_i) = \begin{cases} 1, & z_i \geq \tau; \\ 0, & z_i < \tau, \end{cases}$$

$$z_i = \text{med}(C_i) - \text{med}(C_c),$$

$$\text{med}(C_i) = \text{med}(g(0), \dots, g(l), \dots, g(2l)) = g(l),$$

где $\text{med}(C_i)$ – медианное значение яркостей пикселей фрагмента C_i , элементы которого расположены в порядке возрастания значений яркости $g(j)$, т.е. выполняется условие: $g(0) \leq \dots \leq g(l) \leq \dots \leq g(2l)$, $j = \overline{0, 2l}$, $l = 2r_f(r_f + 1) - \text{индекс медианного элемента}$, $\text{med}(C_c)$ – медианное значение центрального фрагмента C_c . Пороговое значение τ берется большим, чем ноль (например, $\tau = 0,01$), чтобы обеспечить стабильность в расчетах. Выражение для яркости пикселя имеет вид:

$$g(j) = 0,3 \cdot R_j + 0,59 \cdot G_j + 0,11 \cdot B_j, \quad (5)$$

где R_j , G_j и B_j цветные составляющие красного, зеленого и голубого цвета пикселя. Ниже (рис. 3) представлена схема для нахождения $\text{med}(C_i)$, где каждая ячейка – это значение интенсивности пикселя $g(j)$.

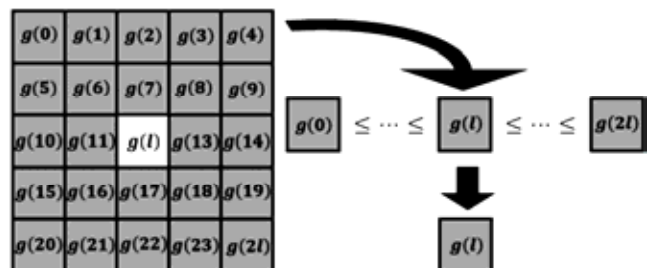


Рис. 3. Поиск медианного значения фрагмента C_i

Дальнейшим развитием ЛБШМП стало изменение формы построения ЛБШ и увеличение числа сравнива-

емых фрагментов (рис. 4). На рис. 4 $b_1 = 8r_f + 4$ и $b_2 = r + 2r_f + 1$ – размеры окрестности фрагментов.

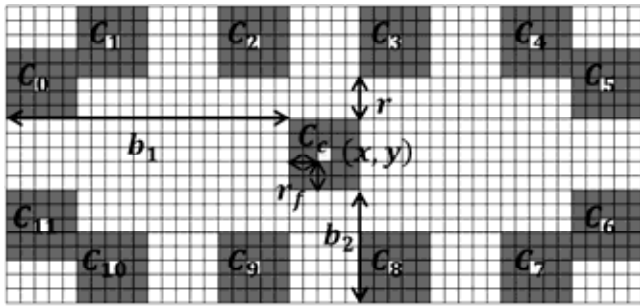


Рис. 4. Схема построения ЛБШМП, с увеличенным числом фрагментов

Для увеличения вероятности правильного распознавания добавлена внешняя структура сравнения. Подобно Four-Patch ЛБШ изменился алгоритм сравнения – фрагменты сравниваются не с центральным блоком, а с фрагментами внешней структуры. Результатом преобразований стала схема, представленная на рис. 5, где $\lceil r/2 \rceil$ – оператор выделения целой части от вещественного выражения. Количество и схема размещения фрагментов данной конструкции могут варьироваться в зависимости от требуемого результата.

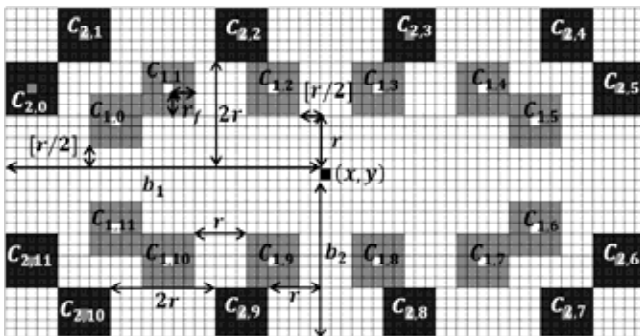


Рис. 5. Схема построения ЛБШМП для $P=12$, $r=5$, $P_f=25$ и $r_f=2$

Для того что бы наглядно увидеть реализацию информативных признаков необходимо: рассчитать для каждого пикселя значение ЛБШ, привести значения ЛБШ к общему множеству (0..256), для этого необходимо нормализовать значения ЛБШ, установить цветовые составляющие красного, зеленого и голубого цвета пикселя равными значению ЛБШ этого пикселя. На рис.6 представлена иллюстрация к ЛБШМП.

В операции распознавания образов используются гистограммы кодовых комбинаций ЛБШ. Гистограмма ЛБШ, построенная одновременно для всего изображения, содержит только текстурные особенности без информации о расположении кодов на изображении. Для учета информации о расположении особенностей, изображение предлагается разбивать по одной из следующих схем (7 а и 7 б) (в первоначальном способе Т.Анопен предлагается разделять изображение на 7x7 блоков), в каждом регионе вычисляется своя гистограмма кодовых комбинаций. Гистограммы для каждой области объединяются в одну общую (рис. 8), при этом общая гистограмма содержит

текстурные особенности и информацию о расположении кодовых комбинаций ЛБШ.

Ниже приведены примеры комбинаций деления изображения лица на регионы и примеры гистограмм.

Таким образом, математическая модель информативного признака гистограммы кодов ЛБШМП со схемы 5 будет иметь вид:

$$H_{P,r,P_f,r_f}(h,k) = V_k \sum_{x=b_1}^{W_k-b_1-1} \sum_{y=b_2}^{H_k-b_2-1} f(M_{P,r,P_f,r_f}(x,y),h), \quad (6)$$

где P – количество рассматриваемых фрагментов, r – ближайшее расстояние от центра до 1-й окрестности фрагментов (рис. 5), P_f – количество рассматриваемых пикселей во фрагменте, r_f – радиус рассматриваемых фрагментов, h – номер столбца гистограммы, $h = \overline{0, 2^p - 1}$, k – номер региона гистограммы, $k = \overline{1, K}$, K – количество регионов разделяющих изображение, V_k – весовой коэффициент k -го региона, x, y – координаты центрального пикселя, $b_1 = 3r + 6r_f + 3$ и $b_2 = 2r + 2r_f + 1$ – размеры 2-й окрестности, а функция f имеет вид:

$$f(M_{P,r,P_f,r_f}(x,y),h) = \begin{cases} 1, & M_{P,r,P_f,r_f}(x,y) = h; \\ 0, & M_{P,r,P_f,r_f}(x,y) \neq h, \end{cases}$$

где $M_{P,r,P_f,r_f}(x,y)$ – функция расчета кодовой комбинации ЛБШМП, основной операцией которой является сравнение фрагментов окрестностей 1 и 2:

$$M_{P,r,P_f,r_f}(x,y) = \sum_{i=0}^{P-1} s(z_i) 2^i, \quad (7)$$

$$s(z_i) = \begin{cases} 1, & z_i \geq \tau; \\ 0, & z_i < \tau, \end{cases}$$

$$z_i = \text{med}(C_{1,i}) - \text{med}(C_{2,i}),$$

$$\text{med}(C_{q,i}) = \text{med}(g(0), \dots, g(l), \dots, g(2l)) = g(l),$$

где z_i – разность медианных значений фрагментов, g – порядковый номер окрестности, $q = \overline{1, 2}$, i – номер шага сравнения фрагментов, $i = \overline{0, P-1}$, τ – пороговое значение берется большим, чем ноль (например, $\tau = 0,01$), $g(j)$ – значение яркости пикселя выражение для нахождения яркости представлено в 5, $\text{med}(C_{q,i})$ – медианное значение фрагмента $C_{q,i}$, элементы которого, расположены в порядке возрастания значений яркости $g(j)$, т.е. выполняется условие: $g(0) \leq \dots \leq g(l) \leq \dots \leq g(2l)$. На рис. 3 представлен пример вычисления $\text{med}(C_{q,i})$.

Алгоритм вычисления гистограмм ЛБШМП:

Шаг 1. Изображение делится на регионы, для каждого региона, в зависимости от его ценности, устанавливаются значения весов V_k (например, для глаз $V_k = 1, 2$, для носа $V_k = 1, 1$, для рта $V_k = 1$).

Шаг 2. Вычисляется медианное значение для каждого фрагмента окрестности 1 и окрестности 2.



а) изображение; б) представление – $LBP_{8,1}$; в) представление – $TP_{8,1,2}$.

Рис. 6. а) изображение; б) представление – $LBP_{8,1}$; в) представление – $TP_{8,1,2}$.

г) Представление - $FP_{8,1,2,1}$ · д) Представление – $M_{12,5,8,1}$

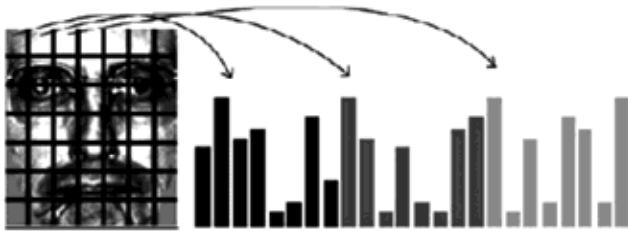
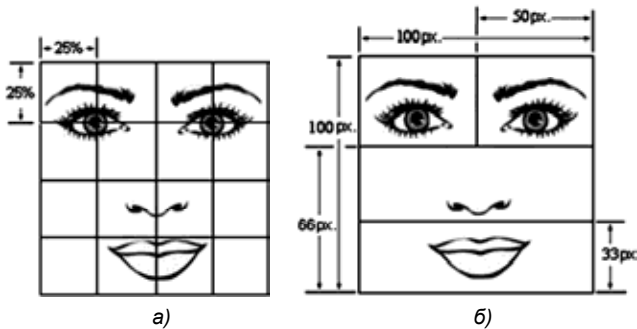
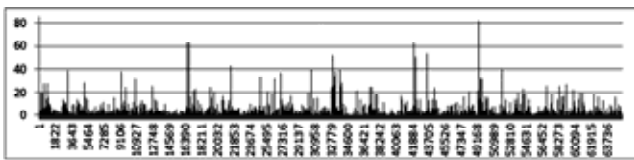


Рис. 7. Пример построения общей гистограммы

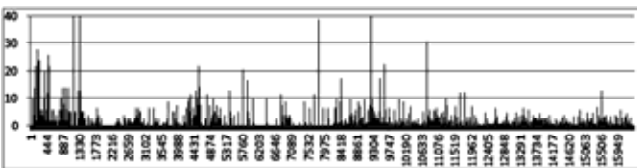


а)

б)



е)



г)

Рис. 8. а) схема из 16 регионов – 4×4 ;

б) схема из 4 регионов – 3×2 , с областями глаз, носа и рта;

в) гистограмма из 16 регионов – 4×4 ;

г) гистограмма из 4 регионов – 3×2 ,

с областями глаз, носа и рта

Шаг 3. Для каждой связки фрагментов, вычисляется z_i . Если z_i меньше порогового τ , то результатом будет – 0, иначе 1.

Шаг 4. Результаты сравнения переводятся в десятичную систему счисления и объединяются в одну общую кодовую комбинацию.

Шаг 5. Полученное число на шаге 4 является номе-

ром столбца h гистограммы H_{P_r, P_f, r_f} , к значению которого добавляется 1.

Шаг 6. Собранные гистограммы регионов объединяются в одну общую гистограмму с учетом веса региона.

Результаты работы

Информативные признаки тестировались на базе данных (далее БД) (табл. 1), полностью включающей изображения баз Калифорнийского технологического института [10], TheIndianFaceDatabase[5] и выборочно FERET[8]. Изображения подбирались таким образом, чтобы избежать перекрытий регионов лица, больших поворотов головы ($>20^\circ$) и сильных эмоций. В базе содержатся изображения лиц людей различных рас, возрастов и мимики, изображения подготавливались с различными условиями съемки – освещением, масштабом и ракурсом. Алгоритм предобработки и приведения изображений к общей форме приведен в [2].

Таблица 1.

	Тестовая выборка	Обучающая выборка
Зарегистрированные	1000 изображений, 100 человек	200 изображений, 100 человек
Отсутствующие в БД	200 изображений различных лиц	

Применение в роли информативного признака гистограмм ЛБШМП, позволило добиться ожидаемых результатов распознавания. Для сопоставления информативных признаков использовалась функция расчета расстояния – ядро пересечения гистограмм. При построении информативных признаков использовалась схема на рис. 7 б. Процент случаев выдачи доступа зарегистрированному пользователю составил 98,9, а процент выдачи доступа человеку, не имеющему допуск – 3,5. Использование информативного признака гистограмм кодов ЛБШМП (рис. 5) увеличивает вероятность правильного распознавания на 13 % и 7 % по сравнению с Three-PatchЛБШ и Four-PatchЛБШ на подобранной БД.

Использование информативного признака гистограмм кодов ЛБШМП (рис. 5) увеличивает скорость работы в 1,7 и 2 раза по сравнению с Three-PatchЛБШ и Four-PatchЛБШ, соответственно. Ниже в табл. 2. выполнен сравнительный анализ скоростных характеристик с аналогами алгоритма, выполняемых на изображении размером 100×100 пикселей.

Таблица 2.

Алгоритм	Время выполнения
Three-Patch ЛБШ	107 мс
Four-Patch ЛБШ	124 мс
ЛБШМП	61 мс

Заключение

В данной работе гистограммы ЛБШМП хорошо зарекомендовали себя в процессе распознавания лиц. Разработанные информативные признаки найдут свое применение в самых различных областях компьютерного зрения, таких как распознавание дорожных знаков, распознавание номеров транспортных средств и т.д.

Основной задачей, которую необходимо решить при использовании гистограмм ЛБШМП, является наиболее корректный подбор параметров информативного признака.

Представленный алгоритм вычисления информативных признаков можно модифицировать: добавлять новые фрагменты или изменять расположение фрагментов, с условием сохранения основной идеи исследования.

Статья подготовлена при поддержке фонда содействия развитию малых форм предприятий в научно-технической сфере в номинации «Информационные технологии» («У.М.Н.И.К.» I полугодие г. Самара 2014).

Литература

1. Ефимов И.Н. Методы повышения достоверности идентификации обучаемых при дистанционном обучении / И.Н. Ефимов // Дистанционное и виртуальное обучение – 2013. – № 3 – 62-73 с.
2. Ефимов И.Н. Биометрическая идентификация в дистанционной подготовке кадров на ж. д. транспорте / И.Н. Ефимов, А.М. Косолапов // Вестник транспорта Поволжья – 2013. – № 1 – 62-66 с.

3. Ahonen T. Face Recognition with Local Binary Patterns, 2004. – 469-481 с.
4. Ahonen T. Face description with local binary patterns: Application to face recognition / T. Ahonen, A. Hadid, M. Pietikäinen // IEEE Trans. Pattern Anal. Mach. Intell. – 2006. – Т. 28 – № 12 – 2037–2041с.
5. Jain V. The Indian face database / V. Jain, A. Mukherjee – 2002.
6. O'Connor B. Facial Recognition using Modified Local Binary Pattern and Random Forest / B. O'Connor, K. Roy // Int. J. Artif. Intell. Appl. – 2013. – Т. 4 – № 5.
7. Ojala T. Multiresolution gray-scale and rotation invariant texture classification with local binary patterns / T. Ojala, M. Pietikainen, T. Maenpaa // IEEE Trans. Pattern Anal. Mach. Intell. – 2002. – Т. 24 – № 7.
8. Phillips P.J. The facial recognition technology (FERET) database / P. J. Phillips // IEEE Trans. Pattern Anal. Mach. Intell. – 2004. – Т. 22.
9. Trefný J. Extended set of local binary patterns for rapid object detection / J. Trefný, J. Matas // Proc. Comput. Vis. Winter – 2010.
10. Weber M. Frontal face dataset, california institute of technology // – 1999.

LOCAL BINARY PATTERNS MEDIAN PIXEL'S – EFFECTIVE TECHNOLOGY OF PATTERN RECOGNITION

Efimov I.N,

In this paper author propose the new informative features local binary patterns median pixel's (LBPMP). The author changed the shape and the algorithm for constructing local binary patterns. Author compared the advantages of the new feature in relation to pattern recognition.

УВАЖАЕМЫЕ КОЛЛЕГИ!

Для тех, кто не успел оформить подписку на первое полугодие 2015 года через ОАО «Роспечать», сохраняется возможность приобретения журналов непосредственно в редакции по адресу: г. Москва, ул. Авиамоторная, дом 8, Научный Центр МТУСИ, ком. 612. Российское научно-техническое общество радиотехники, электроники и связи им. А.С. Попова, метро «Авиамоторная», или оформить Заказ в соответствии с требованиями, выставленными на сайте журнала: www.dsra.ru.

Справки по телефонам:

(8-495) 362-42-75; (+7 903) 221-79-79 (Алексеева Любовь Ильинична).

E-mail: rntores@mail.ru

ПРЕОБРАЗОВАНИЕ ФУРЬЕ В ЗАДАЧЕ ИЗМЕРЕНИЯ РАССТОЯНИЯ ЧАСТОТНЫМ ДАЛЬНОМЕРОМ В ПРОСТРАНСТВЕ С ДИСПЕРСИЕЙ

Давыдочкин В.М., к.т.н., начальник отдела ООО «КОНТАКТ-1», e-mail: skb@kontakt-1.ru.

Ключевые слова: преобразование Фурье, дисперсия шума, спектр.

Введение

Во многих современных системах контроля и управления технологическими процессами существует необходимость точного измерения уровня жидкости в резервуарах. Последние годы в промышленности наблюдается устойчивый рост интереса к радиолокационным методам измерения уровня, в частности, к волноводным уровнемерам с полыми волноводами, в которые свободно заходит жидкий контролируемый продукт через перфорационные отверстия. Основная проблема при таких методах измерения связана с повышенной погрешностью измерения из-за дисперсии в волноводе, вследствие которой возникает нелинейная зависимость времени задержки эхо сигнала от частоты и, соответственно, искажения спектра сигнала разностной частоты (СРЧ). За оценку частоты, линейно связанную с расстоянием, наиболее часто принимают частоту разностного сигнала, соответствующую максимуму спектральной плотности амплитуды (СПА). Основные недостатки такой оценки, обусловленные методической погрешностью преобразования Фурье из-за низкой разрешающей способности и искажающих действий боковых лепестков, во многом устраняются сглаживающими весовыми функциями (ВФ) и увеличением диапазона модуляции. Однако, искажения обусловленные дисперсией, сводят на нет преимущества большого диапазона модуляции, который используется в современных прецизионных радиолокационных уровнемерах, и снижают эффективность применения ВФ. Из-за негативного влияния дисперсии ограничиваются использованием волноводов с длинами, не превышающими единиц метров, в то время как в промышленности часто требуются измерения уровней с применением волноводов с длинами до полутора десятков метров, а иногда и более.

Целью работы является оценка влияния дисперсии в волноводах на искажения спектра СРЧ, разработка и исследование метода снижения этих искажений.

Влияние дисперсии на сигнал разностной частоты

Учтём в традиционной математической модели СРЧ [1] зависимость длины волны в волноводе от его поперечного сечения [2].

$$u(t) = U(t) \cos \Phi(t) = U(t) \cos \left[L \frac{4\pi}{\lambda_{\text{св}}(t)} \sqrt{1 - \lambda_{\text{св}}^2(t)/\lambda_{\text{кр}}^2} - \varphi_{\text{м}} \right], \quad (1)$$

Исследовано влияние дисперсии в волноводе на свойства сигнала разностной частоты. Показано, что дисперсия приводит к возникновению нелинейной частотной модуляции в сигнале разностной частоты, вызывающей искажение его спектра и соответствующее увеличение погрешности измерения расстояния. Предложен вариант преобразования Фурье, позволяющий снизить искажения спектра, и исследованы его свойства.

где $\lambda_{\text{св}}(t)$ – текущая длина волны при распространении в свободном пространстве, $\lambda_{\text{кр}}$ – критическая длина волны в волноводе, L – геометрическая длина волновода до зондируемой поверхности, $U(t)$ – функция, характеризующая паразитную амплитудную модуляцию (ПАМ) СРЧ.

Из (1) следует, что дисперсия в волноводе порождает паразитную нелинейную зависимость разностной частоты $\Omega(t)$ от времени, т.е. его паразитную частотную модуляцию (ПЧМ) при линейной ЧМ передатчика, которая приводит к искажению спектра.

$$\Omega(t) = \Delta\omega 2L(1+at)/cT \sqrt{(1+at)^2 - b^2}, \quad (2)$$

где c – скорость распространения волн в свободном пространстве; T – длительность интервала анализа, совпадающая с длительностью монотонного изменения частоты при модуляции по линейному закону; $t = t_a/T$; t_a – текущее время внутри интервала анализа; $a = \Delta\omega/\omega_0$; $b = \omega_{\text{кр}}/\omega_0$; ω_0 – начальная частота диапазона модуляции при $t = 0$; $\Delta\omega$ – диапазон модуляции в интервале T ; $\omega_{\text{кр}}$ – критическая частота волновода.

Однако среднее значение разностной частоты линейно зависит от расстояния до зондируемой поверхности в волноводе

$$\Omega_{\text{ср}} = \frac{1}{T} \int_0^T \Omega(t) dt = \omega_0 2L \left[\sqrt{(1+a)^2 - b^2} - \sqrt{1-b^2} \right] / cT. \quad (3)$$

Выражение (3) позволяет определить расстояние L до зондируемой поверхности по оценённой каким либо образом средней частоте разностного сигнала.

В дальнейшем используем безразмерные нормированные частоты и расстояния: $x = \Omega T / 2\pi$ – нормированная текущая частота; $x_{\text{ср}} = \Omega_{\text{ср}} T / 2\pi$ – нормированная средняя частота; $x_L = \Delta\omega L / c\pi$ – нормированное расстояние до зондируемого объекта; при $b = 0$ (свободное пространство) $x_{\text{ср}} = x_L$.

Преобразование Фурье для обработки сигнала разностной частоты с паразитной частотной модуляцией

Для снижения искажений спектра используем вариант преобразования Фурье [3], применимый к сигналам с паразитной частотной модуляцией

$$S(x) = \int_0^1 u(t)w[\varphi(t)] \frac{d\varphi(t)}{dt} \exp[-j2\pi x\varphi(t)] dt \quad (4)$$

где $u(t)$ - выделенный СРЧ; $w[\varphi(t)]$ - ВФ;

$$\varphi(t) = \int_0^t x(t)dt / \varphi(t=1); \quad (5)$$

$\varphi(t=1) = \varphi(t)$ при $t=1$.

Оценка разностной частоты в спектральной области предполагает использование методов цифровой обработки сигналов. Выполнение преобразований (4) и (5) даёт вариант интегрально-дискретного преобразования Фурье применимый для практического анализа сигнала (1)

$$S_d(x) = \frac{1}{M} \frac{a}{\sqrt{(1+a)^2 - b^2} - \sqrt{1-b^2}} \times \sum_{m=0}^{M-1} \left\{ u(m) w(m) \frac{1+a(m+0,5)/M}{\sqrt{[1+a(m+0,5)/M]^2 - b^2}} \times \exp \left[-j2\pi x \frac{\sqrt{(1+am/M)^2 - b^2} - \sqrt{1-b^2}}{\sqrt{(1+a)^2 - b^2} - \sqrt{1-b^2}} \right] \right\}, \quad (6)$$

где M – число отсчётов сигнала; $u(m)$ – отсчёт СРЧ на m -ой ступеньке генерируемого зондирующего сигнала; $w(m)$ может быть произвольной.

Учитывая, что все виды ВФ, в том числе и те, которые не выражаются в элементарных функциях, могут быть представлены адаптируемыми ВФ (АВФ) [1, 4, 5] или аппроксимированы АВФ с любой, наперёд заданной точностью, анализ свойств преобразования (6) проведём, используя два семейства АВФ $w(m, b_1, \dots, b_N)$, которые с учётом (4) и (5) для частотной зависимости (2) и для преобразования (6) примут, соответственно, вид

$$w_s(m, b_1, \dots, b_N) = 1 + \sum_{n=1}^N (-1)^n C_{sn}(b_1, \dots, b_N) \cos[2n\pi \Phi(m)] \quad (7)$$

$$w_c(m, b_1, \dots, b_N) = \frac{1}{K} \left\{ \sin[\pi \Phi(m)] + \sum_{n=1}^N (-1)^n C_{cn}(b_1, \dots, b_N) \sin[(2n+1)\pi \Phi(m)] \right\}, \quad (8)$$

где $\Phi(m) = \frac{\sqrt{[1+a(m+0,5)/M]^2 - b^2} - \sqrt{1-b^2}}{\sqrt{(1+a)^2 - b^2} - \sqrt{1-b^2}}$;

$$C_{sn}(b_1, \dots, b_N) = \frac{(-1)^{n+1}}{\cos(nM)} \prod_{i=1}^N \left[\frac{\cos(2nM) - \cos(2b_i M)}{2 \sin^2(b_i M)} \right] \times$$

$$\times \prod_{\substack{k=1 \\ k \neq n}}^N \frac{[1 - \cos(2kM)]}{\cos(2nM) - \cos(2kM)};$$

$$C_{cn}(b_1, \dots, b_N) = \frac{(-1)^{n+1} \sin(0,5\pi/M)}{\sin[(n+0,5)\pi/M]} \times$$

$$\times \prod_{i=1}^N \left\{ \frac{\cos[(2n+1)\pi/M] - \cos(2b_i \pi/M)}{\cos(\pi/M) - \cos(2b_i \pi/M)} \right\} \times$$

$$\times \prod_{\substack{k=1 \\ k \neq n}}^N \frac{\cos(\pi/M) - \cos[(2k+1)\pi/M]}{\cos[(2n+1)\pi/M] - \cos[(2k+1)\pi/M]};$$

$$K = \frac{-2}{M_o} \left\{ \frac{\sin(0,5\pi/M)}{\cos(\pi/M) - 1} + \right.$$

$$\left. + \sum_{n=1}^N C_{cn}(b_1, \dots, b_N) \frac{\cos(n\pi) \sin[(n+0,5)\pi/M]}{\cos[(2n+1)\pi/M] - 1} \right\}; \quad N - \text{число}$$

варьируемых тригонометрических слагаемых весовой функции; n – номер тригонометрического слагаемого; $b_i = \omega_i T / 2\pi$ – нормированная частота, на которой задаётся равенство нулю спектральной плотности амплитуды (СПА) весовой функции.

На рис. 1 показаны нормированные амплитудные спектры СРЧ (1), рассчитанные по стандартному преобразованию Фурье (пунктир) и по преобразованию (6) – (сплошная линия), ВФ – равномерная. Число отсчётов сигнала $M=64$, заданное нормированное расстояние до отражателя по свободному пространству 10, с учётом дисперсии значение нормированной разностной частоты (3) равно $x_{cp} = 14,389$. Здесь и далее в расчётах использованы: диаметр круглого волновода 25 мм; $f_0 = 8,5$ ГГц; использован максимально широкий выделенный диапазон модуляции $\Delta f = 3$ ГГц.

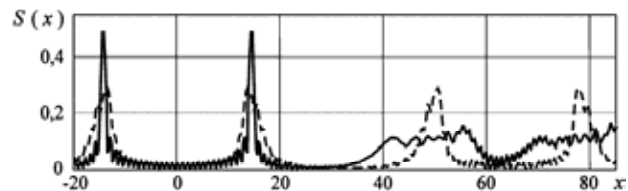
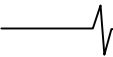


Рис. 1. Нормированные амплитудные спектры СРЧ

Из сравнения видны значительные искажения основного лепестка спектра по стандартному преобразованию Фурье, приводящие к погрешности оценки частоты. В данном примере нормированная погрешность оценки частоты составляет $\Delta x = 0,85$.

Анализируя искажения спектра при использовании преобразования (6), отметим практически неискажённый основной лепесток на частоте x_{cp} (степень искажения оценим ниже) и лепесток на частоте $-x_{cp}$. Но на нормированных частотах $M - x_{cp}$, $-(M - x_{cp})$ (на рисунке не показан) лепестки сильно искажённые. Искажения последующих лепестков нарастают. Причём амплитудные значения лепестков на нормированных частотах $M - x_{cp}$, $-(M - x_{cp})$ уменьшаются.



При отсутствии дисперсии, помех и шума погрешность оценки частоты определяется боковыми лепестками слагаемого спектра из области $\Omega < 0$ (методическая погрешность) [1]. Откуда следует возможность снижения методической погрешности сглаживающими ВФ. Этот вопрос рассмотрим более детально.

При отсутствии дисперсии и шума АВФ, в силу заложенных при их получении спектральных свойств, теоретически обеспечивают получение нулевой погрешности оценки частоты разрешаемого сигнала [4, 5, 6]. В этой связи необходимо выяснить насколько сохраняются спектральные свойства АВФ и какие условия получения сигнала для этого необходимы. При этом необходимо получить ответ на три вопроса.

Во-первых, сохраняется ли возможность задания на произвольной частоте равенства нулю СПА вместе с заданным количеством производных СПА? Важность этого вопроса требует пояснения. С использованием АВФ могут быть устранены погрешности измерения частот, амплитуд и фаз разрешаемых спектральных слагаемых многочастотного сигнала. Это осуществляется заданием формы ВФ варьируемыми параметрами таким образом, чтобы на частоте каждого из слагаемых сигнала СПА других слагаемых, т.е. мешающие слагаемые, были равны нулю вместе с заданным количеством их производных [1, 4, 5, 6].

Во-вторых, насколько искажается основной лепесток спектра? Поскольку искажения основного лепестка, в особенности несимметричные, могут привести к погрешности оценки частоты.

И третий важный вопрос, какое количество отсчётов сигнала необходимо для получения допустимого уровня искажения спектра?

Для ответа на первый вопрос рассмотрим спектры СРЧ (1) в области боковых лепестков при $U(t) = const$, $x_{cp} \ll M$, и при x_{cp} , близкое к $M/2$.

На рис. 2 а и 2 б приведены участки спектров СРЧ с $M = 1024$ в области частот $x \ll M$. В расчётах использовались: заданное нормированное расстояние до отражателя по свободному пространству 41,698; значение нормированной разностной частоты (3), с учётом дисперсии, равно $x_{cp} = 60$; АВФ – (6) с $N = 3$ и тремя нормированными частотами, на которых задаётся равенство нулю погрешности оценки частоты $b_1 = 160,5$, $b_2 = b_3 = 40,5$. Для сравнения на этих же рисунках приведены аналогичные участки спектров СРЧ с нормированной частотой $x_{cp} = 60$, АВФ с такими же параметрами, но для пространства без дисперсии. В этом случае преобразование (6) переходит в стандартное интегрально-дискретное преобразование Фурье при $b = 0$. На рис. 2а различия спектров не обнаруживаются.

Поясним выбор параметров АВФ. Рисунками иллюстрируется возможность получения на нормированной частоте, равной 100,5, нулевого значения слагаемого спектра СРЧ, нормированная частота которого $x_{cp} = 60$. Тем самым, если в зондируемом пространстве будет отражатель, создающий СРЧ с частотой $x_{cp} = 100,5$ то

основной лепесток его спектра не будет искажаться боковыми лепестками обоих слагаемых спектра СРЧ с нормированной частотой $x_{cp} = 60$ и, соответственно, его параметры (частота, фаза и амплитуда) могут быть оценены достаточно точно.

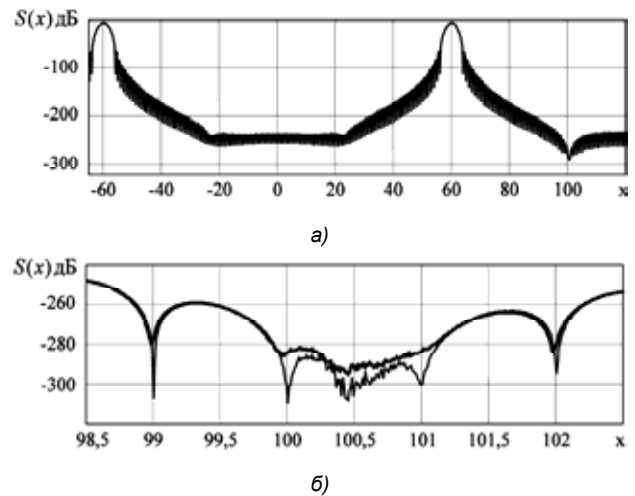


Рис. 2. Оценка искажений спектра в области боковых лепестков

На более детальном рис. 2 б видно, что в рассматриваемом примере глубина подавления боковых лепестков на заданной нормированной частоте теоретически составляет около 290 дБ для преобразования (6) (толстая линия) и около 300 дБ для стандартного преобразования Фурье (тонкая линия) и СРЧ, полученному без дисперсии.

На рис. 3 а и 3 б приведены участки спектров СРЧ в области частот соизмеримых с $M/2$. Заданное нормированное расстояние до отражателя по свободному пространству 300. С учётом дисперсии значение нормированной разностной частоты (3) равно $x_{cp} = 431,671$. При расчете использованы: АВФ (7) с $N = 3$ и тремя нормированными частотами $b_1 = 878,342$, $b_2 = b_3 = 15$; $M = 1024$ (пунктир на рис. 3 а), $M = 1280$ и 2048 (толстые линии сливаются в одну). Как и в первом примере, b_1 минимизирует слагаемое спектра с нормированной частотой $-x_{cp}$. В данном примере минимизируются слагаемые спектра на нормированной частоте превышающей нормированную частоту сигнала x_{cp} на 15.

Из рисунка видно, что при $M = 1024$ сильно искажённый лепесток слагаемого спектра с частотой $M - x_{cp}$ искажает интересующий нас участок спектра, увеличивая уровень боковых лепестков более чем на 100 дБ. Увеличение M до 1280 и тем более до 2048 практически исключило влияние указанного слагаемого.

Для сравнения на этих же рисунках приведены аналогичные участки спектров СРЧ с $M = 1280$ нормированной частотой $x_{cp} = 431,671$ и АВФ (7) для пространства без дисперсии. На рис. 3 а различия не обнаруживаются.

На более детальном рис. 3 б видно, что в рассматриваемом примере глубина подавления боковых лепестков на заданной нормированной частоте для преобразова-

ния (б) составляет более 220 дБ при $M = 1280$ (толстая линия), 250 дБ при $M = 2048$ (тонкая линия) и около 290 дБ для стандартного преобразования Фурье и СРЧ, полученному без дисперсии (тонкий пунктир).

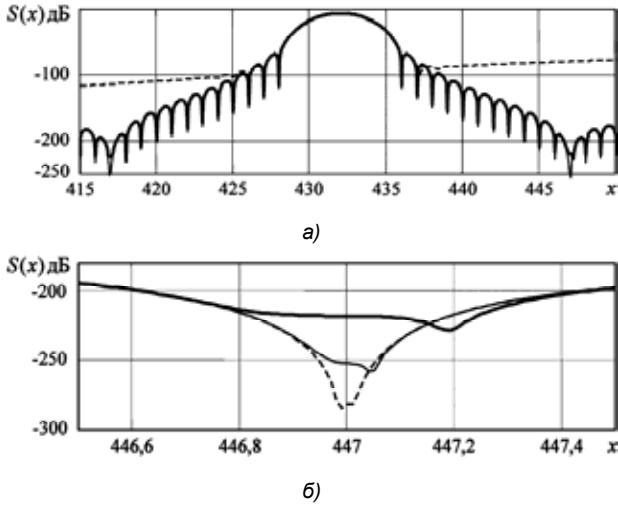


Рис. 3. Оценка искажений спектра в области боковых лепестков

Рассмотренные примеры показывают, что как при $x_{cp} \ll M$, так и при x_{cp} соизмеримом с $M/2$, искажения спектра в области боковых лепестков практически отсутствуют, если частота сигнала на 20...40% ниже частоты Найквиста.

Для ответа на второй вопрос оценим получаемый результат сохранения формы основного лепестка спектра. Для этого определим разницу основных лепестков модулей спектра СРЧ искажённого дисперсией и неискажённого дисперсией, но с заданной центральной частотой (3).

Сравнение проведём для тех же условий, но нормированные частоты $b_1 = b_2 = b_3 = 2x_{cp}$ задаются из условия подавления мешающего слагаемого спектра с нормированной частотой $-x_{cp}$. На рис. 4 а и 4 б приведены результаты расчётов. Из рисунков следует, что как на близком нормированном расстоянии $x_{cp} \ll M$ (рис. 4 а), так и при x_{cp} , соизмеримом с $M/2$ (рис. 4 б), основной лепесток спектра полученный с использованием преобразования (6), несколько уже чем при использовании стандартного преобразования. Для использованной ВФ (7) и $M = 2048$ ширина основного лепестка спектра СРЧ, искажённого дисперсией и вычисленного по преобразованию (6), на 0,007% меньше ширины основного лепестка спектра СРЧ не искажённого дисперсией в максимуме их разницы. Уменьшение числа отсчётов сигнала до $M = 1280$ приводит к увеличению разницы, но симметрия сохраняется как для $x_{cp} \ll M$, так и для $x_{cp} = 431,671$. Однако дальнейшее снижение числа отсчётов до 1024 приводит к нарушению симметрии основного лепестка при $x_{cp} = 431,671$, что, в свою очередь, приводит к погрешности оценки частоты по максимуму спектра. Нормированная погрешность оценки составляет около 0,0001.

Таким образом, не симметричные искажения спектра в области основного лепестков практически отсутствуют, если частота сигнала ниже частоты Найквиста на те же 20...40%.

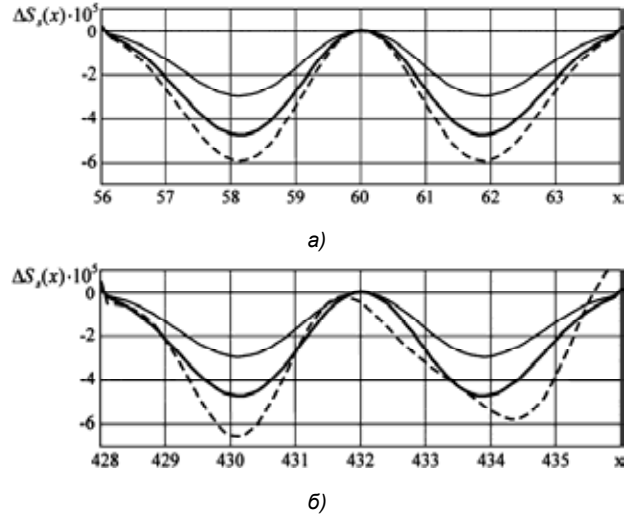


Рис. 4. Оценка искажений спектра в области основного лепестка

Обратим внимание ещё на одну особенность преобразования (6) для сигнала (1), отмеченную на рис. 1. Из-за того, что амплитудные значения лепестков спектра на нормированных частотах $M - x_{cp}$, $-(M - x_{cp})$ и последующих значительно уменьшаются возможна оценка частоты сигнала при $x_{cp} > M/2$. Спектры таких сигналов приведены на рис. 5 а и рис. 5 б. Для расчета спектра использованы 256 отсчётов СРЧ (2) с частотой $x_{cp} = 431,671$ и АВФ (6) с $N = 2$, $b_1 = 2$, $b_2 = 2,6$ (рис. 5 а) и с частотой $x_{cp} = 1438,9$ и той же АВФ (рис. 5 б). Несмотря на то, что x_{cp} первого сигнала почти в 3,5 раза превышает частоту Найквиста, а второго сигнала более чем в 11 раз, их можно обнаружить и оценить параметры сигнала, если требования к погрешности оценок не высокие. В частности, нормированная погрешность оценки частоты в рассмотренном примере для первого сигнала составляет $\Delta x = 0,0415$, а для второго сигнала $\Delta x = 0,0576$. Максимальные значения погрешности (около 0,1) наблюдаются вблизи частоты Найквиста.

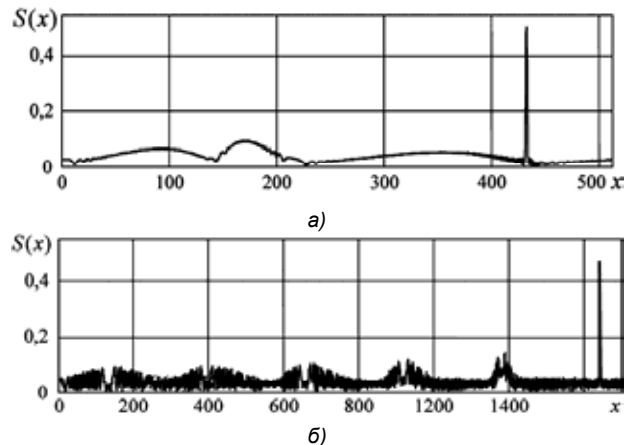
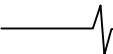


Рис. 5. Нормированные амплитудные спектры СРЧ



Заключение

Наличие дисперсии в зондируемом пространстве приводит к погрешности измерения расстояния, величина которой обычно неприемлема для измерителей расстояния промышленного применения. Получено выражение интегрально-дискретного преобразования Фурье, позволяющее снизить искажения спектра до пренебрежимо малого уровня в области частот ниже частоты Найквиста.

Показано, что искажения спектра, как основного лепестка, так и боковых лепестков пренебрежимо малы, если частота сигнала на 20...40% ниже частоты Найквиста. Показано, что, если требования к погрешности оценок не высокие, то оценку можно выполнить, даже если частота сигнала в несколько раз выше частоты Найквиста.

Литература

1. Прецизионные системы ближней частотной радиолокации промышленного применения. Монография. / Атаянц Б.А., Давыдочкин В.М., Езерский В.В., Паршин В.С., Смольский С.М. – М.: Радиотехника, 2012. – 512 с.
2. Гольдштейн Л.Д., Зернов Н.В. Электромагнитные поля и волны. Изд. 2-е. М.: Сов. Радио. 1971. – 664 с.
3. Давыдочкин В.М. Методы снижения погрешности оценки параметров полигармонических сигналов в ближней частотной радиолокации. // Вестник РГРТА. Рязань, РГРТА, 2006. Вып. 18. – с. 63-70.

4. Давыдочкин В.М., Давыдочкина С.В. Весовые функции для адаптивного гармонического анализа сигналов с многомодовым спектром // ЦОС. 2008. № 4. С. 44-48.

5. Давыдочкин В.М., Давыдочкина С.В. Весовые функции для цифрового адаптивного гармонического анализа сигналов // Радиотехника. 2009. № 9. С. 11-20.

6. Способ гармонического анализа периодического многочастотного сигнала. Патент РФ №2435168. Оpubл. 27.11.2011. Бюлл. – № 33.

APPLICATION OF FOURIER TRANSFORM TO THE PROBLEM OF DISTANCE MEASUREMENT BY FREQUENCY-MODULATED RANGEFINDER IN THE PRESENCE OF DISPERSION.

Davydochkin V.M.

We investigate the influence of waveguide dispersion on the properties of difference frequency signal. The dispersion leads to emergence of nonlinear frequency modulation in the difference frequency signal, which causes distortion of its spectrum and the corresponding increase of the distance measurement error. We propose a version of Fourier transform that allows to reduce this spectrum distortion and investigate its properties.

УВАЖАЕМЫЕ АВТОРЫ!

Редакция научно-технического журнала «Цифровая обработка сигналов» просит Вас соблюдать следующие требования к материалам, направляемым на публикацию:

1) Требования к текстовым материалам и сопроводительным документам:

1. Текст - текстовый редактор *Microsoft Word*.
2. Таблицы и рисунки должны быть пронумерованы. На все рисунки, таблицы и библиографические данные указываются ссылки в тексте статьи.
3. Объем статьи до 12 стр. (ширифт 12). Для заказных обзорных работ объем может быть увеличен до 20 стр.
4. Название статьи на русском и английском языках.
5. Рукопись статьи сопровождается:
 - краткой аннотацией на русском и английском языках;
 - номером УДК;
 - сведениями об авторах (Ф.И.О., организация, должность, ученая степень, телефоны, электронная почта);
 - ключевыми словами;
 - актом экспертизы (при наличии в вашей организации экспертной комиссии).

2) Требования к иллюстрациям:

- Векторные (схемы, графики) - желательно использование графических редакторов *Adobe Illustrator* или *Corel DRAW*.
- Растровые (фотографии, рисунки) - М 1:1, разрешение не менее 300dpi, формат *tiff, jpg*.

УДК 681.3.07

ЭКРАНИРОВАНИЕ В СИСТЕМАХ ЦИФРОВОЙ ОБРАБОТКИ СИГНАЛОВ

*Ботвинёнок А.А., Курпичников А.П., институт проблем управления
им. В.А. Трапезникова РАН, e-mail: snv@ipr.ru.*

В отличие от систем управления на малопотребляющих микроконтроллерах и многих других применений цифровых систем, специфика изделий ЦОС заключается в использовании большого вычислительного ресурса для обработки сигналов в реальном времени в сочетании с наличием аналоговых цепей и датчиков вблизи – как причин и сигналов для последующей обработки. Немаловажным является и то, что для дискретизации используются устройства с широкополосным входом (УВХ АЦП), для которых характерно, помимо прочих проблем помехоустойчивости, перемножение сигнала с помехой, что существенно усложняет ситуацию с обработкой. На помощь приходят сложные алгоритмы и новейшая элементная база с сотнями миллионов транзисторов, нанотехнологическими проектными нормами и низковольтным питанием (0,9 В и ниже).

Однако, наряду с благом возрастающих вычислительных возможностей и уменьшении потребляемой мощности и стоимости, современная элементная база (в части вычислителей) несёт в себе скрытую угрозу – существенно сниженную устойчивость к воздействиям ЭМИ. Расширение рабочей полосы логических схем, снижение паразитных емкостей и пороговых уровней при общем сохранении размеров кристалла дают оценку снижения устойчивости к ЭМИ на 2 порядка и более. Таким образом, цифровая часть и, в частности, элементы ЦОС, из генераторов помех превращаются в хорошие приемники приходящих извне помех (что часто проявляется в статистике сбоев при выполнении алгоритмов).

Кроме того, всё большее развитие приобретают различные беспроводные технологии обмена данными, поэтому зачастую в едином устройстве сосредоточены разнородные модули: радиointерфейсы и вычислители, что приводит к дополнительному воздействию ЭМИ на вычислительные средства.

Решение проблемы, на первый взгляд, кажется очевидным – экранирование, тем более, что большинство литературных источников даже для простого однослойного экрана приводят цифры эффективности 60-80 дБ и более. Но в массе практических случаев это не соответствует действительности.

Рассмотрим простейший случай, когда на поверхность замкнутого экрана (плоская коробочка, припаянная к слою заземления на плате) падает плоская электромагнитная волна. Экранирующий эффект складывается в основном из отражения от внешней поверхности экрана, поглощения внутри экрана и повторного (многократного) отражения от внутренней поверхности экрана [1]:

$$A_{\text{дб}} = A_{\text{отр}} + A_{\text{погл}} + A_{\text{м.отр}} \quad (1)$$

Если рассмотреть каждую составляющую в отдельности, то в соответствии с [2]:

$$A_{\text{отр}} = 20 \log_{10} \left| \frac{\eta_0}{4\eta} \right| = 20 \log_{10} \left| \frac{\sqrt{(\sigma + j\omega\varepsilon)\varepsilon_0}}{4\mu_0 \sqrt{4j\omega\mu}} \right| \quad (2)$$

где ε_0 – электрическая постоянная, μ_0 – магнитная постоянная, ε, μ – диэлектрическая и магнитная проницаемости экрана, σ – удельная проводимость материала экрана, ω – круговая частота электромагнитного поля. При подстановке констант в (2) получим приближительное равенство [1]:

$$A_{\text{отр}} = 20 \log_{10} \left(94.25 \sqrt{\frac{\sigma}{\omega\mu}} \right) \quad (3)$$

Потери на поглощение и многократное отражение при этом составляют [1]:

$$A_{\text{погл}} = 6.16d\sqrt{\omega\mu\varepsilon}, \text{ где } d - \text{толщина экрана} \quad (4)$$

$$A_{\text{м.отр}} = 20 \log_{10} (1 - e^{-2d/\sigma}) = 20 \log_{10} (1 - e^{-d\sqrt{2\omega\mu\sigma}}) \quad (5)$$

При этом многократное отражение здесь является отрицательным, т.е. паразитным эффектом, ухудшающим экранирование.

График зависимости эффективности экранирования для электрического поля медным экраном толщиной 0.5 мм от частоты при расстояниях от источника **более одной длины волны (!)** приведён на рис. 1 [2]. Исходя из приведённого графика, эффект такого экрана (верхняя кривая) всегда больше 100 дБ, чего в большинстве случаев более чем достаточно.

Однако, всё вышеизложенное справедливо только в том случае, если на экран падает плоская электромагнитная волна, а это, как раз, означает, что расстояние от экрана до источника излучения превышает λ_0 . Таким образом, например, для радиомодуля GSM (900 МГц) или Wi-Fi (2.4 ГГц) всё вышесказанное соответствует случаю, когда источник излучения (усилитель мощности, антенна модуля) и экран разнесены более чем на 33 см и 13 см соответственно. Однако, в современных компактных конструкциях даже расстояние от аналоговых цепей датчиков и элементов ЦОС до встроенных радиointерфейсов на плате может быть меньшим. Более того, на практике, при применении беспроводных интерфейсов производится экранирование самого радиомодуля с расстояниями менее 1-2 см от внутренних источников ЭМИ (выходной усилитель, связи) до экрана. Это значит, что для оценки эффективности подобных экранов такие рассуждения не подходят, поскольку источник находится в так называемой «ближней зоне» по отношению к экрану. Отметим здесь, что и большинство рассматриваемых при конструировании изделий ЦОС случаев в отношении помех также подпадают под это определение, и исключение составляют разве что достаточно удалённые от устройства радиостанции и станции сотовой связи.

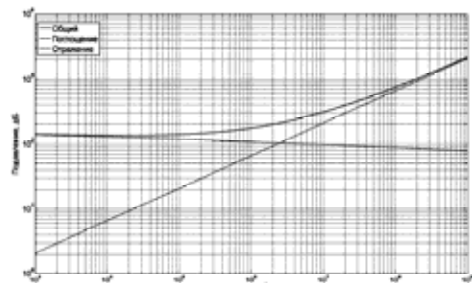


Рис. 1. Зависимость подавления плоской электромагнитной волны медным экраном толщиной 0,5 мм

В простом приближении, опустив рассуждения о переходной зоне, экранирование в ближней зоне отличается импедансами электрического и магнитного полей, которые соответственно равны [2]: $\hat{\eta}_e = 60 \frac{\lambda_0}{r}$, $\hat{\eta}_m = 2369 \frac{r}{\lambda_0}$, где r – расстояние

от источника до экрана, а λ_0 – длина волны излучения. При этом, потери на поглощение электрического поля $A_{\text{погл}}$ остаются такими же, как в дальней зоне, а потери отражения для него составляют ещё больше:

$$A_{\text{отр}} = 20 \log_{10} \left(\frac{\sqrt{\mu}}{4\epsilon r \sqrt{\omega \sigma}} \right) \quad (6)$$

Но главной проблемой в ближней зоне является магнитное поле, подавление которого требует более серьёзных и многослойных экранов. Потери магнитного поля в ближней зоне для упоминавшегося выше медного экрана можно оценить выражением [2, 5]:

$$A_{\text{отр}} = 14.6 + 10 \log_{10} (fd^2) \quad (7)$$

На рис. 2 представлена зависимость подавления магнитной составляющей помехи для медного экрана в зависимости от частоты излучения для различных дистанций до источника. При этом, поскольку источником может быть соседний проводник, для минимального расстояния 2.5 мм показаны также экспериментальные результаты [5] как для медного так и для стального экранов.

Именно магнитное поле может вызывать в системе ЦОС не только наводки в аналоговых цепях (в том числе воздействуя на катушки индуктивности фильтров, дроссели) и индуцировать разность потенциалов на выводах усилителей, но также вывести из строя чувствительные и высокотехнологичные элементы ЦОС. В случае наличия исполнительных устройств с коммутацией тока, внешние наводки от них могут вызвать даже «защёлкивание» КМОП микросхем.

Рассмотрим случай, когда находящийся вблизи платы ключевой МОП транзистор коммутирует ток 20 А (при этом фронт импульса может составлять 20 нс), и пути тока проходят на расстоянии 1 см от параллельной шины процессора с логическими уровнями 3.3 В, ширина разводки которой на РСВ при этом составляет 1 см, а длина, например, 10 см. В этом случае индуцированную ЭДС в шине можно оценить как уровень 5 В и более, что, в случае повторяющегося процесса или ШИМ, может привести не только к помехам на шине, но и к кратковременному (до выключения питания) или полному выходу из строя процессора. Есть повод задуматься о влиянии поля непосредственно на незащищенное ядро процессора ЦОС (под эпоксидным корпусом – кристалл с линейными размерами внутренних связей 0,5 см), где пороговые уровни логики всего 0,3 В!

Всё вышесказанное имеет в своей основе существенную идеализацию – замкнутость экрана. В реальности экраны редко бывают замкнутыми и имеют щели, вентиляционные отверстия, каналы ввода кабелей, плохо экранированные окна под дисплей и т.п. В таком случае, эффективность экрана определяется именно наличием таких отверстий. Количественную оценку эффективности экрана с отверстиями можно получить из следующих выражений [4]:

$$S = A_{\text{отр}} + A_{\text{погл}} + A_{\text{множ.отр}} + K1 + K2 + K3 \quad (8)$$

где $A_{\text{отр}} = 20 \log_{10} [1 + 4K^2] / 4K$ – потери на отражение, а K выбирается по следующим правилам: $K = W / \pi r$ – для низкоимпедансных полей и прямоугольного отверстия длиной W ; $K = D / 3.682r$ – для низкоимпедансных полей и круглыми отверстиями диаметром D ; $K = j6.69 \times 10^{-5} fW$ – для плоской электромагнитной волны и прямоугольного отверстия длиной W на частоте f ; $K = j5.79 \times 10^{-5} fD$ – для плоской электромагнитной волны и круглого отверстия диаметром D на частоте f .

Потери поглощения зависят от формы и размеров отверстия: $A_{\text{погл}} = 27.3 \text{ дБ} / W$ для $f \ll f_c$, прямоугольного отверстия длиной W ; $A_{\text{погл}} = 32 \text{ дБ} / D$ для $f \ll f_c$ – для круглого отверстия

диаметром D . d – глубина отверстия. При этом $f_c = \frac{c}{2W}$ для

прямоугольных отверстий, и $f_c = \frac{c}{3.4D}$ для круглых отверстий.

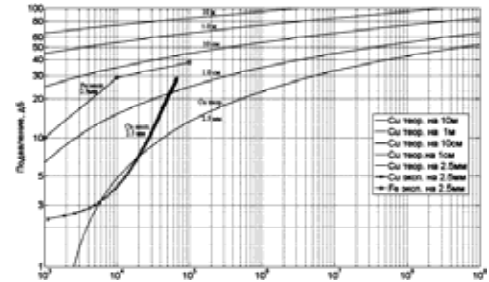


Рис. 2. Зависимость подавления магнитного поля экраном 0,5 мм от частоты для ближней зоны

При эффект множественного отражения также влияет:

$$A_{\text{множ.отр}} = 20 \log_{10} \left[1 - \frac{K-1}{(K+1)^2} 10^{-A_{\text{погл}}/10} \right] \text{ дБ, при } A_{\text{погл}} < 15 \text{ дБ, а}$$

поправочные коэффициенты $K1, K2, K3$ определяются следующим образом: $K1 = -10 \log_{10} an$ – для удалённых источников ($r \gg D, r \gg W$), где a – площадь отверстия, n – количество отверстий на 1 см^2 . Для близких источников этот член следует опускать.

$K2 = -20 \log_{10} (1 + 35p^{-2/3})$ дБ, где p – отношение минимального диаметра (лощины) проводника сетки к глубине скин-слоя в материале.

$K3 = 20 \log_{10} (\coth(A_{\text{погл}} / 8.69))$ дБ – поправочные коэффициенты, описывающий связь между соседними отверстиями экрана, что особенно актуально, например, для рядов вентиляционных отверстий.

Отдельно стоит сказать о вертикальных металлических перегородках. Экранирующий эффект таких незамкнутых экранов обуславливается дифракционными потерями и может быть оценен с достаточной точностью только для случая падения электромагнитной волны т.е. в случае дальней зоны. Дифракционные потери можно оценить, используя зоны Френеля.

Рассматривая излучение и прием помехи, безусловно, следует учитывать также взаимное расположение проводников разводки, как наиболее явных «конструкционных» антенн. Помимо этого, важным эффектом при подавлении помехи является резонанс внутри экрана. Для пустого объема (а это, разумеется, лишь очередное упрощение) резонансная частота определяется выражением [3]:

$$F = 150 \sqrt{(k/l)^2 + (m/h)^2 + (n/w)^2} \text{ МГц,} \quad (9)$$

где l, h, w – размеры корпуса в метрах, а k, m, n – числа, определяющие моду возникающего колебания. При этом, чем выше эффективность отражения поверхности экрана, тем большую добротность будет иметь резонансный контур, увеличивая избирательность системы к ЭМИ и сводя на нет эффект экранирования на волновых максимумах.

Литература

1. «Радиоэлектронные средства и мощные электромагнитные помехи», В.И. Кравченко, А.Е. Болотов, Н.И. Летунова, М.: Радио, 1987г.
2. «Introduction to Electromagnetic Compatibility» Clayton R. Paul, «A Wiley Interscience publication», 1992
3. «ЭМС для разработчиков продукции», Т. Уильямс, перевод с английского Кармашев В.С., Кечиев Л.Н. М.: Технологии, 2003 г.
4. «Principles of Electromagnetic Compatibility», В. Keiser, 3rd edition, Artech House Inc., 1987
5. «Noise reduction techniques in electronic systems» Н.В. Ott, 2nd edition, 1988.



Международная конференция
**Радиоэлектронные устройства и системы
 для инфокоммуникационных технологий**
 (REDS-2015). 27 - 29 мая 2015 г. Москва

INTERNATIONAL CONFERENCE
 THE RADIO-ELECTRONIC DEVICES AND SYSTEMS FOR
 THE INFOCOMMUNICATION TECHNOLOGIES» (REDS-2015)

Конференция посвящена «Дню Радио»



Уважаемые коллеги! В год 70-летия РНТОРЭС им. А.С. Попова

ПРИГЛАШАЕМ ВАС ПРИНЯТЬ УЧАСТИЕ В РАБОТЕ КОНФЕРЕНЦИИ

ОРГАНИЗАТОРЫ:

- Институт радиотехники и электроники имени В.А. Котельникова РАН
- Российское научно-техническое общество радиотехники, электроники и связи им. А.С. Попова
- Российская секция IEEE
- Московский технический университет связи и информатики
- Вневедомственный экспертный совет по проблемам воздушно-космической обороны
- ОАО «Конструкторское бюро – 1»

ПРИ УЧАСТИИ:

- Министерство образования и науки РФ
- Федеральное агентство по промышленности РФ
- ОАО «ГСКБ Концерн ПВО «АЛМАЗ-АНТЕЙ»
- ОАО «Концерн радиостроения «Вега»
- ОАО «Корпорация «Комета»
- ФГУП «НИИР»
- Владимирский государственный университет
- Московский авиационный институт
- Московский энергетический институт
- Московский государственный технический университет им. Н.Э. Баумана
- Нижегородский госуниверситет им. Н.И. Лобачевского
- Рязанский государственный радиотехнический университет
- Санкт-Петербургский государственный университет телекоммуникаций
- Санкт-Петербургский государственный электротехнический университет – («ЛЭТИ»)
- Ульяновский государственный технический университет
- Ярославский государственный университет

ПРОГРАММНЫЙ КОМИТЕТ:

Председатель: академик **Гуляев Юрий Васильевич**

Члены комитета: академик **Бугаев А.С.**, член-корр. **Зубарев Ю.Б.**, академик **Савин А.И.**, академик **Фёдоров И.Б.**

МЕЖДУНАРОДНЫЙ ОРГАНИЗАЦИОННЫЙ КОМИТЕТ:

Председатель: член-корр. **Никитов Сергей Аполлонович**

Сопредседатели: проф. **Аджемов А.С.**, д.т.н. **Ашурбейли И.Р.**, академик **Сигов А.С.**

Члены оргкомитета: к.т.н. **Алёшин В.С.**, проф. **Безруков В.Н.**, д.т.н. **Бутенко В.В.**, проф. **Быховский М.А.**, проф. **Верба В.С.**, проф. **Воробиенко П.П.** (Украина), проф. **Дроздов Б.В.**, проф. **Мисник В.П.**, проф. **Пестряков А.В.**, проф. **Портнов Э.Л.**, проф. **Постников И.И.**, проф. **Потапов В.Т.**, проф. **Пшеничников А.П.**, доц. **Самсонов Г.А.**, д.э.н. **Сеилов Ш.Ж.** (Казахстан), проф. **Скородумов А.И.**, проф. **Сперанский В.С.**, проф. **Степанов С.Н.**, проф. **Хорев А.А.**, к.т.н. **Цыркин В.В.**, проф. **Чупрунов Е.В.**, проф. **Шинаков Ю.С.**

Организация работы в форме пленарных и проблемно-тематических заседаний. Пленарные доклады будут представлены по основным направлениям работы Конференции (г.Москва, Конгресс-центр МТУСИ).

Рабочие языки конференции - *русский и английский*. Доклады, включенные в Программу конференции, будут опубликованы на этих же языках в **сборнике «Доклады» Конференции** к началу ее работы.

Новый уровень производительности модулей ЦОС



Высокопроизводительные
многоядерные процессоры ЦОС
Масштабируемая модульная
многопроцессорная архитектура



Мезонинные модули
ввода-вывода в стандарте FMC
Программируемая логика ПЛИС
PCI Express, cPCI, VPX, VME
Ethernet, Serial Rapid IO, USB 3.0



Закрытое акционерное общество

“Инструментальные Системы”

www.insys.ru, info@insys.ru