





**Главному редактору нашего журнала,  
Председателю Национального организационного комитета  
Международной конференции и выставки  
«Цифровая обработка сигналов и ее применение-DSPA»,  
члену-корреспонденту РАН,  
доктору технических наук, профессору  
Зубареву Юрию Борисовичу 75 лет!**



Свой 75-летний юбилей Зубарев Юрий Борисович встретил 9 февраля 2013 года. Талантливый Ученый, получивший должное признание, как в нашем отечестве, так и за рубежом! Замечательный Человек и Учитель, он по-прежнему востребован российской научно-образовательной общественностью, студенческой молодежью, коллегами по работе в самых различных областях его деятельности.

Особую роль сыграл Юрий Борисович в создании в 1999 году журнала «Цифровая обработка сигналов». Сегодня – это один из ведущих российских научно-технических журналов в области современных информационных технологий. Журнал «Цифровая обработка сигналов» включен в Перечень ведущих рецензируемых научных изданий, в которых по рекомендации Минобрнауки РФ, должны быть опубликованы основные научные результаты диссертаций на соискание ученой степени доктора и кандидата наук.

Большие заслуги Юрия Борисовича связаны с организацией и непосредственным личным участием в работах Международной конференции и выставки «Цифровая

обработка сигналов и ее применение-DSPA», которая уже в 15-й раз будет проходить в марте текущего года в Институте проблем управления РАН.

Следует отметить активную поддержку и непосредственное участие кавалера многих правительственных наград и лауреата государственных премий Юрия Борисовича Зубарева во внедрении в современные локальные и глобальные информационные системы технологий цифровой обработки сигналов, возникшей на стыке электросвязи и компьютерной техники.

Важной областью деятельности профессора Зубарева Ю.Б. является подготовка молодых специалистов и ученых, авторство в более 350 научных трудах в области телевидения, связи и информатики, участие в работе диссертационных и научно-технических Советов, в Экспертном совете ВАК РФ, в качестве председателя и заместителя председателя в настоящее время.

Отмечая огромные заслуги Зубарева Ю.Б. перед мировым научным сообществом, его активную практическую деятельность в объединении научного потенциала российских ученых на пути формирования отечественной школы специалистов в области цифровой обработки сигналов, коллектив нашего журнала поздравляет Юрия Борисовича с 75-летним юбилеем! Крепкого Вам здоровья и долгих лет эффективной работы в области информатики реального времени!

*Заместители Главного редактора*

*профессор В.В. Витязев  
профессор В.П. Дворкович*





# ЦИФРОВАЯ ОБРАБОТКА СИГНАЛОВ

Научно-технический журнал  
№ 1/2013

Издается с 1999 года  
Выходит четыре раза в год

**ГЛАВНЫЙ РЕДАКТОР**  
Ю.Б. ЗУБАРЕВ

**ЗАМЕСТИТЕЛИ ГЛАВНОГО РЕДАКТОРА:**  
В.В. ВИТЯЗЕВ, В.П. ДВОРКОВИЧ

## РЕДАКЦИОННАЯ КОЛЛЕГИЯ:

А.С. Аджемов, Б.А. Алпатов, Б.А. Бабаян,  
Ю.И. Борисов, Ю.А. Брюханов, В.В. Еремеев,  
Г.В. Зайцев, Р.В. Зубарев, А.П. Кирпичников,  
М.И. Кривошеев, Н.А. Кузнецов, В.Г. Мистюков,  
С.Л. Мишенков, А.А. Петровский, Ю.Н. Прохоров,  
В.Г. Санников, В.С. Сперанский, Ю.С. Шинаков

## Адрес редакции:

107031 Москва, Рождественка, 6/9/20, стр. 1  
Тел.: (495) 621-06-10, (495) 621-16-39  
Факс: (495) 621-16-39  
E-mail: rntores@mail.ru  
vityazev.v.v@rsreu.ru  
http://www.dsps.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 13.03.13 г.  
Формат 60x90/8.

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

Заказ № 1357. 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

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

## УЧРЕДИТЕЛИ:

**ЗАО «Инструментальные системы»  
ФГУП «НИИ радио»**

## В НОМЕРЕ:

### Леднов Д.А.

Анализ скрытых траекторных моделей резонаторов  
речевого тракта для систем распознавания фонов ..... 2

### Соловьева Е.Б.

Каскадный предкомпенсатор для линеаризации  
характеристики усилителя мощности ..... 9

### Батенков К.А.

Математические модели модулятора и демодулятора  
с заданным порядком нелинейности ..... 14

### Пиманкин Д.А.

Метод наименьших взвешенных квадратов  
для анализа изображений ..... 22

### Калинин П.В., Сирота А.А.

Моделирование аппликативных искажений с различной  
степенью прозрачности и случайной формой ..... 28

### Минаков Е.И., Калистратов Д.С.

Метод формирования бинарной кадровой маски движения ..... 34

### Самойленко М.В.

Определение местоположений и мощностей источников  
излучения однопозиционной локационной станцией ..... 38

### Паршин А.Ю., Паршин Ю.Н.

Использование максимально правдоподобных оценок  
фрактальной размерности в негауссовских статистиках  
для обнаружения радиосигналов ..... 42

### Мандрикова О.В., Глушкова Н.В., Полозов Ю.А.

Алгоритмы выделения и анализа аномалий в параметрах  
критической частоты ионосферы  $f_oF2$  на основе совмещения  
вейвлет-преобразования и авторегрессионных моделей ..... 47

### Волченков В.А., Витязев В.В.

Методы и алгоритмы детектирования активности речи ..... 54

### Гудкова Н.В.

Приложение принципов цифровой адаптивной фильтрации  
к задаче управления динамическим объектом  
с неизвестной математической моделью ..... 61

### Григоренко Д.В., Ручкин В.Н.

Повышение восстанавливаемости кластерных  
нейропроцессорных систем обработки данных ..... 67

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



## АНАЛИЗ СКРЫТЫХ ТРАЕКТОРНЫХ МОДЕЛЕЙ РЕЗОНАТОРОВ РЕЧЕВОГО ТРАКТА ДЛЯ СИСТЕМ РАСПОЗНАВАНИЯ ФОНЕМ

*Леднов Д.А., к. т. н., старший научный сотрудник, научный консультант научно-технического департамента ООО «Стэл – Компьютерные Системы», г. Москва, e-mail: lednov@stel.ru*

**Ключевые слова:** распознавание фонем, речевой тракт, траекторная модель, резонатор, КИХ-фильтр, линейное предсказание, оптимизация.

### Введение

Согласно классической схеме систем распознавания слитной речи [1], система распознавания фонем является частью системы распознавания речи и, в основном, именно точность распознавания фонем определяет точность системы распознавания в целом.

В основе большинства современных систем распознавания фонем лежит скрытая модель Маркова (СММ), которая описывает вероятности следования состояний, описывающих фонемы, друг за другом. Для описания состояний фонем используются смеси нормальных плотностей распределения (в литературе их чаще называют гауссовыми смесями) векторов наблюдений.

В соответствии с работой [2], средняя точность системы распознавания фонем (по всему множеству фонем английского языка), которая достигнута на современном уровне технологий в условиях не зашумленного сигнала, при частоте оцифровки 8кГц и при использовании СММ, составляет 69.1%. Эта величина не является высокой, поэтому актуальной является задача поиска новых математических моделей для систем распознавания фонем.

В период с 2000 по 2010 год появилась серия публикаций *Li Deng* и его коллег (компания Microsoft) [4-8], в которых описывается новый подход к задаче распознавания фонем, основанный на фундаментальном свойстве артикуляционных органов – их инерции.

Инерция артикуляционных органов отражается в речи двумя свойствами:

- гладкостью изменения ее параметров во времени в процессе произнесения вокализованных звуков;
- коартикуляцией [3].

Гладкость изменений параметров речи проявляется в том, что амплитуды и частоты гармоник линейчатого спектра, порожденного вокализованным звуком, не изменяются скачкообразно. Коартикуляция заключается во влиянии целевых артикуляций соседних звуков друг на друга.

Авторы [4-8] назвали построенный способ распознавания фонем на основе инерции артикуляционных органов «скрытыми траекторными моделями» (СТМ) резонаторов речевого тракта (РРТ). Здесь термин «скрытость»,

*Приводится перевод и анализ оригинальных работ Li Deng и его коллег (компания Microsoft), появившихся в период с 2000 по 2010 год в области фонетического распознавания речи. Основное направление этих публикаций связано с разработкой модели скрытых траекторий параметров резонаторов вокального тракта. В ходе разработки было показано, как функционально зависит динамика коэффициентов линейного предсказания от параметров резонаторов вокального тракта, которые предварительно сглаживаются КИХ-фильтром. Затем, для этой зависимости введена статистическая модель, для которой поставлена и решена оптимизационная задача. Автор настоящей работы дополнил развитую модель уравнением непрерывности, которое позволяет определить характеристики КИХ-фильтров для каждого фонетического состояния и ввел альтернативную оптимизационную схему, позволяющую определять параметры статистической модели.*

как и в модели Маркова, заключается в том, что при регистрации речи мы можем измерить только физические параметры речи (вектора наблюдений), но информационная ее составляющая (последовательность фонем), которую хочет передать говорящий, является для нас скрытой и определяется с некоторой вероятностью в результате анализа.

Работа над пониманием названных выше публикаций выявила, с одной стороны, допущения, сделанные авторами, которые искажают первоначальную идею модели, в результате чего траектории параметров РРТ становятся разрывными. С другой стороны, авторы не указали способ определения начального приближения плотности вероятности параметров модели.

Настоящая работа ставит перед собой цель – изложение основных теоретических положений скрытых траекторных моделей резонаторов речевого тракта и устранение в ней указанных ранее недостатков.

### Генерация скрытых траекторий параметров резонаторов речевого тракта

Схематически идею СТМ параметров РРТ можно представить, как показано на рис. 1. Это схема состоит из двух частей: модели генерации речи; модели восприятия и обработки речи. В стадию обработки речи включены два процесса – обучение (оценка параметров модели) и распознавание (принятие решение о звучащей фонеме или их последовательности).

В соответствии с рис. 1, блок генерации символов преобразует целевые семантические категории в последовательность слов, затем последовательность слов - в последовательность фонем, которые поступают на вход блока управления речевым трактом (РТ).



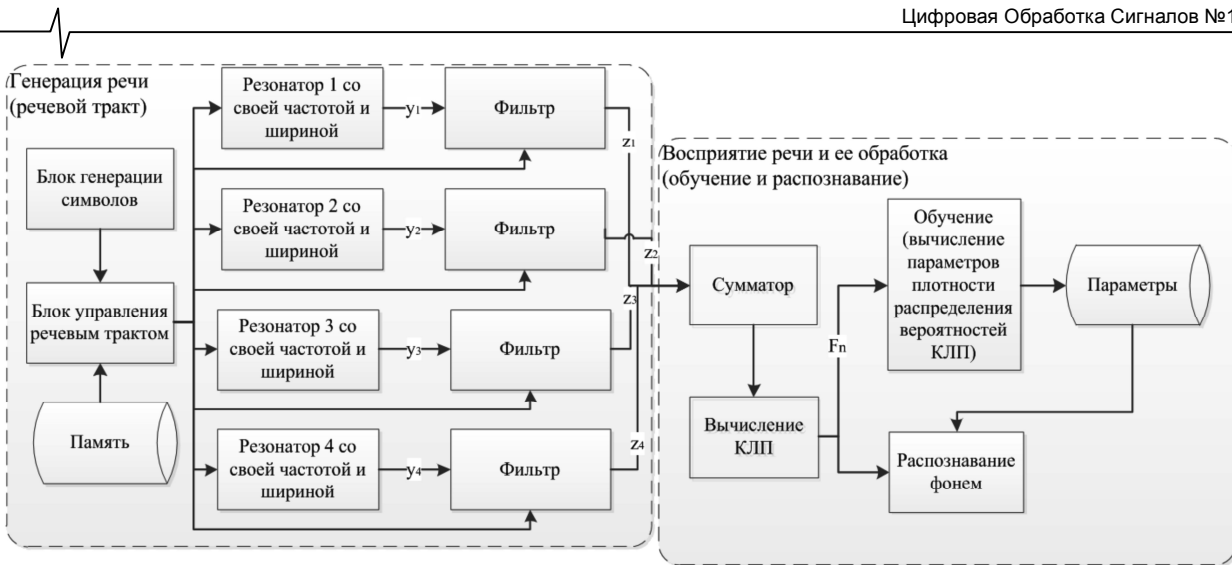


Рис. 1. Схематическое представление идеи STM для PPT

В блоке управления РТ каждой поступающей фонеме ставится в соответствие набор целевых параметров РТ, определяющих его форму (целевые параметры РТ извлекаются из памяти). Блок управления РТ вызывает динамику речевого тракта, целью которой является форма РТ, определяющая звучание фонемы. Заметим, что РТ может и не достигнуть заданной формы, это зависит от темпа смены фонем в их последовательности, поступающей на вход управляющего блока.

В STM речевой тракт представлен четырьмя резонаторами, которые характеризуются частотой и шириной резонанса, а также сглаживающими КИХ-фильтрами одинаковыми для всех резонаторов. При смене одной фонемы другой блок управления изменяет целевые параметры резонаторов и импульсную характеристику сглаживающих фильтров. Источником возбуждения резонаторов является широкополосный белый шум.

С одной стороны, можно считать недостатком то, что модель пренебрегает фактом существования голосового источника, порождающего гармонические линейчатые спектры, но, с другой стороны, это предположение позволяет ввести универсальное описание акустических свойств фонем, независимое от их типа (вокализованные, невокализованные) на основе плотности распределения энергии в спектре акустического сигнала.

Модель восприятия и обработки речи предполагает, что параметризация речевого сигнала происходит с помощью известной процедуры вычисления коэффициентов линейного предсказания (КЛП). Ниже будет показана явная зависимость КЛП от параметров резонаторов и, как следствие, зависимость вектора наблюдения от параметров резонаторов в условиях сглаживания КИХ-фильтрами.

И далее, в соответствии с классической схемой систем распознавания [1], в процессе обучения происходит вычисление параметров (математического ожидания и ковариационной матрицы) плотностей распределения вероятности векторов наблюдений (плотности вероятности предполагаются нормальными), после чего в процессе распознавания происходит вычисление наиболее вероятной цепочки фонем на основе входной последовательности векторов наблюдений.

Формализуем описанный процесс генерации речи. С этой целью составим фонетически зависимый целевой

вектор  $T_s$  (target) из параметров резонаторов, где  $s$ -индекс фонемы.

Условное распределение динамики текущих параметров  $y$  резонаторов, при заданном целевом векторе, определенном звучащей фонемой  $s$ , предположим нормальным:

$$p(y|s) = G(y; \mu_{T_s}, \Sigma_{T_s}),$$

где  $\mu_{T_s}, \Sigma_{T_s}$  – математическое ожидание и ковариационная матрица параметров резонаторов.

Скрытый случайный процесс  $z(t)$  образуется путем фильтрации целевого процесса  $y_s(t)$ . Эта фильтрация выполняется с помощью фильтра с конечной импульсной характеристикой вида:

$$h_s(t) = \begin{cases} g_s \gamma_{s(t)}^{-t}, & -D < t < 0; \\ g_s, & t = 0; \\ g_s \gamma_{s(t)}^t, & 0 < t < D, \end{cases}$$

где  $g_s$  – нормализующая константа,  $\gamma_{s(t)}^t$  – параметр шкалы (лежит в интервале  $[0, 1]$ ).

Скрытый случайный фонетически зависимый процесс  $z_s(t)$  можно представить в виде свертки целевого процесса и импульсной характеристики:

$$z_s(t) = h_s(t) * y_s(t) = \sum_{t=k-D}^{t=k+D} g_s \gamma_{s(t)}^{|t-k|} y_s(t),$$

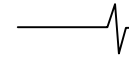
где  $y_s(t)$  текущее значение параметра резонатора при заданной фонеме  $s$ .

Линейная зависимость между скрытым процессом  $z_s(t)$  и целевым процессом  $y_s(t)$  приводят к линейным зависимостям между математическим ожиданиями целевого вектора и скрытого процесса:

$$\mu_{z_s}(t) = \sum_{t=k-D}^{t=k+D} g_s \gamma_{s(t)}^{|t-k|} \mu_{T_s}(t) = a_k \mu_{T_s}.$$

Каждая  $f$ -ая компонента вектора  $\mu_{z_s}(t)$  может быть представлена в виде:





$$\mu_{z(t)}(f) = \sum_{l=1}^{\Phi} a_l(l) \mu_T(l, f), \quad (1)$$

где  $\Phi$  – количество фонетических состояний и  $f = 1, \dots, 8$  для четырех частот и соответствующих им полос пропускания.

Сходным образом для ковариационной матрицы справедливо

$$\Sigma_{z(k)} = \sum_{l=k-D}^{k+D} g_s^2 \gamma_{s(t)}^{2|l-k|} \Sigma_{s(t)}.$$

Аппроксимируем ковариационную матрицу диагональной матрицей

$$\sigma_{z(k)}^2 = \mathcal{G}_k \sigma_T^2$$

или в покомпонентной форме

$$\sigma_{z(t)}^2(f) = \sum_{l=1}^{\Phi} \mathcal{G}_l(l) \sigma_T^2(l, f). \quad (2)$$

В работах [4-6] нет прямых указаний о способе вычисления коэффициентов  $a_k$  и  $\mathcal{G}_k$ . Авторы пишут, что нет краткой математической формы, в которой можно записать выражение для них.

Рассмотрим вопрос, - а можно ли в действительности получить такую математическую форму? Пусть  $k$  это момент времени, в который помещена середина «окна» импульсной характеристики наблюдения размером  $2D+1$ . Необходимо вычислить влияние сегмента речи, заключенного в интервале времени  $d$  (см. рис. 2), на систему в текущий момент времени  $k$ . Можно выделить три различных случая положения границ  $d$  звучания этого сегмента речи ( $L$ -левая граница,  $R$ -правая граница) относительно середины окна наблюдения  $k$ , которые показаны на рис. 2.

Для случая, показанного на рис. 2 а, где  $k-D \leq L < k$  и  $k \leq R \leq k+D$ , справедливо

$$a_k = g \left( \sum_{l=L}^k \gamma^{|l-k|} + \sum_{l=k}^R \gamma^{|l-k|} \right) = \frac{g}{1-\gamma} (2 - \gamma^{k-L+1} - \gamma^{R-k+1}). \quad (3.1)$$

Для случая, показанного на рис. 2 б, где  $k-D \leq L < k$  и  $L \leq R \leq k$ , справедливо

$$a_k = g \gamma^{R-k} \frac{1 - \gamma^{R-L+1}}{1 - \gamma}. \quad (3.2)$$

Для случая, показанного на рис. 2 в, где  $k \leq L < k$  и  $k \leq R \leq k+D$ , справедливо

$$a_k = g \gamma^{k-L} \frac{1 - \gamma^{R-L+1}}{1 - \gamma}. \quad (3.3)$$

Таким образом, для каждого типа положения границ фонемы относительно центра «окна» импульсной характеристики можно найти простую математическую форму записи коэффициентов в (1) и аналогично в (2), причем

$$\sum_{l=k-D}^{k+D} a_l(l) = 1 \quad \forall l. \quad (4)$$

Откуда можно вычислить значение коэффициента  $g = (1-\gamma)/2(1-\gamma^{(D+1)})$ .

Для определения параметра  $\gamma$  можно использовать условие непрерывности, физический смысл которого состоит в том, что изменения значения траектории должны быть малы при малом влиянии фонемы. Пусть в момент времени  $k-D+1$  одна фонема меняет другую и на протяжении времени от  $k-D+1$  до  $k+D$  звучит некоторая определенная последовательность фонем. Затем «окно» импульсной характеристики смещается на шаг вправо, при этом крайняя левая фонема исчезает из поля «окна», а остальная последовательность фонем сохраняется. В этом случае можно записать условие непрерывности

$$g_1 \gamma_1^{1-D} (1 + \gamma_1) \mu_{1T} + \sum_{i=2}^N g_i^{(k)} \mu_{iT} \sum_{l=L_i}^{t=R_i} \gamma_i^{|l-k|} \approx \sum_{i=2}^n g_i^{(k+1)} \mu_{iT} \sum_{l=L_i}^{t=R_i} \gamma_i^{|l-k-1|}. \quad (5)$$

где  $N+1$  – количество фонем, которые укладываются в окне длительностью  $2D+1$ .

Обратим внимание на то, что в работе [4] для всех типов акустических данных используются односторонние КИХ-фильтры, а в работе [5] предполагается, что тип КИХ-фильтра не зависит от фонемы, т.е.  $\gamma = \gamma_s$  для всех  $s$ . В дальнейшем мы используем (5) для определения параметров модели фильтров, метод получения которых не был описан в анализируемых работах.

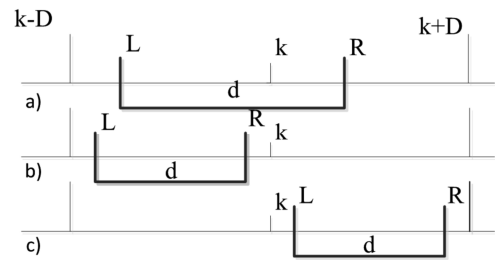


Рис.2. Различные положения границ звучания фонемы относительно центра границ импульсной характеристики

### Генерация акустических данных

В работе [6] получена аналитическая форма коэффициентов линейного предсказания (КЛП) для случая резонансной модели речеобразования.

Рассмотрим комплексные корни передаточной функции речевого тракта:

$$z_m = \exp \left\{ -\pi \frac{b_m}{f_s} + j 2\pi \frac{f_m}{f_s} \right\}, \quad (6)$$

где  $f_s$  – частота оцифровки речевого сигнала,  $b_p(t)$  – ширина резонанса,  $f_p(t)$  – частота резонанса.



Передаточную функцию речевого тракта с  $K$  полюсами можно записать в виде

$$H(t) = G \prod_{k=1}^K \frac{1}{(1 - z_k z^{-1})(1 - z_k^* z^{-1})},$$

или в логарифмической форме

$$\begin{aligned} \log H(z) &= \\ &= \log G - \sum_{k=1}^K \log(1 - z_k z^{-1}) - \sum_{k=1}^K \log(1 - z_k^* z^{-1}). \end{aligned}$$

Если использовать известное разложение логарифма

$$\log(1 - x) = - \sum_{i=1}^{\infty} \frac{x^i}{i},$$

то последнее выражение можно представить в форме

$$\begin{aligned} \log H(z) &= \log G - \sum_{k=1}^K \sum_{i=1}^{\infty} \frac{z_k^i z^{-i}}{i} - \sum_{k=1}^K \sum_{i=1}^{\infty} \frac{z_k^{*i} z^{-i}}{i} = \\ &= \log G + \sum_{i=1}^{\infty} \left[ \sum_{k=1}^K \frac{z_k^i + z_k^{*i}}{i} \right] z^{-i} = c_0 + \sum_{i=1}^{\infty} c_i z^{-i}. \end{aligned}$$

где

$$c_0 = \log G,$$

$$c_i = \sum_{k=1}^K \frac{z_k^i + z_k^{*i}}{i}.$$

Если подставить (6) в формулу для коэффициентов  $c_i$ , то несложно получить

$$c_n = \frac{2}{n} \sum_{k=1}^K \exp \left\{ -\pi n \frac{b_k}{f_s} \right\} \cos \left( 2\pi n \frac{f_k}{f_s} \right).$$

Из последней формулы непосредственно следует, что для  $n$ -компонентного значения вектора наблюдений, порожденного скрытой векторной функцией  $z(t)$ , справедливо

$$\begin{aligned} F_n(z(t) = \{f_p, b_p\}_p) &= \\ &= \frac{2}{n} \sum_{p=1}^P \exp \left\{ -\frac{\pi n b_p(t)}{f_s} \right\} \cos \left( \frac{2\pi n f_p(t)}{f_s} \right), \end{aligned}$$

$P$  – количество резонансов ( $P = 4$ ),  $f_s$  – частота оцифровки.

Остаток от аппроксимации наблюдения можно записать в виде

$$r_s(t) = o(t) - F(\{z_s(t) = \{f_p, b_p\}_p\}).$$

Предположим, что плотность распределения величины остаточного вектора - это случайная величина, заданная нормальным распределением

$$p(r_s(t) | z(t), s) = G(r_s(t); \mu_r, \Sigma_r), \quad (7)$$

$s$  – индекс фонемы.

Из последнего предположения следует, что условное распределение наблюдения можно записать в виде:

$$p(o(t) | z(t), s) = G(o(t); F(z_s(t)) + \mu_r, \Sigma_r), \quad (8)$$

и представить вектор наблюдения в форме:

$$o(t) = F(z_s(t)) + \mu_r + w_s(t),$$

где  $w_s(t)$  – случайная величина с нормальным распределением  $G(w_s(t), 0, \Sigma_r)$ .

### Линеаризация коэффициентов линейного предсказания

Нелинейная функция  $F(z_s(t))$  сложна для дальнейших вычислений, поэтому проведем ее линеаризацию. С помощью разложения в ряд Тейлора с точностью до первого порядка представим функцию  $F(z_s(t))$  в виде:

$$F(z(t)) = F(z_0(t)) + F'(z_0(t))(z(t) - z_0(t)),$$

где компоненты матрицы Якоби могут быть представлены в форме:

$$F'_n(f_p(t)) = \frac{-4\pi}{f_0} \exp \left\{ -\frac{\pi n b_p(t)}{f_0} \right\} \sin \left( \frac{2\pi n f_p(t)}{f_0} \right),$$

$$F'_n(b_p(t)) = \frac{-2\pi}{f_0} \exp \left\{ -\frac{\pi n b_p(t)}{f_0} \right\} \cos \left( \frac{2\pi n f_p(t)}{f_0} \right).$$

Подстановка разложения в ряд Тейлора приведет к новой форме для условной вероятности (8)

$$\begin{aligned} p(o(t) | z(t), s) &= G(o(t); F(z_0(t)) + \\ &+ F'(z_0(t))(z(t) - z_0(t)) + \mu_r, \Sigma_r), \end{aligned}$$

где можно ввести обозначение математического ожидания вектора наблюдения

$$\mu_0(t) = F(z_0(t)) + F'(z_0(t))(z(t) - z_0(t)) + \mu_r(t),$$

и в сокращенном виде записать

$$p(o(t) | z(t), s) = G(o(t); \mu_0(t), \Sigma_r).$$

### Вычисление вероятности параметров РРТ

Рассмотрим условную вероятность возникновения наблюдения при звучании известной фонемы

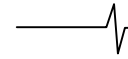
$$\begin{aligned} p(o(t) | s) &= \int p(o(t) | z(t), s) p(z(t) | s) dz \sim \\ &\sim \int G(o(t); \mu_{o(t)}, \Sigma_{r(t)}, s) G(z(t); \mu_{z(t)}, \Sigma_{z(t)}, s) dz = \\ &= G(o(t); \mu_{o(t)}, \Sigma_{o(t)}, s), \end{aligned} \quad (9)$$

где введено новое обозначение

$$\Sigma_{o(t),s}(t) = \Sigma_{r(t),s}(t) + F'(z_0(t)) \Sigma_{z(t)} [F'(z_0(t))]^T.$$

Раскроем введенные обозначения математического ожидания и ковариационной матрицы наблюдения, используя выражения (1), (2), и для сокращения записи выпишем отдельно числитель  $h$  и знаменатель  $v$  экспоненты для  $j$ -ой компоненты вектора наблюдения:

$$\begin{aligned} h &= (o_j - F_j(z_0(t)) - \mu_r(j) - \\ &- F'_j(z_0(t)) \left( \sum_{l=1}^{\Phi} a_l(l) \mu_T(l, f) - z_0(t) \right))^2, \end{aligned}$$



$$v = \sigma_{r(t,s)}^2(j) + \sum_{k=1}^P F'_{jk}(z_o(t)) F'_{kj}{}^T(z_o(t)) \sum_{l=1}^{\Phi} g_l(l) \sigma_T^2(l, k) \quad (11)$$

Из приведенных формул для числителя и знаменателя видно, что неизвестными параметрами модели являются математические ожидания  $\mu_T(l, f)$  и дисперсии  $\sigma_T^2(l, k)$  целевых векторов фонем, математические ожидания  $\mu_{r(t)}(j)$  и дисперсии  $\sigma_{r(t,s)}^2(j)$  остаточного вектора и коэффициенты свертки математических ожиданий  $a_l(l)$  и коэффициенты свертки дисперсий  $g_l(l)$  целевых векторов. Проведем оценку перечисленных параметров.

### Оценка параметров для СТМ

Оценка параметров модели происходит на размеченной речевой базе данных, принципы разметки данных описаны в работе [7].

В оригинальных статьях [4, 8] параметры распределения (9) были разделены на две группы: параметры распределения кепстрального остатка и значения компонент целевых векторов. Метод оценки величин  $\gamma_s$  (или даже величины  $\gamma = \gamma_s \forall_s$ , в соответствии с [5]) авторы не приводят. Здесь сначала приводится оценка для приведенных групп параметров, а затем эта схема будет дополнена.

### Оценка параметров распределения кепстрального остатка

Для определения параметров распределения кепстрального остатка найдем максимум функции логарифмического правдоподобия. Для математического ожидания справедливо уравнение:

$$\frac{\partial \log \prod_{t=1}^{K_s} p(o(t) | s)}{\partial \mu_{r(s)}} = 0,$$

где  $K_s$  - количество дискретных отсчетов времени, в которых наблюдалась фонема  $s$ .

Выполняя дифференцирование последнего уравнения для математического ожидания остатка, получим:

$$\mu_{r(s)} = \frac{\sum_{t=1}^K [o(t) - F'(z_0(t)) \mu_{z(t)} - F(z_0(t)) + F'(z_0(t)) z_0(t)]}{K_s} \quad (12)$$

Аналогично, для диагональной ковариационной матрицы справедливо уравнение

$$\frac{\partial \log \prod_{t=1}^{K_s} p(o(t) | s)}{\partial \sigma_{r(s)}^2} = 0,$$

которое приводит к выражению

$$\sum_{t=1}^{K_s} \frac{\sigma_{r(s)}^2 + q(t) - (o(t) - \mu_{0(s)})^2}{(\sigma_{r(s)}^2 + q(t))^2} = 0, \quad (13)$$

где введено обозначение

$$q(t) = \text{diag} \{ F'(z_0(t)) \Sigma_{z(t)} [F'(z_0(t))]^T \}. \quad (14)$$

Уравнение (13) можно разрешить тремя способами:

- 1) предположить, что  $q(t)$  не зависит от времени;
- 2) методом градиентного спуска;
- 3) методом ограниченного градиентного спуска.

### Оценка параметров основного распределения

Подмножество параметров СТМ состоит из: 1) средних векторов  $\mu_{0(s,t)}(j)$ ; 2) диагональных элементов ковариационной матрицы  $\sigma_{0(t,s)}^2(j)$ .

#### 1) Средние вектора

Ранее была предположена диагональная форма ковариационной матрицы.

Запишем многомерное распределение

$$p(o(t) | s(t)) = \prod_{j=1}^J \frac{1}{\sqrt{2\pi\sigma_{o(t,s)}^2(j)}} \exp \left\{ -\frac{(o_t(j) - \mu_{o(s,t)}(j))^2}{\sigma_{o(t,s)}^2(j)} \right\},$$

$\sigma_{0(t,s)}^2(j)$  -  $j$ -ая компонента наблюдения во фрейме в момент времени  $t$ , полученный от фонемы  $s$ .

Функция логарифмического правдоподобия относительно среднего вектора  $\mu_{0(s,t)}(j)$  может быть записана в виде

$$P = \sum_{t=1}^{K(s)} \sum_{j=1}^J \left\{ \frac{(o_t(j) - \mu_{o(s,t)}(j))^2}{\sigma_{o(t,s)}^2(j)} \right\} = \sum_{t=1}^{K(s)} \sum_{j=1}^J \left\{ \frac{\sum_f F'(z_0(t), j, f) \sum_l a_l(l) \mu_T(l, f) - d_t(j)}{\sigma_{o(t,s)}^2(j)} \right\}, \quad (15)$$

где  $l$  и  $f$  - индексы фонем и PPT компонент соответственно, и

$$d_t(j) = o_t(j) - F(z_0(t), j) + \sum_f F'(z_0(t), j, f) z_0(t, f) - \mu_{r(s,t)}(j).$$

Если выполнить дифференцирование (15)

$$\frac{\partial P}{\partial \mu_T(l_0, f_0)} = 0$$

и перегруппировать полученное выражение так, чтобы слагаемые, содержащие  $\mu_T(l, f)$  были слева, а все прочие элементы справа, то получим уравнение

$$\sum_f \sum_l A(l, f, l_0, f_0) \mu_T(l, f) = \sum_t \left\{ \sum_j \frac{F'(z_0(t), j, f_0) d_t(j)}{\sigma_{o(t,s)}^2(j)} \right\} a_l(l_0), \quad (16)$$



где  $f_0 = 1, 2, \dots, 8$  для каждой размерности  $l_0 = 1, 2, \dots, 58$  для каждого фонетического элемента

$$A(l, f, l_0, f_0) = \sum_{t,j} \frac{F'(z_0(t), j, f_0) F'(z_0(t), j, f_0)}{\sigma_{o(t,s)}^2(j)} a_t(l) a_t(l_0).$$

**2) Матрица ковариаций**

Для матрицы ковариации введем функцию правдоподобия

$$L \sim - \sum_{t=1}^K \sum_{j=1}^J \left\{ \frac{(o_t(j) - \mu_{o(s,t)}(j))^2}{\sigma_{o(t,s)}^2(j) + q(t, j)} - \log(\sigma_{z(t,s)}^2(j) + q(t, j)) \right\}$$

где  $q(t, j)$  – элемент вектора, определенный в (14).

Используя метод градиентного спуска, получим

$$\frac{\partial L}{\partial \sigma_T^2(l, f)} = \sum_{t=1}^K \sum_{j=1}^J \left\{ \frac{(o_t(j) - \mu_{o(s,t)}(j))^2 F_{ff}'' \mathcal{G}_t(l)}{(\sigma_{r(s)}^2(j) + q(t, j))^2} - \frac{F_{ff}'' \mathcal{G}_t(l)}{\sigma_{r(s)}^2(j) + q(t, j)} \right\}$$

$$\sigma_T^2(l, f) \leftarrow \sigma_T^2(l, f) + \frac{\partial L}{\partial \sigma_T^2(l, f)}. \quad (17)$$

Резюмируем приведенные вычисления. Схема поиска параметров, предложенная в работах [4-6], состоит в рекуррентном решении уравнений (12), (13) и (16), (17) с помощью классических методов градиентного спуска. К сожалению, в работах не обсуждается проблема выбора начального приближения для этих рекуррентных вычислений.

**Альтернативная схема оценки параметров**

Прежде всего, решим вопрос относительно начального приближения. Для вычисления начальных приближений средних частот и ширины полосы резонансов выберем всевозможные спектры, принадлежащие определенной фонеме  $s$  из обучающей речевой базы. Частотный диапазон спектров разобьем на перекрывающиеся участки с границами, указанными в работе [6] (табл. 1).

Таблица 1. Список спектральных границ для вычисления начального приближения

Номер резонанса	(rb) Правая граница (Гц)	(lb) Левая граница(Гц)
1.	200	900
2.	600	2800
3.	1400	3800
4.	1700	5000

**Шаг 1.** В качестве частоты  $i$ -го резонанса в начальном приближении используем среднее значение частоты в  $i$ -ом спектральном диапазоне

$$f_i^{(0)} = \frac{1}{K_s} \sum_{t=1}^{K_s} \frac{\int_{lb_i}^{rb_i} S_t(\omega) \omega d\omega}{\int_{lb_i}^{rb_i} S_t(\omega) d\omega},$$

а в качестве ширины полосы частот  $i$ -го резонанса среднеквадратическое отклонение

$$b_i^{(0)} = \sqrt{\frac{1}{K_s} \sum_{t=1}^{K_s} \frac{\int_{lb_i}^{rb_i} S_t(\omega) (\omega - f_i^{(0)})^2 d\omega}{\int_{lb_i}^{rb_i} S_t(\omega) d\omega}},$$

где  $S_t(\omega)$  спектр фонемы, наблюдаемый в момент времени  $t$ . Таким образом, построены начальные приближения целевых векторов фонем  $\mu_s^{(0)} = \left( \{f_i^{(0)}\}_s, \{b_i^{(0)}\}_s \right)^T$ .

**Шаг 2.** Рассмотрим последовательности фонем, которые удовлетворяют условиям построения уравнений (5). Из всевозможных последовательностей будем выбирать такие подпоследовательности, на основе которых можно строить разрешимые нелинейные системы уравнений (5). Каждая из таких систем будет определять множество решений  $\{\gamma\}$  для фонем, входящих в подпоследовательность. Очевидно, что одна и та же фонема может входить в различные подпоследовательности и для нее может быть получено множество решений. Для полученных решений построим гистограмму их встречаемости и выберем наиболее часто встречающиеся решения. Так поступим с множествами решений для каждой фонемы.

**Шаг 3.** Предположим, что

$$z_0(t) = \sum_{l=1}^{\Phi} a_t^{(0)}(l) \mu_T^{(0)}(l, f).$$

Подстановка этого равенства в (10) и (11) и дифференцирование правдоподобия по параметру математического ожидания кепстрального остатка и его дисперсии дает нулевое приближение этих параметров:

$$\mu_{r(t)}^{(0)}(j) = \frac{1}{K_s} \sum_{t=1}^{K_s} [o_t(j) - F_j(z_0(t))];$$

$$\left( \sigma_{r(t,s)}^{(0)}(j) \right)^2 = \left( o_t(j) - F_j(z_0(t)) - \mu_{r(t)}^{(0)}(j) \right)^2 - \sum_{k=1}^P F_{jk}'(z_0(t)) F_{jk}'^T(z_0(t)) \sum_{l=1}^{\Phi} \mathcal{G}_t^{(0)}(l) \left( \sigma_T^{(0)}(l, k) \right)^2.$$

**Шаг 4.** Расчет первого приближения параметров резонансов вокального тракта происходит на основании уравнений (16) и (17).

**Шаг 5.** На основании нового приближения параметров резонансов вокального тракта возвращаемся к шагу 2 и решаем системы уравнений (5) с новыми параметрами.

Необходимо отметить, что в последнем слагаемом (10) в выражении в скобках, свертка  $\sum_{l=1}^{\Phi} a_t(l) \mu_T(l, f)$

от  $z_0(t)$  отличается лишь тем, что она вычисляется на основании параметров, вычисленных на одну итерацию выше.

**Экспериментальные результаты**

Экспериментальные результаты были получены с помощью корпуса речи TIMIT (частота оцифровки 16кГц). Для обучения были выделены 538 предложений [9]. Эти предложения были фонетически размечены и найдены формантные траектории, на основе которых проведена оценка параметров модели (12), (14), (16) и (17). На остальной части корпуса речи TIMIT было проведено сравнение точностей распознавания фонем, полученных с помощью скрытой модели Маркова и скрытой траекторной модели. Скрытая модель Маркова была реализована в следующей форме: в качестве вектора признаков использовались кепстральные коэффициенты линейного



предсказания; модель фонем представлена трифонами, включающими в себя три состояния; использовалась биграммная модель языка. Точности, показанные этими системами, приведены в табл. 2.

Таблица 2. Сравнение точностей распознавания фонем скрытой модели Маркова и скрытой траекторной модели

Тип фонем	Сонорные (гласные, звонкие согласные, нозальные)	Взрывные	Фрикативные	Смычные
Количество фонем	3814	889	1252	1578
СММ	64.05	72.10	75.74	88.72
СТМ	72.42	76.27	75.74	90.94

Из табл. 2 видно, что по сравнению со скрытой моделью Маркова модель СТМ дает существенное преимущество в точности.

### Заключение

Как известно, речь обладает существенной вариативностью, и поведение резонаторов речевого тракта при реализации той или иной фонемы может изменяться от одного говорящего к другому [11]. Если предположить, что такие различные типы поведения речевого тракта можно объединять в группы и каждая группа обладает сходными характеристиками и динамикой речевого тракта, то можно заключить, что распределение плотности вероятности (7) справедливо только в рамках одной группы. Тогда рассуждения, которые приводят к выражению (9), остаются справедливыми, если допустить, что нам априори известна принадлежность голоса говорящего к той или иной группе. Если же такой информации нет, то (9) можно записать в форме:

$$p(o(t) | s) = \sum_{i=1}^{m(s)} p_i(s) G(o(t); \mu_{o(t),i}; \sum_{o(t),i} s), \quad (17)$$

где  $m$  – количество групп голосов для каждой фонемы,  $p_i(s)$  вероятность появления  $i$ -ой группы фонемы  $s$ . С точки зрения автора, введенная гауссова смесь (17) способна в полном объеме придать системе распознавания фонем свойство дикторонезависимости и описать вариативность речи говорящих.

Таким образом, в работе описана модель скрытых траекторий параметров резонаторов речевого тракта, введены поправки к исходной модели, а также предложен путь развития модели на основе описания фонем – как гауссовых смесей скрытых траекторий.

Поиск информации о продолжении работы в направлении развития СТМ для РРТ после 2010 года не принес результатов.

### Литература

1. Jelinek F. Continuous speech recognition by statistical methods.// In: Proceedings of the IEEE, 1976, v. 64, № 4, p. 532–556.
2. Thomas, S.; Ganapathy, S. & Hermansky, H. (2008), Spectro-Temporal Features for Automatic Speech Recognition using Linear Prediction in Spectral Domain//in 'EUSIPCO'
3. Кодзасов С.В., Кривнова О.Ф. Общая фонетика: Учеб-

ник. М.: Рос.гос. гуманит. ун-т, 2001. 592с.

4. Li Deng, Dong Yu and Alex Acero, Structured Speech Modeling// IEEE Transactions on Audio, Speech, and Language Processing, Vol. 14, No. 5, Sep. 2006
5. Li Deng and et al., Acoustic Models with Structured Hidden Dynamics with Integration Over Many Possible Hidden Trajectories// US Patent No. 7,565,284 B2, Date of Patent Jul. 21, 2009
6. Li Deng, Alex Acero and Issam Bazzi, Tracking Vocal Tract Resonances Using a Quantized Nonlinear Function Embedded in a Temporal Constraint // IEEE Transactions on Audio, Speech, and Language Processing, Vol. 14, No. 2, March 2006
7. Leo Jingyu Lee, Hidden Dynamic Models for Speech Processing Applications// A thesis presented to the University of Waterloo Waterloo, Ontario, Canada, 2004
8. Li Deng, Dong Yu, Xiaolong Li and Alex Acero, A long-contextual-span model of resonance dynamics for speech recognition: parameter learning and recognizer evaluation// IEEE Workshop on Automatic Speech Recognition and Understanding (ASRU '05), Nov 27 – Dec 1, Cancun, Mexico, 2005.
9. Li Deng, Xiaodong Cui, Robert Pruvencok, Jonathan Huang, Safiyya Momen, Yanyi Chen and Abeer Alwan A Database of Vocal Tract Resonance Trajectories for Research in Speech Processing// Proceedings of the IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP 2006), May, 2006
10. Потапова Р. К., Потапов В. В. Речевая коммуникация: От звука к высказыванию – М.: Языки славянских культур, 2012. – 464 с.

## ANALYSIS OF HIDDEN TRAJECTORY MODELS (HTM) OF VOCAL TRACT RESONANCE (VTR) FOR PHONEMES RECOGNITION SYSTEMS

Lednov D.A.

In this paper there is given the translation to Russian and drawn the analysis of original In this paper there is given the translation to Russian and drawn the analysis of original works of Li Deng and his colleagues devoted to phoneme recognition published from 2000 to 2009. The main direction of these publications is the development of hidden trajectory models of vocal tract resonance. There has been illustrated the functional dependence between linear prediction coefficients and resonators parameters which are preliminarily smoothed by the FIR-filter. Optimization of a statistical model introduced for this functional dependence is performed. In this paper the above model is completed with continuity equation which enables to determine the FIR-filter features and an alternative optimization scheme is introduced which allows to estimate the statistical model parameters.

The main direction of these publications is the development of hidden trajectory models of vocal tract resonance. There has been illustrated the functional dependence between linear prediction coefficients and resonators parameters which are preliminarily smoothed by the FIR-filter. Optimization of a statistical model introduced for this functional dependence is performed. In this paper the above model is completed with continuity equation which enables to determine the FIR-filter features and an alternative optimization scheme is introduced which allows to estimate the statistical model parameters.



УДК 621.391.82

## КАСКАДНЫЙ ПРЕДКОМПЕНСАТОР ДЛЯ ЛИНЕАРИЗАЦИИ ХАРАКТЕРИСТИКИ УСИЛИТЕЛЯ МОЩНОСТИ

*Соловьева Е.Б., д.т.н., профессор, заведующая кафедрой теоретических основ электротехники Санкт-Петербургского государственного электротехнического университета "ЛЭТИ" им. В.И. Ульянова (Ленина), e-mail: selenab@hotmail.ru*

**Ключевые слова:** усилитель мощности, цифровой предкомпенсатор, линеаризация, модель Винера-Гаммерштейна, модель Вольтерры, персептронная сеть.

### Введение

С развитием мобильной связи все более жесткие требования предъявляются к передающим СВЧ-устройствам, содержащим усилители мощности (УМ). Увеличение скорости передачи данных приводит к расширению полосы частот сигналов и, как следствие, к появлению нелинейных искажений в выходных сигналах УМ [1], [2]. Совершенствование методов синтеза адаптивных предкомпенсаторов (предысказателей, digital predistorter, DPD) для линеаризации УМ является актуальной задачей. Наряду с модификацией полиномиальных моделей DPD развиваются методы синтеза DPD в виде нейронных сетей [3]. Нейронные модели DPD могут быть существенно проще полиномиальных, что важно при аппаратной реализации DPD.

В работе предлагается использовать функционально связанную искусственную нейронную сеть (functional link artificial neural network, FLANN) [3-5] и полиномиальную персептронную сеть (polynomial perceptron network, PPN) [5], [6] для синтеза каскадного DPD с целью повышения уровня подавления нелинейных искажений сигналов в УМ. Выполняется сравнительный анализ указанных моделей на примере синтеза DPD, компенсирующего нелинейные искажения в модели Винера-Гаммерштейна УМ на классе GSM-сигнала с четырьмя несущими.

### Адаптивный DPD с каскадной структурой

Задача DPD – внести нелинейные предыскажения, которые позволят компенсировать нелинейные искажения усилителя (линеаризовать УМ). Блок-схема включения адаптивного DPD с применением прямого алгоритма его обучения [7] изображена на рис. 1. Здесь DPD содержит два каскадно-соединенных блока В1 и В2.

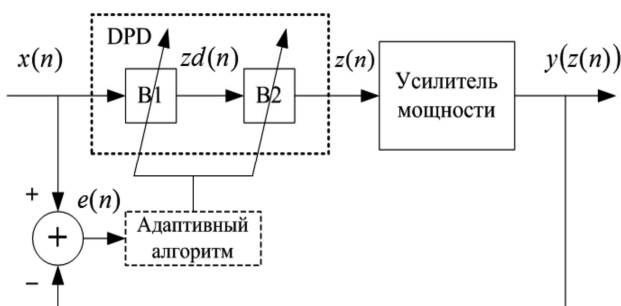


Рис. 1

*Предложена каскадная структура нелинейного цифрового предкомпенсатора, синтезируемого прямым алгоритмом обучения. Показано, что наибольшую точность линеаризации усилителя мощности с моделью Винера-Гаммерштейна обеспечивает каскадный предкомпенсатор, включающий полиномиальную персептронную сеть и радиально ограниченную модель Вольтерры.*

Блоки В1 и В2 описываются соответствующими операторными уравнениями:

$$zd(n) = S_1[x(n)],$$

$$z(n) = S_2[zd(n)],$$

где  $n$  – нормированное дискретное время,  $x(n)$ ,  $z(n)$  – входной и выходной сигналы DPD,  $zd(n)$  – выходной сигнал блока В1 предкомпенсатора,  $S_1[\cdot]$  и  $S_2[\cdot]$  – нелинейные операторы.

Параметры моделей  $S_1[\cdot]$  и  $S_2[\cdot]$  находим в метрике  $L_2$  последовательно (вначале – параметры модели  $S_1[x(n)]$ , затем модели  $S_2[zd(n)]$ ) с помощью итерационного метода (метода простых итераций [8]) при решении задачи аппроксимации

$$\|y(z(n)) - x(n)\| \Rightarrow \min_{n \in [0, N_x]},$$

где  $y(z(n))$  – выходной сигнал усилителя мощности,  $N_x$  – длительность сигнала  $x(n)$ . В качестве моделей  $S_1[\cdot]$  и  $S_2[\cdot]$  могут выступать как полиномиальные конструкции, так и нейронные сети.

Рассмотрим преобразования, предшествующие методу простых итераций. Введем приближенное равенство

$$y(z(n)) \cong x(n), \quad (1)$$

преобразуем его следующим образом:

$$x(n) - y(z(n)) \cong 0,$$

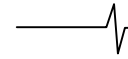
$$z(n) \cong z(n) + (x(n) - y(z(n))) = P[z(n)],$$

где  $P$  – нелинейный оператор, отображающий множество сигналов  $Z$  в себя. Оператор  $P$  называется оператором сжатия, если выполняется условие Липшица [9]

$$\|P[z_2] - P[z_1]\| \leq \alpha \|z_2 - z_1\|,$$

где  $\forall z_1, z_2 \in Z$ ,  $\alpha$  – множитель, не зависящий от  $z_1, z_2$ ,  $0 < \alpha < 1$ .

Согласно принципу сжатых отображений, оператор  $P$  задает метод последовательных приближений (простых итераций [8], [9])



$$z_{k+1}(n) = z_k(n) + (x(n) - y(z_k(n))) = z_k(n) + e_k(n), \quad k \geq 1,$$

где  $k$  – номер итерации,  $e_k(n)$  – ошибка приближения сигналов из выражения (1) на  $k$ -й итерации расчета,  $e_k(n) = x(n) - y(z_k(n))$ .

Усилитель мощности является слабонелинейным устройством, его модель можно представить в виде сходящегося функционального ряда Вольтерры (предкомпенсатор также слабонелинейен). Следовательно, сходимость итерационной процедуры синтеза DPD определяется слабой нелинейностью усилителя, выходной сигнал  $y(z_k(n))$  которого рассчитывается на итерациях.

Сигнал  $zd(n)$  формируем итерационной процедурой. Расчет можно выполнить двумя способами:

$$zd_{k+1}(n) = z_k(n) + e_k(n) \quad (2)$$

или

$$zd_{k+1}(n) = zd_k(n) + e_k(n). \quad (3)$$

На практике применяем формулу (3), при которой параметры моделей блоков B1 и B2 находятся с использованием выходных сигналов соответствующих блоков. В результате параметры моделей B1 и B2 сходятся к истинным значениям с большей точностью, чем при равенстве (2).

При  $k = 1$  полагаем  $z_1(n) = zd_1(n) = x(n)$ .

В рассмотренном подходе параметры моделей блоков B1 и B2 DPD находятся в два этапа. В результате:

- точность синтеза DPD повышается, так как решение, полученное на первом этапе аппроксимации (блок B1), уточняется на втором этапе (блок B2);

- частично снимается проблема плохой обусловленности, поскольку размерность задач, решаемых на каждом этапе аппроксимации, может быть меньше размерности задачи аппроксимации, решаемой при типовой (одноблочной) структуре DPD.

В качестве моделей операторов  $S_1[\cdot]$  и  $S_2[\cdot]$  каскадно-соединенных блоков DPD используем FLANN [3-5] и PPN [5], [6].

### Функционально связанная искусственная нейронная сеть (FLANN)

FLANN является однослойной сетью (в ней отсутствует внутренний слой) [3-5]. Алгоритм обучения такой сети включает меньшее число преобразований и обеспечивает более быструю сходимость к решению задачи аппроксимации по сравнению с традиционными нейронными сетями.

Модель FLANN имеет вид

$$y(n) = f\left(\sum_{i=1}^G w_i \varphi_i(\vec{X}(n))\right) = f\left(\vec{W}^T \vec{\Phi}(\vec{X}(n))\right) \quad (4)$$

где  $f$  – нелинейная функция активации,

$\vec{X}(n)$  – вектор воздействий,  $\vec{X}(n) = [x_1(n), x_2(n), \dots, x_Q(n)]^T$ ,  $T$  – знак транспонирования,  $\vec{W}(n)$  – вектор весов сети,  $\vec{W} = [w_1, w_2, \dots, w_G]^T$ ,  $\vec{\Phi}(\vec{X}(n))$  – вектор функций  $\varphi_i$ ,  $i = 1, 2, \dots, G$ ,

$\vec{\Phi}(\vec{X}(n)) = [\varphi_1(\vec{X}(n)), \varphi_2(\vec{X}(n)), \dots, \varphi_G(\vec{X}(n))]^T$ ,  $y(n)$  – выходной сигнал модели.

Функции  $\varphi_i$ ,  $i = 1, 2, \dots, G$  преобразуют входные сигналы в базисные функции с помощью, например, тригонометрического полинома, полинома Лежандра, Чебышева, а также выполняют многомерное преобразование базисных функций. Формирование базисных функций в модели (4) служит для понижения числа обусловленности при решении задачи аппроксимации с высокой степенью нелинейности.

Структурная схема FLANN показана на рис. 2.

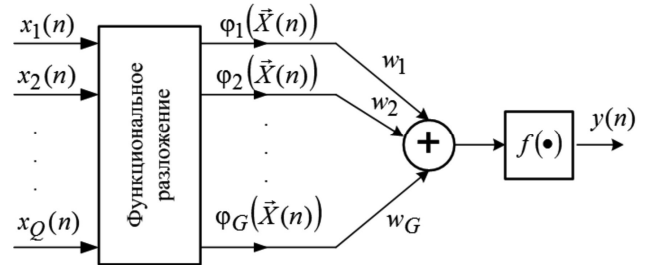


Рис. 2

Для синтеза DPD используем модель (4) с линейной функцией активации  $f$  и базисными функциями – полиномами Чебышева  $T_i(\vec{X}(n))$  степени  $i$ ,  $i = 1, 2, \dots, P$  [3], [10]. В этом случае модель (4) имеет вид

$$y(n) = \vec{W}^T \vec{\Phi}(\vec{X}(n)), \quad (5)$$

где

$$\vec{\Phi}(\vec{X}(n)) = [1, T_1(x_1(n)), \dots, T_Q(x_Q(n)), \dots, T_2(x_1(n)), \dots, T_2(x_Q(n))^T, \dots, T_1(x_1(n))T_1(x_2(n)), \dots, T_1(x_{Q-1}(n))T_1(x_Q(n)), \dots, T_P(x_1(n)), \dots, T_P(x_Q(n))]^T.$$

Блок «Функциональное разложение» (рис. 2) представляется структурой, показанной на рис. 3, включающей формирование полиномов Чебышева и их многомерное преобразование.

FLANN с чебышевскими базисными функциями будем называть чебышевской функционально связанной искусственной нейронной сетью (Chebyshev functional link artificial neural network, CFLANN). Отметим, что CFLANN соответствует двухслойной перцептронной нейронной сети [10].

CFLANN в DPD должна формировать выходной сигнал со спектром, расположенным в полосе пропускания усилителя на частотах входного сигнала и комбинационных частотах интермодуляционных составляющих, поро-

ждаемых нелинейностью УМ [11]. Составляющие модели CFLANN конструируются при перемножении каждого сомножителя из верхнего столбца табл. 1–4 на каждый сомножитель из соответствующего нижнего столбца. В таблицах знак \* – знак комплексного сопряжения, переменная  $i$  – смещение сигнала во времени. В итоге образуется вектор  $\vec{\Phi}(\vec{X}(n))$ , который умножается на вектор коэффициентов согласно формуле (5). Таким образом, конструируется модель CFLANN с нечетными членами.

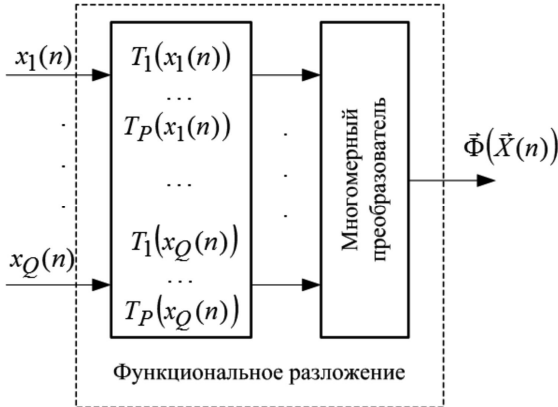


Рис. 3

Таблица 1. Множители составляющих 3-й степени модели CFLANN

Первый сомножитель	$2x^2(n) - 1$
	$x(n)x(n-i),$
	$2x^2(n-i) - 1$
Второй сомножитель	$x^*(n)$
	$x^*(n-i)$

Таблица 2. Множители составляющих 5-й степени модели CFLANN

Первый сомножитель	$4x^3(n) - 3x(n)$
	$4x^3(n-i) - 3x(n-i)$
	$(2x^2(n) - 1)x(n-i)$
	$(2x^2(n-i) - 1)x(n)$
Второй сомножитель	$2(x^*(n))^2 - 1$
	$2(x^*(n-i))^2 - 1$
	$x^*(n)x^*(n-i)$

Таблица 3. Множители составляющих 7-й степени модели CFLANN

Первый сомножитель	$8x^4(n) - 8x^2(n) + 1$
	$8x^4(n-i) - 8x^2(n-i) + 1$
	$(2x^2(n) - 1)(2x^2(n-i) - 1)$
	$x(n)(4x^3(n-i) - 3x(n-i))$
	$x(n-i)(4x^3(n) - 3x(n))$

Второй сомножитель	$4(x^*(n))^3 - 3x^*(n)$
	$4(x^*(n-i))^3 - 3x^*(n-i)$
	$(2(x^*(n))^2 - 1)x^*(n-i)$
	$(2(x^*(n-i))^2 - 1)x^*(n)$

Таблица 4. Множители составляющих 9-й степени модели CFLANN

Первый сомножитель	$16x^5(n) - 20x^3(n) + 5x(n)$
	$16x^5(n-i) - 20x^3(n-i) + 5x(n-i)$
	$x(n)(8x^4(n-i) - 8x^2(n-i) + 1)$
	$x(n-i)(8x^4(n) - 8x^2(n) + 1)$
	$(4x^3(n) - 3x(n))(2x^2(n-i) - 1)$
	$(4x^3(n-i) - 3x(n-i))(2x^2(n) - 1)$
Второй сомножитель	$8(x^*(n))^4 - 8(x^*(n))^2 + 1$
	$8(x^*(n-i))^4 - 8(x^*(n-i))^2 + 1$
	$(2(x^*(n))^2 - 1)(2(x^*(n-i))^2 - 1)$
	$(4(x^*(n))^3 - 3x^*(n))x^*(n-i)$
	$(4(x^*(n-i))^3 - 3x^*(n-i))x^*(n)$

**Полиномиальная персептронная сеть (PPN)**

PPN – однослойная сеть (без внутреннего слоя), что обеспечивает простоту алгоритма ее обучения и высокую скорость сходимости к решению задачи аппроксимации.

Модель PPN имеет вид [5], [6]

$$y(n) = f(\vec{W}^T \vec{F}(\vec{X}(n))), \tag{6}$$

где  $f$  – нелинейная функция активации,  $\vec{X}(n)$  – вектор воздействий,  $\vec{X}(n) = [x_1(n), x_2(n), \dots, x_Q(n)]^T$ ,

$T$  – знак транспонирования,  $\vec{W}(n)$  – вектор весов нейронной сети,  $\vec{W} = [w_1, w_2, \dots, w_G]^T$ ,  $\vec{F}(\vec{X}(n))$  – вектор функций, содержащий элементы с многомерным преобразованием,

$$\vec{F}(\vec{X}(n)) = [1, x_1(n), \dots, x_Q(n), \dots, x_1^2(n), \dots, x_Q^2(n), \dots, x_1(n)x_2(n), \dots, x_{Q-1}(n)x_Q(n), \dots, x_1^P(n), \dots, x_Q^P(n)]^T,$$

$P$  – степень элемента вектор-функции,  $y(n)$  – выходной сигнал модели.

Структурная схема PPN изображена на рис. 4.

Анализ PPN и FLANN показывает, что PPN (рис. 4) может быть преобразована во FLANN (рис. 2) путем дополнительного преобразования входных сигналов в базисные функции на основе полиномов Лежандра, Чебышева и т.д.



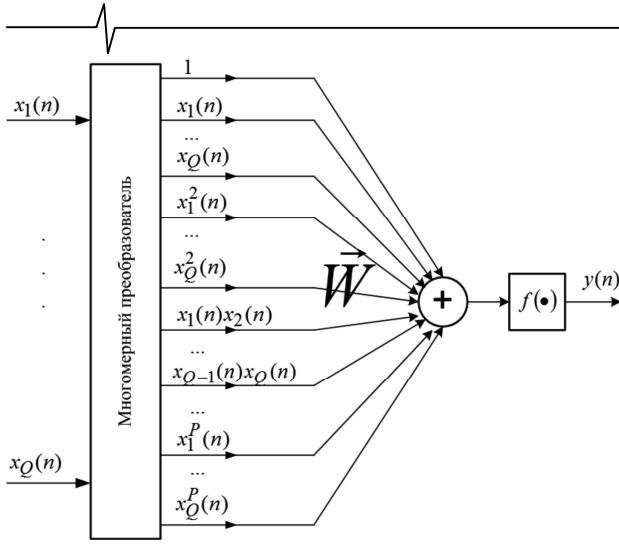


Рис. 4

При синтезе DPD введем в модель (6) линейную функцию активации. В итоге запишем

$$y(n) = \vec{W}^T \vec{F}(\vec{X}(n)). \quad (7)$$

Вектор воздействий  $\vec{X}(n)$  сформируем с помощью линии задержки и выполним многомерное преобразование в  $\vec{F}$  модели (7) с учетом правил конструирования интермодуляционных спектральных составляющих в выходном сигнале DPD [11]. В результате получим используемую для синтеза DPD модель PPN вида

$$y(n) = y_1(n) + y_2(n), \quad (8)$$

где

$$\begin{aligned}
y_1(n) = & \sum_{i=0}^M x(n-i) \sum_{k=0}^{(P-1)/2} w_{2k+1}^{(1)} |x(n-i)|^{2k} + \\
& \sum_{i=1}^M x(n-i) \sum_{k=1}^{(P-1)/2} w_{2k+1}^{(2)} |x(n)|^{2k} + \\
& + \sum_{i=0}^M x^*(n-i) x^2(n) \sum_{k=0}^{(P-3)/2} w_{2k+3}^{(3)} |x(n)|^{2k} + \\
& + \sum_{i=1}^M x^*(n) x^2(n-i) \sum_{k=0}^{(P-3)/2} w_{2k+3}^{(4)} |x(n-i)|^{2k} + \\
& + \sum_{i=1}^M x(n) \sum_{k=1}^{(P-1)/2} w_{2k+1}^{(5)} |x(n-i)|^{2k} + \\
& + x^2(n) \sum_{i=0}^M x^*(n-i) \sum_{k=0}^{(P-3)/2} w_{2k+3}^{(6)} |x(n-i)|^{2k} + \\
& + |x(n)|^2 \sum_{i=0}^M x(n-i) \sum_{k=0}^{(P-3)/2} w_{2k+3}^{(7)} |x(n-i)|^{2k} + \\
& + \sum_{i=0}^M (x^*(n))^2 x^3(n-i) \sum_{k=0}^{(P-5)/2} w_{2k+5}^{(8)} |x(n-i)|^{2k}. \quad (9)
\end{aligned}$$

\* – знак комплексного сопряжения,  $M$ ,  $P$  – длина памяти и степень модели соответственно,

$$y_2(n) = \sum_{i=1}^M (x^*(n-i))^2 x^3(n) \sum_{k=0}^{(P-5)/2} w_{2k+5}^{(9)} |x(n)|^{2k} +$$

$$\begin{aligned}
& + \sum_{i=1}^M x^*(n) x^2(n-i) \sum_{k=1}^{(P-3)/2} w_{2k+3}^{(10)} |x(n)|^{2k} + \\
& + \sum_{i=1}^M (x^*(n-i))^2 x^3(n) \sum_{k=1}^{(P-5)/2} w_{2k+5}^{(11)} |x(n-i)|^{2k} + \\
& + \sum_{i=1}^M (x^*(n))^2 x^3(n-i) \sum_{k=1}^{(P-5)/2} w_{2k+5}^{(12)} |x(n)|^{2k} + \\
& + |x(n)|^2 \sum_{i=1}^M x(n) \sum_{k=1}^{(P-3)/2} w_{2k+3}^{(13)} |x(n-i)|^{2k} + \\
& + \sum_{i=1}^M |x(n-i)|^2 x(n) \sum_{k=1}^{(P-3)/2} w_{2k+3}^{(14)} |x(n)|^{2k}. \quad (10)
\end{aligned}$$

Отметим, что выражение (9) является радиально ограниченной моделью Вольтерры (radially pruned Volterra model, RPVM). Модель RPVM – регрессионная форма усеченного ряда Вольтерры. Ядра Вольтерры в RPVM строятся на сетке  $n$ -мерного куба (гиперкуба), радиальные направления выбираются на базе ядра 3-го порядка [12], [13].

#### Компенсация нелинейных искажений в модели Винера-Гаммерштейна УМ

Низкочастотная модель Винера-Гаммерштейна УМ класса АВ описывает каскадное соединение следующих блоков [13], [14]:

– линейной динамической цепи с передаточной функцией

$$H_1(z) = \frac{1 + 0,5z^{-2}}{1 - 0,2z^{-1}};$$

– безынерционной нелинейности

$$w(n) = b_1 v(n) + b_3 v(n) |v(n)|^2 + b_5 v(n) |v(n)|^4,$$

где  $b_1 = 1,0108 + 0,0858j$ ,  $b_3 = 0,0879 - 0,1583j$ ,  $b_5 = -1,0992 - 0,8891j$ ,  $v(n)$  – выходной сигнал динамической цепи с  $H_1(z)$ ;

– линейной динамической цепи с передаточной функцией

$$H_2(z) = \frac{1 - 0,1z^{-2}}{1 - 0,4z^{-1}}.$$

Воздействие низкочастотной модели УМ – комплексная огибающая GSM-сигнала с четырьмя несущими в частотной полосе 20 МГц, расположенной относительно центральной частоты 1,845 ГГц. Частота дискретизации огибающей GSM-сигнала  $f_d = 184,32$  МГц.

Для линеаризации описанной модели УМ построены адаптивные DPD с применением модели 9-й степени CFLANN, RPVN (9) при  $P = 7$ , модели (8) PPN при  $P = 7$ , а также каскадных соединений указанных моделей. Отметим, что в выражениях (9), (10) верхние пределы операторов суммирования равнялись  $(P-1)/2$ .

Длина памяти исследуемых моделей равнялась четырем ( $M = 4$  в (9), (10)).

Точность линеаризации УМ оценивалась с помощью нормированной среднеквадратичной погрешности (normalized mean-square error, NMSE, дБ), рассчитываемой по формуле

$$NMSE = 10 \log_{10} \left( \frac{\sum_{n=0}^{N_x-1} |y(z(n)) - x(n)|^2}{\sum_{n=0}^{N_x-1} |x(n)|^2} \right),$$

где  $x(n)$  – входной сигнал длиной  $N_x$  отсчетов ( $N_x = 106\,339$ ) каскадного соединения DPD и PA, изображенного на рис. 1,  $y(z(n))$  – выходной сигнал УМ.

Значения NMSE, полученные на 45-й итерации адаптации DPD, а также число параметров  $Q$  в моделях DPD, представлены в табл. 5.

Таблица 5. NMSE и число параметров  $Q$  в моделях DPD

	NMSE, $Q$	NMSE	$Q$
	Модель 9-й степени CFLANN	-65,04	226
	RPVM (9) при $P = 7$	-72,87	145
	Модель (8) PPN при $P = 7$	-75,50	221
Каскадное соединение	Модель 9-й степени CFLANN и модель 7-й степени CFLANN	-69,42	375
	RPVM (9) при $P = 7$ и RPVM (9) при $P = 7$	-78,14	290
	Модель (8) PPN при $P = 7$ и RPVM (9) при $P = 7$	-79,14	366

Из анализа табл. 5 следует:

- модель PPN обеспечивает более высокую точность линеаризации УМ, чем RPVM и модель CFLANN,
- применение каскадной структуры DPD повышает точность линеаризации УМ. Большая точность линеаризации достигается при каскадной структуре DPD с моделью PPN.

### Заключение

В работе синтезированы адаптивные DPD с каскадной структурой на основе прямого алгоритма обучения. В качестве моделей блоков DPD использованы: функционально связанная нейронная сеть, полиномиальный перцептрон, радиально ограниченная модель Вольтерры.

В каскадном DPD задача аппроксимации нелинейного оператора решалась поэтапно. При таком подходе частично снимается проблема плохой обусловленности, характерная для задач с высокой степенью нелинейности моделей, а также повышается точность линеаризации УМ за счет лучшего приближения к оптимальным параметрам DPD.

Компенсация нелинейных искажений сигналов в УМ с моделью Винера-Гаммерштейна на классе GSM-сигнала с четырьмя несущими показала, что более точную линеаризацию характеристики УМ обеспечивает каскадный DPD, состоящий из PPN и RPVM, худший результат дает

DPD на основе моделей CFLANN.

### Литература

1. RF power amplifier behavioral modeling / D. Schreurs, M. O'Droma, A. A. Goacher, M. Gadringer.– N. Y.: Cambridge University Press, 2009.
2. Legarda J. Feedforward amplifiers for wideband communication systems.– The Netherlands, Dordrecht: Springer, 2006.
3. Li M., Liu J., Jiang Y., Feng W. Complex-Chebyshev functional link neural network behavioral model for broadband wireless power amplifiers // IEEE Trans. MTT.– 2012.– Vol. 60, № 6, Part 2.– P. 1979–1989.
4. Identification of nonlinear dynamic systems using functional link artificial neural networks / J. C. Patra, R. N. Pal, B. N. Chatterji, G. Panda // IEEE Trans. Systems, Man and Cybernetics – Part B: Cybernetics.– 1999.– Vol. 29, № 2.– P. 254–262.
5. Nonlinear channel equalization for QAM signal constellation using artificial neural networks / J. C. Patra, R. N. Pal, R. Baliarsingh, G. Panda // IEEE Trans. Systems, Man and Cybernetics – Part B: Cybernetics.– 1999.– Vol.29, № 2.– P. 262–271.
6. Xiang Z., Bi G., Le-Ngoc T. Polynomial-perceptrons and their applications to fading channel equalization and co-channel interference suppression // IEEE Trans. SP.– 1994.– V.42, № 9.– P.2470–2480.
7. Zhou D., DeBrunner V. E. Novel adaptive nonlinear predistorters based on the direct learning algorithm // IEEE Trans. SP.–2007.– Vol. 55.– № 1.– P. 120–133.
8. Бахвалов Н. С., Жидков Н. П., Кобельков Г. М. Численные методы.– М.: БИНОМ. Лаборатория знаний, 2003.
9. Лебедев В. И. Функциональный анализ и вычислительная математика.– М.: Физматлит, 2000.
10. Lee T. T., Jeng J. T. The chebyshev polynomial based unified model neural networks for function approximations // IEEE Trans. Systems, Man and Cybernetics – Part B: Cybernetics.– 1998.– Vol. 28, № 6.– P. 925–935.
11. On the baseband representation of a bandpass nonlinearity / G. T. Zhou, H. Qian, L. Ding, R. Raich // IEEE Trans. SP.– 2005.– Vol. 53, № 8, P. 2953–2957.
12. A new approach to pruning Volterra models for power amplifiers / C. Crespo-Cadenas, J. Reina-Tosina, M. J. Madero-Ayora, J. Munoz-Cruzado // IEEE Trans. SP.– 2010.– Vol. 58.– № 4.– P. 2113–2120.
13. Соловьева Е. Б. Предкомпенсатор Вольтерры с девиацией динамики для линеаризации усилителя мощности // Известия вузов. Радиоэлектроника.– 2011.– Т.54, № 10.– С. 29–36.
14. A robust digital baseband predistorter constructed using memory polynomials / L. Ding, G. T. Zhou, D. R. Morgan, Z. Ma, J. S. Kenney, J. Kim, C. R. Giardina // IEEE Trans. Com.– 2004.– Vol. 52.– № 1.– P. 159–165.

## CASCADE PREDCOMPENSATOR FOR POWER AMPLIFIER LINEARIZATION

*Solovyeva E. B.*

Cascade structure of nonlinear digital predcompensator synthesized by the direct learning algorithm is proposed. It is shown that the most accuracy of linearization of power amplifier described by Wiener-Hammerstein model is provided with the cascade predcompensator including the polynomial perceptron network and radially pruned Volterra model.

## МАТЕМАТИЧЕСКИЕ МОДЕЛИ МОДУЛЯТОРА И ДЕМОДУЛЯТОРА С ЗАДАНЫМ ПОРЯДКОМ НЕЛИНЕЙНОСТИ

*Батенков К.А., Академия ФСО, г. Орел, e-mail: pustur@yandex.ru*

**Ключевые слова:** ряды Вольтерра, математическая модель, нелинейная модуляция, демодуляция, базисные функции, оценка сложности.

### Введение

Широкое использование в современных системах передачи информации цифровой обработки сигналов неуклонно увеличивает возможности подобных систем с точки зрения количества и качества передаваемой информации. При этом достаточно узким местом оказывается непосредственно процедура преобразования аналогового сигнала в цифровой и обратно, неизбежно приводящая к информационным потерям. Именно поэтому разработка преобразований, позволяющих минимизировать подобные негативные моменты, является достаточно обоснованной и насущной задачей. В этой связи следует отметить, что наиболее рациональный путь снижения потерь – это переход от линейных систем обработки сигналов к нелинейным, поскольку подобная трансформация, по сути, осуществляет обобщение линейных частных случаев отображений сигналов на более широкий класс преобразований, включающих как линейные, так и нелинейные эффекты.

В данной работе исследуются процедуры нелинейной модуляции и демодуляции, осуществляющие преобразование конечномерных сигналов в континуальную форму и обратно. Причем делается попытка рассмотреть подобные трансформации в общем виде, не прибегая к фиксации типа преобразований как каких-либо строгих функциональных зависимостей (например, экспоненциальных, степенных, гармонических и т.п.). В качестве основы подобных операций положены функциональные ряды Вольтерра, обладающие достаточно широкими возможностями как с позиции последующей реализации получаемых алгоритмов, так и с позиции удобства пользования математическим аппаратом.

Так, с одной стороны, существующая теория рядов Вольтерра позволяет достаточно полно описать возможные нелинейные эффекты, возникающие при преобразовании аналоговых сигналов в дискретную форму, что делает возможным аналитическое оценивание качества синтезируемых систем и устройств. С другой стороны, использование обобщенных рядов Фурье с заданным базисом (например, Котельникова) для разложения функциональных ядер Вольтерра приводит к реализуемости операций многопараметрической свертки на базе методов цифровой обработки сигналов, а значит, появляется возможность реализации процедур нелинейной модуляции и демодуляции на современных процессорах

*Синтезированы математические модели модулятора и демодулятора с заданным порядком нелинейности на основе функционального ряда. Проведена оценка сложности процедур обработки сигналов на уровне количества используемых базисных функций.*

обработки сигналов. Следует отметить, что в данном случае неизбежно появляется некоторая погрешность вследствие оперирования лишь с конечным числом коэффициентов разложения ряда Фурье. Тем не менее, зачастую подобные ошибки аппроксимации оказываются пренебрежимо малыми, а в пределе стремятся вообще к нулю, что также позволяет варьировать качество синтезируемых систем за счет усложнения выполняемых ими операций.

Таким образом, основной задачей данной статьи является разработка аналитической модели нелинейного модулятора и демодулятора в достаточно обобщенном виде (на базе функциональных рядов Вольтерра), позволяющей ограничивать порядок нелинейности операций модуляции и демодуляции, что в свою очередь делает возможным задавать порог вычислительной сложности подобных процедур. В результате, главным побудительным мотивом постановки задачи синтеза модулятора и демодулятора в данной формулировке оказывается попытка снизить потери передаваемой информации по каналу связи за счет усложнения используемых преобразований, то есть путем перехода от существующих линейных к новым нелинейным формам модулятора и демодулятора. В то же время следует учитывать неизбежный рост сложности реализации синтезируемых систем, что приведет к повышению требований к процессорам цифровой обработки сигналов как по производительности, так и по объему оперативной памяти. В связи с этим в работе аналитически оценивается сложность процедур обработки сигналов в зависимости от порядка нелинейности операторов модуляции и демодуляции, а также от размерности обрабатываемых сигналов. В итоге разработчики получают возможность предварительно, до непосредственной реализации алгоритмов, оценить требуемые вычислительные ресурсы для синтеза тех или иных модуляторов и демодуляторов и сопоставить с ресурсами доступных аппаратных средств.

Данные обстоятельства и обосновывают актуальность решения поставленной в работе задачи, поскольку с одной стороны в настоящее время сохраняется тенденция к увеличению количества и качества передаваемой информации, а с другой в доступной литературе нет изложения теории нелинейных преобразований применительно к задачам модуляции и демодуляции сигналов,



которые и позволяют значительно снизить потери информации в каналах связи.

### Описание операторов нелинейной модуляции и демодуляции

Следует отметить, что в каноническом виде ряд Вольтерра определяет нелинейные преобразования между бесконечномерными пространствами, в то время как операции модуляции  $\Phi$  и демодуляции  $\Phi'$  задают отношения между конечномерными и бесконечномерными пространствами. Таким образом, необходимо несколько видоизменить ряды Вольтерра путем представления сигналов на соответствующих конечномерных концах в виде рядов из взвешенных дельта-функций, то есть сигналы на входе модулятора и выходе демодулятора можно записать в непрерывном виде, как следующие суммы:

$$x(t_j, \mathbf{r}_j) = \sum_{k_j=1}^N x_{k_j} \delta(k_j, t_j) \delta(t_j \mathbf{1}_{n_a} - \mathbf{r}_j), \quad j = \overline{1, \infty}, \quad (1)$$

$$x(t', \mathbf{r}') = \sum_{k_j=1}^{N'} x'_{k_j} \delta(k_j, t') \delta(t' \mathbf{1}_{n_b} - \mathbf{r}'), \quad (2)$$

где  $\mathbf{x} = (x_1, x_2, \dots, x_N)^T$ ,  $\mathbf{x}' = (x'_1, x'_2, \dots, x'_{N'})^T$  – вектора на входе и выходе дискретного канала связи (на входе модулятора и выходе демодулятора) размерности  $N$  и  $N'$  соответственно;  $\delta(t)$  – дельта-функция Дирака;  $\mathbf{1}_n$  – единичный вектор размерности  $n$ ;  $n_a, n_b$  – число измерений пространства на входе и выходе непрерывного многопараметрического канала связи соответственно.

При этом в функциональном ряде ядра Вольтерра трактуются в смысле некоторых функций (базисных или несущих), то есть по сути являются системными характеристиками, но не канала, а модулятора и демодулятора и обозначаются  $\varphi_i(t, \mathbf{r}, t_1, \dots, x_i, \mathbf{r}_1, \dots, \mathbf{r}_i)$  и  $\varphi'_i(t, \mathbf{r}, t_1, \dots, x_i, \mathbf{r}_1, \dots, \mathbf{r}_i)$  соответственно. Подстановка непрерывного вида сигналов на входе модулятора (1) в ряд Вольтерра, а также учет нелинейности, не превышающей величины  $N_a$  (число членов ряда Вольтерра ограничено данным значением), и замена переменных делают возможным следующее представление:

$$\begin{aligned} x(t, \mathbf{r}) = & \\ = & \sum_{i=1}^{N_a} \int_{t_1} \dots \int_{t_i} \dots \int_{\mathbf{r}_1} \dots \int_{\mathbf{r}_i} \left\{ \prod_{j=1}^i \sum_{k_j=1}^N x_{k_j} \delta(k_j, t_j) \delta(t_j \mathbf{1}_{n_a} - \mathbf{r}_j) \right\} \times \quad (3) \\ & \times \varphi_i(t, \mathbf{r}, t_1, \dots, x_i, \mathbf{r}_1, \dots, \mathbf{r}_i) dt_1 \dots dt_i d\mathbf{r}_1 \dots d\mathbf{r}_i. \end{aligned}$$

Раскрытие произведения, перегруппировка слагаемых внутри него, а также использование фильтрующего свойства дельта-функции трансформируют операцию нелинейной модуляции к следующему виду:

$$\begin{aligned} x(t, \mathbf{r}) = & \sum_{i=1}^{N_a} \left[ \sum_{k_1=1}^N \sum_{k_2=1}^N \dots \sum_{k_i=1}^N \left\{ \prod_{j=1}^i x_{k_j} \right\} \right] \times \quad (4) \\ & \times \varphi_i(t, \mathbf{r}, k_1, \dots, k_i, k_1 \mathbf{1}_{n_a}, \dots, k_i \mathbf{1}_{n_a}). \end{aligned}$$

Отсюда следует важное условие осуществимости нелинейной модуляции, выражающееся в обязательном превышении или по крайней мере равенстве размерно-

сти сигналов на входе модулятора  $N$  степени нелинейности модулятора  $N_a$ , то есть  $N_a \leq N$ . В обратном случае ( $N_a > N$ ) произведения в фигурной скобке (4) не существуют, что означает идентичность операций модуляции с нелинейностью большей  $N_a$  операции модуляции с нелинейностью  $N_a$ .

Следует также заметить, что вследствие свойства коммутативности операции умножения (то есть  $\prod_{j=1}^i x_{k_j} = \prod_{j'=1}^i x_{k'_j}$ , даже при условии, что  $\exists k_j \neq k'_j$ ) число суммируемых базисных функций можно существенно сократить. Количество одинаковых произведений соответствует числу перестановок упорядоченного множества индексов сигнала на входе модулятора  $\{k_j\}$ ,  $j = \overline{1, i}$ ,  $k_j = k'_j | j < j', j' = \overline{1, i}$ , а само данное множество перестановок целесообразно обозначить как  $P_i$ , каждый элемент которого  $P_{i,p} = \{k_{j_p}\}$  представляет собой уникальную перестановку исходного упорядоченного множества. Следовательно, базисные функции модуляции являются функциями только временных и пространственных выходных координат, так как входные координаты дискретны и имеют следующий вид:

$$\varphi_{k_1, \dots, k_i}(t, \mathbf{r}) = \sum_{\{k_{j_p}\} \in P_i} \varphi_i(t, \mathbf{r}, k_{i_1} \mathbf{1}_{n_a}, \dots, k_{i_i} \mathbf{1}_{n_a}),$$

$$j_p = \overline{1, i}, \quad p = \overline{1, i!}, \quad i = \overline{1, N_a}, \quad (5)$$

где число элементов множества перестановок  $P_i$  равно числу всех возможных перестановок, то есть  $i!$  [1]. Перегруппировка слагаемых в (4) и использование (5) приводят к формуле нелинейной модуляции в виде

$$x(t, \mathbf{r}) = \sum_{i=1}^{N_a} \left[ \sum_{k_1=1}^N \sum_{k_2=1}^N \dots \sum_{k_i=1}^N \left\{ \prod_{j=1}^i x_{k_j} \right\} \varphi_{k_1, \dots, k_i}(t, \mathbf{r}) \right]. \quad (6)$$

Следует отметить, что по сравнению с (4) выражение (6) содержит значительно меньшее количество слагаемых, что бесспорно упрощает процедуру реализации нелинейной модуляции. При этом число элементов сумм соответствует количеству базисных функций модуляции  $N_\Phi$ . Так, каждая сумма по  $k_j$  в (4) содержит  $N$  элементов:

$$N'_{\Sigma_{k_j}} = N, \quad j = \overline{1, i}. \quad (7)$$

В результате все суммы по  $k_j$  имеют  $N^i$  членов, как число размещений с повторениями

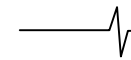
$$N'_{\Sigma_{k_i}} = N^i. \quad (8)$$

Тогда количество слагаемых в сумме по  $i$ , соответствующее числу базисных функций модуляции  $N'_\Phi$ , вычисляется как сумма геометрической прогрессии  $N_a$  элементов со знаменателем и первым членом равными  $N$ , то есть

$$N'_\Phi = \sum_{i=1}^{N_a} N \cdot N^{i-1} = \frac{N(N^{N_a} - 1)}{N - 1}. \quad (9)$$

В то время как в (6) крайняя правая сумма по  $k_i$  имеет только

$$N'_{\Sigma_{k_i}} = k_{i-1} + 1 = C_{k_{i-1}+1}^1 \quad (10)$$



элементов, как число сочетаний без повторов

$$C_b^a = \frac{b!}{a!(a-b)!}, \quad a \leq b, \quad a, b = \overline{0, \infty} \quad [1], \quad \text{по } k_{i-1}:$$

$$N_{\Sigma k_{i-1}} = \sum_{k_{i-1}=0}^{k_{i-2}} C_{k_{i-1}+1}^1 = C_{k_{i-2}+2}^2 \quad (11)$$

членов, в соответствии с формулой для сумм биномиальных коэффициентов [2]. Следовательно, по индукции сумма по  $k_1$  содержит

$$N_{\Sigma k_1} = \sum_{k_1=0}^{N-1} C_{k_1+i-1}^{i-1} = C_{N+i-1}^i \quad (12)$$

слагаемых. В итоге количество элементов в сумме по  $i$  вычисляется как число сочетаний из  $N_a+N$  по  $N$

$$N_\varphi = \sum_{i=1}^{N_a} C_{N+i-1}^i = C_{N_a+N}^{N_a} \quad (13)$$

Таким образом, в (4) количество элементов сумм (9) показательно зависит от степени нелинейности модулятора  $N_a$  и полиномиально (с показателем степени  $N_a$ ) от размерности входного сигнала  $N$ , то есть

$$N'_\varphi \approx N^{N_a}, \quad N \gg 1. \quad (14)$$

В (6) зависимости числа слагаемых подобны, однако величина масштабирующего коэффициента отлична от единицы и зависит от степени нелинейности модулятора  $N_a$ . Так, использование приближенной формулы Стирлинга [1] и сокращение подобных множителей преобразуют выражение для числа слагаемых (13) в (6) к виду

$$N_\varphi \approx \frac{(N_a + N)^{N_a+N+1/2}}{N_a^{N_a+1/2} N^{N_a+1/2} \sqrt{2\pi}} \dots \quad (15)$$

Деление числителя и знаменателя на  $N^{N_a+N+1/2}$ , внесение обратной величины в выражение внутри скобки, сокращение и применение свойства произведения показательных функций с одинаковыми основаниями трансформируют данное выражение в следующую форму:

$$N_\varphi \approx \frac{(N_a/N + 1)^{N_a+1/2} (N_a/N + 1)^N N^{N_a}}{N_a^{N_a+1/2} \sqrt{2\pi}} \quad (16)$$

Применение следствия второго замечательного предела при  $N \gg 1$ , а также приближенное равенство единице первого выражения в скобке, так как  $N_a/N \rightarrow 0$ , приводит к асимптотическому равенству

$$N_\varphi \approx \frac{e^{N_a}}{\sqrt{2\pi} N_a^{N_a+1/2}}, \quad N \gg 1, \quad (17)$$

которое доказывает полиномиальную зависимость числа базисных функций модулятора  $N_\varphi$  от достаточно больших размерностей сигналов на входе модулятора  $N$  с показателем степени, соответствующим степени нелинейности  $N_a$ , и масштабирующим множителем

$$\frac{e^{N_a}}{\sqrt{2\pi} N_a^{N_a+1/2}}$$

меньшим единицы при  $N_a > 1$  и умень-

шающимся по мере роста  $N_a$ , что свидетельствует о значительном сокращении числа базисных функций моду-

ляции, определяемых выражением (6) по сравнению с (4). Для примера на рис. 1 представлены полученные на основе (9), (13) и (17) зависимости количества базисных функций модуляции  $N_\varphi$  от размерности сигналов на входе модулятора  $N$  при различных степенях нелинейности  $N_a$ . Данные графики подтверждают достаточно высокую точность приближения при  $N \gg 1$ , а также снижение числа слагаемых, определяемое масштабирующим коэффициентом в (17).

При исследовании поведения числа элементов в суммах (4) и (6) при увеличении степени нелинейности модулятора  $N_a$  следует учитывать условие осуществимости подобного типа модуляции, то есть неравенство  $N_a \leq N$ . В результате при больших степенях нелинейности  $N_a$  они становятся сопоставимы с размерностями сигналов  $N$ , что выражается в форме  $N_a \approx N$ ,  $N \gg 1$ . Таким образом (9) преобразуется к виду

$$N_{\Sigma_i} \approx N_a^{N_a} = 2N_a, \quad N_a \approx N, \quad (18)$$

известному как степенная башня порядка два, или тетрация  $2N_a$ . Количество базисных функций модуляции определяется на основе (15) при условии  $N_a \approx N$ . В результате после сокращения подобных множителей данное выражение приводится к форме

$$N_\varphi \approx \frac{4^{N_a}}{\sqrt{\pi} N_{N_a}}, \quad N_a \approx N. \quad (19)$$

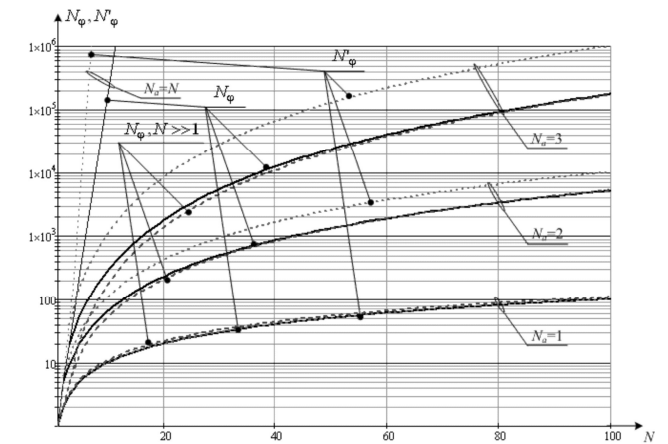


Рис. 1. Зависимости количества базисных функций модуляции  $N_\varphi$  от размерности сигналов на входе модулятора  $N$  при различных степенях нелинейности  $N_a$

Таким образом, согласно (19), число базисных функций модуляции растет даже медленнее, чем по показательному закону, в то время как в соответствии с (18) увеличение числа слагаемых подчиняется гипероператорному закону с высотой два, значительно превосходящему показательный по скорости нарастания функции. Для примера на рис. 2 представлены полученные на основе (9), (13) и (19) зависимости числа базисных функций модуляции  $N_\varphi$  от степени нелинейности  $N_a$  при различных размерностях сигналов на входе модулятора  $N$ . Данные графики подтверждают достаточно высокую точность приближения при  $N_a \approx N$ , а также значительное снижение числа слагаемых, определяемое различием

между ростом тетрации  $2N_a$  и полиномиально-показательной функции (19).

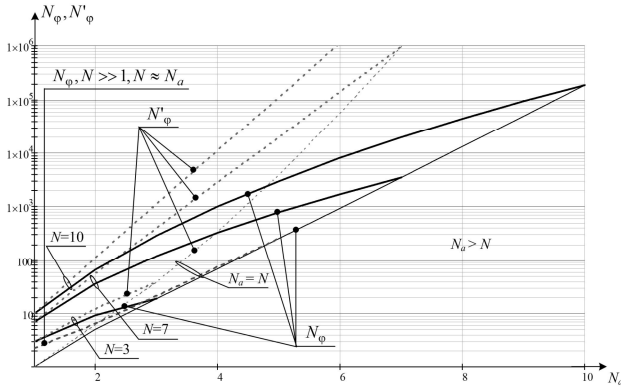


Рис. 2. Зависимости числа базисных функций модуляции  $N_\phi$  от степени нелинейности  $N_a$

при различных размерностях сигналов на входе модулятора  $N$

Следует также отметить, что, по сути, определенное число базисных функций модуляции  $N_\phi$  в операторе (13) задает максимально возможное их количество, поскольку в общем случае возможно использование меньшего числа членов ряда (6), вплоть до одного. Именно поэтому формулы (13), (17) и (19) позволяют вычислять верхние границы сложности нелинейной модуляции, причем последние две используют достаточно простые степенные зависимости в отличие от первой, требующей вычисления факториалов, что не всегда оказывается возможным. Данное обстоятельство оказывается в ряде случаев довольно критичным при реализации алгоритмов модуляции программно-аппаратными средствами.

Для описания оператора нелинейной модуляции  $\Phi'$  также следует использовать непрерывный вид сигналов, но на выходе демодулятора (2). Его подстановка в ряд Вольтерра, учет нелинейности, не превышающей величины  $N_b$  (число членов ряда Вольтерра ограничено данным значением), и замена переменных, делают возможным следующее представление:

$$\sum_{k=1}^{N'} x'_k \delta(k-t') \delta(t' \mathbf{1}_{n_b} - \mathbf{r}') = \sum_{i=1}^{N_b} \int \dots \int \dots \int \left\{ \prod_{j=1}^i x'(t'_j, \mathbf{r}'_j) \right\} \times \varphi'_i(t', \mathbf{r}', t'_1, \dots, t'_i, \mathbf{r}'_1, \dots, \mathbf{r}'_i) \times dt'_1 \dots dt'_i \dots dt'_i d\mathbf{r}'_1 \dots d\mathbf{r}'_i. \quad (20)$$

Использование свойства дельта-функции (отличия от нуля только в случае равенства нулю аргумента) позволяет представить выражение для координат сигнала на выходе демодулятора в виде

$$x'_k = \sum_{i=1}^{N_b} \int \dots \int \dots \int \left\{ \prod_{j=1}^i x'(t'_j, \mathbf{r}'_j) \right\} \times \varphi'_i(k, k \mathbf{1}_{n_b}, t'_1, \dots, t'_i, \mathbf{r}'_1, \dots, \mathbf{r}'_i) \times dt'_1 \dots dt'_i \dots dt'_i d\mathbf{r}'_1 \dots d\mathbf{r}'_i, \quad k = \overline{1, N'} \quad (21)$$

Естественно, что в отличие от модулятора соотношения между размерностью выходного сигнала  $N'$  и степенью нелинейности демодулятора  $N_b$  оказываются

произвольными. Следует также заметить, что равенства между размерностями сигналов на входе модулятора  $N$  и выходе демодулятора  $N'$ , а также между степенями нелинейности модулятора  $N_a$  и демодулятора  $N_b$  также являются не обязательными, то есть в общем случае  $N \neq N'$  и  $N_a \neq N_b$ . Следовательно, единственным ограничением на данные величины остается условие осуществимости нелинейной модуляции ( $N_a \leq N$ ).

Таким образом, базисные функции демодуляции являются функциями временных и пространственных входных координат, так как выходные координаты дискретны, и подобные функции имеют следующий вид:

$$\varphi'_{i,k}(t'_1, \dots, t'_i, \mathbf{r}'_1, \dots, \mathbf{r}'_i) = \varphi'_i(k, k \mathbf{1}_{n_b}, t'_1, \dots, t'_i, \mathbf{r}'_1, \dots, \mathbf{r}'_i), \quad i = \overline{1, N_a}, \quad k = \overline{1, N'}. \quad (22)$$

На основе этого выражения операция нелинейной демодуляции преобразуется к форме

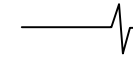
$$x'_k = \sum_{i=1}^{N_b} \int \dots \int \dots \int \left\{ \prod_{j=1}^i x'(t'_j, \mathbf{r}'_j) \right\} \times \varphi'_{i,k}(t'_1, \dots, t'_i, \mathbf{r}'_1, \dots, \mathbf{r}'_i) \times dt'_1 \dots dt'_i \dots dt'_i d\mathbf{r}'_1 \dots d\mathbf{r}'_i, \quad k = \overline{1, N'} \quad (23)$$

В результате число базисных функций демодуляции (аналог количества базисных функций модуляции  $N_\phi$ ) определяется как произведение размерности сигналов на выходе демодулятора  $N'$  на степень его нелинейности  $N_b$ :

$$N_{\phi'} = N' N_b. \quad (24)$$

Несмотря на их малое по сравнению с базисными функциями модуляции количество, имеющее линейную зависимость как от размерности сигналов на выходе демодулятора, так и от степени его нелинейности, сложности операций нелинейной модуляции и демодуляции оказываются сопоставимыми. В первую очередь данное обстоятельство обусловлено увеличенным количеством переменных в функциях демодуляции порядка нелинейности большего двух по сравнению с модуляционными.

Снижение числа базисных функций демодуляции вполне возможно, также вплоть до одной, поэтому (24) определяет верхнюю границу их количества. Однако в этом случае, подобно ситуации с нелинейным модулятором, сужается класс исследуемых операторов демодуляции. В то же время снижение сложности демодуляции допустимо на основе учета свойств входящих в оператор функционалов. Так, все данные функционалы являются  $\{i(n'+1)\}$  – линейными, поскольку по каждой из пространственных или временной координатам они являются линейными [3]. Кроме того, подобные функционалы оказываются также и однородными со степенью  $i$  по каждой из координат при фиксированных остальных, что предполагает возможность их представления с симметричными ядрами в рамках определенных координат [3]. В результате, функционалы рассматриваются как регулярные однородные, в которых регулярность следует трактовать как симметричность базисных функций  $\varphi'_{i,k}(t'_1, \dots, t'_i, \mathbf{r}'_1, \dots, \mathbf{r}'_i)$  относительно переменных



$t'_1, \dots, t'_i$ , описывающих временную координату, при всех  $\mathbf{r}'_1, \dots, \mathbf{r}'_i$ , описывающих пространственные координаты, а также относительно переменных  $r'_{j,1}, \dots, r'_{j,i}$ ,  $j = \overline{1, n'}$ , описывающих  $j$ -ю пространственную координату, при фиксированных  $t'_1, \dots, t'_i$ ,  $r'_{j,1}, \dots, r'_{j,i}$ ,  $j \neq j'$ ,  $j' = \overline{1, n'}$ ,  $\mathbf{r}'_1, \dots, \mathbf{r}'_i$ , описывающих временную и остальные пространственные координаты. Таким образом, сам оператор нелинейной демодуляции (23) образует функциональный полином степени  $N_b$  при нулевых начальных условиях [3]. Использование же свойства симметричности базисных функций демодуляции позволяет определять их вид только для части области определения, граница которой задается осью симметрии –  $t'_1 \leq t'_2 \leq \dots \leq t'_i$ ,  $\mathbf{r}'_1 \leq \mathbf{r}'_2 \leq \dots \leq \mathbf{r}'_i$ . Естественно, что вследствие однородности функционалов в (23) возможна произвольная нумерация как пространственных, так и временных переменных. Для оставшейся области определения значения базисных функций демодуляции вычисляются из условия симметрии

$$\varphi'_{i,k}(P'_{i,p}) = \varphi'_{i,k}(P'_{i,p'}), \quad p, p' = \overline{1, i!}, \quad (25)$$

где  $P'_{i,p} = \{t'_{j,p}, \mathbf{r}'_{j,p}\}$  – представляет собой уникальную перестановку исходного упорядоченного множества временных и пространственных переменных  $\{t'_{j,p}, \mathbf{r}'_{j,p}\}$ ,  $j = \overline{1, i!}$ .

### Структурные схемы нелинейных модулятора и демодулятора

Для представления структурной схемы дискретного канала связи, включающего нелинейные модулятор и демодулятор необходимо несколько видоизменить операторы нелинейных модуляции (6) и демодуляции (23). Так, оператор модуляции целесообразно рассматривать не как множество сумм произведений координат входных сигналов  $x_k$ ,  $k = \overline{1, N}$ , на базисные функции модуляции  $\varphi_{k_1, \dots, k_i}(t, \mathbf{r})$ ,  $i = \overline{1, Na}$ ,  $k_j = \overline{1, N}$ ,  $j = \overline{1, i}$ ,  $k_j < k_{j'}, |j < j'|$ ,  $j' = \overline{1, i!}$ , а как всего одну сумму произведений вектора различных упорядоченных произведений координат входных сигналов и вектор-функции упорядоченных базисных функций.

Первоначально следует рассмотреть случай линейной модуляции ( $i = 1$ ), для которого в операторе модуляции (6) сохраняется всего лишь одна сумма, достаточно просто представляемая в матричном виде

$$\sum_{k_1=1}^N x_{k_1} \varphi_{k_1}(t, \mathbf{r}) = \mathbf{x}_1^T \Phi_1(t, \mathbf{r}), \quad (26)$$

где  $\mathbf{x}_1 = \mathbf{x}$  – вектор сигнала на входе дискретного канала связи (на входе модулятора) размерности  $N$ ;  $\Phi_1(t, \mathbf{r}) = [\varphi_1(t, \mathbf{r}), \varphi_2(t, \mathbf{r}), \dots, \varphi_N(t, \mathbf{r})]^T$  – вектор базисных функций модуляции первого порядка.

Операцию модуляции исключительно второго поряд-

ка ( $i = 2$ ) целесообразно представить в форме рекуррентного соотношения

$$\begin{aligned} \sum_{k_1=1}^N x_{k_1} \sum_{k_2=k_1}^N x_{k_2} \varphi_{k, k_2}(t, \mathbf{r}) &= \\ &= \sum_{k_1=1}^N x_{k_1} \mathbf{x}_{1, k_1}^T \Phi_{2, k_1}(t, \mathbf{r}) = \mathbf{x}_{1, k_1}^T \Phi_2(t, \mathbf{r}), \end{aligned} \quad (27)$$

где  $\mathbf{x}_{1, k} = (x_k, x_{k+1}, \dots, x_N)^T$ ,  $k = \overline{1, N}$ , – усеченный вектор сигнала на входе модулятора размерности  $N-k+1$  (очевидно, что  $\mathbf{x}_1 = \mathbf{x}_{1,1} = \mathbf{x}$ , а  $\mathbf{x}_{1, N} = x_N$ );  $\mathbf{x}_2 = (x_1, \mathbf{x}_{1,1}^T x_{2,2}, \dots, x_N \mathbf{x}_{1, N}^T)^T$  – вектор произведений второго порядка сигнала на входе модулятора размерности  $C_{N+1}^2$ , согласно (12);

$\Phi_{2, k}(t, \mathbf{r}) = [\varphi_{k, k}(t, \mathbf{r}), \varphi_{k, k+1}(t, \mathbf{r}), \dots, \varphi_{k, N}(t, \mathbf{r})]^T$  – усеченный вектор базисных функций модуляции второго порядка (очевидно, что  $\Phi_{2, N}(t, \mathbf{r}) = \varphi_{N, N}(t, \mathbf{r})$ );  $\Phi_2(t, \mathbf{r}) = [\Phi_{2,1}^T(t, \mathbf{r}), \Phi_{2,2}^T(t, \mathbf{r}), \dots, \Phi_{2, N}^T(t, \mathbf{r})]^T$  – вектор базисных функций модуляции второго порядка.

Операция модуляции исключительно третьего порядка ( $i = 3$ ) также представима в форме рекуррентного соотношения

$$\begin{aligned} \sum_{k_1=1}^N \sum_{k_2=k_1}^N \sum_{k_3=k_2}^N \left\{ \prod_{j=1}^3 x_{k_j} \right\} \varphi_{k_1, k_2, k_3}(t, \mathbf{r}) &= \\ &= \sum_{k_1=1}^N x_{k_1} \sum_{k_2=k_1}^N x_{k_2} \mathbf{x}_{1, k_2}^T \Phi_{3, k_1, k_2}(t, \mathbf{r}) = \\ &= \sum_{k_1=1}^N x_{k_1} \mathbf{x}_{2, k_1}^T \Phi_{3, k_1}(t, \mathbf{r}) = \mathbf{x}_3^T \Phi_3(t, \mathbf{r}), \end{aligned} \quad (28)$$

где  $\mathbf{x}_{2, k} = (x_k, \mathbf{x}_{1, k}^T, x_{k+1} \mathbf{x}_{1, k+1}^T, \dots, x_N \mathbf{x}_{1, N}^T)^T$ ,  $k = \overline{1, N}$ , – усеченный вектор произведений второго порядка сигнала на входе модулятора;  $\mathbf{x}_3 = (x_1, \mathbf{x}_{2,1}^T, x_2 \mathbf{x}_{2,2}^T, \dots, x_N \mathbf{x}_{2, N}^T)^T$  – вектор произведений третьего порядка сигнала на входе модулятора размерности  $C_{N+2}^3$ , согласно (12);

$\Phi_{3, k_1, k_2}(t, \mathbf{r}) = [\varphi_{k_1, k_2, k_2}(t, \mathbf{r}), \varphi_{k_1, k_2, k_2+1}(t, \mathbf{r}), \dots, \varphi_{k_1, k_2, N}(t, \mathbf{r})]^T$  – усеченный по третьему индексу вектор базисных функций модуляции третьего порядка;  $\Phi_{3, k_1}(t, \mathbf{r}) = [\Phi_{3, k_1, k_1}^T(t, \mathbf{r}), \Phi_{3, k_1, k_1+1}^T(t, \mathbf{r}), \dots, \Phi_{3, k_1, N}^T(t, \mathbf{r})]^T$  – усеченный по второму индексу вектор базисных функций модуляции третьего порядка;  $\Phi_3(t, \mathbf{r}) = [\Phi_{3,1}^T(t, \mathbf{r}), \Phi_{3,2}^T(t, \mathbf{r}), \dots, \Phi_{3, N}^T(t, \mathbf{r})]^T$  – вектор базисных функций модуляции третьего порядка.

Далее по индукции операция модуляции исключительно  $i$ -го порядка имеет рекуррентный вид

$$\begin{aligned} \sum_{k_1=1}^N \sum_{k_2=k_1}^N \dots \sum_{k_i=k_{i-1}}^N \left\{ \prod_{j=1}^i x_{k_j} \right\} \varphi_{k_1, \dots, k_i}(t, \mathbf{r}) &= \mathbf{x}_i^T \Phi_i(t, \mathbf{r}), \\ i = \overline{1, Na}, \end{aligned} \quad (29)$$



где  $\mathbf{x}_i = (x_1 \mathbf{x}_{i-1}^T, x_2 \mathbf{x}_{i-1}^T, \dots, x_N \mathbf{x}_{i-1}^T)^T$  – вектор произведений  $i$ -го порядка входных сигналов, каждый элемент которого является произведением упорядоченных по возрастанию координат входного сигнала, размерность данного вектора определяется на основе (12);  $\Phi_i(t, \mathbf{r}) = [\Phi_{i,1}^T(t, \mathbf{r}), \Phi_{i,2}^T(t, \mathbf{r}), \dots, \Phi_{i,N}^T(t, \mathbf{r})]^T$  – вектор базисных функций модуляции  $i$ -го порядка.

Таким образом, оператор нелинейной модуляции порядка  $N_a$  (6) представим в виде суммы операций модуляции меньших порядков (26)–(29):

$$x(t, \mathbf{r}) \sum_{i=1}^{N_a} \mathbf{x}_i^T \Phi_i(t, \mathbf{r}). \quad (30)$$

Следовательно, структурную схему модулятора целесообразно представить в виде блоков, каждый из которых выполняет исключительно операции модуляции некоторого порядка. Между же блоками модуляции следует разместить устройства коммутации, формирующие вектора произведений соответствующего порядка сигнала на входе модулятора. В результате структурная схема нелинейного модулятора имеет вид, изображенный на рис 3.

Оператор демодуляции также следует разбивать на составляющие, выполняющие операции демодуляции определенного порядка. Так, операция демодуляции первого порядка на основе свойства линейности операций интегрирования и суммирования представима в виде

$$x'_k = \int \int x'(t'_1, \mathbf{r}'_1) \{ \varphi'_{1,k}(t'_1, \mathbf{r}'_1) + x'_k(t'_1, \mathbf{r}'_1) \} dt'_1 d\mathbf{r}'_1, \quad (31)$$

$k_j = \overline{1, N'}$

где  $x'_k(t'_1, \mathbf{r}'_1) = \sum_{i=2}^{N_b} \int \dots \int \dots \int \left\{ \prod_{j=2}^i x'(t'_j, \mathbf{r}'_j) \right\} \times \varphi'_{1,k}(t'_1, \dots, t'_i, \mathbf{r}'_1, \dots, \mathbf{r}'_i) dt'_2, \dots, dt'_i d\mathbf{r}'_2, \dots, d\mathbf{r}'_i$  – сигнал на выходе блока демодуляции второго порядка.

Операция демодуляции второго порядка имеет подобную рекуррентную форму

$$x'_k(t'_1, \mathbf{r}'_1) = \int \int x'(t'_2, \mathbf{r}'_2) \{ \varphi'_{2,k}(t'_1, t'_2, \mathbf{r}'_1, \mathbf{r}'_2) + x'_k(t'_1, t'_2, \mathbf{r}'_1, \mathbf{r}'_2) \} dt'_2 d\mathbf{r}'_2 \quad (32)$$

где  $x'_k(t'_1, \mathbf{r}'_1) = \sum_{i=3}^{N_b} \int \dots \int \dots \int \left\{ \prod_{j=3}^i x'(t'_j, \mathbf{r}'_j) \right\} \times \varphi'_{1,k}(t'_1, \dots, t'_i, \mathbf{r}'_1, \dots, \mathbf{r}'_i) dt'_3, \dots, dt'_i d\mathbf{r}'_3, \dots, d\mathbf{r}'_i$  – сигнал на выходе блока демодуляции третьего порядка.

По индукции операция демодуляции  $m$ -го порядка представима в рекуррентном виде

$$x'_k(t'_1, \dots, t'_{m-1}, \mathbf{r}'_1, \dots, \mathbf{r}'_{m-1}) = \int \int x'(t'_m, \mathbf{r}'_m) \times \{ \varphi'_{m,k}(t'_1, \dots, t'_m, \mathbf{r}'_1, \dots, \mathbf{r}'_m) + x'_k(t'_1, \dots, t'_m, \mathbf{r}'_1, \dots, \mathbf{r}'_m) \}, \quad m = \overline{1, N_b}, \quad (33)$$

где  $x'_k(t'_1, \dots, t'_m, \mathbf{r}'_1, \dots, \mathbf{r}'_m) = \sum_{i=m+1}^{N_b} \int \dots \int \dots \int \left\{ \prod_{j=m+1}^i x'(t'_j, \mathbf{r}'_j) \right\} \times \varphi'_{i,k}(t'_1, \dots, t'_i, \mathbf{r}'_1, \dots, \mathbf{r}'_i) \times dt'_{m+1} \dots dt'_i d\mathbf{r}'_{m+1} \dots d\mathbf{r}'_i$  – сигнал на выходе блока демодуляции  $m$ -го порядка.

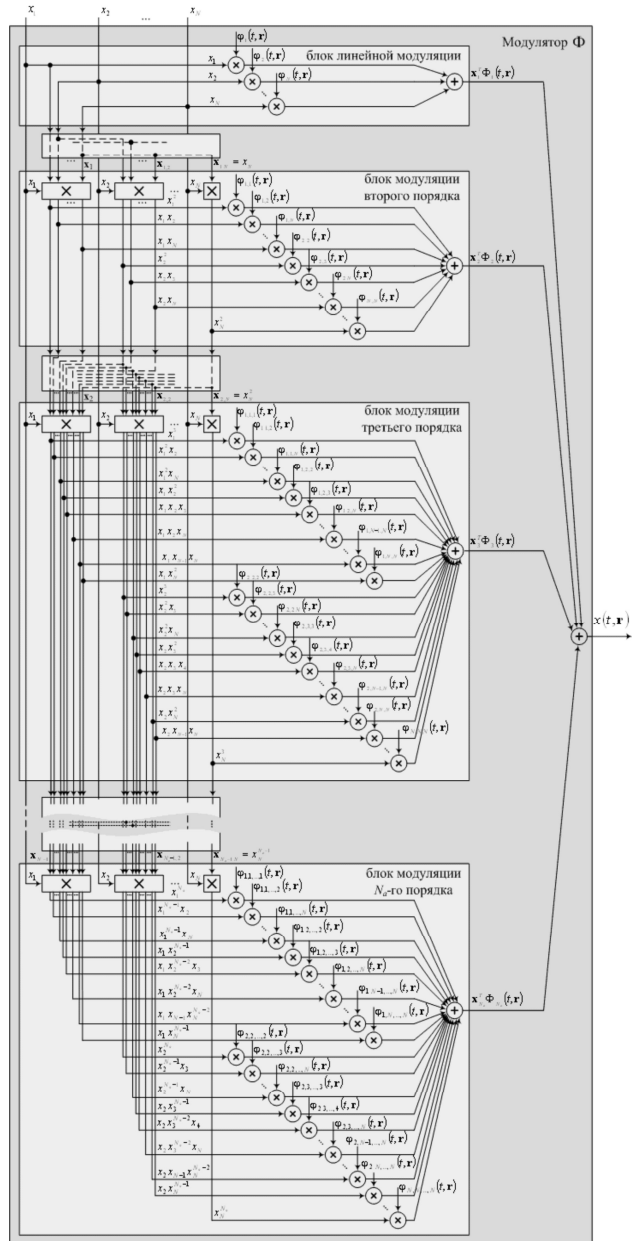


Рис. 3. Структурная схема нелинейного модулятора

Также, по индукции, операция демодуляции  $N_b$ -го порядка задается в форме интеграла произведения входного сигнала и базисной функции демодуляции  $N_b$ -го порядка

$$\begin{aligned}
 & x'_k(t'_1, \dots, t'_{N_b-1}, \mathbf{r}'_1, \dots, \mathbf{r}'_{N_b-1}) = \\
 & = \int_{t'_{N_b}} \int_{\mathbf{r}'_{N_b}} x'(t'_{N_b}, \mathbf{r}'_{N_b}) \times \\
 & \varphi'_{N_b, k}(t'_1, \dots, t'_{N_b}, \mathbf{r}'_1, \dots, \mathbf{r}'_{N_b}) dt'_{N_b} d\mathbf{r}'_{N_b}
 \end{aligned} \quad (34)$$

Таким образом, структурную схему нелинейного демодулятора целесообразно рассматривать как соединение блоков демодуляции определенного порядка, каждый выход которых является входом блоков меньшего на единицу порядка нелинейности. В результате структурная схема нелинейного демодулятора принимает вид, изображенный на рис. 4.

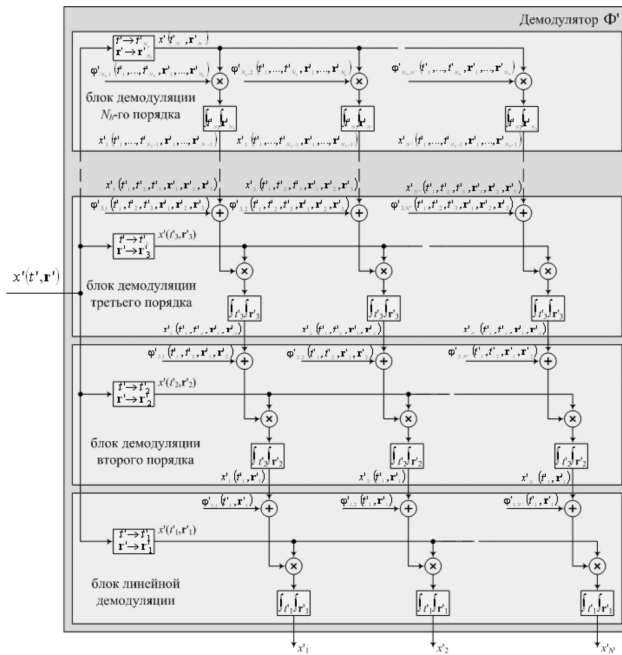


Рис. 4. Структурная схема нелинейного демодулятора

### Аналитическое моделирование нелинейных схем обработки сигналов

В качестве примера и для демонстрации технического эффекта от введения нелинейных схем обработки сигналов в ходе исследований было проведено аналитическое моделирование (с использованием программной среды MathCad) модулятора и демодулятора с нелинейностями второго порядка ( $N_a = N_b = 2$ ) двумерных сигналов ( $N = N' = 2$ ), применительно к каналу с аддитивным белым гауссовским шумом. В качестве критерия оптимального синтеза системы обработки был использован минимум среднеквадратического отклонения передаваемого сигнала на входе модулятора относительно принятого на выходе демодулятора  $\sigma^2$  [3]. Разложение базисных функций модуляции и демодуляции (ядер Вольтерра) осуществлялось в ряд Котельникова. Рассматривалось три центрированных источника (двумерные распределения на входе модулятора) – равномерно распределенное трехточечное созвездие (несимметричное относительно осей координат), равномерно распределенное четырехточечное созвездие (соответствующее сигналам четырехточечной квадратурной амплитудной модуляции – КАМ-4) и гауссовский источник. Результаты аналитического моделирования представлены на рис. 5.

В случае равномерно распределенного трехточечного созвездия (рис. 5,а) применение нелинейной демодуляции оказывается достаточным для получения выигрыша относительной линейной схемы обработки порядка 13 дБ. Добавление нелинейного модулятора дополнительно повышает энергетический эффект на 0,5 дБ. Для созвездия же КАМ-4 ситуация несколько иная – только совместное использование нелинейного модулятора и демодулятора приводит к снижению энергетических затрат, также примерно на 13,5 дБ (рис. 5, б). Подобное различие вызвано тем, что все моменты нечетного порядка у сигналов КАМ-4 равны нулю, в то время как у равномерно распределенного трехточечного созвездия данные моменты отличны от нуля. В результате, нелинейный демодулятор даже с линейным модулятором может использовать дополнительную информацию о высших порядках распределения, а соответственно и снижать ошибку оценки.

У гауссовского распределения нечетные моменты также равны нулю, однако в отличие от сигналов КАМ-4, значения четных моментов таковы, что не позволяют осуществить коррекцию среднеквадратического отклонения при введении нелинейного демодулятора (рис. 5, в). Следует отметить, что применение нелинейного модулятора совместно с линейным демодулятором не приводит к энергетическому выигрышу, поскольку в данном случае демодулятор способен обрабатывать лишь информацию о максимально вторых моментах распределений, а нелинейный модулятор добавляет информацию о более высоких моментах, перераспределяя часть энергии с линейного преобразователя.

В результате, применение нелинейных модулятора и демодулятора позволяет снизить энергетические затраты, в случае равномерно распределенного трехточечного созвездия и созвездия КАМ-4, на величину порядка 13,5 дБ по сравнению с линейными системами обработки. В случае же гауссовского источника подобное повышение энергетического эффекта не представляется возможным, что согласуется с результатами нелинейной фильтрации [3].

### Заключение

Таким образом, в наиболее обобщенном структурированном виде уравнения (6) и (23), являющиеся, по сути, операторами нелинейной модуляции  $\Phi$  и демодуляции  $\Phi'$  соответственно, задают дискретное отображение непрерывного многопараметрического канала связи. Выражения же (30) и (31)-(34) делают возможным представление данных операций в виде структурных схем, реализующих последовательно процедуры модуляции и демодуляции от низкого порядка нелинейности к более высокому. Следует подчеркнуть еще одно немаловажное достоинство именно структурного типа дискретного отображения. Так, в наиболее общем виде подобное отображение предполагает учет нелинейности произвольно большой степени, вплоть до бесконечной. В результате и реализация и синтез оптимальных операций как модуляции, так и демодуляции оказываются весьма затруднительными, за исключением случаев определенной заданной формы используемых базисных функций.

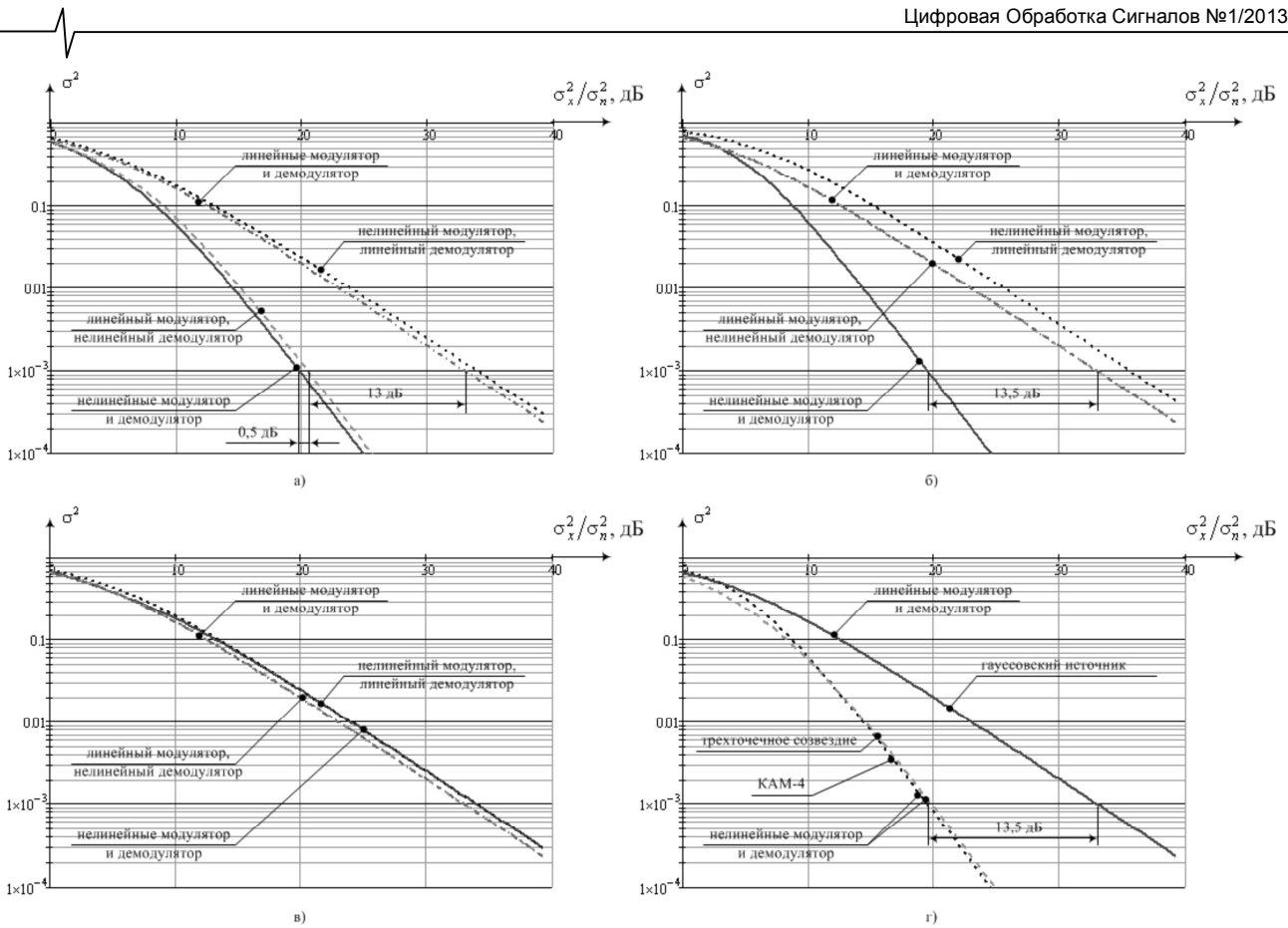


Рис. 5. Зависимости минимума среднеквадратического отклонения передаваемого сигнала на входе модулятора относительно принятого на выходе демодулятора  $\sigma^2$  от отношения сигнал–шум  $\sigma_x^2/\sigma_n^2$  в канале связи с аддитивным белым гауссовским шумом для различных источников: а) равномерно распределенное трехточечное созвездие; б) КАМ-4; в) гауссовский источник; г) оптимальные системы обработки для всех видов источников

Однако в данное обстоятельство является достаточно сильным ограничением, поскольку способно существенно сократить класс операторов отображения. Представление же операций в параметризованном виде (6) и (23) делает возможным как выбор заданного числа базисных функций в рамках определенного порядка нелинейности, так и введение ограничения на максимальную нелинейность оператора дискретного отображения. В итоге, решение задач синтеза оптимальных модулятора и демодулятора возможно как в общем виде (при устремлении степени нелинейности к бесконечности), что должно быть подобно оптимальным нелинейным системам, как правило трудно реализуемым на практике, так и при достаточно произвольных ограничениях на число и порядок нелинейности базисных функций, что приводит к сравнительно простым реализациям вышеописанных процедур.

**Литература**

1. Бронштейн И.Н., Семендяев К.А. Справочник по математике для инженеров и учащихся втузов. 13-е изд., исправленное. – М.: Наука. Главная редакция физико-математической литературы, 1986. – 544 с.

2. Градштейн И.С., Рыжик И.М. Таблицы интегралов, сумм, рядов и произведений. – М.: Наука, 1963. – 1 100 с.  
 3. Пупков К.А., Капалин В.И., Ющенко А.С. Функциональные ряды в теории нелинейных систем. – М.: Наука, 1976. – 448 с.  
 4. Репин В.Г., Тартаковский Г.П. Статистический синтез при априорной неопределенности и адаптация информационных систем. – М.: Советское радио, 1977. – 432 с.  
 5. Репин В.Г., Тартаковский Г.П., Слепокуров А.А. Проверка многоальтернативных гипотез в условиях априорной неопределенности // Проблемы передачи информации. 1991. Т. 27. Вып. 3. С. 39–56.

**MATHEMATICAL MODELS OF MODULATOR AND DEMODULATOR WITH NONLINEAR ORDER PRESCRIBED**

*Batenkov K.A.*

Mathematical models of modulator and demodulator with nonlinear order prescribed are synthesized on basis of functional series. Signal processing complexity evaluation is analyzed at level of basis function quantity.

## МЕТОД НАИМЕНЬШИХ ВЗВЕШЕННЫХ КВАДРАТОВ ДЛЯ АНАЛИЗА ИЗОБРАЖЕНИЙ

*Пиманкин Д.А., аспирант (ассистент) кафедры «Компьютерные технологии в проектировании и производстве» Нижегородского государственного технического университета им. Р.Е. Алексеева, e-mail: denis@pimankin.com*

**Ключевые слова:** метод наименьших взвешенных квадратов, ортогональное преобразование, признаковое описание изображения, сопоставление ключевых точек.

### Введение

Для успешного решения целого ряда задач компьютерного зрения и распознавания образов необходимо осуществление отображения исходного множества измерений в новое множество признаков. На сегодняшний день проблема генерации признаков, особенно на основе нелинейных преобразований, остается открытой. Целью такой генерации является сокращение информации об анализируемом объекте до «значимой», необходимой для решения конкретной задачи.

Под извлечением признака будем понимать отображение  $f: P \rightarrow Q_f$ , где  $P$  – множество измерений (множество всех изображений),  $Q_f$  – множество допустимых значений признака. Достаточно хорошо изучены методы извлечения на основе различных ортогональных преобразований, таких как дискретное преобразование Фурье (ДПФ), дискретное косинусное преобразование (ДКП) и др., входного сигнала (изображения).

Количество признаков стремятся уменьшать по ряду причин. Одна из них состоит в следующем. Часто признаковое описание изображения состоит в представлении его в виде точки в  $N$ -мерном метрическом пространстве. Для эффективного поиска по большой базе данных используются специфические структуры данных, например  $M$ -дерева. С увеличением размерности пространства скорость обработки запросов типа ближайший сосед или  $k$ -ближайших соседей сокращается [3, 4].

### Извлечение признаков на основе ортогональных преобразований

Известным методом анализа сигнала является его анализ в частотной области [1].

Рассмотрим некоторое ортогональное преобразование входного сигнала в самом общем случае. Пусть  $x_n$  – анализируемый сигнал. Разложим  $x_n$  по базисным векторам  $b_{k,n}$ . Если сигнал состоит из  $N$  отсчетов, то он может быть полностью описан  $N$  коэффициентами разложения. Иначе говоря,  $x_n$  однозначно восстановим по  $N$  коэффициентам. Пусть базис  $B$  ортонормирован. Т. е.  $\|b_n\| = 1$  и  $\langle b_m, b_n \rangle = 0$  ( $m \neq n$ ). Тогда проекцию сигнала на  $k$ -й базисный вектор можно записать следующим образом:

*Предлагается метод построения признакового описания областей изображений на основе анализа ориентаций векторов градиента изображений с помощью метода наименьших взвешенных квадратов (метод WLSF). Показано, что для широкого класса изображений предлагаемый подход может быть с успехом применен для решения задачи сопоставления ключевых точек. Показано преимущество метода WLSF перед методом SURF при анализе зашумленных изображений.*

$$X_k = \sum_{n=0}^{N-1} x_n b_{k,n}. \quad (1)$$

Очевидно, что при выше описанных условиях проекцию  $X_k$  можно представить как:

$$X_k = \arg \min_{X_k} \left( \sum_{n=0}^{N-1} (x_n - X_k b_{k,n})^2 \right). \quad (2)$$

То есть при разложении сигнала по ортогональному базису минимизируется сумма квадратов разностей компонент сигнала и базисного вектора, умноженного на коэффициент разложения.

В качестве компонент искомого вектора признаков обычно используются  $K$  низкочастотных коэффициентов разложения.

### Метод наименьших взвешенных квадратов

При обычном ортогональном разложении сигнала все отсчеты входят в преобразование с одинаковыми коэффициентами, т. е. все отсчеты равнозначны. Идеей предлагаемого метода является построение такого преобразования, в котором различные отсчеты входили бы с различными коэффициентами [1]. Наибольший вес должен быть у наиболее информативных участков сигнала, наименьший – у наименее информативных.

Зависимость величины вклада некоторого участка сигнала от информативности этого участка достигается тем, что каждый из коэффициентов разложения  $X_k$  находится путем минимизации не суммы квадратов отклонений отсчетов сигнала  $x_n$  и базисного вектора  $b_{k,n}$ , а путем минимизации суммы взвешенных квадратов отклонений – методом наименьших взвешенных квадратов (МНВК):

$$X_k = \arg \min_{X_k} \left( \sum_{n=0}^{N-1} w_n (x_n - X_k b_{k,n})^2 \right). \quad (3)$$

Рис. 1 иллюстрирует проекцию двумерного сигнала на базисный вектор при одинаковых и разных ( $w_0 < w_1$ ) весах компонент  $x_1$  и  $x_2$ .



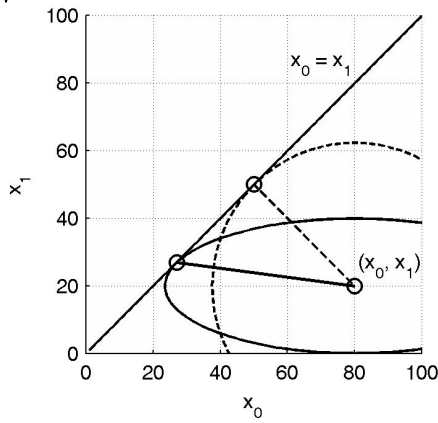


Рис. 1. Иллюстрация расстояния в МНК (пунктирная линия) и МНБК (сплошная линия)

Проанализируем, является ли преобразование (3) обратимым. Распишем выражение (3):

$$X_k = \arg \min_{X_k} \left( L^2 \sum_{n=0}^{N-1} \left( \frac{x_n \sqrt{w_n}}{L} - X_k \cdot \frac{b_{k,n} \sqrt{w_n}}{L} \right)^2 \right) = \quad (4)$$

$$= \frac{1}{L} \sum_{n=0}^{N-1} b_{k,n} x_n \sqrt{w_n},$$

где:

$$L^2 = \sum_{n=0}^{N-1} (x_n \sqrt{w_n})^2. \quad (5)$$

Из (4) и условий, накладываемых на  $b_{k,n}$ , следует, что в общем случае преобразование не является обратимым. Невозможность получения исходного сигнала по его отображению возникает при равенстве хотя бы одного весового коэффициента нулю. В остальных случаях преобразование является обратимым.

Распишем метод наименьших взвешенных квадратов для суммы первых двух функций Хаара (Уолша). В данном случае будем минимизировать сумму взвешенных квадратов отклонений анализируемого сигнала и суммы базисных векторов, умноженных на коэффициенты разложения  $X_0$  и  $X_1$ . Обозначим  $B_{0,n} = X_0 \cdot |b_{0,n}|$  и  $B_{1,n} = X_1 \cdot |b_{1,n}|$ .

$$\sum_{n=0}^{N/2-1} w_n (x_n - B_{0,n} - B_{1,n})^2 + \quad (6)$$

$$+ \sum_{n=N/2}^{N-1} w_n (x_n - B_{0,n} + B_{1,n})^2 \rightarrow \min.$$

Продифференцируем левую часть выражения (6) по  $B_{0,n}$ :

$$-2 \sum_{n=0}^{N/2-1} w_n (x_n - B_{0,n} - B_{1,n}) - \quad (7)$$

$$-2 \sum_{n=N/2}^{N-1} w_n (x_n - B_{0,n} + B_{1,n}) = 0,$$

после раскрытия скобок имеем

$$B_{0,n} \sum_{n=0}^{N-1} w_n + B_{1,n} \left( \sum_{n=0}^{N/2-1} w_n - \sum_{n=N/2}^{N-1} w_n \right) = \sum_{n=0}^{N-1} w_n x_n. \quad (8)$$

Продифференцируем левую часть выражения (6) по  $B_{1,n}$ :

$$-2 \sum_{n=0}^{N/2-1} w_n (x_n - B_{0,n} - B_{1,n}) + \quad (9)$$

$$+ 2 \sum_{n=N/2}^{N-1} w_n (x_n - B_{0,n} + B_{1,n}) = 0,$$

после раскрытия скобок имеем

$$B_{0,n} \left( \sum_{n=0}^{N/2-1} w_n - \sum_{n=N/2}^{N-1} w_n \right) + B_{1,n} \sum_{n=0}^{N-1} w_n = \quad (10)$$

$$= \sum_{n=0}^{N/2-1} w_n x_n - \sum_{n=N/2}^{N-1} w_n x_n.$$

Из (8) и (10) видно, что  $X_0$  и  $X_1$  для МНБК (в отличие от МНК) независимы только в случае, если

$$\sum_{n=0}^{N/2-1} w_n = \sum_{n=N/2}^{N-1} w_n. \quad (11)$$

То есть в общем случае минимум СКО сигнала и каждого из векторов  $X_k \cdot b_{k,n}$  по отдельности не соответствует минимуму СКО сигнала и суммы этих векторов. Этот вывод важен для понимания последующих рассуждений.

На рис. 2 показан пример дискретного сигнала, его некоторой весовой функции  $w_n$  и его аппроксимации первыми двумя функциями Хаара (Уолша), умноженными на коэффициенты разложения по формулам (2) и (3).

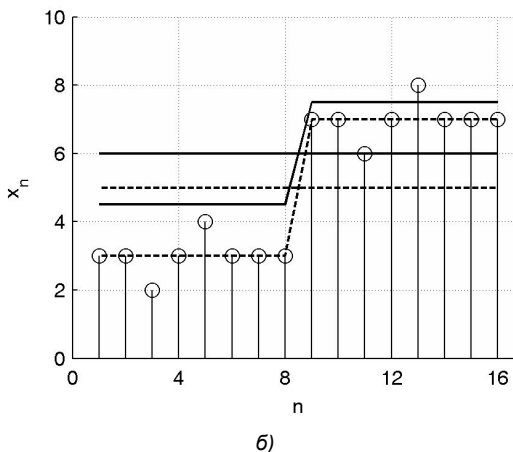
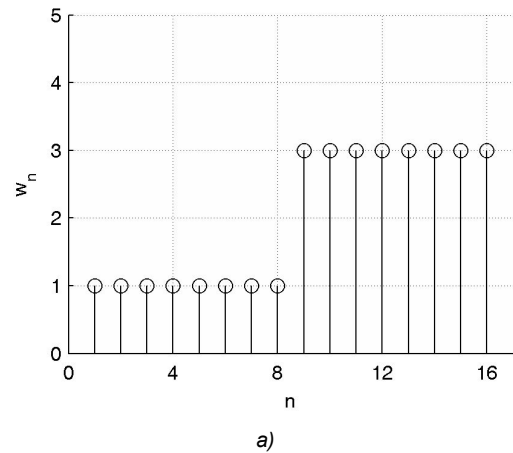
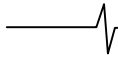


Рис. 2. а) график  $w_n$ ; б) график  $x_n$  и его разложение МНК (пунктирная линия) и последовательным МНБК (сплошная линия)



При разложении МНВК, перед нахождением  $X_1$ , из анализируемого сигнала вычитается  $X_0 \cdot b_{k,n}$  (последовательная схема разложения). Очевидно, что при разложении в соответствии с (2) вычитание излишне. Здесь  $w_n = 1$  при  $n = 1, 2, \dots, 8$ ;  $w_n = 3$  при  $n = 9, 10, \dots, 16$ .

### Применение МНВК для анализа изображений

Определим изображение как двумерную функцию  $f_{m,n}$  пространственных координат  $m$  и  $n$ . Значение функции в каждой точке, задаваемой парой координат, соответствует интенсивности изображения в этой точке. В цифровом представлении координаты  $m$  и  $n$  принимают дискретные значения:  $m = 0, 1, \dots, H - 1$ ,  $n = 0, 1, \dots, W - 1$ .  $W$  и  $H$  – ширина и высота изображения в пикселях, соответственно.

Для функции изображения в некоторой точке  $(m, n)$  можно определить градиент как двумерный вектор-столбец. Компонентами этого вектора являются частные производные функции  $f_{m,n}$  по  $m$  и  $n$ . Значение вектора градиента в каждой точке можно представить в виде комплексного числа  $g_{m,n} = gn_{m,n} + j \cdot gm_{m,n}$ . Здесь  $gm$  соответствует частной производной вдоль оси  $m$ , а  $gn$  – частной производной вдоль оси  $n$ . Таким образом, функции  $f_{m,n}$  соответствует комплексная функция  $g_{m,n}$ . Обозначим соответствующую функции  $g_{m,n}$  комплексную матрицу градиентов как  $G$ . Операции пространственного дифференцирования в цифровой обработке изображений осуществляются при помощи сверток с дифференцирующими масками. Очевидно, что если размер дифференцирующей маски  $S \times S$ , то размер матрицы  $G$  –  $(H - S + 1) \times (W - S + 1)$ . Значение градиента в каждой точке можно также записать через его длину и угол (действительные функции двух переменных  $a_{m,n}$  и  $\varphi_{m,n}$ ).

Для простоты анализа эффекты перекрытия при движении дифференцирующих масок учитываться в дальнейшем не будут. На рис. 3 приведено тестовое изображение и визуализация амплитуды поля градиента.

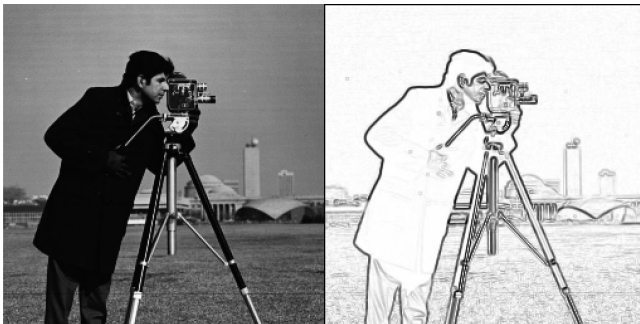


Рис 3. Тестовое изображение (слева), визуализация амплитуды градиента (справа)

Разделим комплексную функцию значений градиента изображения от координат на «сигнальную» составляющую, в качестве которой будем использовать  $\varphi_{m,n}$ , и «весовую» составляющую, в качестве которой будем использовать  $a_{m,n}$ . Идея анализа ориентаций градиентов для выделения дескрипторов областей изображений не нова, различные подходы описаны в работах [2, 5-8].

Теперь необходимо, во-первых, адаптировать МНВК для разложения аргумента градиента, т.е. для разложения сигнала с областью значений  $[-\pi, \pi)$ , а во-вторых,

т.к. полутоновое изображение является двумерным сигналом, описанное выше преобразование обобщить на двумерный случай.

Попытаемся построить некоторое отображение функции  $g_{m,n}$ , компактно описывающее распределение ориентаций градиента в пределах исследуемой области изображения. При этом должна быть учтена и амплитуда градиента: подобласти с высоким значением амплитуды градиента должны вносить больший вклад, чем подобласти с малым значением.

Для начала рассмотрим одномерный сигнал  $x_n$ .

Будем искать выходной вектор искомого отображения в виде

$$X_k = A_k \cdot \exp(j\Phi_k). \quad (12)$$

Внесем несколько замечаний по поводу используемых в дальнейшем обозначений. Здесь и далее будем обозначать  $a_k = |x_k|$ ,  $\varphi_k = \arg(x_k)$ ,  $A_k = |X_k|$ ,  $\Phi_k = \arg(X_k)$ . Также будем считать каждое комплексное число  $x_k$  эквивалентным вектору:

$$\mathbf{v}_k = (\Re(x_k), \Im(x_k)). \quad (13)$$

В качестве базисных векторов могут быть использованы базисные вектора ДКП, дискретного преобразования Уолша и др. Также можно использовать метод главных компонент. Наиболее простыми, с точки зрения обоснования, являются функции Хаара, их и будем использовать. На рис. 4 (слева) показаны графики первых 4-х функций Хаара.

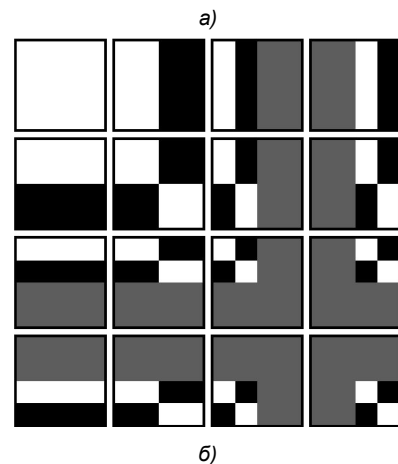
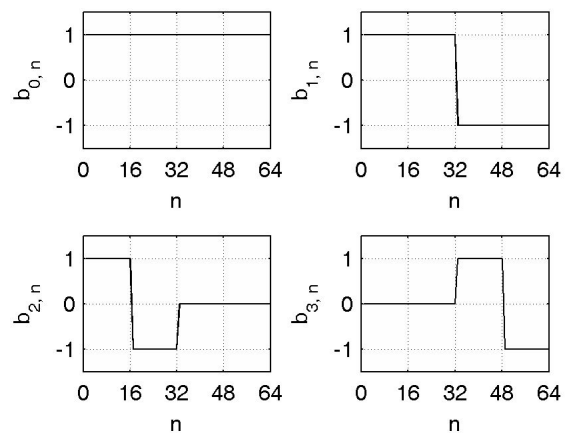


Рис 4. а) графики первых 4-х функций Хаара; б) двумерный вариант графиков первых 4-х функций Хаара

**Найдем аргумент  $\Phi_k$ .** Для начала необходимо определение адекватной метрики для множества комплексных чисел с единичным модулем. Воспользуемся представлением комплексных чисел в виде векторов в декартовой системе координат и определим расстояние между двумя числами, через скалярное произведение. Найдем такое значение угла вектора, при котором максимизируется среднее значение скалярных произведений анализируемых векторов и единичного вектора

$$\mathbf{u} = \arg \max_{\mathbf{u}} \left( \frac{1}{N} \sum_{n=0}^{N-1} \langle \mathbf{v}_n, \mathbf{u} \rangle \right). \quad (14)$$

Приведем (14) к виду формулы (3):

$$\begin{aligned} \theta &= \arg \min_{\theta} \left( \sum_{n=0}^{N-1} a_n (1 - \cos(\varphi_n - \theta)) \right) = \\ &= \arg \min_{\theta} \left( \sum_{n=0}^{N-1} a_n (\sqrt{1 - \cos(\varphi_n - \theta)})^2 \right), \end{aligned} \quad (15)$$

где  $\theta = \arg(\mathbf{u})$ .

Тогда выражение для расстояния между двумя числами (метрики) запишется следующим образом:

$$\begin{aligned} D(\varphi_0, \varphi_1) &= \sqrt{1 - \cos(\varphi_1 - \varphi_0)} = \\ &= \sqrt{2} \cdot \sin\left(\frac{\varphi_1 - \varphi_0}{2}\right). \end{aligned} \quad (16)$$

Для любого  $K \leq N$  можно найти вектор  $\Phi^K = (\Phi_{K-1,0}, \Phi_{K-1,1}, \dots, \Phi_{K-1,K-1})$ :

$$\Phi^K = \arg \min_{\Phi^K} \left( \sum_{n=0}^{N-1} a_n D^2(\varphi_n, s_n(\Phi^K)) \right), \quad (17)$$

где

$$s_n(\Phi^K) = \sum_{k=0}^{K-1} \text{haar}_{k,n} \Phi_{K-1,k}, \quad (18)$$

$\text{haar}$  – функция Хаара с множеством значений  $\{0, 1, -1\}$ .

Введем для удобства следующие величины:

$$z_k^+ = \frac{\sum_{n=0}^{N-1} x_n (\text{sgn}(\text{haar}_{k,n} - 1) + 1)}{\sum_{n=0}^{N-1} \text{sgn}(\text{haar}_{k,n} - 1) + 1}, \quad (19)$$

$k = 0, 1, \dots, K-1$ ;

$$z_k^- = \frac{\sum_{n=0}^{N-1} x_n (\text{sgn}(\text{haar}_{k,n} + 1) - 1)}{\sum_{n=0}^{N-1} \text{sgn}(\text{haar}_{k,n} + 1) - 1}, \quad (20)$$

$k = 1, 2, \dots, K-1$ .

Из (17) нас в первую очередь будут интересовать  $\Phi_{k,k}$ . Аргумент  $\Phi_{0,0}$  может быть расписан следующим образом:

$$\Phi_0 = \Phi_{0,0} = \arg(z_0^+). \quad (21)$$

Аргумент  $\Phi_{k,k}$ ,  $k \neq 0$ , запишется следующим образом:

$$\Phi_k = \Phi_{k,k} = \arg(z_k^+) - \arg(z_k^-). \quad (22)$$

**Найдем модуль  $A_k$ .** Введем вектор  $\Psi^K = (\Psi_{K-1,0}, \Psi_{K-1,1}, \dots, \Psi_{K-1,K-2})$ :

$$\begin{aligned} \Psi^K &= \\ &= \arg \min_{\Psi^K} \left( \sum_{n=0}^{n-1} a_n D^2(\varphi_n, s_n((\Psi^K \Phi_{K-1,K-1} + \Delta\phi))) \right). \end{aligned} \quad (23)$$

Теперь определим модуль  $A_k$  как

$$\begin{aligned} A_k &= \\ &= \lim_{\Delta\phi \rightarrow 0} \left( \frac{d \left( \frac{1}{N} \sum_{n=0}^{N-1} a_n D^2(\varphi_n, s_n((\Psi^k \Phi_{k-1,k-1} + \Delta\phi))) \right)}{d(\Delta\phi)} \right) \times \\ &\times \frac{d(\Delta\phi)}{d(D(0, \Delta\phi))}. \end{aligned} \quad (24)$$

Из формулы (22) следует, что при расчете  $\Phi_k$  сигнал может быть эквивалентным образом представлен двумя комплексными числами. Это также следует из области значений функций Хаара. Найдем для двух произвольных комплексных чисел коэффициенты  $\Phi_0$  и  $\Phi_1$ . Изменим  $\Phi_1$  на бесконечно малую величину  $\Delta\phi$  и зафиксируем. Найдем новое значение для  $\Phi_0$  выполнением условия (17) при условии неизменности  $\Phi_1$ . Тогда  $\Delta\varphi_k = \Delta\varphi_k^+ + \Delta\varphi_k^-$ . Зафиксируем  $\Delta\varphi_k$ . Из (16) и (23) следует

$$\begin{aligned} &a_k^+ \cos(\Delta\varphi_k^+) + a_k^- \cos(\Delta\varphi_k^-) = \\ &= a_k^+ \cos(\Delta\varphi_k^+) + a_k^- \cos(\Delta\varphi_k - \Delta\varphi_k^+) \rightarrow \max \end{aligned} \quad (25)$$

Продифференцировав по  $\Delta\varphi_k^+$ , получим

$$-a_k^+ \sin(\Delta\varphi_k^+) + a_k^- \sin(\Delta\varphi_k - \Delta\varphi_k^+) = 0. \quad (26)$$

Так как  $\Delta\varphi_k$  бесконечно мало, то в пределе получим

$$\begin{aligned} &\lim_{\Delta\varphi \rightarrow 0} \left( -a_k^+ \sin(\Delta\varphi_k^+) + a_k^- \sin(\Delta\varphi_k - \Delta\varphi_k^+) \right) = \\ &= -a_k^+ \Delta\varphi_k^+ + a_k^- (\Delta\varphi_k - \Delta\varphi_k^+) = 0 \end{aligned} \quad (27)$$

Отсюда и из условия  $\Delta\varphi_k = \Delta\varphi_k^+ + \Delta\varphi_k^-$  получаем выражения для  $\Delta\varphi_k^+$  через  $\Delta\varphi_k$ :

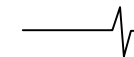
$$\Delta\varphi_k^+ = \Delta\varphi_k \cdot \frac{a_k^-}{a_k^+ + a_k^-}, \quad (28)$$

аналогично для  $\Delta\varphi_k^-$  получим

$$\Delta\varphi_k^- = \Delta\varphi_k \cdot \frac{a_k^+}{a_k^+ + a_k^-}. \quad (29)$$

Теперь подставим  $\Delta\varphi_k^+$  и  $\Delta\varphi_k^-$  в левую часть (25) и продифференцируем по  $\Delta\varphi_k$ . В пределе получим

$$\begin{aligned} &\lim_{\Delta\varphi_k \rightarrow 0} \left( -a_k^+ \cdot \frac{a_k^-}{a_k^+ + a_k^-} \cdot \sin(\Delta\varphi_k) - \right. \\ &\left. - a_k^- \cdot \frac{a_k^+}{a_k^+ + a_k^-} \cdot \sin(\Delta\varphi_k) \right) = -\Delta\varphi_k \cdot \frac{2a_k^+ a_k^-}{a_k^+ + a_k^-}. \end{aligned} \quad (30)$$



Сравнивая (16), (24) и (30), получаем выражение для  $A_k$ :

$$A_0 = A_{0,0} = |z_k^+|; \quad (31)$$

$$A_k = A_{k,k} = \frac{|z_k^+| \cdot |z_k^-|}{|z_k^+| + |z_k^-|}. \quad (32)$$

С учетом введенных выше обозначений предлагаемое преобразование можно записать следующим образом:

$$X_0 = z_0^+; \quad (33)$$

$$X_k = \frac{|z_k^+| \cdot |z_k^-|}{|z_k^+| + |z_k^-|} \cdot \exp(j(\arg(z_k^+) - \arg(z_k^-))). \quad (34)$$

В качестве компонент искомого вектора признаков используются  $K$  низкочастотных коэффициентов описанного разложения. Т. е. каждому анализируемому изображению ставится в соответствие  $K$ -компонентный вектор признаков  $(X_1, X_2, \dots, X_K)$  или точка в  $K$ -мерном метрическом пространстве. Коэффициенты  $|X_k^q|$  перед расчетом  $D$  можно отнормировать.

$$W_k = \frac{A_k}{\sqrt{\sum_{n=0}^{N-1} A_k^2}}. \quad (35)$$

В качестве меры близости двух областей изображений будем использовать следующую величину:

$$D_f(x^0, x^1) =$$

$$= \sum_{k=0}^{K-1} W_k^0 \cdot W_k^1 \cdot D^2(\arg(X_k^0), \arg(X_k^1)). \quad (36)$$

Обобщение преобразования на двумерный случай заключается лишь в замене функций Хаара на их двумерные аналоги. Рис. 4 б иллюстрирует визуализацию 16 первых двумерных функций Хаара. В дальнейшем будем называть описанное преобразование методом WLSF (Weighted Least Squares Features).

#### Экспериментальные исследования и сравнительный анализ

На основе описанного преобразования был реализован алгоритм дескрипторного описания областей изображений в окрестностях ключевых точек для задачи их сопоставления. Сравнение проводилось с методом SURF [2].

Эксперимент ставился следующим образом. В качестве входных данных выступали изображения – оригинальные и зашумленные, с различной дисперсией гауссова шума. Гауссов шум был выбран как наиболее часто встречающийся на практике.

Примеры оригинальных и зашумленных ( $\sigma = 0.07$  и  $\sigma = 0.15$ ) изображений приведены на рис. 5.

Ключевые точки, размеры и ориентации анализируемых окрестностей искались по оригинальному изображению методом, описанным в работе [2]. Отметим, что основным предметом тестирования являлось качество предложенного описания окрестностей точек, поэтому для чистоты эксперимента (чтобы исключить влияние алгоритма поиска ключевых точек) точки на исходном и зашумленном изображениях одни и те же – найденные по оригинальному изображению.

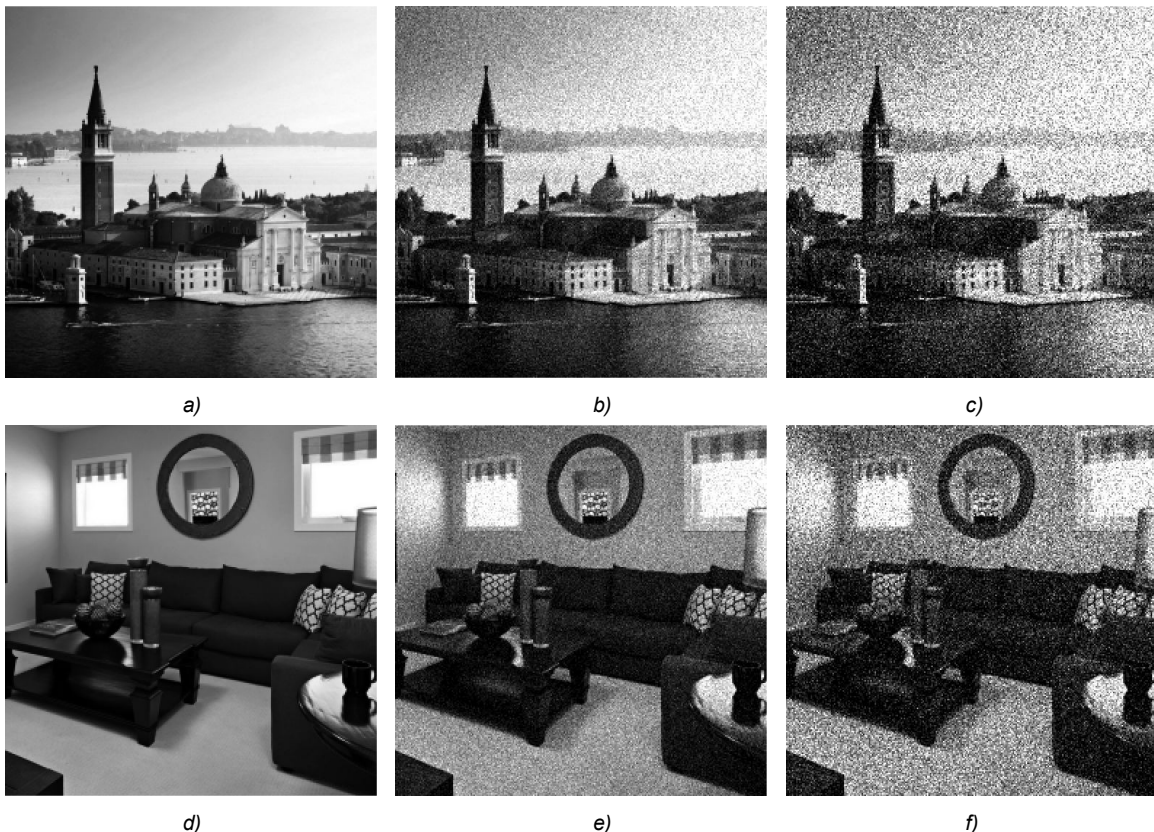


Рис. 5. Пример искажений тестовых изображений

Вектора признаков окрестностей ключевых точек в первом случае получались методом SURF, а во втором случае методом, предлагаемым в данной работе ( $K = 16$ ).

Примеры найденных ключевых точек представлены на рис. 6.

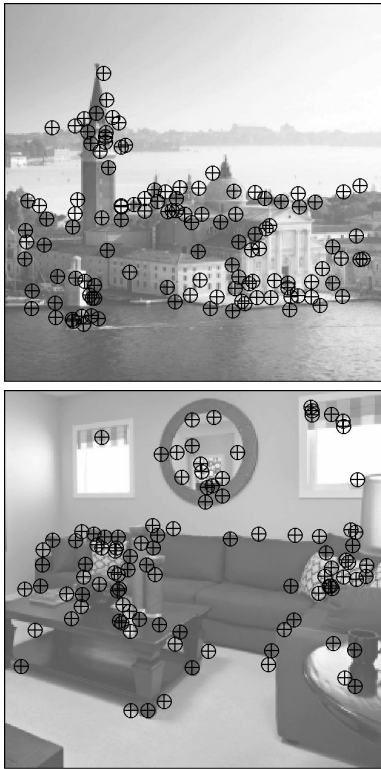


Рис. 6. Тестовые изображения и ключевые точки, полученные методом SURF

Графики зависимостей процента неправильно сопоставленных ключевых точек от среднеквадратического отклонения гауссова шума для методов SURF и WLSF приведены на рис. 7.

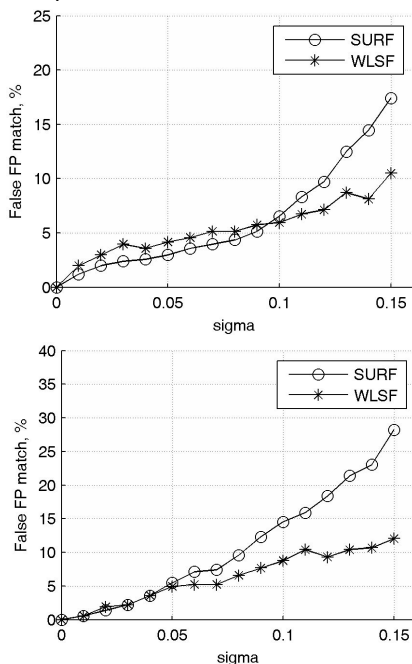


Рис. 7. Графики зависимостей процента неправильно сопоставленных ключевых точек от среднеквадратического отклонения гауссова шума

Из графиков видно, что наибольший выигрыш точности от использования предложенного метода достигается при анализе сильно зашумленных изображений ( $\sigma > 0.1$ ). При слабом уровне шумов методы SURF и WLSF в среднем показывают схожие результаты.

Итак, в ходе экспериментов был выявлен ряд позитивных и негативных особенностей метода. К минусам метода можно отнести повышенную чувствительность к изменению масштаба по одной из осей (при повороте 3D сцены). К плюсам – слабую чувствительность к гауссову шуму и к различным градационным преобразованиям изображений.

Описанный метод с успехом может быть использован для построения признакового описания областей изображений в различных задачах компьютерного зрения, в том числе в задаче поиска изображений по содержанию.

## Литература

1. Пиманкин Д. А. Построение признакового описания изображения на основе анализа распределения ориентаций векторов градиента. // Информационные системы и технологии-2012: Тезисы докл. – Нижний Новгород, 2012.
2. Bay, H., Tuytelaars, T., and Van Gool, L. (2006). SURF: Speeded up robust features. In Ninth European Conference on Computer Vision (ECCV 2006), pp. 404-417.
3. Berchtold S., Keim, D. A., and Kriegel H.-P. (1996) The X-Tree: An Index Structure for High-Dimensional Data. Proceedings of the 22nd International Conference on Very Large Data Bases (VLDB), Bombay, India: 28–39.
4. Ciaccia, P., Patella, M., and Zezula, P. (1997). M-tree An Efficient Access Method for Similarity Search in Metric Spaces. Proceedings of the 23rd VLDB Conference Athens, Greece. pp. 426–435.
5. Ke, Y., and Sukthankar, R. (2004). PCA-SIFT: a more distinctive representation for local image descriptors. In IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'2004), pp. 506–513, Washington, DC.
6. Lowe, D.G. (1999). Object recognition from local scale-invariant features. In Seventh International Conference on Computer Vision (ICCV'99), pp. 1150–1157, Kerkyra, Greece.
7. Lowe, D. G. (2004). Distinctive image features from scale-invariant keypoints. International Journal of Computer Vision, 60(2):91–110.
8. Mikolajczyk, K. and Schmid, C. (2005). A performance evaluation of local descriptors. IEEE Transactions on Pattern Analysis and Machine Intelligence, 27(10):1615–1630.

## WEIGHTED LEAST SQUARES FOR IMAGE ANALYSIS

*Pimankin D.*

Feature description build-up method for image areas based on image gradient vectors orientation analysis by weighted least squares (WLSF-method) is proposed. It is shown, that this approach can be successfully applied for the task of feature points matching for a large class of images. The advantage of the WLSF-method over SURF-method in the case of noisy image analysis is shown.



## МОДЕЛИРОВАНИЕ АППЛИКАТИВНЫХ ИСКАЖЕНИЙ С РАЗЛИЧНОЙ СТЕПЕНЬЮ ПРОЗРАЧНОСТИ И СЛУЧАЙНОЙ ФОРМОЙ

*Калинин П.В., аспирант Воронежского государственного университета, e-mail: kalinin\_pv@sc.vsu.ru;  
Сирота А.А., д.т.н., профессор Воронежского государственного университета, e-mail: sir@cs.vsu.ru*

**Ключевые слова:** обработка изображений, аппликативные искажения, моделирование, случайная форма, область закрытия.

### Введение

К числу основных задач, решаемых [1, 2, 3] в системах компьютерного зрения, относятся [0,0,0]: задачи сегментации изображений на компоненты, представляющие различные объекты сцены; задачи повышения визуального качества изображений; задачи объединения множества изображений одной и той же сцены в одно панорамное изображение; задачи получения 3D реконструкций на основе множества снимков одного и того же объекта с разных ракурсов; задачи классификации и распознавания объектов на изображениях.

Эффективному решению вышеописанных задач препятствует то, что изображения, формируемые в различных информационных системах, подвергаются воздействию помех, что затрудняет не только их восприятие человеком, но и обработку (контурный анализ, распознавание объектов) в автоматическом режиме. Как правило, возникающие на изображениях помехи носят аддитивный и импульсный характер. Импульсные помехи являются частным случаем аппликативных помех, действие которых проявляется в затенении (замещении) отдельных участков полезного изображения фрагментами со случайной амплитудой, текстурой, формой. Аппликативные помехи (АП) связаны с появлением различного рода неоднородностей, локальных областей аномальных значений, областей закрытия и пораженных участков.

Модели аппликативных помех могут использоваться как при разработке алгоритмов, направленных на их устранение, так и при решении других задач, например, задачи сегментации облаков на изображениях [4]. При моделировании аппликативных помех важны следующие характеристики: площадь генерируемых закрытий, форма помех (изотропная, анизотропная), нерегулярность и связность формы, прозрачность помехи. Недостатками одних алгоритмов [5, 6] является отсутствие возможности определить степень прозрачности закрытия, заранее задать его площадь, а также ограниченность возможностей по моделированию формы. Другие алгоритмы моделирования аппликативных помех, например [7], позволяют заранее определять площадь локальных закрытий, однако для больших значений площадей они генерируют локальные закрытия, близкие по форме к эллипсу или кругу, то есть не позволяют получать помехи с достаточно случайной формой. В существующей литературе в недостаточной степени рассматривался вопрос генера-

*Разработана обобщенная модель аппликативных помех на изображениях в виде локальных областей закрытия с различной степенью прозрачности и случайной формой. Рассмотрены различные варианты реализации алгоритмов моделирования аппликативных помех с возможностью управления параметрами интенсивности, связности, прозрачности и степени нерегулярности формы.*

ции полупрозрачных аппликативных помех. С целью решения описанных задач в данной работе рассматривается обобщенная модель генерации аппликативных помех, позволяющая создавать локальные закрытия произвольной площади и случайной формы с одновременным достижением эффекта полупрозрачности.

### Обобщенная модель генерации аппликативных помех случайной формы с различной степенью прозрачности

Аппликативное искажение характеризуется воздействием в отдельных областях исходного изображения и приводит к «замещению» ее элементов неинформативными фрагментами другой природы. В качестве исходной модели наблюдаемого изображения в присутствии АП будем рассматривать его представление как реализации случайного поля, заданного на квадратной дискретной сетке:

$$f(x, y) = (1 - \eta(x, y)) [s(x, y) + v(x, y)] + \eta(x, y) w(x, y), \quad x, y = \overline{1, m}, \quad (1)$$

где  $s(x, y)$  – исходное случайное поле;  $\eta(x, y)$  – бинарное случайное поле, определяющее расположение локальных областей закрытия с любыми параметрами периодичности, площади и формы ( $\eta(x, y) \in \{0, 1\}$ );  $v(x, y)$  – случайное поле аддитивного шума ( $M\{v(x, y)\} = 0$ ,  $M\{v(x, y)v(x', y')\} = R_v(x - x', y - y')$ );  $w(x, y)$  – случайное поле, характеризующее значение яркостных характеристик элементов областей закрытия (аппликативной помехи). Конкретное определение АП предполагает задание модели размещения и формы областей закрытия, а также статистические характеристики  $w(x, y)$ . Поле  $w(x, y)$  может быть выбрано произвольным (например, как гауссовское случайное поле), поэтому в дальнейшем мы сосредоточим внимание на процедуре генерации поля  $\eta(x, y)$ .

Будем рассматривать реализацию АП как совокупность локальных областей закрытия (ЛОЗ) – пятен некоторой формы и текстуры, замещающих фрагменты изображения и уничтожающих полезную информацию. При генерации каждой реализации АП предлагается выде-

лить два этапа: этап генерации так называемых *порождающих точек* – условных центров ЛОЗ, и этап генерации формы и текстуры АП в пределах каждой ЛОЗ.

Процесс появления *порождающих точек* будем описывать пуассоновским потоком ЛОЗ с заданной интенсивностью  $\Lambda$ . В соответствии с этим, количество таких точек в области изображения площадью  $V$  подчиняется распределению

$$P(S) = \frac{(\Lambda V)^S}{S!} \exp(-\Lambda V), \quad V = (N\Delta x) \times (N\Delta y),$$

$$x, y = \overline{1, m}, \quad (2)$$

где  $\Delta x, \Delta y$  – дискреты задания координатной сетки. При этом вероятность появления порождающей точки в каждом элементе изображения определяется как

$$P_\Lambda \cong \Lambda \Delta x \Delta y + O(\Delta x \Delta y),$$

$$\lim_{\Delta x \rightarrow 0, \Delta y \rightarrow 0} \frac{O(\Delta x \Delta y)}{\Delta x \Delta y} = 0. \quad (3)$$

Физический смысл подобной модели очевиден: порождающие точки равновероятно распределены в области изображения; при этом параметр  $\Lambda$  позволяет управлять количеством таких точек и соответствующих им ЛОЗ.

Далее для каждой порождающей точки введем случайную целочисленную величину  $n$ , значения которой определяют площади ЛОЗ, образующихся в окрестности порождающих точек. При задании  $n$  может использоваться любой закон распределения, например в [7, 8] предлагается использовать пуассоновский закон распределения с заданной интенсивностью  $\lambda$ . При этом с учетом присутствия порождающей точки

$$P(n) = \frac{\lambda^{n-1}}{(n-1)!} e^{-\lambda}, \quad n \geq 1,$$

$$E[n] = \sum_{n=1}^{\infty} n \frac{\lambda^{n-1}}{(n-1)!} e^{-\lambda} = \lambda + 1 = \frac{S}{\Delta x \Delta y}, \quad (4)$$

где  $S$  – имеет смысл средней площади ЛОЗ. С учетом этого результирующий процесс образования АП может рассматриваться как маркированный, или сложный пуассоновский процесс, где с каждой точкой простого потока связана случайная величина, определяющая площадь локальной области закрытия.

Теперь подробно рассмотрим предлагаемый обобщенный алгоритм генерации формы аппликативной помехи для каждой полученной порождающей точки. Алгоритм применяется для генерации формы каждого ЛОЗ *отдельно*, то есть помеха представляется как суперпозиция независимо сгенерированных локальных закрытий. Алгоритм базируется на последовательном наращивании искаженной области вокруг каждого центра локального закрытия путем выбора новых искаженных пикселей, граничащих с уже искаженными. В стандартном режиме генерации АП [7, 8], в соответствии с моделью (1), для каждого пикселя вводятся состояния: искажен – не искажен (1 или 0). Для того чтобы получить более сложные формы ЛОЗ, обладающие свойством полупрозрачности, относительные значения которой варьируются в интервале  $[0, 1]$ , данный подход необходимо

расширить за счет усложнения структуры модели. Предполагается, что генерация формы помехи производится на бесконечной квадратной решетке. Каждому узлу соответствует его значение прозрачности  $\hat{I}(i, j) \in [0, 1]$ , на основе которого формируется поле  $\eta(x, y)$ .

В процессе генерации иметь дело со значениями прозрачности напрямую не всегда удобно, в частности, из-за того, что значения меньше 0 и больше 1 не допускаются. Решить данную проблему можно введя дополнительную величину  $I(i, j) \in R$ , однозначно связанную с  $\hat{I}(i, j)$  некоторым отображением  $M$ . Например, в роли  $I(i, j)$  может выступать концентрация вещества физической среды, закрывающей изображение в данном узле решетки. Так, для случая закрытия изображений облачными образованиями, значения  $I$  могут определяться значениями высоты (толщины) облака по каждой конкретной координате. Можно предположить следующее: если в некоторой точке сконцентрировано значительное количество вещества, то данная ситуация должна повышать ожидаемую концентрацию вещества и для близлежащих точек. Для учета указанных факторов  $I(i, j)$  определяется как случайная величина с условным распределением  $I(i, j) \sim p_I(x | I, i, j)$ , отражающим зависимость значений прозрачности АП в данной точке от окружающих точек.

В процессе наращивания области закрытия в итеративном режиме на каждом шаге производится выбор следующего узла для искажения. При этом вероятность выбора считается ненулевой для еще неискаженных узлов, которые попадают в некоторую маску расширения  $U$  (рис. 1). Маска используется следующим образом. Своим центром (черный цвет) она помещается на уже искаженный пиксель. Все неискаженные пиксели, попадающие в зону расширения маски (серый цвет), становятся кандидатами на искажение. Применение данной маски ко всем искаженным пикселям позволяет получить координаты всех потенциальных кандидатов на искажение на данном шаге итерации.

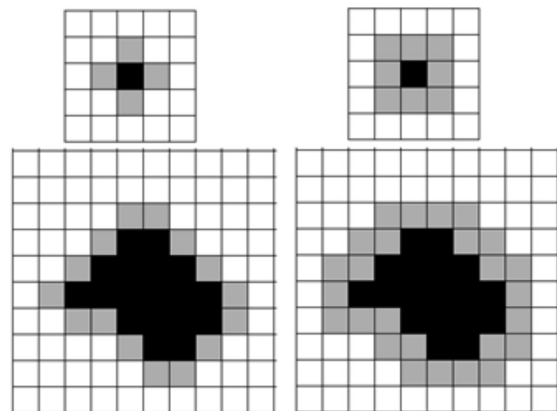
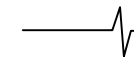


Рис. 1. Примеры маски расширения  $U$  (вверху) и результаты ее применения к искаженным пикселям на некотором этапе итерации (внизу): слева - маска  $U_4$ , справа - маска  $U_9$

Простейший способ определения следующего искажаемого узла состоит в случайном выборе одного из узлов множества потенциальных кандидатов, считая их равновероятными. Но, так как данный подход является



ограниченным и ведет к снижению степени нерегулярности формы ЛОЗ, логичнее воспользоваться другим подходом. Каждому элементу множества кандидатов на искажение ставится в соответствие обновляемая на каждом шаге величина  $p_{distortion}(i, j)$ , определяющая вероятность его выбора. Значения данного распределения могут пересчитываться на основе значений поля  $I$ , однако наибольшая гибкость достигается, если отделить значения, влияющие на прозрачность, от значений, влияющих на форму. С этой целью вводится поле  $D$ , определенное в узлах решетки и содержащее всю информацию, требуемую для пересчета  $p_{distortion}(i, j | D)$ .

Одним из простейших способов задания  $D$  является следующий подход. Можно положить, что каждый искаженный элемент решетки генерирует вокруг себя направленное или ненаправленное искажающее воздействие, увеличивающее вероятности искажения окружающих его элементов. Тогда, для каждого узла  $D(i, j)$  может включать как одно значение, характеризующее силу ненаправленного искажающего воздействия, так и несколько, характеризующих силу и направление искажающего воздействия. Далее в конкретных алгоритмах, приведенных в работе, используется одно значение, т.е.  $D$  является скалярным. Значения  $D(i, j)$  определяются на вероятностной основе:  $D(i, j) \sim p_D(x | D, I(i, j), i, j)$ . Данная запись отражает то, что сила искажающего воздействия узла может зависеть только от значения концентрации «вещества»  $I(i, j)$  в данной точке. Зависимость от остальных значений  $I$  в других точках не учитывается. При этом сила и направление искажающего воздействия в данном узле могут зависеть от соответствующих величин в соседних узлах, для чего вводится зависимость от значений поля  $D$  остальных узлов. Координаты  $(i, j)$  в приведенных условных распределениях можно использовать не только для индексации позиции узлов, но и для моделирования неоднородности введенных случайных полей.

С учетом приведенных обоснований, общий алгоритм генерации формы локальных закрытий относительно порождающей точки (в соответствии с разработанной моделью) приведен на блок-схеме рис. 2. Данный алгоритм позволяет одновременно моделировать прозрачность, форму, а также зависимость между формой и прозрачностью аппликативной помехи.

В общем случае, уже сгенерированные значения  $I, D$  для искаженных узлов могут меняться в процессе последующего выполнения алгоритма. Однако это вносит дополнительные сложности и увеличивает объем вычислений за счет необходимости пересчета значений уже искаженных пикселей на каждом шаге алгоритма. Поэтому в предлагаемой модели данные изменения игнорируются и значения  $I, D$  считаются неизменными от шага к шагу.

Для инициализации алгоритма необходимо выбрать порождающую точку – условный центр помехи с координатами  $(i, j)$ , относительно которой будет произ-

водиться генерация формы локального закрытия, начальные распределения  $I(i, j) \sim p_I^0(x | i, j)$  и  $D(i, j) \sim p_D^0(x | I(i, j), i, j)$  для центра помехи, а также условие останова, например, значение максимального количества искаженных пикселей (площадь ЛОЗ).

По завершении работы алгоритма возможно выполнение различной постобработки, например: сглаживание, масштабирование, повороты, с тем чтобы достичь желаемых статистических характеристик.

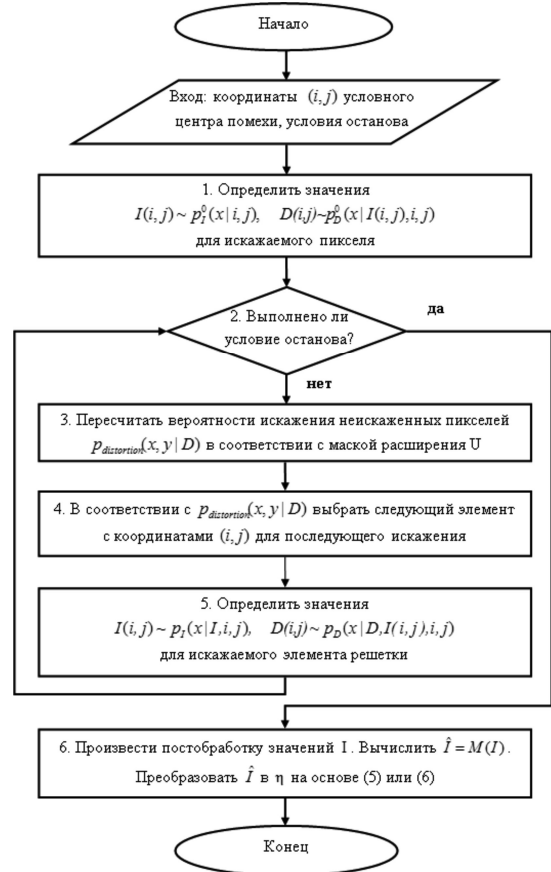


Рис. 2. Блок-схема обобщенного алгоритма генерации аппликативных помех на изображениях

Значения  $\eta(i, j)$  на основе значений прозрачности  $\hat{I}(i, j)$  можно определить двумя различными способами. Согласно первому способу, значение вероятности искажения устанавливается либо 0, либо 1, на основе сравнения  $\hat{I}(i, j)$  со значением равномерно распределенной случайной величины, принимающей значения в интервале  $[0, 1]$ :

$$\eta(i, j) = \begin{cases} 1, & p(\eta(i, j) = 1) = \hat{I}(i, j); \\ 0, & p(\eta(i, j) = 0) = 1 - \hat{I}(i, j) \end{cases} \quad (5)$$

В данном случае помеха полностью замещает значения искаженных пикселей. Согласно второму способу, помеха просто смешивается с изображением:

$$\eta(i, j) = \hat{I}(i, j). \quad (6)$$

Второй способ является стандартным подходом в компьютерной графике и широко используется в различных средах (например, в VisualStudio) для моделирова-

ния прозрачности. Как правило, он реализуется путем добавления четвертого канала (к основным трем каналам) – канала альфа-значений.

Следует отметить, что второй способ не совсем точно соответствует классическому определению аппликативной помехи, однако применение данного понятия оправдывается тем, что он является предельным случаем первого способа при уменьшении размера ячеек дискретной решетки. При этом интенсивности изображения и помехи начинают сливаться, и достигается эффект, характерный для первого способа моделирования искажения. Иллюстрацией этого могут служить приведенные далее иллюстративные примеры. Приближая и удаляя изображения, можно видеть, что при достаточной степени удаленности, различные способы дают примерно одинаковый эффект.

**Частные алгоритмы генерации аппликативных помех с различной степенью прозрачности**

В [7] был предложен алгоритм формирования изотропных пятен случайной формы, площадь которых подчиняется заданному распределению. На первом шаге для порождающей точки ЛОЗ искажается один ближайший к ее центру пиксель. Далее, на каждом шаге искажается один примыкающий к ранее полученному образованию пиксель, который выбирается равновероятно среди всех точек, образующих границу этого образования. Указанный процесс продолжается до достижения заданного значения площади ЛОЗ.

Даная модель может быть получена как частный случай предлагаемого обобщенного алгоритма, если положить:  $p_i^0(x|i, j) = p_i(x|I, i, j) = \delta(x-1)$ , где  $\delta(x)$  – дельта функция Дирака. Вероятность  $p_{distortion}(i, j|D)$  можно считать одинаковой для всех неискаженных пикселей, соседствующих (в соответствии с маской расширения  $U$ ) с уже искаженными пикселями:

$$p_{distortion}(i, j|D) = \frac{1}{N^{(t)}}, \tag{7}$$

где  $N^{(t)}$  – количество элементов-кандидатов на искажение на шаге  $t$ . Поле  $D(i, j)$  полагается скалярным, а его значения определяют, искажен или нет заданный пиксель. Остальные параметры игнорируются. Пример генерации помех, в соответствии с данным алгоритмом, приведен на рис. 3а.

Рассмотрим различные комбинации параметров данного частного алгоритма. Маску расширения будем считать фиксированной и положим ее равной  $U_4$  (рис. 1, слева). В исходном варианте алгоритма не выполняется никакой постобработки. Введем дополнительную постобработку, заключающуюся в сглаживании получаемого поля  $I$  гауссовым фильтром. Это позволяет добиться эффекта прозрачности на краях искажения. Пример сглаживания помехи при размере окна фильтра 15x15 и дисперсии гауссова фильтра, равного 1.5, приведен на рис. 3б. Заметим, что значение 1 соответствует полной непрозрачности, 0 – полной прозрачности. Соответственно, чем ярче пиксель, тем ниже степень его прозрачности.

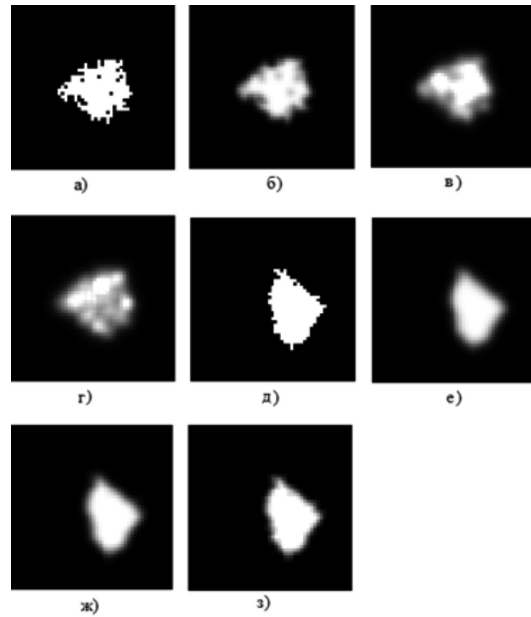


Рис. 3. Примеры генерируемых локальных областей закрытия (площадь – 250 пикселей)

До текущего момента считалось, что  $p_i(x|I, i, j)$  дает ненулевую вероятность только значению 1. Попробуем усложнить данное распределение.

Пусть  $p_i^0(x|i, j) = p_i(x|I, i, j) = \lambda e^{-\lambda \cdot x}$ . Положим  $\lambda = 1$ . В данном случае значения  $I(i, j)$  больше не принадлежат интервалу  $[0,1]$ . Поэтому, после выполнения сглаживания гауссовым фильтром, все значения, превышающие 1, положим равными 1. Результат генерации показан на рис. 3в.

Пример другого варианта выбора параметра, заключающийся в генерации значений на основе смеси экспоненциально распределенной и модуля гауссовой случайной величин, приведен на рис. 3г, где  $\alpha = 0.5$ ,  $\lambda = 1$ ,  $\sigma = 1$ ,  $\mu = 0$ :

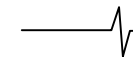
$$I(i, j) = \alpha \cdot \xi_1 + (1 - \alpha) \cdot |\xi_2|, \tag{8}$$

где

$$\xi_1 \sim \lambda e^{-\lambda \xi_1}, \xi_1 \geq 0, \xi_2 \sim \frac{1}{\sigma \sqrt{2\pi}} e^{-\frac{(\xi_2 - \mu)^2}{2\sigma^2}}.$$

Отметим, что в вышеописанных примерах поле  $D$  использовалось только для обозначения координат искаженных пикселей. Следствием этого упрощения является то, что при больших значениях площади закрытие принимает форму круга или эллипса, в соответствии с заданными параметрами. Для моделирования реальных помех это не всегда желательно, так как известно, что их форма обычно нерегулярна. В то же время, при генерации по данным алгоритмам форму помех с небольшой площадью можно считать случайной. В совокупности с тем, что форма помехи должна быть связанной, наличие изолированных искаженных пикселей не допускается. Задача получения нерегулярной формы для больших площадей закрытия достаточно сложна.

Для получения помех случайной формы с большей площадью, в соответствии с приведенным обобщенным



алгоритмом, будем использовать следующий подход. Как и ранее, условием останова будем считать накопленную площадь искажения. Степень анизотропии локальных закрытий определяется конкретной процедурой определения поля  $D$  (в данном случае скалярного) и выбора следующего искажаемого пикселя

$p_{distortion}(i, j | D)$ . Положим

$$p_I^0(x) = p_I(x | I, i, j) = \delta(x - 1), \quad (9)$$

$$p_D^0(x | I(i, j), i, j) = p_D(x | D, I(i, j), i, j) = p_D(x).$$

Будем использовать маску расширения  $U_4$ . Конкретные значения распределения  $p_D(x)$  и алгоритм определения  $p_{distortion}(i, j | D)$  определяют итоговую форму помехи. При проведении экспериментов удачным оказался следующий вариант: вместо явного задания распределения  $p_D(x)$  использовалась генерация значения  $x$  на основе следующего выражения:

$$x = \lambda^n, p(\lambda) = \alpha \exp(-\alpha\lambda), \quad (10)$$

где  $M\{\lambda\} = \alpha^{-1}$ . Данная случайная величина подчиняется распределению

$$p(x) = \frac{\alpha}{n} x^{n-1} \exp(-\alpha^n \sqrt{x}), \quad (11)$$

с математическим ожиданием

$$M\{x\} = \alpha^{-n} \Gamma(1 + n), \quad (12)$$

где  $\Gamma(x)$  – гамма функция. Достаточно хорошие результаты при моделировании формы помехи были получены при  $n = 8, \alpha = 2$ .

Вероятность  $p_{distortion}(i, j | D)$  будем определять по следующему алгоритму. Пусть  $C$  – массив координат уже искаженных пикселей. Для каждого пикселя с координатами  $(i, j)$ , являющегося соседним к искаженным пикселям, в соответствии с маской  $U_4$  определим фактор, пропорциональный вероятности его выбора на следующем шаге, вычисляемый по формуле

$$F(i, j | D) = \sum_{(x,y) \in C} D(i, j) \cdot e^{-\beta \sqrt{(x-i)^2 + (y-j)^2}}. \quad (13)$$

После нормализации данного фактора получается искомое распределение  $p_{distortion}(i, j | D)$ , позволяющее выбрать следующий искажаемый пиксель. В ходе экспериментов было установлено, что достаточно высокая степень нерегулярности для ранее выбранных параметров достигается при  $\beta = 1.5$ .

Примеры генерации формы помехи в соответствии с данным алгоритмом без сглаживания и с различными параметрами сглаживания (дисперсия гауссова фильтра равна 1.5, размер окна 15x15, 5x5, 3x3 соответственно) показаны на рис. 3д, е, ж, з. Как видно из рисунков, при использовании данного подхода наблюдается более высокая реалистичность помехи, чем при использовании подхода из [7]. Отметим, что данный подход позволяет

получать достаточно случайную форму ЛОЗ с большим количеством пикселей (более 10 000). Однако, при необходимости получения формы локального закрытия, содержащего значительно большее количество пикселей, наиболее логичным подходом будет использование 6-го шага алгоритма – генерации помехи относительно небольшой площади и последующее ее масштабирование до желаемых размеров. Дополнительные примеры закрытий, сформированных по данному алгоритму и выбранных случайным образом, с площадью искажения 2500 пикселей, приведены на рис. 4. Результаты применения различных параметров прозрачности и сглаживания помехи на тестовом изображении приведены на рис. 5, 6. Интенсивность значений помехи на данных рисунках полагалась равной 127.

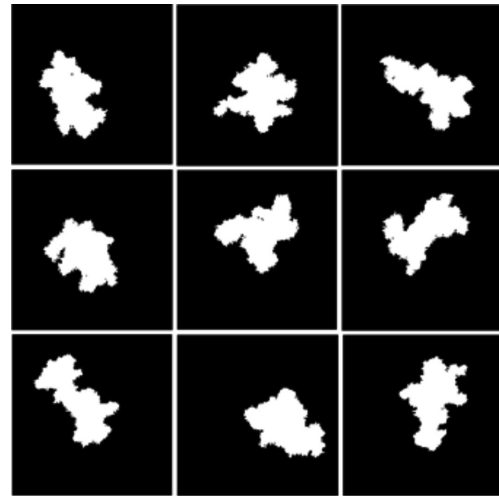


Рис. 4. Примеры генерации помех со случайной формой. Размер пятна – 2500 пикселей

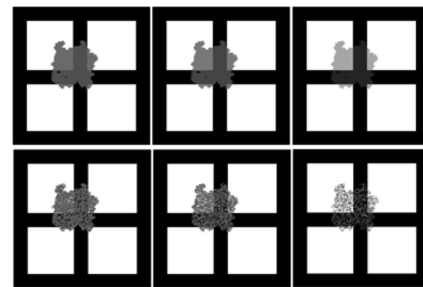


Рис. 5. Примеры генерируемых локальных областей закрытия без размытия с различными параметрами прозрачности – 10%, 20%, 50% (слева направо): сверху – прозрачность на основе формулы (6), внизу – на основе формулы (5)

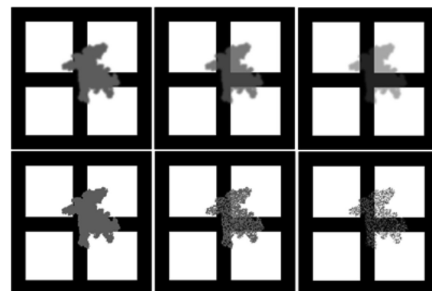


Рис. 6. Примеры генерируемых локальных областей закрытия с различными параметрами прозрачности – 0%, 20%, 50% (слева направо): сверху – прозрачность на основе формулы (6) (с размытием), внизу – на основе формулы (5) (без размытия)



## Заключение

Предложенный подход позволяет достаточно эффективно моделировать полупрозрачные случайные аппликативные помехи. По сравнению с существующими методами он дает конкретные алгоритмы, генерирующие закрытия с произвольной случайной формой и площадью. Приведенные алгоритмы моделирования закрытий с различной степенью прозрачности позволяют эффективно достигать визуально правдоподобного результата.

## Литература

1. Методы компьютерной обработки изображений/под ред. В. А. Сойфера. – М.: ФИЗМАТЛИТ, 2003. – 784 с.
2. Гонсалес Р., Вудс Р. Цифровая обработка изображений. М.: Техносфера, 2005. – 1072 с.
3. Szeliski R. Computer Vision: Algorithms and Applications. Springer, 2011. – 832p.
4. Ветров А.А., Кузнецов А.Е. Сегментация облачных объектов на панхроматических изображениях земной поверхности // Цифровая обработка сигналов. – 2011. – №3.
5. Robidoux N., Stelldinger P., Cupitt J. Simple random generation of smooth connected irregular shapes for cognitive studies // C3S2E '11 Proceedings of The Fourth

International C\* Conference on Computer Science and Software Engineering. – 2011. – P. 17-24.

6. Auer T., Held M. Heuristics for the Generation of Random Polygons // Proceedings of the 8th Canadian Conference on Computational Geometry. – 1996. - P.38-43.

7. Сирота А.А., Попов В.Г. Статистические и нейросетевые алгоритмы восстановления цифровых изображений в условиях аппликативных искажений // Вестник ВГУ. Сер. Системный анализ и информационные технологии. - 2006. - №1. – С. 54-62.

8. Алгазинов Э.К., Сирота А.А. Анализ и компьютерное моделирование информационных процессов и систем. М.: Диалог-МИФИ, 2009. – 416 с.

## MODELING OF APPLICATIVE NOISE WITH RANDOM SHAPE AND VARIOUS DEGREES OF OPACITY

*Kalinin P.V., Sirota A.A.*

A new generalized model of applicative noise on images represented as a set of local occlusions with various degrees of opacity and random shapes is proposed. Different implementations of applicative noise modeling algorithms that allow to control intensity, connectivity, opacity, and irregularity of shapes of occlusions are considered.



## SAMSUNG ELECTRONICS ПРИГЛАШАЕТ НА РАБОТУ ПРОГРАММИСТОВ, МАТЕМАТИКОВ, ИНЖЕНЕРОВ

**Samsung Electronics Co., Ltd.** – мировой лидер в разработке цифрового, телекоммуникационного и полупроводникового оборудования – приглашает на работу разработчиков, программистов и инженеров в исследовательские центры компании в Республике Корея (Южная Корея).

Особый интерес представляют специалисты следующих направлений:

- **Разработка программного обеспечения, информационная безопасность, криптография** (C/C++, Linux, Embedded Linux, Android, Java, RTOS, UI, RTOS, Security)
- **Разработка мобильных устройств, мультимедийных систем, цифровая обработка сигналов** (Smartphone, Messaging, Multimedia, Streaming, RIL, DSP)
- **Разработка систем цифрового телевидения** (Digital TV, Set-Top Box, FPGA, Verilog HDL, SoC, 3D Display, Metamaterials)
- **Разработка антенных устройств, радиоэлектронных систем, устройств для телекоммуникации и сетевых технологий** (Antenna Design, ASIC Design, Integrated Circuits, FPGA, HW / SW Co-Design, System LSI, TCP / IP Protocol, VoIP, OFDM-MIMO)

Требования к кандидатам:

- Образование: высшее / ученая степень кандидата технических / физико-математических наук в области программирования, кибернетики, прикладной математики, телекоммуникаций, радиоэлектроники, полупроводниковой физики, наук о материалах, химии;
- Опыт работы: научная / исследовательская работа / прикладные разработки (**по одному из вышеперечисленных направлений**) НЕ МЕНЕЕ 3~5 ЛЕТ;
- Английский язык: хорошее владение английским языком (устным и письменным).

Дополнительная информация: <http://JOB.samsung.ru>

Электронная почта: [JobFair@samsung.com](mailto:JobFair@samsung.com)

## МЕТОД ФОРМИРОВАНИЯ БИНАРНОЙ КАДРОВОЙ МАСКИ ДВИЖЕНИЯ

*Минаков Е. И., д.т.н., профессор кафедры радиоэлектроники Тульского государственного университета, e-mail: EMinakov@bk.ru;*

*Калистратов Д. С., аспирант кафедры радиоэлектроники Тульского государственного университета*

**Ключевые слова:** маска движения, кадр, информация, кодирование изображений, видеокодек.

### Введение

Задача разделения кадрового изображения на подвижную и неподвижную составляющие играет значимую роль в области видео кодирования, поскольку кодированию подвергается только подвижная часть и от того, насколько качественно она отделена от фона, в серьёзной степени зависит как время обработки кадра, так и качество декодируемого изображения. Визуально легко отделить подвижный объект от фона (рис. 1), однако для технического устройства задача распознавания движения осложняется многими факторами, в числе которых в первую очередь можно выделить неравномерность внешнего освещения и не идеальность применяемых технических устройств. Для обозначения подвижных частей изображения современные видеокодеки [2] используют бинарную кадровую маску движения, которая показывает наличие или отсутствие движения в каждом отдельно взятом пикселе текущего кадра. Неподвижные пикселы обозначаются логическими нулями, а подвижные – логическими единицами. Для вычисления маски движения используется текущий кадр и один или несколько дополнительных кадров, называемых ссылочными. В процессе компенсации движения кадровое изображение разбивается на фрагменты, называемые макроблоками.

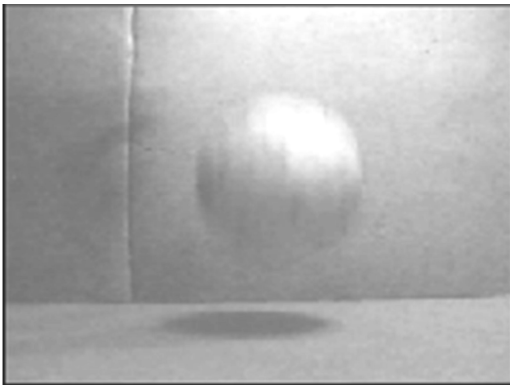


Рис.1. Кадр динамической видеопоследовательности

По маске движения проводится сортировка макроблоков, в результате которой часть блоков признаётся подвижной, а часть неподвижной. Для подвижных блоков кодером проводится поиск их наилучших прогнозов на ссылочных кадрах. Векторы смещений найденных прогнозов кодируются и передаются декодеру, чтобы тот по ним мог восстановить текущий кадр, имея в наличии

*Рассматривается метод формирования бинарной кадровой маски движения на основе принципов контекстно-зависимого арифметического кодирования. Предлагается метод формирования бинарной кадровой маски движения на основе динамической регенерации подвижной части маски, проводится структурный синтез и сравнительный анализ моделей видеокодеков, использующих указанные методы.*

предыдущий. Поиск хороших блочных прогнозов сопровождается большими временными затратами. В этой связи, в целях повышения временной эффективности видеокодека, актуальной является задача формирования качественной бинарной кадровой маски движения.

### Формирование бинарной кадровой маски движения на основе принципов контекстно-зависимого арифметического кодирования

Современные видеокодеки в процессе формирования кадровых масок движения проводят предварительное построение маски по паре последовательных кадров с её последующей коррекцией на основе статистического анализа полученной маски. Руководствуясь принципами контекстно-зависимого арифметического кодирования [2], можно записать следующую формулу для предварительного формирования кадровой маски движения:

$$M_{0,i,j} = \begin{cases} 0, & \text{при } |Y_{t,i,j} - Y_{t-1,i,j}| = 0; \\ 1, & \text{при } |Y_{t,i,j} - Y_{t-1,i,j}| \neq 0, \end{cases} \quad (1)$$

$$(i = 0..(j = 0..cs - 1).ls - 1).$$

где:  $M$  – массив масок;  $Y$  – массив яркостных кадровых матриц с представлением их элементов относительными единицами яркости из диапазона [0..255];  $t$  – индекс кадра;  $i, j$  – индексы пикселов в кадре;  $ls$  – количество строк в кадре;  $cs$  – количество столбцов в кадре.

Согласно этой формуле, каждый пиксел может быть признан подвижным или неподвижным в зависимости от абсолютного отклонения между значениями яркостей пиксела в текущем и предыдущем кадрах. Коррекция маски движения может быть проведена по формуле:

$$M_{1,i,j} = \begin{cases} 0, & \text{при } 0 \leq P_{1,i,j} \leq p; \\ 1, & \text{при } p < P_{1,i,j} \leq 1, \end{cases} \quad (2)$$

$$(i = 0..(j = 0..cs - 1).ls - 1),$$

где:

$$P_{1,i,j} = \frac{1}{S} \sum_{k=-1}^1 \sum_{l=-1}^1 (M_{0,i+k,j+l})$$

$$с ограничениями \begin{cases} ((i+k) > 0); \\ ((j+l) > 0); \\ ((i+k) < (ls-1)); \\ ((j+l) < (cs-1)). \end{cases}$$

$P_{1i,j}$  – массив вероятностей движения пикселей, значения которого лежат в вероятностном диапазоне [0..1];  $p$  – порог вероятности движения, разделяющий вероятностный диапазон на два поддиапазона с границами [0.. $p$ ] и ( $p$ ..1];  $S$  – количество соседей текущего обрабатываемого пиксела, включая сам пиксел;  $k, l$  – индексы соседних пикселей.

При проведении коррекции маски вычисляются вероятности движения пикселей в контексте движения их пространственных соседей. Конечное заключение относительно движения пиксела делается по попаданием рассчитанных вероятностей в один из двух вероятностных поддиапазонов. Таким образом, каждый пиксел может поменять свой статус на противоположный. Результат работы алгоритма построения маски движения (рис. 2), использовавшего указанные формулы со значением порога вероятности движения равным  $p = 0.5$  и количеством контекстных пикселей  $S = 9$ , представлен далее. Анализируя полученную маску движения, можно отметить, что эффект влияния помех и слабых отклонений параметров освещения не был устранён. Это означает, что при фрагментации изображения обнаружится много подвижных макроблоков, что, в свою очередь, потребует серьёзных затрат машинного времени на поиск их прогнозов. Уменьшение порога вероятности движения усугубляет действие помех, а при его увеличении вместе с зашумлёнными областями начинают выключаться из обработки действительно подвижные части изображения.

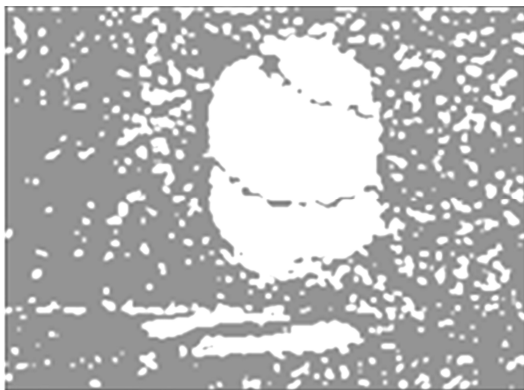


Рис.2. Бинарная кадровая маска движения, полученная на основе принципов контекстно-зависимого арифметического кодирования

**Формирование бинарной кадровой маски движения на основе динамической регенерации подвижной части маски**

С точки зрения временной эффективности кодека представляет интерес метод, который без существенных потерь сможет выделить подмножество действительно подвижных пикселей и исключить из обработки подмножество пикселей, движение которых вызвано помехами и

слабыми отклонениями параметров внешнего освещения. Предлагаемый метод также использует начальное формирование маски с её последующей коррекцией. Стартовый вид маски движения определяется формулой:

$$M_{0,i,j} = \begin{cases} 0, & \text{при } |Y_{t,i,j} - Y_{t-1,i,j}| \leq L; \\ 1, & \text{при } |Y_{t,i,j} - Y_{t-1,i,j}| > L, \end{cases} \quad (3)$$

$$(i = 0..(j = 0..cs - 1)..ls - 1),$$

где  $L$  – условный максимум значений помех.

Для исследования эффективности предлагаемого метода использовалось значение этого максимума, равное  $L = 25$  отн. ед. яркости. В соответствии с представленной формулой на этапе формирования начального вида маски действительно подвижными признаются только те пиксели, абсолютное отклонение сигналов которых от значений в предыдущем кадре превышает условный максимум помех. При этом (рис. 3) отсекаются как помехи, так и часть действительно подвижных областей кадра. Дальнейшая задача метода состоит в том, чтобы обеспечить динамическую реконструкцию подвижных частей кадра, исключив помехи.



Рис.3. Начальный вид маски движения с учётом условного максимума значений помех

С этой целью организуется цикл коррекции маски движения, в котором маска пересчитывается заранее фиксированное число раз. Динамическая реконструкция маски ведётся по формуле

$$M_{n,i,j} = \begin{cases} 0, & \text{при } 0 \leq P_{1i,j} \leq p; \\ 1, & \text{при } p < P_{1i,j} \leq 1, \end{cases} \quad (4)$$

$$(n = 1..(i = 0..(j = 0..cs - 1)..ls - 1)..ns),$$

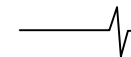
где

$$P_{n,i,j} = \frac{(k1 * P1_{n,i,j} + k2 * P2_{n,i,j})}{2}$$

с условием  $k1 + k2 = 2$ ;

$$P1_{n,i,j} = \begin{cases} \frac{|Y_{t,i,j} - Y_{t-1,i,j}|}{L}, & \text{при } |Y_{t,i,j} - Y_{t-1,i,j}| \leq L; \\ 1, & \text{при } |Y_{t,i,j} - Y_{t-1,i,j}| > L \end{cases}$$

$$P2_{n,i,j} = \frac{1}{S} \sum_{k=-1}^1 \sum_{l=-1}^1 (M_{n-1,i+k,j+l})$$



$$\text{с ограничениями} \left\{ \begin{array}{l} ((i+k) > 0); \\ ((j+l) > 0); \\ ((i+k) < (ls-1)); \\ ((j+l) < (cs-1)), \end{array} \right.$$

$n$  – порядковый номер пересчёта маски;  $ls$  – количество пересчётов маски;  $P1$  – массив вероятностей движения, определяемых по абсолютным отклонениям сигналов;  $P2$  – массив вероятностей движения, определяемых по пикселям предыдущей маски;  $k1, k2$  – весовые коэффициенты вероятностных компонентов.

С каждым последующим пересчётом подвижная часть маски начинает расширяться. Это расширение носит не хаотический, а целенаправленный характер. Суммарная вероятность вовлечения очередного пикселя в подвижное подмножество маски зависит как от абсолютного отклонения сигнала этого пикселя, так и от поведения его пространственных соседей. Приоритет между факторами определения движения регулируется весовыми коэффициентами вероятностных компонентов. В роли динамического параметра выступает порядковый номер пересчёта маски. Динамика восстановления кадровой маски движения (рис. 4) со значением порога вероятности движения равным  $p = 0.1$ , количеством пересчётов маски  $ls = 10$  и при значениях весовых коэффициентов  $k1 = k2 = 1$  показывает, что происходит существенное ослабление влияния помех на конечный вид кадровой маски движения.

Вместе с тем, наблюдается довольно качественная реконструкция подвижных частей изображения. Конечный вид маски движения после десятого пересчёта (рис. 5) предполагает сравнительно малое количество подвижных макроблоков при фрагментации изображения и, как следствие, существенное сокращение времени обработки кадра в целом.

### Структурный синтез и сравнительный анализ моделей видеокодеков, использующих рассмотренные методы формирования бинарной кадровой маски движения

Для исследования эффективности предложенного метода формирования кадровой маски движения, на основе современных стандартов и принципов видеокодирования [1-3], были синтезированы две модели видеокодеков, идентичные по своей структуре (рис. 6). Источник (SRC) и приёмник (RCV) сигналов хранят последова-

$n = 3$



$n = 6$



$n = 9$



Рис. 4. Динамика восстановления кадровой маски движения



Рис. 5. Окончательный вид блочной кадровой маски движения после применения разработанного метода

тельности обрабатываемых кадров. Кодер образован следующей цепочкой элементов: прямой компенсатор движения (MCP), прямой дискретно-косинусный преобразователь (DCT), квантователь (Q), кодер серий переменной длины (RLE) и энтропийный кодер (VLE). В состав декодера входят: энтропийный декодер ( $VLE^{-1}$ ), декодер серий переменной длины ( $RLE^{-1}$ ), деквантователь ( $Q^{-1}$ ), обратный дискретно-косинусный преобразователь ( $DCT^{-1}$ ), обратный компенсатор движения ( $MCP^{-1}$ ). Различия моделей заключаются только в способах формирования кадровой маски движения. Первая модель использует бинарную кадрową маску движения, получаемую на основе принципов контекстно-зависимого арифметического кодирования. Вторая модель использует бинарную кадрową маску, получаемую на основе предложенного метода динамической регенерации подвижной части маски.

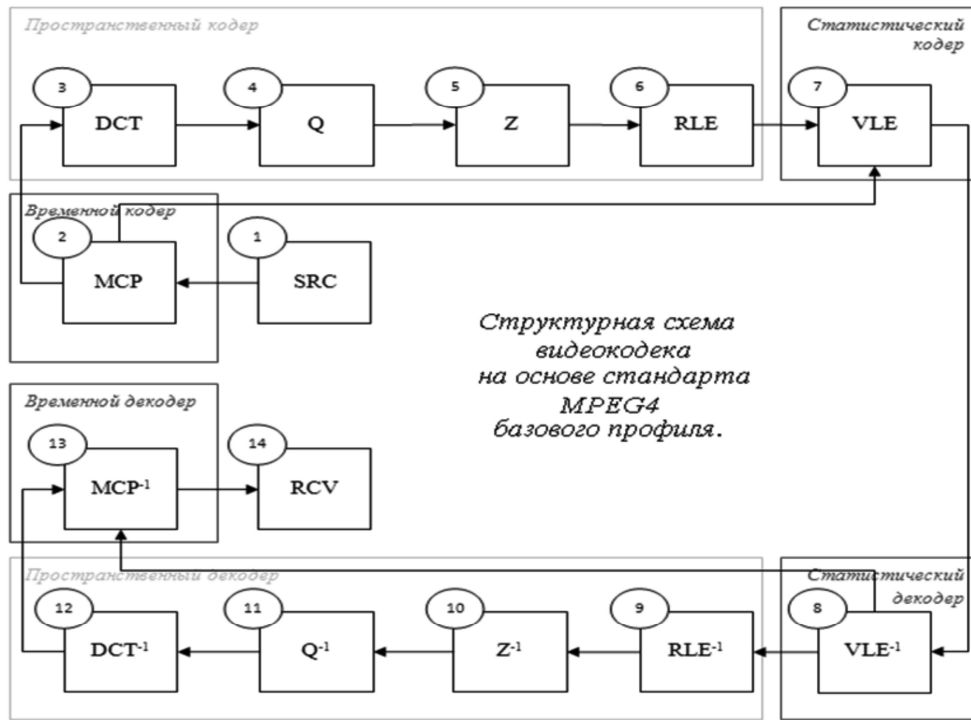


Рис.6. Структурная схема экспериментальных моделей видеокodeков

Для обеих моделей использовались одни и те же значения фундаментальных параметров видеокodeка (табл. 1). Формат кадра, размерность области поиска прогноза, шаг поиска прогноза и размерность макроблока задавались в пикселах. Шаг квантования задавался в относительных единицах яркости.

Таблица.1. Значения ключевых параметров использовавшихся моделей видеокodeков

Параметр	Значение
Режим цветности	чёрно-белый
Формат кадра	256 × 192
Размерность области поиска прогноза	24
Шаг поиска прогноза	2
Размерность макроблока	8
Шаг квантования	4

В ходе сравнительного анализа были получены следующие экспериментальные данные, характеризующие эффективность сравниваемых моделей видеокodeков по выборке из десяти кадров.

Таблица.2. Средние значения параметров качества сравниваемых моделей видеокodeков

Время обработки кадра, T, мс	Объём закодированного кадра, V, бит	Среднее абсолютное отклонение сигнала, E, отн. ед. яркости
Модель 1.	Модель 1.	Модель 1.
162	57138	2.32
Модель 2.	Модель 2.	Модель 2.
141	21342	3.11

По результатам экспериментов (табл. 2) можно сделать вывод, что предложенный метод формирования кадровой маски движения на основе её динамической регенерации обеспечивает существенно лучшие временные и информационные показатели эффективности видеокodeка по сравнению с существующими методами построения маски с применением принципов контекстно-

зависимого арифметического кодирования, но уступает им по качеству декодируемого изображения.

### Заключение

В ходе исследования был рассмотрен метод формирования бинарной кадровой маски движения на основе принципов контекстно-зависимого арифметического кодирования, предложен метод формирования бинарной кадровой маски движения на основе динамической регенерации подвижной части маски, а также проведён структурный синтез и сравнительный анализ моделей видеокodeков, использующих указанные методы. Экспериментально было показано, что предложенный метод построения кадровой маски движения способен обеспечить большую временную и информационную эффективность видеокodeка по сравнению с существующими методами при незначительном снижении качества декодируемого изображения.

### Литература

- Вернер М. Основы кодирования. – М.: Техносфера, 2004. – 288с.
- Ричардсон Я. Видеокodирование. H.264 и MPEG-4 – стандарты нового поколения. – М.: Техносфера, 2005. – 368с.
- Сэломон Д. Сжатие данных, изображений и звука.– М.:Техносфера, 2004. – 368с.

### THE METHOD OF THE ALFA-MASK CONSTRUCTION

*Minakov E.I., Kalistratov D.S.*

This article presents methods of the alfa-mask construction. Existing and proposed methods are discussed here. It is also contains results of the comparative analysis of the videocodec's models, which use considered methods.

## ОПРЕДЕЛЕНИЕ МЕСТОПОЛОЖЕНИЙ И МОЩНОСТЕЙ ИСТОЧНИКОВ ИЗЛУЧЕНИЯ ОДНОПОЗИЦИОННОЙ ЛОКАЦИОННОЙ СТАНЦИЕЙ

Самойленко М.В., к.т.н., доцент Московского авиационного института (национального исследовательского университета), e-mail: [Samoi.Mar@mail.ru](mailto:Samoi.Mar@mail.ru)

**Ключевые слова:** местоположение, источник излучения, приемная антенна, элемент разрешения, отображение, восстановление, псевдообратная матрица.

### Введение

Определение пространственных координат (местоположения) наблюдаемого объекта – источника первичного или вторичного излучения, является классической задачей локации. Достаточно прост способ определения направления на источник излучения по направлению оси диаграммы направленности (ДН), при котором на выходе приемной антенны фиксируется максимальная мощность. Однако для определения местоположения необходимо определить еще и наклонную дальность источника [1]. В активной локации для этого можно использовать дополнительный доплеровский канал. В пассивной локации такой принцип нереализуем. Альтернативный и единственно возможный для пассивной локации метод – это использование двухпозиционных (многопозиционных) локационных систем. Однако организация таких систем не всегда возможна или желательна.

В статье предлагается новый метод определения местоположений источников излучения однопозиционной приемной системой, который позволяет определять одновременно и мощности источников. Метод основан на специальной цифровой обработке измеренной мощности на выходе приемной антенны. Новизна и преимущества метода подтверждены патентом на изобретение [2]. Предлагаемый метод обладает свойством сверхразрешения и может быть использован в системах как активной, так и пассивной локации.

### Постановка задачи

Пусть  $p_{уст}(\mathbf{r})$  – функция, описывающая пространственное распределение мощностей источников излучения, где  $\mathbf{r}$  – координатный вектор. Обозначим ДН приемной антенны по мощности при направляющем векторе ее оси  $\mathbf{e}$  как  $w(\mathbf{r}, \mathbf{e})$ . Тогда мощность на выходе приемной антенны определится выражением

$$p_{вых}(\mathbf{e}) = \int_{\Omega} w(\mathbf{r}, \mathbf{e}) p_{уст}(\mathbf{r}) d\mathbf{r}, \quad (1)$$

где  $\Omega$  – контролируемая область пространства, включающая источники излучения.

Ставится задача восстановить пространственную функцию  $p_{уст}(\mathbf{r})$  по измеренной мощности на выходе

Представлен новый метод, который позволяет определять пространственные координаты и мощности источников излучения по измеренным значениям мощности на выходе однопозиционной приемной антенны посредством специальной цифровой обработки. Приведены математическое изложение метода и результаты компьютерных экспериментов.

приемной антенны  $p_{вых}(\mathbf{e})$ . Восстановленная функция  $\hat{p}_{уст}(\mathbf{r})$  будет определять распределение мощностей излучения по пространству. В идеале при нескольких точечных источниках это будет  $\delta$ -образная функция, выбросы которой характеризуют мощности излучения источников при соответствующим им пространственным координатам  $\mathbf{r}$ .

Восстановление функции, входящей в подынтегральное выражение, по множеству значений интеграла при известной весовой функции  $w(\mathbf{r}, \mathbf{e})$  является задачей многоканальной томографии, основанной на томографическом подходе к обработке сигналов и изложенной в [3].

В терминах томографического подхода  $p_{вых}(\mathbf{e})$  – это элемент интегрального отображения функции-оригинала  $p_{уст}(\mathbf{r})$  при характеристиках системы отображения, описываемых весовой функцией  $w(\mathbf{r}, \mathbf{e})$ . Задача состоит в восстановлении функции-оригинала по совокупности элементов ее отображения.

### Математическое обоснование метода

Разобьем контролируемую область пространства на элементарные объемы – элементы разрешения по местоположению. Пронумеруем их и зафиксируем местоположения координатными векторами  $\mathbf{r}_1, \mathbf{r}_2, \dots, \mathbf{r}_K$ , где  $K$  – число элементов разрешения в контролируемой области пространства. Начало системы координат разместим в точке расположения антенны. Качественно такое разбиение представлено на рис. 1.

Дискретизируем подынтегральное выражение в (1) по элементам разрешения и перейдем к интегральной сумме

$$p_{вых}(\mathbf{e}) = \sum_{k=1}^K w(\mathbf{r}_k, \mathbf{e}) p_{уст}(\mathbf{r}_k) \Delta\Omega,$$

где  $\Delta\Omega$  – размер элемента разрешения, который будем полагать равным единице,  $w(\mathbf{r}_k, \mathbf{e})$  – коэффициенты усиления приемной антенны по мощности для каждого элемента разрешения, которые полагаем известными при заданном направлении оси ДН  $\mathbf{e}$ .



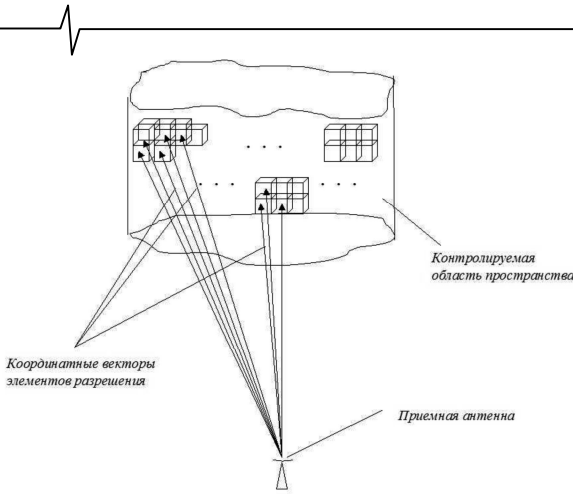


Рис. 1. Разбиение контролируемой области пространства на элементы разрешения

Полученное выражение перепишем в векторной форме

$$p_{\text{вых}}(\mathbf{e}) = \mathbf{w}^T(\mathbf{e})\mathbf{p}_{\text{ист}}, \quad (2)$$

где  $\mathbf{w}(\mathbf{e}) = [w(\mathbf{r}_1, \mathbf{e}) \ w(\mathbf{r}_2, \mathbf{e}) \ \dots \ w(\mathbf{r}_K, \mathbf{e})]^T$  – весовой вектор,  $\mathbf{p}_{\text{ист}} = [p_{\text{ист}}(\mathbf{r}_1) \ p_{\text{ист}}(\mathbf{r}_2) \ \dots \ p_{\text{ист}}(\mathbf{r}_K)]^T$  – вектор-оригинал, индекс  $T$  обозначает транспонирование.

Вектор-оригинал представляет собой дискретизированную по элементам разрешения искомую функцию  $p_{\text{ист}}(\mathbf{r})$ . Компоненты его равны либо нулю, если в соответствующем элементе разрешения нет источника, либо мощности источника, если он там есть.

Согласно томографическому подходу, сформируем отображение вектора-оригинала как совокупность элементов отображения (2), полученных при различающихся условиях интегрирования. Последние будем изменять, меняя направления оси ДН. Совокупность измененных мощностей и будет искомым отображением

$$\begin{aligned} p_{\text{вых}}(\mathbf{e}_1) &= \mathbf{w}^T(\mathbf{e}_1)\mathbf{p}_{\text{ист}}, \\ p_{\text{вых}}(\mathbf{e}_2) &= \mathbf{w}^T(\mathbf{e}_2)\mathbf{p}_{\text{ист}}, \\ &\dots \quad \dots \\ p_{\text{вых}}(\mathbf{e}_N) &= \mathbf{w}^T(\mathbf{e}_N)\mathbf{p}_{\text{ист}}, \end{aligned} \quad (3)$$

где  $N$  – число измерений.

Перепишем систему уравнений (3) в векторно-матричной форме

$$\mathbf{p}_{\text{вых}} = \mathbf{W}\mathbf{p}_{\text{ист}}, \quad (4)$$

где  $\mathbf{p}_{\text{вых}} = [p_{\text{вых}}(\mathbf{e}_1) \ p_{\text{вых}}(\mathbf{e}_2) \ \dots \ p_{\text{вых}}(\mathbf{e}_N)]^T$  – вектор-отображение,  $\mathbf{W} = [\mathbf{w}(\mathbf{e}_1) \ \mathbf{w}(\mathbf{e}_2) \ \dots \ \mathbf{w}(\mathbf{e}_N)]^T$  – матрица отображения размером  $N \times K$ .

В уравнении отображения (4) вектор отображения  $\mathbf{p}_{\text{вых}}$  известен: он объединяет измеренные и оцифрованные мощности на выходе приемной антенны. Известна и матрица отображения  $\mathbf{W}$ , которая включает коэффициенты усиления антенны по мощности для всех элементов разрешения при каждом направлении оси ее ДН  $\mathbf{e}_1, \mathbf{e}_2, \dots, \mathbf{e}_N$ .

Задача состоит в восстановлении вектора-оригинала  $\mathbf{p}_{\text{ист}}$  по его отображению (4). Решение может быть получено известными способами с использованием цифровой обработки, например, методом псевдообращения [4]:

$$\hat{\mathbf{p}}_{\text{ист}} = \mathbf{W}^+ \mathbf{p}_{\text{вых}}, \quad (5)$$

где индекс  $+$  обозначает операцию псевдообращения матрицы.

Восстановленный вектор (5) является решением поставленной задачи в дискретизированной форме. Компоненты его представляют собой распределение оценок мощностей излучения по элементам разрешения, что является оценкой искомой функции  $p_{\text{ист}}(\mathbf{r})$  с точностью элемента разрешения по местоположению. Местоположения источников определяются координатами элементов разрешения, которые соответствуют максимумам в полученном распределении, а мощности источников – соответствующими значениями этого распределения.

Таким образом, специальная цифровая обработка измеренной выходной мощности приемной антенны позволяет решить поставленную задачу с точностью до размера элемента разрешения по местоположению.

### Результаты компьютерных экспериментов

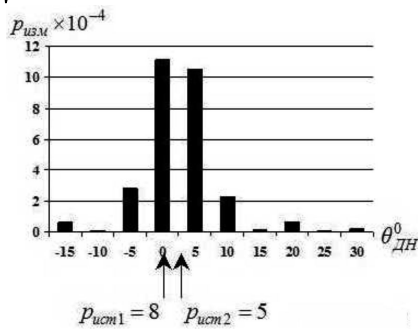
В качестве антенны в экспериментах использовалась 10-элементная линейная фазированная антенная решетка (ФАР). Контролируемая область пространства задавалась координатами граничных точек по оси  $x$ , совпадающей с поверхностью ФАР и по оси  $y$ , перпендикулярной ей:  $x_{\text{нач}} = 10\text{ м}$ ,  $x_{\text{кон}} = 70\text{ м}$ ,  $y_{\text{нач}} = 100\text{ м}$ ,  $y_{\text{кон}} = 1020\text{ м}$ . Число элементов разрешения определяется в экспериментах числом разбиений  $K$  контролируемой области пространства по оси  $x$ . За координаты элементов разрешения принимались координаты их центров. Во всех экспериментах в контролируемой области пространства присутствуют два источника излучения с одной и той же ординатой  $y = 1010\text{ м}$  и разными значениями абсцисс. Мощности источников составляют  $p_{\text{ист1}} = 8$ ,  $p_{\text{ист2}} = 5$ . Диапазон изменения направлений оси ДН при проведении измерений принят от  $-10^\circ$  до  $+30^\circ$ .

Проведено три эксперимента.

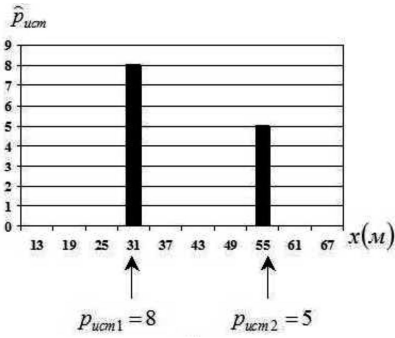
В первом эксперименте (рис. 2) число измерений  $N$  и число элементов разрешения  $K$  равны  $N = K = 10$ . В этом случае матрица отображения квадратная, и при линейно независимых строках ее псевдообратная матрица равна обратной. Тогда выражение (5) приобретает вид

$$\hat{\mathbf{p}}_{\text{ист}} = \mathbf{W}^{-1} \mathbf{p}_{\text{вых}}.$$

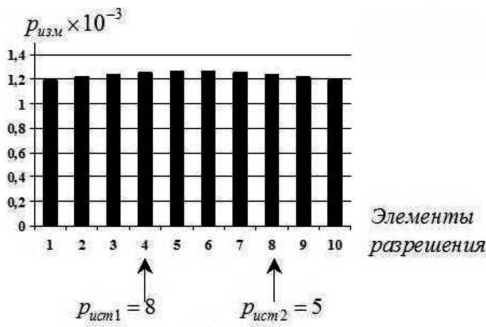
Размер элемента разрешения в первом эксперименте составил  $\Delta x = \frac{x_{\text{кон}} - x_{\text{нач}}}{10} = 6\text{ м}$ . Координаты источников заданы равными: 31 м, 1010 м и 55 м, 1010 м, и совпадают с координатами элементов разрешения. На приведенных на рис. 2 диаграммах соответствующие угловые положения источников показаны стрелками.



а)



б)



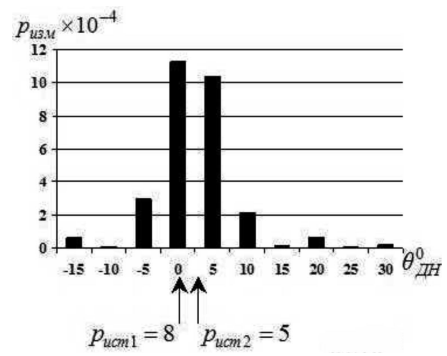
в)

Рис. 2. Измеренные мощности на выходе антенны (а), восстановленное по ним распределение мощностей источников по элементам разрешения (б) и результат сканирования по элементам разрешения (в) при  $x_{ист1} = 31$  м,  $x_{ист2} = 55$  м,  $N=10$ ,  $K=10$

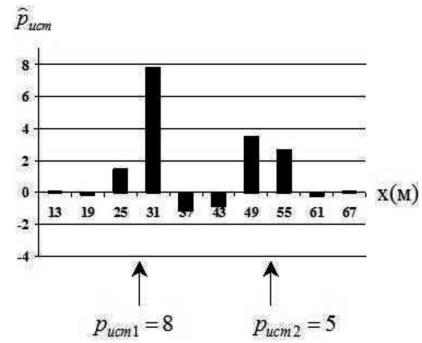
Как видно из диаграммы рис. 2, а), непосредственные измерения мощностей на выходе антенны при разных направлениях оси ее ДН  $\theta^0_{ДН}$  не дают представления о мощностях источников: они на 3-5 порядков меньше их истинных мощностей, что обусловлено удаленностью источников. На этой же диаграмме видно, что источники по непосредственным измерениям на выходе приемной антенны не разрешаются.

Предложенный метод позволяет по этим измерениям восстановить истинные мощности и координаты источников. Результаты восстановления показаны на диаграмме рис. 2, б). Как явствует из полученных результатов, метод позволил в первом эксперименте точно восстановить местоположения и мощности источников.

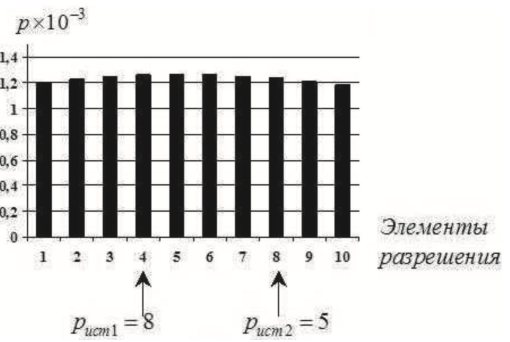
Следует заметить, что шаг изменения направлений оси ДН составлял  $5^0$ , в то время как направления на источники было  $1,8^0$  и  $3,1^0$ . Таким образом, результаты первого эксперимента демонстрируют сверхразрешающие возможности метода.



а)



б)



в)

Рис. 3. Измеренные мощности на выходе антенны (а), восстановленное по ним распределение мощностей источников по элементам разрешения (б) и результат сканирования по элементам разрешения (в) при  $x_{ист1} = 29,5$  м,  $x_{ист2} = 52,5$  м,  $N=10$ ,  $K=10$

Для наглядного подтверждения сверхразрешающих возможностей метода на рис. 2, в) приведены результаты поэлементного сканирования всех элементов разрешения. Как видно из диаграммы, источники и в этом случае не разрешаются, а мощности получаются на 3 порядка меньше действительных.

Второй эксперимент (рис. 3) отличается от первого положением источников: теперь они не совпадают с координатами элементов разрешения. Как видно из диаграммы на рис. 3, б), восстановленное распределение в этом случае «размывается» вблизи истинных положений источников, которые, однако, легко определяются по максимумам в полученном распределении. Что касается мощностей источников, то они восстановились хотя и не точно, но весьма близко к истинным значениям и на 3-4 порядка превосходят измеренные значения по степени близости к истинным значениям. Размер элемента разрешения в эксперименте по-прежнему составляет 6 м по оси  $x$ .

Дополнительный результат по сканированию всех элементов разрешения, приведенный на рис. 3, в), показывает, как и в первом эксперименте, невозможность даже самыми детальными прямыми измерениями оценить мощности или разрешить источники.

В третьем эксперименте (рис. 4) демонстрируются два дополнительных весьма полезных свойства предлагаемого метода: возможность повысить точность определения местоположений источников за счет уменьшения размера элемента разрешения и возможность сократить число измерений. В эксперименте принято число измерений  $N=7$  и число элементов разрешения  $K=16$ . Размер элемента разрешения составляет при этом

$$\Delta x = \frac{x_{кон} - x_{нач}}{16} = 3,75 \text{ м.}$$

Положения источников остались такими же, как во втором эксперименте.

На рис. 4, а) показаны измеренные мощности при соответствующих направлениях ДН антенны, а на рис. 4, б) – восстановленная по ним оценка распределения мощностей источников по элементам разрешения.

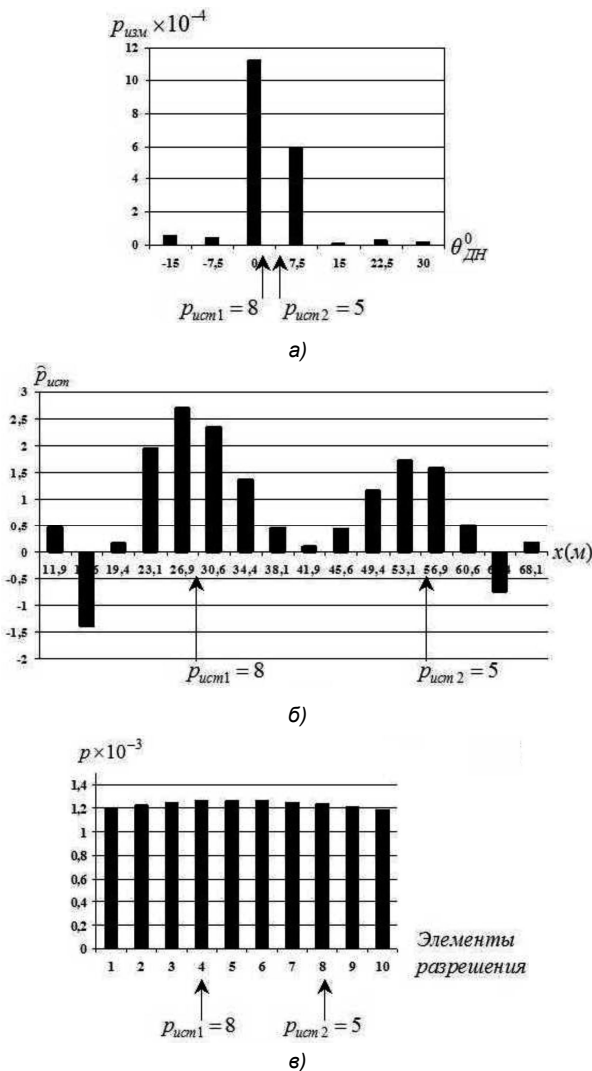


Рис. 4. Измеренные мощности на выходе антенны (а), восстановленное по ним распределение мощностей источников по элементам разрешения (б) и результат сканирования по элементам разрешения (в) при  $x_{ист1} = 29,5 \text{ м}$ ,  $x_{ист2} = 52,5 \text{ м}$ ,  $N=7$ ,  $K=16$ .

Теперь точность определения местоположения источников возросла: вместо 6м в предыдущих экспери-

ментах, она составляет 3,75 м. Что касается восстановления мощностей, то, как и во втором эксперименте, они размыты, однако по суммарному значению позволяют оценить истинные мощности источников, в то время как даже сканирование по элементам разрешения (рис. 4, в) не позволяет ни разрешить источники, ни определить хотя бы порядок их мощностей.

**Заключение**

Проведенные исследования позволяют сделать следующие выводы.

1. Математическое обоснование доказывает возможность определения местоположения источников излучения предложенным методом с точностью до элемента разрешения, размер которого задается априорно, используя цифровую обработку измеренной мощности на выходе приемной антенны,
2. Метод позволяет помимо местоположений оценивать мощности источников излучения, что является важной информацией для решения ряда прикладных задач, например, распознавания объектов.
3. Важным преимуществом метода является возможность определять местоположения и мощности источников, используя однопозиционную приемную систему, которая может быть пассивной.
4. Метод может использоваться для определения местоположений и мощностей сразу нескольких источников. При этом он не содержит поисковых процедур, что весьма привлекательно в плане сокращения времени оценивания и организации измерений.
5. Проведенные компьютерные эксперименты подтверждают работоспособность метода и демонстрируют его свойство сверхразрешения источников излучения, а также возможность не только повышать точность оценивания местоположений за счет уменьшения размера элемента разрешения, но и делать это при одновременном уменьшении числа измерений, что на практике может оказаться весьма полезным.

**Литература**

1. Сайбель А.Г. Основы радиолокации. – М.: Изд-во «Советское радио», 1961.
2. Самойленко М.В. Способ определения местоположения и мощностей источников излучения. // Патент № 2444740, Бюл. № 7 от 10.03.2012.
3. Самойленко М.В. Томография и аэрокосмические антенные системы. – М., изд-во МАИ, 2011.
4. Гантмахер Ф.Р. Теория матриц.– 4-е изд.– М.: Наука. Гл. ред. физ.-мат. лит., 1988.

**DETERMINATION OF LOCATIONS AND RADIATION POWER OF SOURCES WITH A SINGLE RADAR**

*Samoilenko M.V.*

The paper presents a new method which allows to determine spatial coordinates and radiation power of the sources by the measurements of the output power of the receiving antenna with the help of special digital processing. The mathematical statement of the method and the results of the computer experiments are given.

## ИСПОЛЬЗОВАНИЕ МАКСИМАЛЬНО ПРАВДОПОДОБНЫХ ОЦЕНОК ФРАКТАЛЬНОЙ РАЗМЕРНОСТИ В НЕГАУССОВСКИХ СТАТИСТИКАХ ДЛЯ ОБНАРУЖЕНИЯ РАДИОСИГНАЛОВ

*Паршин А.Ю., аспирант Рязанского государственного радиотехнического университета, e-mail: alex90fox@gmail.com*

*Паршин Ю.Н., д.т.н., профессор, зав. кафедрой радиотехнических устройств Рязанского государственного радиотехнического университета, e-mail: parshin.y.n@rsreu.ru*

**Ключевые слова:** обнаружение объектов, радиолокационное изображение, фрактальная размерность, гауссовская аппроксимация, максимальное правдоподобие.

### Введение

Современные способы маскирования подвижных и неподвижных радиолокационных объектов привели к появлению серьезных трудностей в обнаружении малоконтрастных целей на фоне подстилающей поверхности по энергетическим признакам. Кроме того, создание новых радиолокационных систем, к которым предъявляются повышенные требования по точности определения координат, производительности, а также более широкому кругу решаемых задач, привело к возникновению необходимости применения новых моделей и алгоритмов обработки принимаемых сигналов для обеспечения максимальной автоматизации процесса. Одним из них является фрактальный подход, широко используемый для решения различных радиотехнических задач [1-3].

Данный подход основан на принципе самоподобия и дробной меры природных процессов и объектов, а также связанных с ними сигналов и изображений [1-3]. При этом исследуемые явления рассматриваются не как простая совокупность отдельных элементов с определенными характеристиками, а как некоторая структура, обладающая внутренними топологическими связями между элементами и характеризующая сложный объект в целом. Оценка сложности структуры основывается на фрактальной размерности, которая является основным количественным показателем фрактальных структур. Особенностью фрактальной размерности  $D$  является дробный характер, что отличает ее от целой топологической размерности  $T$ .

Размерность имеет несколько определений, каждое из которых характеризует различные способы ее расчета [4]. При выполнении обнаружения данные представляются в виде временных рядов отсчетов сигнала, отраженного от объекта или подстилающей поверхности. Для обработки данных наиболее удобной является корреляционная размерность [4]. Корреляционная размерность имеет определенные вычислительные преимущества перед остальными способами определения фрактальной размерности и ее удобно использовать при решении задачи обнаружения объектов.

*Статья посвящена решению задачи обнаружения объектов по их радиолокационным изображениям, полученным РЛС с синтезированной апертурой на основе их фрактальных свойств. Показано, что достаточной статистикой для решаемой задачи является максимально правдоподобная оценка корреляционной размерности. Синтезирован фрактальный обнаружитель радиосигналов, проведен его анализ с применением гауссовской аппроксимации статистики и на основе степенного распределения. Выполнен синтез оптимального комплексированного энергетико-фрактального обнаружителя, получены аналитические выражения для характеристик обнаружения.*

### Постановка задачи

Как правило, задачу обнаружения необходимо решать в реальном времени, иногда допускается небольшая временная задержка. Количество данных, предоставляемых радиолокационной станцией в таком случае, достаточно ограничено, что приводит к неизбежным ошибкам при вычислении корреляционной размерности, которое требует бесконечное число отсчетов для безошибочной работы. Поэтому в реальных условиях целесообразно применять оценивание размерности. Наиболее точные результаты обеспечиваются при использовании алгоритма максимального правдоподобия [5].

Обычно наблюдается одномерная последовательность значений по одной координате. Для оценивания размерности сигнала, заданного временным рядом, требуется восстановить значения остальных координат, наиболее полно задающих состояние динамической системы. В работе [9] предложен способ реконструкции динамической системы путем использования задержанных во времени значений наблюдаемой компоненты в качестве значений ненаблюдаемых компонент  $\mathbf{x}_k = \{x(t_k), x(t_{k+1}), \dots, x(t_{k+T-1})\}$ , где  $T$  – размерность пространства вложения. Применение этого метода для вычисления корреляционной размерности  $d$  известно под названием алгоритма Грассбергера-Прокачия. Формирование данных для выполнения обнаружения производится по упорядоченной выборке отсчетов, полученной с учетом теоремы Такенса [9], в соответствии с которой из имеющейся упорядоченной последовательности  $N_S$  независимых отсчетов радиолокационного сигнала формируются векторы, число которых  $N = \frac{N_S}{T}$  зависит от выбранной размерности вложения.

Формирование векторов производится путем последовательного использования значений отсчетов в качестве соответствующих координат вектора. Измерение расстояний между векторами производится в соответствии с евклидовой метрикой по формуле

$$l_i = \sqrt{(x_1(t_k) - x_1(t_m))^2 + (x_2(t_k) - x_2(t_m))^2 + (x_3(t_k) - x_3(t_m))^2},$$

$$i = 1, \dots, N(N-1)/2,$$

$$k = 1, \dots, N-1, m = k+1, \dots, N.$$

Число неповторяющихся расстояний между векторами равно  $M = \frac{1}{2}(N-1)N$ , что позволяет считать их независимыми. При условии нормировки расстояний  $r_k = l_k / l_{\max}$  закон распределения вероятностей для расстояний между векторами задается степенной зависимостью  $F(r) = r^d$ , а функция плотности распределения вероятности имеет вид [5]

$$w(r) = \frac{dF(r)}{dr} = d \times r^{d-1}, \quad 0 < x < 1. \quad (1)$$

Полагая измеренные значения расстояний  $\mathbf{R} = \{r_1, r_2, \dots, r_M\}$  между векторами в фазовом пространстве статистически независимыми, запишем многомерную плотность распределения вероятностей

$$w(\mathbf{R}, d) = \prod_{i=1}^M w(r_i) = \prod_{i=1}^M d \times r_i^{d-1}. \quad (2)$$

Оценка максимального правдоподобия корреляционной размерности получена в работе [5]

$$\hat{d} = -\frac{M}{\sum_{i=1}^M \ln r_i}, \quad (3)$$

а минимальная дисперсия ошибки оценивания определяется из неравенства Рао-Крамера

$$D_{err} = -\frac{1}{\mathbf{M} \left\{ \frac{\partial^2 \ln w(\mathbf{R}, d)}{\partial d^2} \right\}} = \frac{d^2}{M}. \quad (4)$$

Для обнаружения и различения сигналов и объектов достаточной статистикой является отношение правдоподобия или его логарифм, который для распределения (1) имеет вид

$$z = \ln \Lambda = \ln \frac{w_1(\mathbf{r})}{w_0(\mathbf{r})} = \sum_{i=1}^M [\ln w_1(r_i) - \ln w_0(r_i)] =$$

$$= \sum_{i=1}^M \left[ \ln \frac{D_1}{D_0} + (D_1 - D_0) \ln r_i \right]. \quad (5)$$

**Синтез и анализ фрактального обнаружителя**

При наличии априорной информации о корреляционной размерности объекта  $d_1$  и фона  $d_0$ , оптимальный алгоритм обнаружения получается на основе сравнения логарифма отношения правдоподобия с порогом

$$\hat{\theta} = \begin{cases} 1, & z = (d_1 - d_0) \sum_{i=1}^M \ln r_i > h \\ 0, & z = (d_1 - d_0) \sum_{i=1}^M \ln r_i \leq h \end{cases}, \quad (6)$$

где  $h$  - значение порога обнаружения, выбираемого из заданной вероятности ложной тревоги.

Сравнивая статистику обнаружения в алгоритме (6) с выражением для оценки максимального правдоподобия корреляционной размерности (3), можно заключить, что данная оценка также является достаточной статистикой обнаружения. Алгоритм обнаружения получается подстановкой оценки  $\hat{d}$  в алгоритм максимального правдоподобия

$$\hat{\theta} = \begin{cases} 1, & z = -\frac{M}{\hat{d}}(d_1 - d_0) > h \\ 0, & z = -\frac{M}{\hat{d}}(d_1 - d_0) \leq h. \end{cases} \quad (7)$$

После пересчета порога можно перейти к виду  $z = \hat{d}$ .

На рис.1 приведена функциональная схема синтезированного фрактального обнаружителя радиосигналов, отражающая последовательность выполнения необходимых операций формирования решающей статистики.

**Гауссовская аппроксимация распределения статистики**

В случае достаточно большого числа  $M$  независимых расстояний  $r_i$  распределение статистики  $z = \hat{d}$  имеет асимптотически гауссовский вид

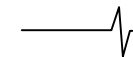
$$w(z) = \frac{1}{\sqrt{2\pi}} e^{-(z-m_z)/2D_z}, \quad \text{где математическое ожидание}$$

$m_z = d$ , дисперсия  $D_z = \frac{d^2}{M}$ , а отношение сигнал-

шум после обработки  $q_z = m_z^2 / D_z = M$ .



Рис. 1. Функциональная схема фрактального обнаружителя радиосигналов



Результатами анализа фрактального обнаружителя являются характеристики обнаружения – вероятность ложной тревоги  $F$  и вероятность правильного обнаружения  $P$ , в зависимости от параметров сигналов. Для случая гауссовской аппроксимации характеристики имеют вид

$$P = \frac{1}{2} \operatorname{erfc} \left( \frac{h - m_{z1}}{\sqrt{2D_{z1}}} \right), \quad F = \frac{1}{2} \operatorname{erfc} \left( \frac{h - m_{z0}}{\sqrt{2D_{z0}}} \right), \quad (8)$$

где  $\operatorname{erfc}(x) = \frac{2}{\sqrt{\pi}} \int_x^{\infty} e^{-t^2} dt$ ,  $m_{zi}$ ,  $D_{zi}$  – математическое ожидание и дисперсия статистики при наблюдении объекта ( $i = 1$ ) и фона ( $i = 0$ ).

При подстановке рассчитанных значений математического ожидания и дисперсии оценки для  $\alpha = 0$ , а также после расчета порога обнаружения по заданной вероятности ложной тревоги получается следующее выражение для вероятности правильного обнаружения [7]:

$$P = \frac{1}{2} \operatorname{erfc} \left( \frac{d_0}{d_1} \left( \sqrt{\frac{M}{2}} + \operatorname{erfc}^{-1}(2F) \right) - \sqrt{\frac{M}{2}} \right), \quad (9)$$

где  $\operatorname{erfc}^{-1}(x)$  – функция обратная  $\operatorname{erfc}(x)$ .

#### Негауссовское распределение статистики

Для небольшого числа расстояний  $M$  гауссовская аппроксимация распределения статистики неприменима, поэтому необходимо более точно определить вероятностные характеристики суммы слагаемых  $z = \sum_{i=1}^M \ln r_i = \sum_{i=1}^M y_i$ . Для каждого из слагаемых методом замены переменных с учетом (1) получена плотность распределения вероятностей  $w_y(y) = de^{dy}$ , а для суммы соответствующее распределение получено индуктивным методом [6]

$$w_z(z) = (-1)^{M-1} d^M \frac{z^{M-1}}{(M-1)!} e^{dz}, \quad -\infty < z < 0.$$

Расчет вероятности правильного обнаружения и ложной тревоги проведем с использованием равенства  $(M-1)! = \Gamma(M)$ :

$$P = \int_h^{\infty} (-1)^{M-1} d_1^M \frac{z^{M-1}}{(M-1)!} e^{d_1 z} dz = \frac{\gamma_u(d_1 h, M)}{\Gamma(M)},$$

$$F = \frac{\gamma_u(d_0 h, M)}{\Gamma(M)}, \quad (10)$$

где  $\gamma_u(d_i h, M) = \int_{-d_i h}^{\infty} t^{M-1} e^{-t} dt$ ,  $i = 0; 1$  – верхняя

неполная гамма-функция, а порог обнаружения  $h$  рассчитывается по заданной вероятности ложной тревоги методом дихотомии.

На рис. 2 приведены характеристики обнаружения

фрактального обнаружителя для различных аппроксимаций решающей статистики при различных значениях вероятности ложной тревоги и различных значениях параметра  $q_{fr} = \frac{d_1}{d_0}$ ;  $M = 500$ .

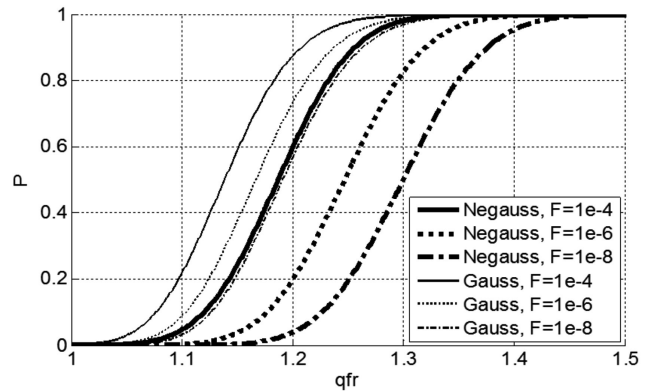


Рис. 2. Характеристики обнаружения фрактального обнаружителя

Расчеты показали, что при  $M > 500$  гауссовская аппроксимация дает достаточно точные результаты, а при малых значениях  $M$  гауссовская аппроксимация дает завышенные значения вероятности правильного обнаружения.

#### Комплексированный оптимальный обнаружитель

В радиолокационную систему могут входить несколько устройств обработки информации, решающих одну и ту же задачу. При этом возникает проблема их оптимального объединения в единый комплекс обработки информации. В рассматриваемом случае объединение производится на уровне решений, принимаемых каждым из обнаружителей – фрактальным и энергетическим, с целью повышения качества обнаружения объекта.

Принцип работы энергетического обнаружителя 1 (рис. 3) основан на использовании энергетических свойств наблюдаемого сигнала. При использовании модели наблюдаемого сигнала в виде некоррелированных гауссовских отсчетов с нулевым математическим ожиданием и дисперсией  $D_x$  статистика оптимального обнаружения пропорциональна энергии выборки наблюдаемого сигнала  $z_1 = \sum_{i=1}^{N_S} |x_i|^2$ , а обнаружение производится

путем сравнения с порогом полученной статистики [7]:

$$\delta_1 = \begin{cases} 0, & z_1 \geq h_1 \\ 1, & z_1 < h_1 \end{cases}.$$

При гауссовской аппроксимации параметры статистики обнаружения равны  $m_{z1} = \sqrt{D_x} N_S$ ,  $D_{z1} = 2N_S D_x$ , при точном описании статистики, заданном законом  $\chi^2$  [6]

$$w(z_1) = \frac{1}{2^{\frac{N_S}{2}} \Gamma\left(\frac{N_S}{2}\right) D_x^{\frac{N_S}{4}}} z_1^{\frac{N_S}{2}-1} e^{-\frac{z_1}{2D_x}}, \quad (11)$$



$0 < z_1 < \infty$ .

Характеристики обнаружения энергетического обнаружителя рассчитываются по выражениям, аналогичным (9) для гауссовской статистики и (10) для негауссовской статистики:

$$P_1 = \frac{\gamma_u(h/2\sqrt{D_{x1}}, N_S/2)}{\Gamma(N_S/2)},$$

$$F_1 = \frac{\gamma_u(h/2\sqrt{D_{x0}}, N_S/2)}{\Gamma(N_S/2)}. \quad (12)$$

Фрактальный обнаружитель 2 (рис. 3) выполняет преобразование наблюдаемого сигнала в совокупность расстояний между векторами  $\mathbf{R}$ , а затем формирует решающую статистику (5) и после сравнения с пороговым значением, в соответствии с алгоритмом (6), формирует решение  $\delta_2$ .

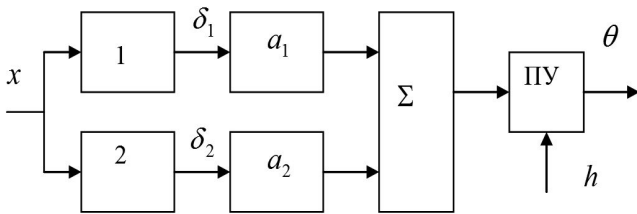


Рис. 3. Структурная схема комплексированного обнаружителя

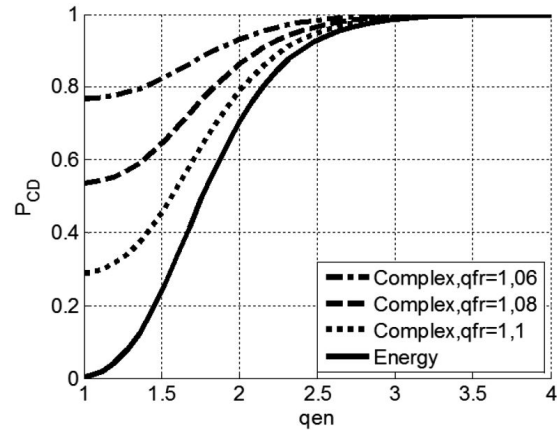
Оптимальный по критерию максимального правдоподобия комплексированный алгоритм обнаружения [7] имеет вид  $z = \delta_1 a_1 + \delta_2 a_2$ , где параметры  $a_1, a_2$  принимают значения  $a_i = \ln \frac{P_i(1-P_i)}{F_i(1-F_i)}$ ,  $i = 1; 2$ , а результаты каждого из обнаружителей  $\delta_i = [0, 1]$ .

Таким образом, статистика  $z$  комплексированного обнаружителя может принимать только 4 значения:  $z = 0$  при  $\delta_1 = 0$  и  $\delta_2 = 0$ ,  $z = a_1 + a_2$  при  $\delta_1 = 1$  и  $\delta_2 = 1$  или, если одна из величин  $\delta_i$  равна 1, а вторая 0, то сумма равна  $a_1$  или  $a_2$ . Вероятности превышения порога  $h$  при расположении его в различных областях рассчитываются в зависимости от результатов срабатывания обнаружителей  $\delta_1 \delta_2 - 00, (01) \vee (10), (11), (01) \vee (10) \vee (11)$ .

Расчет характеристик обнаружения производится при различных значениях отношения мощностей  $q_{en} = \frac{D_{x1}}{D_{x0}}$  и отношения корреляционных размерностей  $q_{fr} = \frac{d_1}{d_0}$  сигналов. Вероятность ложной тревоги обнаружителей принята одинаковой и равной  $F_1 = F_2 = \sqrt{F}$  при  $\delta_1 \delta_2 = 11$  и  $F_1 = F_2 = F/2$  при  $\delta_1 \delta_2 = (01) \vee (10)$  и  $\delta_1 \delta_2 = (01) \vee (10) \vee (11)$ , где  $F$  – вероятность лож-

ной тревоги комплексированного обнаружителя. В дальнейшем при расчете числа расстояний  $M$  считаем, что размерность пространства вложения равна  $T = 3$ .

На рис. 4 приведены характеристики обнаружения комплексированного обнаружителя при использовании негауссовских распределений (10), (12) решающих статистик фрактального и энергетического обнаружителей соответственно при различных значениях  $q_{fr}$ .



$N=150, \delta_1 \delta_2 = (01) \vee (10) \vee (11), F=0,01$

Рис. 4. Характеристики обнаружения комплексированного обнаружителя при использовании негауссовского распределения статистик

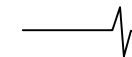
Их графиков характеристик (рис. 4) видно, что применение комплексированного энергетико-фрактального обнаружителя обеспечивает выигрыш в вероятности правильного обнаружения по сравнению с одиночным энергетическим обнаружителем. Выбор пороговой области существенно влияет на качество обнаружения и зависит от значений параметров  $q_{fr}, q_{en}$ . В большинстве случаев целесообразно выбирать  $\delta_1 \delta_2 = (01) \vee (10) \vee (11)$ , что обеспечивает наибольшую вероятность правильного обнаружения и выигрыш за счет использования фрактальных свойств сигналов.

**Заключение**

Проведенный синтез фрактального обнаружителя по критерию максимального правдоподобия позволяет производить обнаружение малоcontrastных объектов на фоне подстилающей поверхности. Одной из основных задач практической реализации фрактального обнаружителя является получение совокупности независимых расстояний между векторами. Выполнение условия независимости позволяет минимизировать вероятности ошибочных решений при ограниченном объеме наблюдаемых данных.

Получены расчетные соотношения для характеристик обнаружения с использованием аналитических выражений для дисперсии ошибок оценивания корреляционной размерности как с учетом эффекта усечения корреляционного интеграла, так и без него. Показано, что максимально правдоподобные оценки корреляционной размерности также являются достаточной статистикой для задачи обнаружения.

Разработан алгоритм комплексирования энергетического и фрактального обнаружителей и проведен анализ



характеристик обнаружения. Результаты анализа свидетельствуют о возможности обнаружения мало контрастных объектов путем оптимального использования его радиоярких и фрактальных свойств.

#### Литература

1. Потапов А.А., Герман В.А. Фрактальный непараметрический обнаружитель радиосигналов // Радиотехника. 2006. № 5. С. 30 – 36.
2. Гуляев Ю.В., Никитов С.А., Потапов А.А., Герман В.А. Идеи скейлинга и дробной размерности в схеме фрактального обнаружителя радиосигналов // Радиотехника и электроника. 2006. Т. 51, № 8. С. 968 – 975.
3. Сосулин Ю.Г., Русскин А.Б. Фрактальное обнаружение протяженных мало контрастных объектов на изображениях // Радиотехника. 2009. № 12. С. 48-57.
4. Паркер Т.С., Чжуа Л. О. Введение в теорию хаотических систем для инженеров // ТИИЭР, 1987, Т. 75, № 8. С. 6-40.
4. Luciana De Luca, Dario Luzio, Massimo Vitale. A ML Estimator of the Correlation Dimension for Left-Hand Truncated Data Samples // Pure and applied geophysics, V.159, № 11-12, 2002. P. 2789-2803
5. Справочник по теории вероятностей и математической статистике / В.С. Королюк, Н.И. Портенко, А.В. Скороход, А.Ф. Турбин. – М.: Наука. Главная редакция физико-математической литературы, 1985. – 640 с.

6. Сосулин Ю.Г. Теоретические основы радиолокации и радионавигации: Учеб. пособие для вузов – М.: Радио и связь, 1992.

7. Потапов А.А. Фракталы в радиофизике и радиолокации: топология выборки. – М.: Университетская книга, 2005. – 847 с.

8. Takens F. Detecting strange attractors in turbulence // Lecture Notes in Mathematics, V. 898, 1981. P. 366-381.

#### USAGE OF FRACTAL PROPERTIES AND MAXIMUM LIKELIHOOD ESTIMATIONS IN NON-GAUSSIAN STATISTICS FOR RF SIGNALS DETECTION

*Parshin A. Yu.*

The article deals with problem of objects detection by its radar data, obtained from RCA. Method is based on fractal properties of this data. Algorithm of fractal dimension estimation by the method of maximum likelihood is proposed. Fractal detector of RF signals is synthesized, analysis with Gaussian approximation of correlation dimension estimation is made, computation of precise distribution of estimation is done. Complex energy-fractal detector is synthesized, its analysis is made and expressions for detection characteristics are got.



## 11<sup>th</sup> IEEE EAST-WEST DESIGN & TEST SYMPOSIUM (EWDTS 2013)

состоится в городе Ростов на Дону, Россия,  
27-30 сентября 2013 года

Цель симпозиума – расширение международного сотрудничества и обмен опытом между ведущими учеными Западной и Восточной Европы, Северной Америки и других стран в области автоматизации проектирования, тестирования и верификации электронных компонентов и систем. Симпозиум проводится, как правило, в странах бассейнов Черного и Балтийского морей, Центральной Азии. Оргкомитет приглашает ученых, аспирантов и студентов принять участие в работе EWDTS'13.

Симпозиум будет проходить в Ростове на Дону – крупнейшем научном и образовательном центре Южного федерального округа России.

**Важные даты:**                      **Срок подачи докладов:** 15 июля, 2013  
**Итоги рецензирования:** 1 августа, 2013

**Детальная информация и регистрация докладов:** <http://www.ewdtest.com/conf>

**Адрес оргкомитета:** Проф. Владимир Хаханов, кафедра Автоматизации проектирования вычислительной техники Харьковского национального университета радиоэлектроники, пр. Ленина 14, Харьков, 61166, Украина. Тел.: +380-57-702-13-26,  
E-mail: [hahanov@kture.kharkov.ua](mailto:hahanov@kture.kharkov.ua), [www.ewdtest.com/conf/](http://www.ewdtest.com/conf/)

УДК 519.65

## АЛГОРИТМЫ ВЫДЕЛЕНИЯ И АНАЛИЗА АНОМАЛИЙ В ПАРАМЕТРАХ КРИТИЧЕСКОЙ ЧАСТОТЫ ИОНОСФЕРЫ $f_0F2$ НА ОСНОВЕ СОВМЕЩЕНИЯ ВЕЙВЛЕТ-ПРЕОБРАЗОВАНИЯ И АВТОРЕГРЕССИОННЫХ МОДЕЛЕЙ

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

*Глушкова Н.В., м.н.с. Института космических исследований и распространения радиоволн ДВО РАН, аспирант Камчатского государственного технического университета, e-mail: nv.glushkova@ya.ru*

*Полозов Ю.А., научный сотрудник Института космических исследований и распространения радиоволн ДВО РАН, доцент кафедры систем управления Камчатского государственного технического университета, e-mail: up\_agent@mail.ru*

**Ключевые слова:** анализ аномалий, частоты ионосферы, многокомпонентное моделирование, вейвлет-преобразование, авторегрессия.

### Введение

Одной из важных задач обработки и анализа ионосферных данных является задача контроля состояния ионосферы и автоматическое выделение и интерпретация аномалий, возникающих в периоды ионосферных возмущений [1-4]. Априорная неопределенность знаний о структуре регистрируемых ионосферных параметров, отсутствие формальной модели их описания, а также априорная неопределенность условий их формирования делают поставленную задачу весьма сложной. Предметом данных исследований являются регистрируемые временные ряды критической частоты ионосферы  $f_0F2$ . На фоне регулярных суточных и сезонных изменений в данных  $f_0F2$  наблюдаются аномалии, длительностью от несколько десятков минут до нескольких часов [3-9]. Данные аномалии обусловлены активностью Солнца, в сейсмоактивных областях они могут возникать в периоды повышения сейсмической активности [3-7, 9]. Сложная структура аномалий не позволяет использовать для их выделения и анализа традиционные методы анализа временных рядов [4-7, 9]. Предложенные в данной статье методы и алгоритмы основаны на применении вейвлет-преобразования [10, 11, 13]. Ввиду наличия быстрых алгоритмов преобразования данных и большого разнообразия базисных функций с компактными носителями этот аппарат позволяет детально изучить внутреннюю структуру данных и выделить локальные особенности различной формы и длительности [10, 13]. На основе совместного применения вейвлет-преобразования с моделями авторегрессии – проинтегрированного скользящего среднего (АРПСС), авторами статьи предложен метод многокомпонентного моделирования данных  $f_0F2$ . Идентификация моделей основана на применении операции кратномасштабного анализа и представлении данных в виде аппроксимирующей и детализирующих компонент. На основе конструкции дискретного вейвлет-преобразования, предложены алгоритмы, позволяющие в автоматиче-

Представлен метод многокомпонентного моделирования ионосферных данных, основанный на совмещении вейвлет-преобразования с моделями авторегрессии – проинтегрированного скользящего среднего, позволяющий выполнить анализ и прогноз данных. Выделены особенности, связанные с солнечной активностью, а также возникающие в периоды сильных землетрясений на Камчатке. Предложены алгоритмы анализа ионосферных данных, основанные на конструкции дискретного вейвлет-преобразования, позволяющие автоматически выделить аномалии в ионосферной плазме и оценить их параметры. Для оценки метода и алгоритмов использовались данные станции «Паратунка» (п-ов Камчатка) за период 2001-2011гг.

ском режиме выделить аномалии, оценить их временную протяженность и интенсивность.

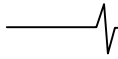
Выполнено моделирование и анализ данных, изучена их внутренняя структура. Выделены характерные структуры данных и аномалии, возникающие в периоды ионосферных возмущений. Анализ полученных результатов показал, что данные аномалии возникают в периоды повышенной солнечной активности и могут наблюдаться накануне сильных землетрясений на Камчатке.

### Многокомпонентное моделирование, прогнозирование и анализ данных на основе совмещения кратномасштабного анализа и моделей АРПСС

В качестве базового пространства регистрируемых дискретных данных рассмотрим замкнутое пространство  $V_j = \text{clos}_{L^2(R)}(2^j \varphi(2^j t - k)) : k \in Z$  с разрешением  $j = 0$ , порожденное скейлинг-функцией  $\varphi \in L^2(R)$  [10, 13], где  $L^2(R)$  – пространство Лебега,  $Z$  – множество целых чисел. Используя кратномасштабное разложение [13] до уровня  $m$ , исходную непрерывную функцию  $f(t)$ , полученную по совокупности регистрируемых дискретных данных, представим в виде

$$f_0(t) = \sum_{j=-1}^{-m} (g_j(t) + e_j(t)) + f_{-m}(t), \quad (1)$$

где  $f_{-m}(t) \in V_{-m}$ ,  $g_j \in W_j$ ,  $W_j$  – пространство с разрешением  $j$ , порожденное вейвлет-базисом



$\Psi_{j,k}(t) = 2^{j/2} \Psi(2^j t - k)$ , аппроксимирующая компонента  $f_{-m}(t) = \sum_k c_{-m,k} \phi_{-m,k}(t)$ , где коэффициенты разложения  $c_{-m,k} = \langle f, \phi_{-m,k} \rangle$ , детализирующие компоненты  $g_j(t) = \sum_k d_{j,k} \Psi_{j,k}(t)$ , где коэффициенты разложения  $d_{j,k} = \langle f, \Psi_{j,k} \rangle$ ;  $e_j(t)$  – шумовые составляющие (предполагается, что шум белый).

В работе [8] показано, что, имея представление данных в виде (1), подавление шума может быть выполнено на основе применения пороговой функции

$$P_T(x) = \begin{cases} x, & \text{если } |x| \geq T \\ 0, & \text{если } |x| < T \end{cases} \quad (2)$$

для каждой детализирующей компоненты  $g_j(t)$ , где порог  $T = \sigma^2$ ,  $\sigma^2$  – дисперсия шума. Следуя работам [8, 10], дисперсию шума  $\sigma^2$  можно оценить на основе соотношения:

$$\sigma^2 \approx \text{Med} \left| \langle f, \Psi_{j,k} \rangle \right|_{0 \leq k < N},$$

где  $\text{Med}$  – медиана,  $j = -1$ ,  $N$  – длина компоненты. В работе [11] показано, что данный способ подавления шума позволяет получить почти оптимальные минимаксные оценки.

Для выделения компонент конструкции (1), описывающих *характерные особенности* данных, и идентификации их параметров применим следующие операции:

1. Используя традиционные подходы, предложенные в работе [12], для уровня разложения  $m$  выберем модели из класса моделей авторегрессии-проинтегрированного скользящего среднего для аппроксимации каждой из компонент  $f_{-m}(t) = \sum_k c_{-m,k} \phi_{-m,k}(t)$  и

$$g_j(t) = \sum_k d_{j,k} \Psi_{j,k}(t), \quad j = \overline{-1, m}, \text{ и оценим параметры моделей. Получим представление компонент в виде:}$$

$$f_{-m}(t) = \sum_k s_{-m,k}(t) \phi_{-m,k}(t), \quad (3)$$

где  $s_{-m,k}(t) = \sum_{l=1}^p \gamma_{-m,l} \omega_{-m,k-l}(t) - \sum_{n=1}^h \theta_{-m,n} a_{-m,k-n}(t)$  – оценочное значение аппроксимирующей компоненты,  $\omega_{-m,k} = \nabla^V c_{-m,k}$ ,  $\nabla^V$  – оператор взятия разности порядка  $V$ ,  $p, \gamma_{-m,l}$  – порядок и параметры авторегрессии аппроксимирующей компоненты,  $h, \theta_{-m,n}$  – порядок и параметры скользящего среднего аппроксимирующей компоненты,  $a_{-m,k}$  – остаточные ошибки модели аппроксимирующей компоненты,

$$g_j(t) = \sum_k s_{j,k}(t) \Psi_{j,k}(t), \quad (4)$$

где  $s_{j,k}(t) = \sum_{l=1}^{p_j} \gamma_{j,l} \omega_{j,k-l}(t) - \sum_{n=1}^{h_j} \theta_{j,n} a_{j,k-n}(t)$  – оценочное значение детализирующей компоненты;  $\omega_{j,k} = \nabla^{V_j} d_{j,k}$ ,  $\nabla^{V_j}$  – оператор взятия разности порядка  $V_j$ ,  $p_j, \gamma_{j,l}$  – порядок и параметры авторегрессии детализирующей компоненты с разрешением  $j$ ,  $h_j$ ;  $\theta_{j,n}$  – порядок и параметры скользящего среднего детализирующей компоненты с разрешением  $j$ ;  $a_{j,k}$  – остаточные ошибки модели детализирующей компоненты с разрешением  $j$ .

2. Выполним диагностические проверки полученных моделей компонент. Если диагностические проверки модели компоненты подтверждают ее адекватность, то будем считать, что модель компоненты готова к использованию и данная компонента является характерной.

3. Объединим модели выделенных характерных компонент вида (3) и (4) в общую многокомпонентную конструкцию. Получим *многокомпонентную модель* вида

$$f(t) = \sum_{\mu=1, M} \sum_{k=1, N_j^\mu} s_{j,k}^\mu(t) b_{j,k}^\mu(t), \quad (5)$$

где  $s_{j,k}^\mu(t) = \sum_{l=1}^{p_j^\mu} \gamma_{j,l}^\mu \omega_{j,k-l}^\mu(t) - \sum_{n=1}^{h_j^\mu} \theta_{j,n}^\mu a_{j,k-n}^\mu(t)$  – оценочное значение  $\mu$ -ой компоненты;  $\gamma_{j,l}^\mu$  – параметры авторегрессии компоненты с номером  $\mu$ ,  $\omega_{j,k}^\mu(t) = \nabla^{V_j} \beta_{j,k}^\mu(t)$ ,  $\beta_{j,k}^1 = c_{j,k}$ ,  $\beta_{j,k}^\mu = d_{j,k}$ ,  $\mu = \overline{2, M}$ ;  $p_j^\mu$  – порядок авторегрессионной модели компоненты с номером  $\mu$ ;  $h_j^\mu, \theta_{j,n}^\mu$  – порядок модели и параметры скользящего среднего модели компоненты с номером  $\mu$ ;  $a_{j,k}^\mu$  – остаточные ошибки модели компоненты с номером  $\mu$ ,  $M$  – количество выделенных на основе п. 2 характерных компонент;  $N_j^\mu$  – длина компоненты с номером  $\mu$ ;  $b_{j,k}^1 = \phi_{j,k}$  – скейлинг-функция,  $b_{j,k}^\mu = \Psi_{j,k}$ ,  $\mu = \overline{2, M}$  – вейвлет-базис компоненты с номером  $\mu$ ,  $j$  – разрешение.

Если данные содержат аномалию, то произойдет изменение их структуры. Поэтому процедура выделения аномалий в компонентах разрешения  $j$  может быть построена на обработке остаточных ошибок  $a_{j,k}^\mu$  моделей компонент при выполнении операции прогнозирования. Данная процедура предполагает следующие операции:

1. Прогнозирование значения  $s_{j,k+q}^\mu$ ,  $q \geq 1$  определяет прогноз  $s_{j,k}^\mu$  в момент  $t = k$  с упреждением  $q$ . Значение  $s_{j,k+q}^\mu$  на основе модели (5) определяется как

$$s_{j,k+q}^{\mu}(t) = \sum_{l=1}^{p_j^{\mu}} \gamma_{j,l}^{\mu} \omega_{j,k+q-l}^{\mu}(t) - \sum_{n=1}^{h_j^{\mu}} \theta_{j,n}^{\mu} a_{j,k+q-n}^{\mu}(t).$$

2. Остаточные ошибки компоненты модели с номером  $\mu$  разрешения определяются как разность между прогнозными и фактическими значениями данных в момент времени  $t = k + q$ :  $a_{j,k+q}^{\mu}(t) = s_{j,k+q,прогноз}^{\mu}(t) - s_{j,k+q,фактич}^{\mu}(t)$ .

3. Обнаружение аномалии в компоненте с номером  $\mu$  разрешения  $j$  можно выполнить на основе проверки условия:

$$D_{U_j} = \frac{1}{U_j} \sum_{q=1}^{U_j} (a_{j,k+q}^{\mu}(t))^2 > T_{A_j}, \quad (6)$$

где  $T_{A_j}$  – некоторое наперед заданное пороговое значение, определяющее наличие в данных аномалии разрешения  $j$ ,  $U_j$  - длина окна наблюдения для разрешения  $j$ .

### Выделение аномалий и оценка их параметров на основе дискретного вейвлет-преобразования

На основе дискретного вейвлет-преобразования для функции  $f \in L^2(R)$  получаем следующее представление в виде вейвлет-схемы [10]:

$$f(t) = \sum_{j,k} \alpha_{j,k} \Psi_{j,k}(t), \quad (7)$$

где  $\{\Psi_{j,k}\}_{(j,k) \in \mathbb{Z}^2}$  – ортонормированный базис пространства  $L^2(R)$ . Коэффициенты  $\bar{\alpha}_j = \{\alpha_{j,k}\}_{k \in \mathbb{Z}}$  являются результатом отображения  $f$  в пространство с разрешением  $j$ ,  $\alpha_{j,k} = \langle f, \Psi_{j,k} \rangle$ .

Важность представления  $f$  в виде (7) состоит в том, что разномасштабные составляющие данных сложной структуры рассортировываются и хранятся в различных пространствах  $W_j$  с разрешением  $j$ .

Процедуру выделения аномальных особенностей построим на основе нелинейных отображений вида:

$$f_A(t) = \sum_{(j,k) \in I_A} \alpha_{j,k} \Psi_{j,k}(t), \quad (8)$$

где  $f_A$  – проекция  $f$  на  $A$  векторов, индексы которых содержатся в некотором множестве  $I_A$ ;  $\{\alpha_{j,k}\}_{(j,k) \in I_A}$  – набор коэффициентов, описывающих аномальные особенности данных.

Будем считать, что функция  $f$  в окрестности некоторой точки  $\xi$  имеет аномальную особенность разрешения  $j$ , если в окрестности точки  $\xi$  выполняется условие:

$$|\alpha_{j,k}| \geq T_j, \quad (9)$$

где  $T_j$  – пороговое значение на масштабе  $j$ .

Если вейвлет  $\Psi$  имеет компактный носитель, равный  $[-C, C]$ , то множество пар точек  $(j, k)$  таких, что точка  $\xi$  содержится в носителе  $\Psi_{j,k}$ , определяют конус влияния точки  $\xi$  [10]. Так как носитель  $\Psi_{j,k}$  для разрешения  $j$  равен  $[k - C * 2^j, k + C * 2^j]$ , то конус влияния точки  $\xi$  для разрешения  $j$  определяется неравенством

$$|k - \xi| \leq C * 2^{-j}, j = -1, -2, \dots, -J. \quad (10)$$

Поэтому временная длительность аномалии для разрешения  $j$  будет определяться конусом влияния точки  $\xi$  и в соответствии с (15) равна

$$L_j = C * 2^{-j}, j = -1, -2, \dots, -J.$$

Интенсивность аномалии для разрешения  $j$  в окрестности точки  $\xi$  определим как

$$E_{j,v} = \max_k |\alpha_{j,k}|, \text{ где } k : |k - \xi| \leq C * 2^{-j}.$$

### Результаты экспериментов

В процессе исследований использовались данные критической частоты  $f_0F2$  за период 2001 – 2011 гг.; регистрацию данных один раз в час ведет Институт космических исследований и распространения радиоволн ДВО РАН (Камчатский край, с. Паратунка). Данные содержат пропуски, что существенно затрудняет процесс их моделирования и анализа. С целью уменьшения погрешности получаемых результатов были выбраны временные периоды с наименьшим количеством пропусков. Учитывая сезонный характер ионосферного процесса, данные предварительно были разделены на сезоны и анализировались отдельно.

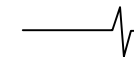
**Результаты моделирования данных  $f_0F2$ .** В качестве базисных функций использовались ортогональные вейвлеты Добеши порядка 3, которые, как показала статистика, обеспечивают наименьшую погрешность аппроксимации данных  $f_0F2$  [6].

На основе кратномасштабных разложений до уровня  $m = 3$  включительно было получено представление данных в виде

$$f_0(t) = \sum_{j=-1}^{-3} (g_j(t) + e_j(t)) + f_{-3}(t),$$

где  $g_j(t) = \sum_k d_{j,k} \Psi_{j,k}(t)$  – детализирующие компоненты разложения,  $d_{j,k} = \langle f, \Psi_{j,k} \rangle$ ;  $f_{-3}(t) = \sum_k c_{-3,k} \phi_{-3,k}(t)$  – аппроксимирующая компонента,  $c_{-m,k} = \langle f, \phi_{-m,k} \rangle$ ;  $e_j(t)$  – шумовые составляющие.

Уровень разложения определялся статистически и основывался на результатах работы [4], в которой показано, что исходные данные  $f_0F2$ , аппроксимирующие



компоненты первого и второго уровней разложения, не могут быть аппроксимированы моделью АРПСС (наблюдалась существенная автокорреляция остатков модели). Наилучшие результаты при диагностике моделей выделенных компонент были получены для уровня разложения  $m = 3$ .

Далее, на основе операции (2) были подавлены шумовые составляющие  $e_j(t)$ , результаты оценки дисперсии шума представлены в табл. 1 (анализировались периоды разных временных сезонов с различной солнечной активностью).

Результаты моделирования, выполненного на основе описанных выше операций, показали, что аппроксимирующая и детализирующая компоненты 3-го уровня разложения содержат характерные признаки данных и их первая разность имеет второй порядок авторегрессии. Полученные при оценке близкие значения параметров моделей компонент позволили получить для данных  $f_0F2$  общую многокомпонентную модель вида

$$f(t) = \sum_{\mu=1,2} \sum_{k=1, N_3^{\mu}} s_{3,k}^{\mu}(t) b_{3,k}^{\mu}(t),$$

где  $s_{3,k}^{\mu}(t) = \sum_{l=1}^2 \gamma_{3,l}^{\mu} \omega_{j,k-l}^{\mu}(t) + a_{3,k}^{\mu}(t) = (1 + 0.9B)^2 \times (1 - B) \omega_{3,k}^{\mu}(t) + a_{3,k}^{\mu}(t)$  – полученные при оценке значения  $\mu$ -ой компоненты,  $B^l \omega_{3,k}^{\mu}(t) = \omega_{3,k-l}^{\mu}(t)$ ,  $B$  – оператор сдвига назад,  $a_{3,k}^{\mu}(t)$  – остаточные ошибки модели компоненты с номером  $\mu$ .

В табл. 2, на примере 2002г., показаны параметры моделей компонент, полученные при идентификации для данных различных временных сезонов. На рис.1, в качестве примера, показаны результаты моделирования аппроксимирующей и детализирующей компонент за период с 13.12.2010 по 31.12.2010 г.

Процедура выделения аномалий в полученных компонентах  $f_0F2$  была основана на операциях (1)-(5) и

определении дисперсии ошибки прогноза (величина  $D_{U_j}$  в соотношении (6)). На рис.2-4. д, е представлены результаты расчета дисперсии ошибки прогноза в скользящем временном окне, равном 24 часа.

Таблица 1. Результаты оценки дисперсии шума в данных  $f_0F2$

Анализируемый период	Дисперсия шума
01.12.1971-18.12.1971	0.223
10.02.1972-26.02.1972	0.1704
05.02.1973-28.02.1973	0.1927
04.02.1975-25.02.1975	0.1623
07.12.1979-22.12.1979	0.2126
23.01.1981-06.02.1981	0.3152
07.02.1981-21.02.1981	0.3272
07.02.1982-28.02.1982	0.2705
07.02.1983-23.02.1983	0.2108
22.01.1987-10.02.1987	0.1671
15.02.1987-28.02.1987	0.1512
01.01.1991-26.01.1991	0.2968
12.01.1992-05.02.1992	0.3418
11.02.1999-28.02.1999	0.1996
01.12.2000-22.12.2000	0.1938
16.01.2002-25.02.2002	0.2068
01.01.2002-31.02.2002	0.1929
01.04.2002-31.05.2002	0.1511
29.12.2005-14.01.2006	0.2071
15.01.2006-04.02.2006	0.1768
01.01.2006-28.02.2006	0.1787
10.12.2010-31.12.2010	0.1716
01.12.2010-31.03.2011	0.1674
01.04.2011-31.05.2011	0.2746
01.06.2011-25.06.2011	0.1241

**Результаты выделения аномалий в ионосферных параметрах на основе дискретного вейвлет-преобразования.** Следуя результатам работы [6], в качестве базисных функций использовались ортогональные вейвлеты Добеши порядка 3. Разложение данных выполнялось до разрешения  $j = 150$ .

Выделение аномалий было основано на проверке условия (9). При построении аппроксимаций (см. соотношение (8)) пороговые значения  $T_j$  определялись в про-

Таблица 2. Результаты моделирования данных  $f_0F2$

Анализируемый период	Параметры моделей аппроксимирующих компонент		Параметры моделей детализирующих компонент	
	Значение первого параметра	Значение второго параметра	Значение первого параметра	Значение второго параметра
01.01.2002-31.03.2002	-0.9875	-0.9918	-0.9942	-0.9689
01.04.2002-31.05.2002	-0.8424	-0.8451	-0.998	-0.9068
01.06.2002-31.08.2002	-0.7019	-0.685	-0.9635	-0.9153
01.09.2002-30.11.2002	-0.9643	-0.9679	-1.021	-0.9591

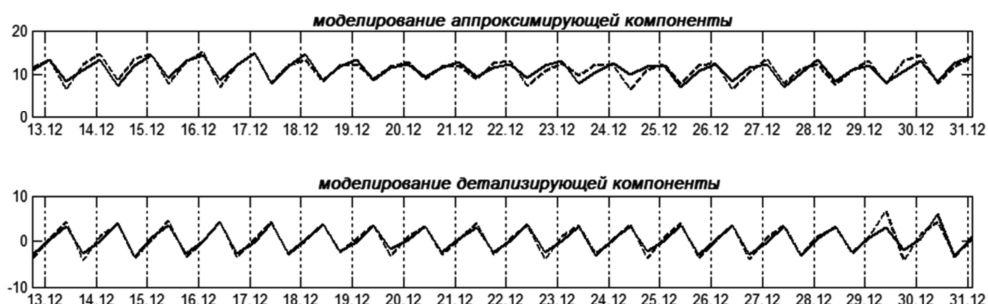


Рис.1. Результаты моделирования компонент (пунктир – исходные данные, сплошная линия – значения прогноза)



процессе работы алгоритма в скользящем окне, равном 168 отсчетов (соответствует временному интервалу 2 недели), по следующей формуле:

$$T_j = \text{med} \left| \alpha_{j,k} \right|_{1 \leq k \leq V} + \Omega * St_j,$$

где  $\text{med}$  - медиана,  $St_j = \sqrt{\frac{1}{V-1} \sum_{k=1}^V (\alpha_{j,k} - \bar{\alpha}_{j,k})^2}$ ,

$\bar{\alpha}_{j,n}$  - среднее значение, определяемое в пределах анализируемого скользящего временного окна. Для анализа возмущений различной интенсивности коэффициент  $\Omega$  изменялся в диапазоне:  $2 \leq \Omega \leq 3.5$ . Выделенные частотно-временные интервалы, содержащие аномалии, показаны на рис. 2-4.а, б, в оттенками серого цвета. При сопоставлении результатов обработки для определения степени магнитосферной возмущенности использовались данные К-индекса (рис. 2-4. г).

**Анализ результатов.** Полученные картины распределения аномальных особенностей в пространстве (масштаб, время), представленные на рис. 2-4. а, б, в, показывают их разномасштабный характер и неравномерное распределение и по времени, и по масштабам. Анализ интенсивности и частоты появления аномалий в данных за разные годы (рис. 2-4. а, б, в), свидетельствует об их зависимости от уровня солнечной и магнитной активности. С ростом солнечной и магнитной активности интенсивность и частота появления аномалий значительно увеличивается. Данный факт также подтверждается результатами расчета дисперсии ошибки прогноза, полученными на основе моделирования (рис. 2-4 д, е), и согласуется с результатами работы [1], где представлены данные исследований ионосферных возмущений в верхней атмосфере Земли на основе GPS-мониторинга. Анализ результатов моделирования данных

(рис. 2-4 д, е) также показывает, что в периоды с высокой активностью Солнца размах дисперсии ошибок прогноза увеличивается для аппроксимирующих компонент в 7-8 раз и для детализирующих компонент в 4-5 раз по сравнению с периодами с низкой солнечной активностью.

Сопоставление полученных результатов с данными каталога землетрясений показывает, что в сейсмически спокойные периоды времени (рис. 3. д, е.) возрастание ошибок моделей наблюдается во время магнитных бурь, особенно для аппроксимирующих компонент. В периоды повышенной сейсмической активности (рис. 2,4. д, е.) характер процесса меняется, и существенное увеличение ошибок моделей наблюдается в периоды возникновения сильных землетрясений. В 2002 г. накануне землетрясения, произошедшего зимой (15.02.02), наблюдается увеличение ошибок прогноза, полученных на основе модели детализирующей компоненты разложения. В периоды землетрясений, произошедших в конце апреля и начале мая (26.04.02 и 08.05.02), увеличивается дисперсия ошибок прогноза аппроксимирующей компоненты. Накануне и после землетрясений, произошедших осенью (08.10.02 и 16.10.02), наблюдается увеличение ошибок моделей и аппроксимирующей, и детализирующей компонент разложения. В 2011 г. накануне землетрясений, произошедших летом (07.06.11, 12.06.11 и 14.06.11), наблюдаются аномальные особенности и в результатах моделирования и в результатах дискретного вейвлет-преобразования. Возрастание ошибок весной (март, апрель 2002г и 2011г), возможно, связано с переходными процессами в ионосфере, характерными для данного периода времени. Также следует отметить, что анализируемые сейсмические явления в большинстве случаев наблюдаются на фоне повышенной солнечной активности.

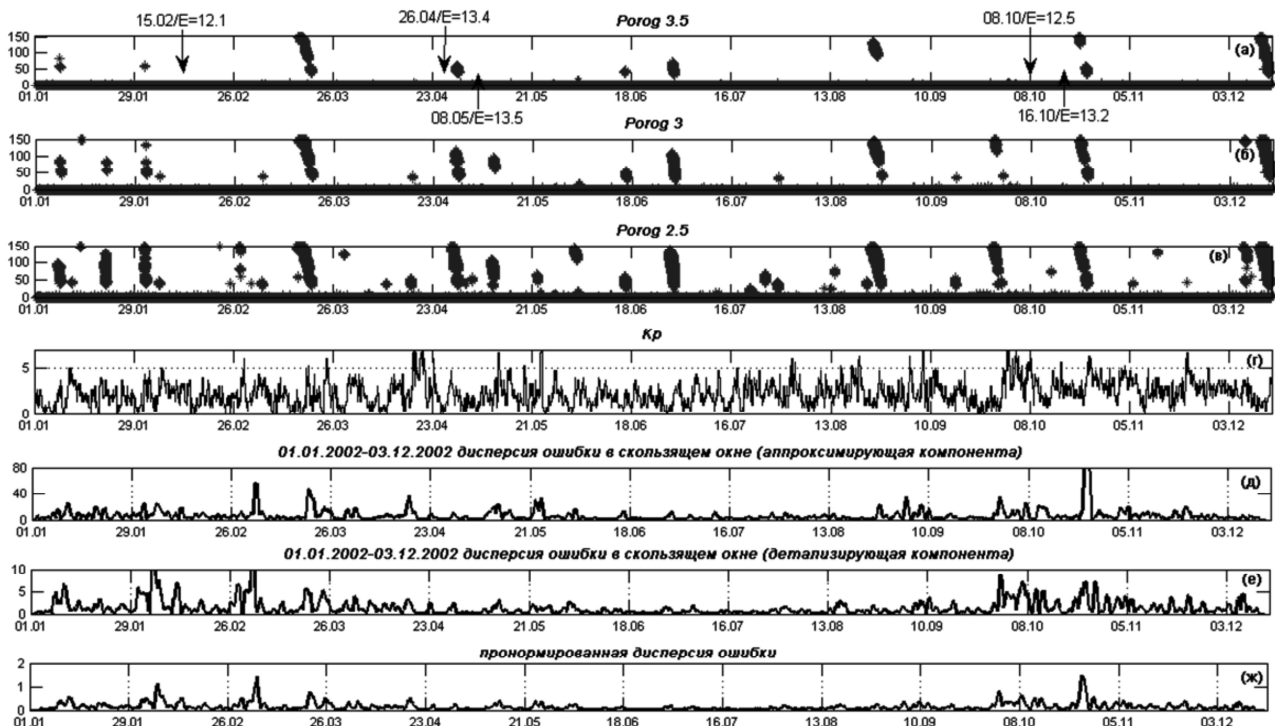


Рис. 2 Результаты моделирования данных критической частоты за период 01.01.2002-14.12.2002 г. Стрелками отмечены моменты возникновения сейсмических событий



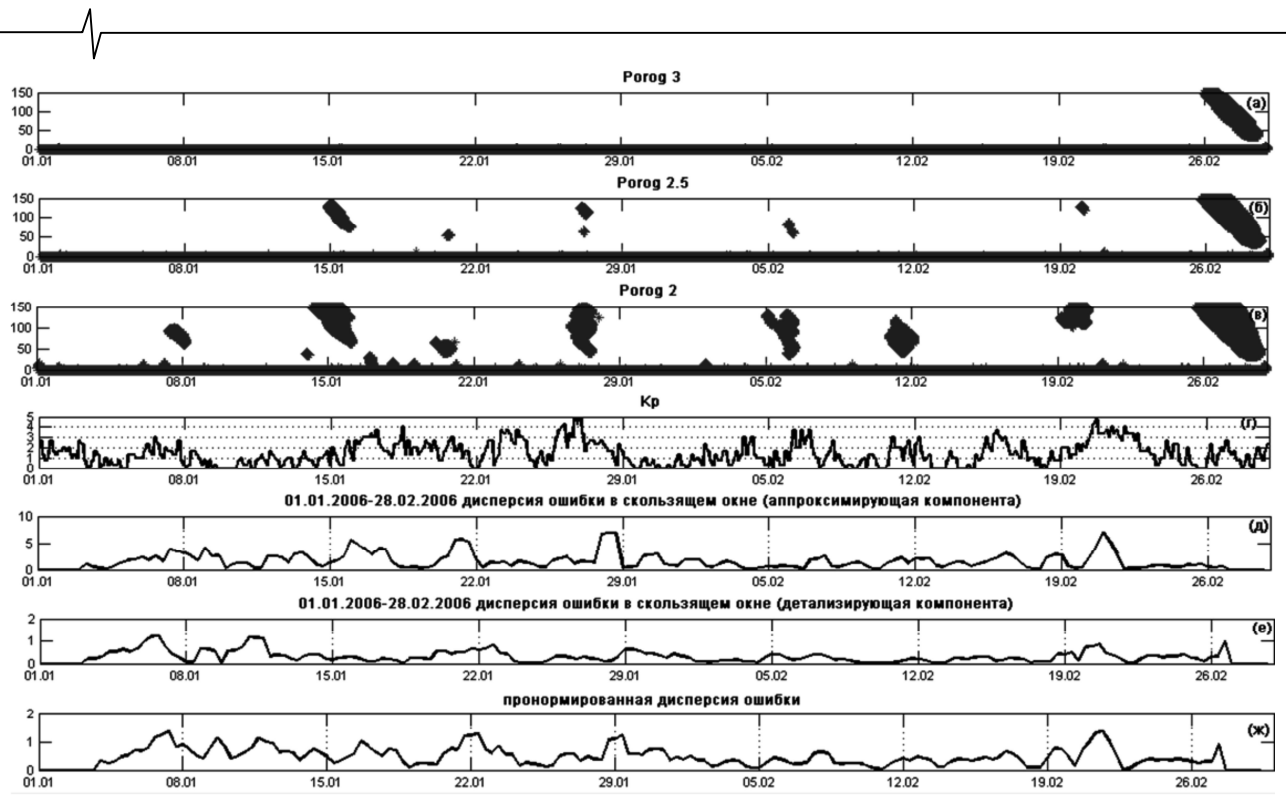


Рис.3 Результаты моделирования данных критической частоты за период 01.01.2006-28.02.2006 г.  
Стрелками отмечены моменты возникновения сейсмических событий

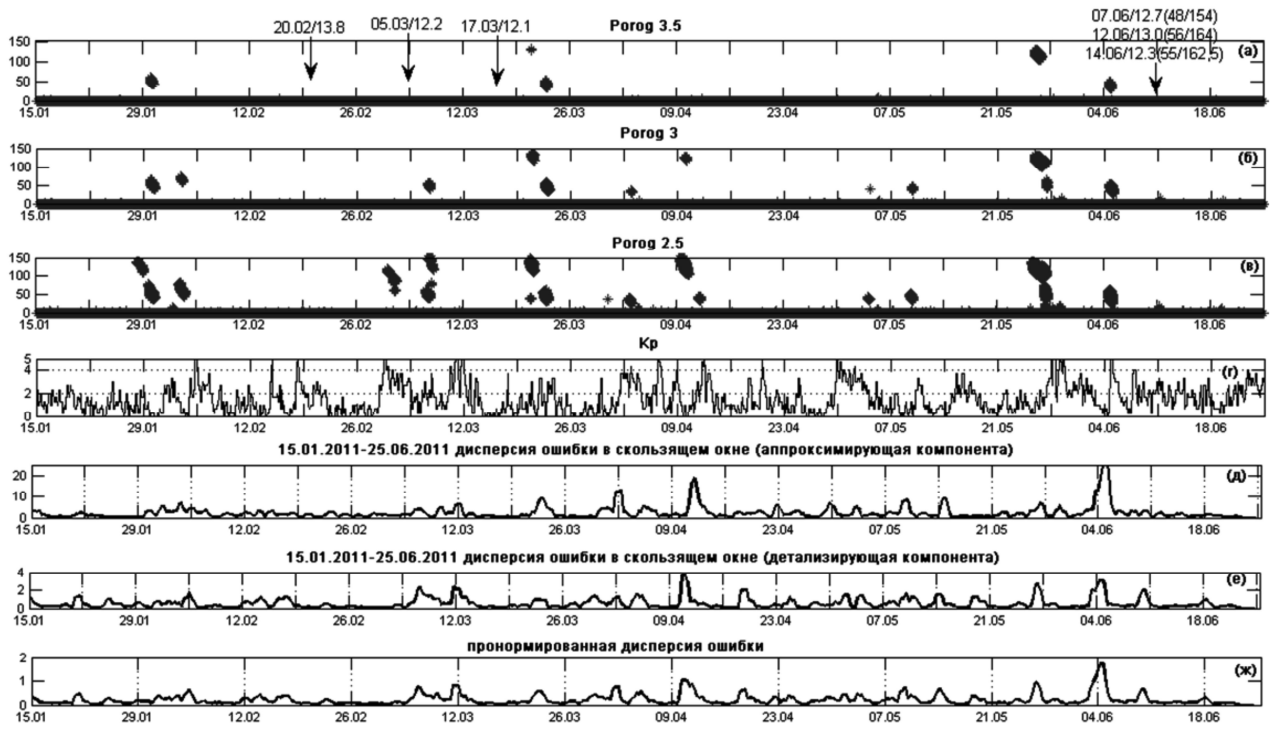


Рис.4. Моделирование данных критической частоты за период 15.01.2011-25.06.2011г.  
Стрелками отмечены моменты возникновения сейсмических событий

### Заключение

Предложенные в статье алгоритмы выделения и анализа аномальных особенностей в ионосферных данных позволили проанализировать вариации параметров в периоды с различной активностью Солнца и в периоды возникновения сильных землетрясений на Камчатке. Проведенный анализ подтвердил сложную нестационарную структуру данных  $f_{oF2}$ , её существенное изме-

нение при возрастании солнечной и сейсмической активности. Детальный анализ данных на основе дискретного вейвлет-преобразования показал, что в периоды возрастания солнечной или сейсмической активности в вариациях ряда  $f_{oF2}$  могут возникать аномалии, имеющие неравномерное распределение и по времени и по масштабам. Оценка интенсивности, длительности и масштаба аномалий позволяет изучить их особенности и характер

протекания на всем анализируемом временном интервале.

Работа поддержана грантом Президента РФ МД-2199.2011.9, грантом РФФИ – ДВО РАН №11-07-98514-р\_восток\_а и грантом «У.М.Н.И.К.» - №8283р/10269 от 30.06.2010 г.

Данные сейсмического каталога любезно предоставлены Камчатским филиалом геофизической службы РАН (г. Петропавловск-Камчатский).

#### Литература

1. Афраймович Э.Л. GPS- мониторинг верхней атмосферы Земли / Э.Л.Афраймович, Н.П. Перевалова - Иркутск: ГУ НУ РВХ ВСНЦ СО РАМН, 2006.-480с.
2. Дёмин М.Г. Ионосфера Земли. Плазменная гелиогеофизика. – М.: Физматлит, 2008. -Т.II, -С.92-163.
3. Липеровская Е.В., Липеровский В.А., Похотелов О.А. О возмущениях в F-области ионосферы перед землетрясениями // Геофизические исследования. -2006. -№ 6. -С. 51-58.
4. Мандрикова О.В., Глушкова Н.В. Метод моделирования данных критической частоты на основе совмещения вейвлет-преобразования и моделей авторегрессии-проинтегрированного скользящего среднего // Научные ведомости Белгородского государственного университета: Белгород. 2011г. -№ 19.-С.59-63.
5. Мандрикова О.В. Многокомпонентная модель сигнала со сложной структурой // Проблемы эволюции открытых систем. 2008. -Вып. 10. -Т. 2. -С.161-172.
6. 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.
7. Мандрикова О.В., Глушкова Н.В. Моделирование и анализ данных критической частоты ионосферы на основе совмещения вейвлет-преобразования и моделей авторегрессии// 14-я Международная конференция «Цифровая обработка сигналов и ее применение – DSPA-2012», Москва. Т. 2. -С. 346-349.
8. Мандрикова О.В., Горева Т.С. Метод идентификации структурных компонентов сложного природного сигнала на основе вейвлет-пакетов. – Москва: Цифровая обработка сигналов.-2010 г.-№1.- С. 45-50.

9. Mandrikova O.V., Polozov Yu.A., Zaliaev T.L. Methods of analysis and interpretation of ionospheric critical frequency  $F_{o}f_2$  data based on wavelet transform and neural networks / European Seismological Commission 33-rd General Assembly (GA ESC 2012) 19.08.2012 – 24.08.2012 / [http://www.esc2012-moscow.org/files/GA\\_ESC\\_2012-Program\\_13.08.2012.xls](http://www.esc2012-moscow.org/files/GA_ESC_2012-Program_13.08.2012.xls).

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

11. Donoho D., Johnstone I. Minimax estimation via wavelet // Annals of Statistics, 1998, pp. 879-921.

12. Бокс Дж., Дженкинс Г. Анализ временных рядов прогноз и управление. -М.: Мир, 1974. -604 с.

13. Чуи К. Введение в вейвлеты. Пер. с английского, М.: Мир.– 2001.

#### ALGORITHMS OF SELECTION AND THE ANALYSIS OF ANOMALIES IN PARAMETERS OF CRITICAL FREQUENCY OF AN IONOSPHERE OF FOF2 ON THE BASIS OF COMBINATION OF WAVELET-TRANSFORMATION AND AUTOREGRESSIVE MODELS

*Mandrikova O.V., Glushkova N.V., Polozov Yu.A.*

Work is devoted to creation of technologies and program systems on studying of processes in an ionosphere and to selection of the anomalies arising in ionospheric plasma during the periods of increased solar or seismic activity. The method of multicomponent model operation of the ionospheric data, based on combination of wavelet-transformation with autoregressive models – the integrated moving average, allowing to analyse and a forecast of data is offered. At model operation of data the features bound to solar activity, and also arising during the periods of the strong earthquakes on Kamchatka are allocated. The analysis algorithms of ionospheric data based on a design of discrete wavelet-transformation are offered, allowing automatically to allocate anomalies in ionospheric plasma and to estimate their parameters. For an assessment of a method and algorithms data of station «Paratunka» (Kamchatka) during the period 2001-2011years were used.

### *Уважаемые коллеги!*

Для тех, кто не успел оформить подписку на первое полугодие 2013 года через ОАО «Роспечать», сохраняется возможность приобретения журналов непосредственно в редакции по адресу:  
107031, г. Москва, Рождественка, 6\9\20, стр. 1,  
Российское научно-техническое общество радиотехники, электроники и связи им. А.С. Попова, или оформить заказ в соответствии с требованиями, выставленными на сайте журнала: [www.dsra.ru](http://www.dsra.ru).

*Справки по телефонам: (495) 621-71-08, 621-06-10.*

*Факс: (495) 621-16-39. E-mail: [rntores@mail.ru](mailto:rntores@mail.ru)*

## МЕТОДЫ И АЛГОРИТМЫ ДЕТЕКТИРОВАНИЯ АКТИВНОСТИ РЕЧИ

*Волченков В.А., аспирант кафедры телекоммуникаций и основ радиотехники Рязанского государственного радиотехнического университета, e-mail: volchenkov.rzn@yandex.ru*

*Витязев В.В., д.т.н., профессор, зав. кафедрой телекоммуникаций и основ радиотехники Рязанского государственного радиотехнического университета*

**Ключевые слова:** детектор активности речи, обнаружение пауз, методы и алгоритмы, сравнительный анализ.

### Введение

В голосовой связи речь может быть охарактеризована как прерывистый носитель информации из-за пауз, которые являются характерными особенностями данного вида сигналов по сравнению с другими мультимедийными сигналами, такими как видео, аудио и данные [1]. Интервалы, где голосовая информация присутствует, называют участками активной речи, а паузы между речевыми «всплесками» – участками пауз. Пример, иллюстрирующий участки активной речи и пауз для речевого сигнала, представлен на рис. 1 [1].

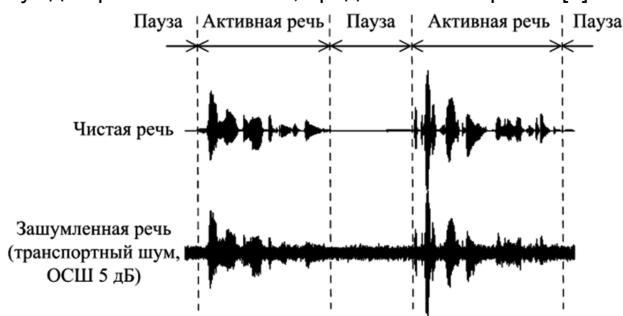


Рис. 1. Участки активной речи и паузы

Детектор активности речи (voice activity detector – VAD) – это алгоритм, предназначенный для различения интервалов активной речи и пауз. Когда система детектирует участки пауз, передача, как правило, прекращается и передается только общее описание информации о фоне. На другом конце декодер восстанавливает интервалы пауз посредством генерации комфортного шума (comfort noise generation – CNG), который дает естественный фоновый шум с плавными переходами от речевых всплесков к паузе и наоборот. Для повышения естественности генерируемого фонового сигнала анализатор фонового шума (comfort noise insertion module – CNI) регулярно передает информацию о среднем уровне фонового шума. Общая структурная схема сжатия интервалов пауз, использующая VAD, CNG и CNI, представлена на рис. 2 [3].

Системы передачи речи, которые используют VAD для сжатия участков пауз, обладают различными преимуществами, особенно полезными для каналов передачи с ограниченной пропускной способностью. Из основных преимуществ детектора активности речи можно перечислить следующие:

Проводится сравнительный анализ эффективности методов и алгоритмов детектирования участков активности речи и пауз между ними. Приведена общая информация о стандартизированных методах детектирования активности речи и оценка их производительности. Представлен новый способ обнаружения пауз в речи.

– сокращение помех от соседних каналов в сотовой связи. Уменьшая мощность передачи во время неактивных интервалов (речевые паузы), возможно подавить помехи от соседних каналов в сотовых системах беспроводной связи.

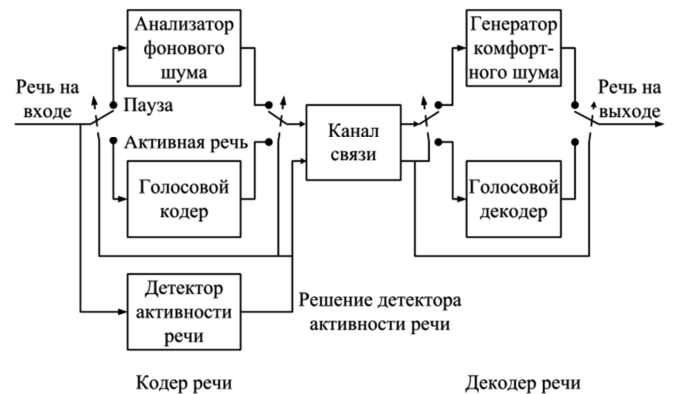


Рис. 2. Общая структурная схема кодера речи со сжатием речевых пауз

– увеличение пропускной способности канала связи в системе множественного доступа с кодовым разделением (code division multiple access – CDMA). Теоретическая пропускная способность системы CDMA, как правило, определяется возможными комбинациями расширяющего кода. Тем не менее, из-за помех от других пользователей, емкость CDMA ограничена значением меньшим, чем этот теоретический предел, то есть из-за помех от других пользователей коэффициент ошибок приема у некоторых пользователей может быть слишком высоким, чтобы иметь возможность осуществить точное декодирование. При снижении мощности передачи во время пауз интерференционные помехи в эфире могут быть уменьшены, что автоматически позволит находиться большему числу пользователей в системе, т.е. происходит увеличение емкости системы CDMA.

– энергосбережение для мобильных терминалов. Мобильные терминалы не должны передавать радиосигналы во время речевых пауз. Таким образом, срок службы батареи терминалов может быть увеличен за счет сохранения энергии во время пауз.

– увеличение пропускной способности канала путем статистического мультиплексирования. Канал может

быть предоставлен только на время передачи активной речи и должен быть освобожден во время пауз. Полученный пользователем канал занимает до конца участка активной речи и освобождается сразу после последнего интервала активной речи. Чтобы вновь получить доступ к каналу, устройство пользователя посылает запрос в начале следующего речевого всплеска. Таким образом, каналный ресурс может быть использован более эффективно посредством применения схемы статистического мультиплексирования, которая позволяет нескольким пользователям общаться в одно и то же время, с учетом ограниченной пропускной способности канала связи. Примечание: в статистическом мультиплексировании существует вероятность того, что когда пользователь делает запрос, свободных слотов в канале не окажется. В этом случае новому пользователю может быть отказано после тайм-аута, что может привести к потере информации и, как следствие, ухудшению качества связи.

– снижение потери пакетов при передаче голоса по сетям пакетной передачи данных. Если количество пакетов больше, чем может быть обработано, то система пакетной передачи данных будет перегружена. Перегрузку системы пакетной передачи данных голосовой связи можно уменьшить, передавая пакеты только с активной речью и вырезая пакеты с паузами.

– снижение скорости передачи. В дополнение к сокращению скорости передачи, достигаемому за счет использования методов сжатия речи, использование VAD вместе со сжатием пауз (удаление интервалов речевых пауз) дает дополнительное снижение скорости передачи независимо от применяемых кодеров речи.

Детектор активности речи обычно производит двоичное решение для обрабатываемого речевого фрейма (как правило, длиной 10-20 мс), указывая на наличие или отсутствие речи, что довольно легко при отсутствии фонового шума [1, 2]. Например, проверяя энергетический уровень входного сигнала, можно получить высокую производительность обнаружения речи / неречи (паузы). Однако, в реальных условиях входной сигнал может быть смешан с шумом, характеристики которого могут быть неизвестны и изменяться со временем. В некоторых случаях, когда уровень фонового шума очень высок, речь может быть заглушена этим шумом. Особенно глухие звуки, которые являются важными для разборчивости речи, могут быть неправильно определены в такой шумной обстановке. Рис. 1 показывает пример для зашумленного сегмента речи с шумом от транспорта, отношение сигнал-шум (ОСШ) – 5 дБ. Как видно из рисунка, некоторые части речи с малой энергией полностью погружены в шум, что делает эти речевые всплески трудно различимыми даже визуально. Неправильная классификация этих речевых всплесков может привести к обрезанию звуков и, как следствие, к значительному ухудшению качества речи. С другой стороны, увеличение ложного обнаружения «не речи» лишает потенциальной выгоды от сжатия пауз. Существует компромисс в производительности VAD – увеличение показателя правильного обнаружения активной речи при минимизации показателя ложных обнаружений интервалов пауз.

## Стандартизированные методы детектирования активности речи

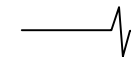
Чтобы использовать преимущества сжатия речевых пауз, было предложено множество алгоритмов VAD, некоторые из них были отобраны организациями по стандартизации, включая: Международный консультационный комитет по телефонии и телеграфии (International Telecommunication Union - Telecommunication sector – ITU-T), Европейский институт по стандартизации в области телекоммуникаций (European Telecommunications Standards Institute – ETSI), Ассоциация телекоммуникационной промышленности США (Telecommunications Industry Association – TIA) и Альянс отраслей электронной промышленности (Electronics Industries Alliance – EIA).

Международный комитет ITU-T выпустил кодеры G.729 Annex B (G.729B) [3] и G.723.1 Annex A (G.723.1A) [4] в качестве расширения для 8 кбит/с G.729 [5] и 5.3/6.3 кбит/с G.723.1 [6] кодеров речи для возможности выполнения прерывистой передачи (Discontinuous transmission – DTX). Европейский институт стандартизации ETSI рекомендовал GSM-FR, -HR и -EFR методы детектирования активности речи для европейских систем цифровой сотовой связи [7-9]. Затем ETSI представил еще два детектора активности речи: адаптивный многоскоростной VAD вариант 1 (кодер AMR1) и вариант 2 (кодер AMR2) [10], с тем, чтобы использовать их в сетях третьего поколения мобильной связи UMTS. Североамериканские организации по стандартизации TIA и EIA представили два алгоритма VAD: один для кодера IS-96 [11], а другой для кодеров IS-127 [12] и IS-733 [13] (алгоритмы VAD для IS-127 и IS-733 имеют одинаковую структуру). В таблице 1 представлены стандартизированные алгоритмы VAD, классифицированные по принципу того, что они анализируют. Главным образом – это энергии поддиапазонов и спектральная форма сигнала [1]. Например, алгоритмы VAD, предложенные TIA и EIA, используют предварительное разбиение сигнала на небольшое количество поддиапазонов, в то время как алгоритм VAD для кодера IS-96 анализирует общую энергию сигнала. С другой стороны, алгоритм VAD кодеров IS-127 и IS-733 также раскладывает сигнал, но только на два поддиапазона.

Таблица 1. Классификация стандартизированных методов VAD в зависимости от способа анализа сигнала (в скобках указано количество спектральных поддиапазонов)

Способ анализа	VAD
Анализ спектральной формы	GSM-FR, GSM-HR, GSM-EFR
Анализ энергии поддиапазонов	IS-96 (1), IS-127 (2), IS-733 (2), AMR1 (9), AMR2 (16)
Другое	G.729B, G.723.1A

Традиционно методы VAD Европейского института стандартизации ETSI были основаны на более точном анализе – анализе спектральной формы входного сигнала. Причиной этого является то, что энергия ошибки кодирования с предсказанием увеличивается, когда спектральные формы фона и входного сигнала не совпадают



(например, в случае активной речи). Тем не менее, в последнем стандарте AMR приняты два вида алгоритмов VAD, каждый из которых основан на анализе спектральной энергии поддиапазонов, а не на более точном анализе формы спектра. Стандартизированные методы VAD в кодерах G.729B и G.723.1A ведут обнаружение с помощью четырех различных способов, включая как анализ спектральной формы, так и анализ энергии поддиапазонов.

### Алгоритмы VAD кодеров G.729B/G.723.1A

В качестве расширения к кодеру речи G.729 международный комитет ITU-T SG16 выпустил кодер G.729 Annex B с целью поддержки прерывистой передачи DTX посредством детектирования активности речи VAD, анализа фонового шума CN1 и генерации комфортного шума CNG. Кодер G.729B делит речь на интервалы по 10 мс и вырабатывает решение о наличии или отсутствии речи для каждого фрейма, оценивая при этом четыре параметра [3, 5]:

– разность энергий всего диапазона –  $\Delta E_f = \overline{E_f} - E_f$ ,

– разность энергий диапазона НЧ –  $\Delta E_l = \overline{E_l} - E_l$ ,

– искажение спектра –  $\Delta LSF = \sum_{i=0}^9 (\overline{LSF_i} - LSF_i)^2$ ,

– разность частоты переходов через ноль –  $\Delta ZC = \overline{ZC} - ZC$ ,

где  $E_f$  – энергия всего диапазона,  $E_l$  – энергия диапазона НЧ,  $LSF_i$  –  $i$ -я частота спектра сигнала и  $ZC$  – частота переходов через ноль входного сигнала,  $\overline{E_f}$ ,  $\overline{E_l}$ ,  $\overline{LSF_i}$ ,  $\overline{ZC}$  – параметры, характеризующие шум и обновляемые посредством анализа фонового шума.

Блок схема алгоритма VAD кодера G.729B представлена на рис. 3 [5]. Входные параметры для анализа VAD могут быть получены из входного сигнала или из промежуточных значений речевого кодера. Затем рассчитываются параметры разницы между параметрами входного сигнала и шума  $\Delta E_f$ ,  $\Delta E_l$ ,  $\Delta LSF$  и  $\Delta ZC$ . Решение о наличии речи получают путем анализа интервалов речевого сигнала по четырем параметрам, которые поступают на схему анализа предыдущих решений. Блок обновления параметров шума основан на схеме авторегрессии первого порядка. Они обновляются, если разница энергии всего диапазона меньше заданного фиксированного порога. Алгоритм VAD кодера G.723.1A имеет аналогичную структуру. Отличие заключается в том, что в кодер G.723.1A делит речь на интервалы по 30 мс.

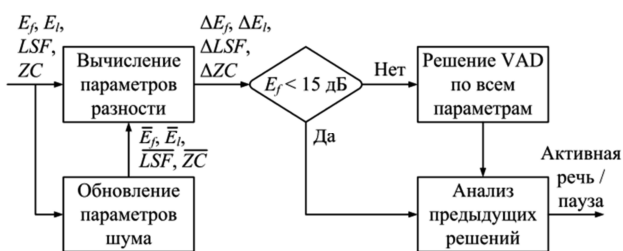


Рис. 3. Алгоритм VAD кодера G.729B

### Алгоритмы VAD кодеров GSM-FR/HR/EFR

Алгоритмы VAD кодеров GSM-FR/HR/EFR, предложенные ETSI, имеют схожую структуру. Здесь предсказанная остаточная энергия сравнивается с адаптивным порогом. Предсказанная остаточная энергия вычисляется с использованием значений действительной и сглаженной автокорреляции, которые описывают спектральные характеристики сигнала. Предполагается, что если сигнал является только фоновым шумом, который обычно считается стационарным, то средняя спектральная форма будет подобна форме спектра текущего фрейма, что приведет к меньшей остаточной энергии сигнала. Порог решения алгоритма VAD обновляется только в течение интервалов «неречи», используя самые последние сигналы шума, чтобы отражать актуальные шумовые характеристики. Блок схема алгоритмов VAD кодеров GSM-FR/HR/EFR представлена на рис. 4 [9].

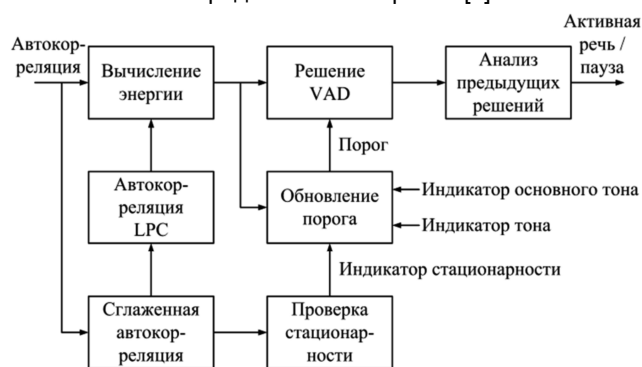


Рис. 4. Алгоритм VAD кодера GSM-EFR

### Алгоритмы VAD кодеров AMR

Кодер AMR1 раскладывает входной сигнал на девять неравномерных поддиапазонов с помощью банка фильтров, где нижние полосы частот имеют меньшую полосу пропускания, а более высокие частотные диапазоны – большую пропускную способность. Затем он вычисляет энергию в каждом поддиапазоне, сопровождаемую ее соответствующей оценкой ОСШ. Энергия фонового шума, используемая в расчетах ОСШ, вычисляется адаптивным методом на основе авторегрессионной модели первого порядка и внутренней логической схемы VAD. Наконец, решение VAD принимается путем сравнения суммы ОСШ поддиапазонов с адаптивным порогом и поступает на схему анализа предыдущих решений. Блок схема алгоритма VAD кодера AMR1 представлена на рис. 5 [10].



Рис. 5. Алгоритм VAD кодера AMR1

Структура кодера AMR2 похожа на AMR1 в том, что детектирование речевой активности осуществляется с использованием информации об энергии в поддиапазонах вместе с энергией фонового шума. Однако алгоритм VAD кодера AMR2 трансформирует входной сигнал в частотную область, используя БПФ вместо банка фильтров, применяемого в AMR1, и затем вычисляет энергию в каждом поддиапазоне. Число полос равняется 16, ширина полос также является неравномерной, как и в AMR1. Впоследствии, по спектрам входного сигнала и фонового шума, вычисляется ОСШ для каждого поддиапазона. Энергия фонового шума для каждой полосы изменяется во время интервалов «неречи» с помощью авторегрессионной схемы первого порядка. Чтобы предотвратить возможность возникновения чрезмерной чувствительности к условиям нестационарного фонового шума, AMR2 увеличивает порог принятия окончательного решения алгоритма VAD для высокофлуктуирующих сигналов, оцениваемых по дисперсии их мгновенных межкадровых ОСШ. Кроме того, адаптация уровня шума путем измерения девиации спектра может проводиться неточно, если уровень энергий поддиапазонов меняется быстро. Таким образом, кодер AMR2 изменяет порог VAD адаптивным способом, ориентируясь также на уровень всплеска и принятые решения по прошлым интервалам. Контроль за принятыми решениями осуществляется путем измерения отношения пикового значения сигнал/шум к среднему, где среднее значение ОСШ рассчитывается при помощи авторегрессионной адаптации с увеличенным мгновенным ОСШ. Другими словами, для увеличения отношения пикового значения сигнал/шум к среднему, уменьшаются расчеты по прошлым значениям и уровням всплеска, наряду с увеличением порога алгоритма VAD. Блок-схема алгоритма VAD кодера AMR2 представлена на рис. 6 [10].

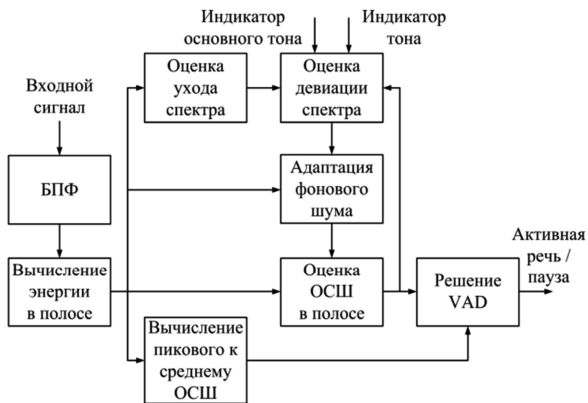


Рис. 6. Алгоритм VAD кодера AMR2

### Алгоритмы VAD кодеров IS-127/133

Основанные на CDMA цифровые системы сотовой связи имеют естественную структуру для встраивания алгоритма VAD, называемого алгоритмом определения скорости (rate determination algorithm – RDA), который дает существенное увеличение канальной емкости путем контроля мощности радиопередачи с целью ослабления межканальной интерференции. Ассоциации TIA/EIA представили два вида алгоритмов RDA для стандартов кодеров IS-96 и IS-127, называемые кодером с линейным предсказанием с мультикодовым управлением Qualcomm (Qualcomm code-excited linear prediction – QCELP) 8 кбит/с и кодером с расширенной изменяющейся скоростью (enhanced variable rate codec – EVRC), соответственно. В Северной Америке стандарт CDMA RDA IS-127 поддерживает три скорости: 1, 1/2 и 1/8. Интервалы, в которых содержится речь, кодируются со скоростью 1 или 1/2, а фоновый шум со скоростью 1/8. Алгоритм RDA кодера IS-733 называют 13 кбит/с QCELP, он такой же, как и алгоритм RDA кодера IS-127.

В качестве входных параметров алгоритм RDA кодера IS-127 использует энергии двух поддиапазонов с увеличением долгосрочного предсказания. Сначала, с помощью авторегрессионной модели первого порядка, вычисляется сглаженная энергия в поддиапазонах. Затем производится адаптация энергий сигнала и шума в каждом поддиапазоне в зависимости от усиления долгосрочного предсказания. Другими словами, энергия сигнала активно адаптируется к текущему значению на входе, если коэффициент усиления предсказания относительно высок. С другой стороны, если коэффициент усиления является небольшим, это позволяет увеличить скорость адаптации шума. ОСШ в каждом из двух поддиапазонов вычисляется с использованием информации об энергиях сигнала и шума в каждой подполосе. Окончательная скорость определяется путем сравнения отношений сигнал/шум со значениями адаптивных порогов, зависящих от уровня фонового шума и ОСШ предыдущего фрейма, поступая затем на схему анализа предыдущих решений. Блок-схема RDA кодера IS-127 представлена на рис. 7 [12].

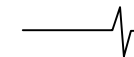
Блок-схема алгоритма RDA кодера IS-127 представлена на рис. 7 [12].



Рис. 7. Алгоритм RDA кодера IS-127

### Сравнение производительности алгоритмов VAD, используемых в различных стандартах

В [1] была проведена сравнительная оценка пяти стандартизированных алгоритмов VAD на предмет количества ошибок обнаружения для интервалов активной речи и паузы. Тестовым сигналом была речь длительностью 96 секунд, преобразованная с помощью модифицированной системы промежуточного отклика, а затем смешанная с транспортным шумом с ОСШ: 5, 10, 15, 20 и 25 дБ. Интервалы активной речи и паузы были отмечены вручную. Пропорции между неактивными и активными участками речи были 0,43 и 0,57, соответственно. Решение алгоритмом VAD для кодеров G.729B и AMR2 принимается каждые 10 мс, для кодеров GSM-EFR, AMR1 и IS-127 – каждые 20 мс. При незначительной модификации исходного кода AMR2, результаты можно получать каждые 10 мс, т.к. в своей основе AMR2 принимает решения каждые 10 мс и затем возвращает решение по интервалу длительностью 20 мс, анализируя логическую комбинацию решений по двум интервалам длительностью 10 мс. Что касается многоскоростного кодера IS-127, две верх-



них скорости (1 и 1/2) применяются для кодирования активной речи, а нижняя скорость (1/8) используется для кодирования «неречи».

Производительность работы алгоритмов в среде транспортного шума показана на рис. 8 и 9 [1]. Осциллограммы тестового сигнала и результатов работы детекторов для ОСШ 15 дБ представлены на рис. 10 [1].

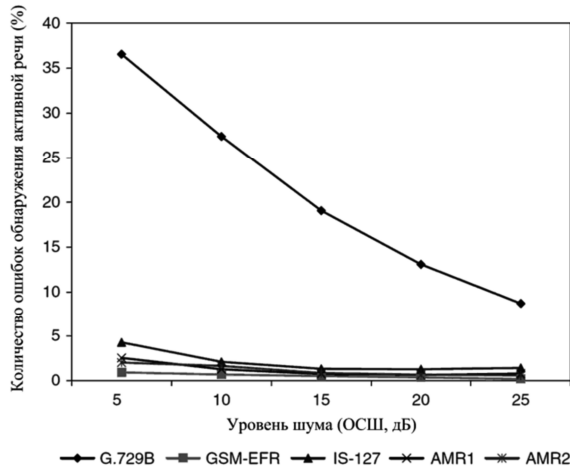


Рис. 8. Количество ошибок обнаружения активной речи по отношению к уровню транспортного шума

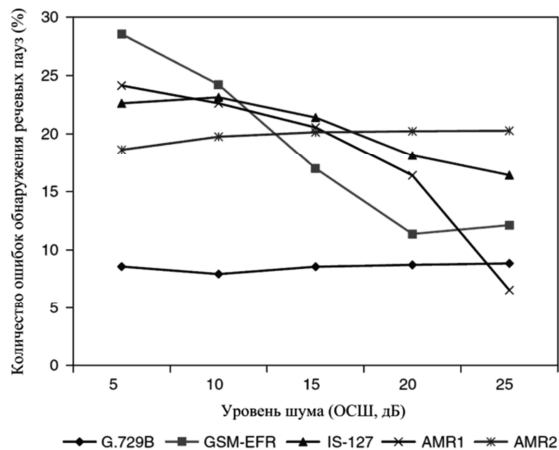


Рис. 9. Количество ошибок обнаружения речевых пауз по отношению к уровню транспортного шума

Кодер G.729B показывает хороший результат по сравнению с другими методами при обнаружении речевых пауз. Однако он демонстрирует очень высокий уровень ошибок детектирования речи, что может повлечь резкое обрезание сигнала речи. Кодер IS-127 показывает относительно высокое количество ошибок при детектировании активной речи в сравнении с алгоритмами VAD, предложенными ETSI. Методы VAD Европейского института стандартизации ETSI, т.е. алгоритмы VAD кодеров GSM-EFR, AMR1 и AMR2, показывают довольно схожие результаты в обнаружении интервалов активной речи, в то время как результаты детектирования интервалов пауз сильно разнятся. Кодер GSM-EFR демонстрирует наилучшие результаты для сравнительно высокого уровня ОСШ (больше 15 дБ). Тем не менее, количество ошибок обнаружения интервалов пауз при уменьшении уровня ОСШ существенно растет. Кодер AMR2 показывает относительно последовательные результаты, не смотря на изменение уровня шума, при

обнаружении участков пауз для речевого сигнала с транспортным шумом. По производительности AMR1 находится между GSM-EFR и AMR2.

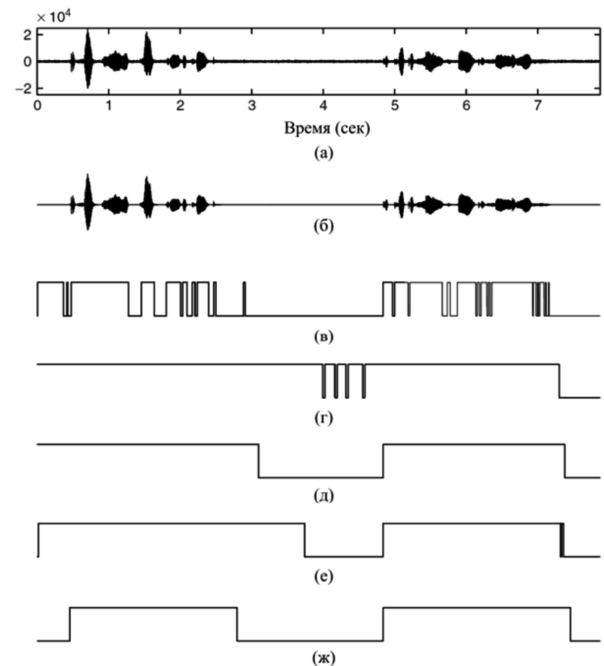


Рис. 10. Сравнение результатов работы алгоритмов VAD для речевого сигнала с транспортным шумом (ОСШ 15 дБ): а) входной зашумленный речевой сигнал; б) чистая речь; в) G.729B; г) IS-127; д) GSM-EFR; е) AMR1; ж) AMR2

Как видно из приведенных графиков, ни один метод не демонстрирует высокой достоверности и точности обнаружения активной речи и речевых пауз одновременно. Таким образом, встает проблема разделения речевого сигнала на интервалы активной речи и речевых пауз с высокой степенью достоверности и наименьшими потерями.

#### Новый способ обнаружения пауз в речевых сигналах

В настоящей работе предложен детектор активности речи, обеспечивающий существенное повышение точности правильного разделения речевых сигналов на периоды активной речи и паузы [14].

Структурная схема детектора активности речи изображена на рис. 11. Речевой сигнал с выхода электроакустического преобразователя усиливается селективным усилителем и подается на вход сумматора. На второй вход сумматора подается сигнал с выхода генератора измерительного сигнала. Суммарный сигнал с выхода сумматора поступает на вход усилителя-ограничителя 1, где происходит усиление сигнала, а затем ограничение по амплитуде. Аналогичная операция проводится над сигналом, поступающим с выхода селективного усилителя на вход усилителя-ограничителя 2. Сигнал с выхода усилителя-ограничителя 1 подается на первый вход перемножителя. На второй вход перемножителя подается сигнал с выхода усилителя-ограничителя 2. Сигнал с выхода перемножителя поступает на вход накопителя-усреднителя (интервал накопления 10 мс), где происходит вычисление сигнала, по амплитуде которого принимают решение о наличии периода активного речевого сигнала или паузы в пороговом устройстве.





Рис. 11. Структурная схема детектора активности речи

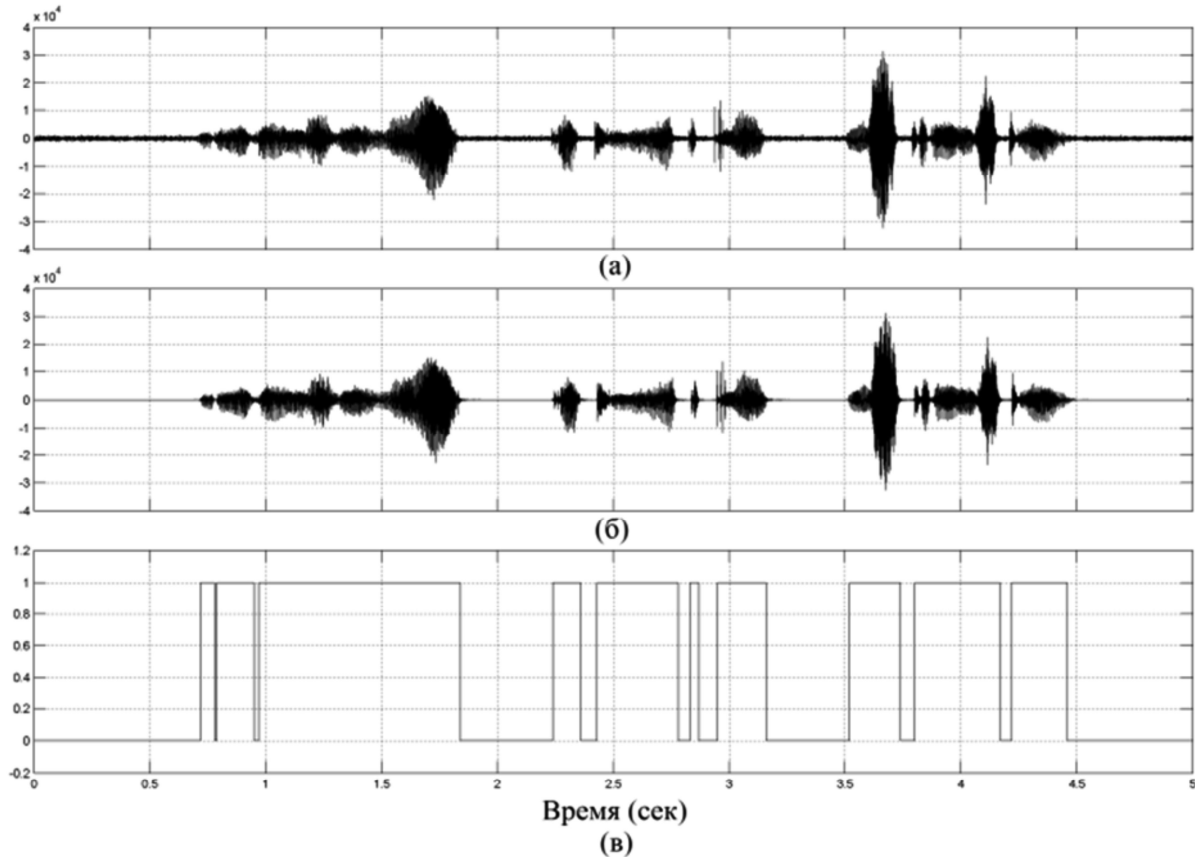


Рис. 12. Осциллограмма тестовой фразы с добавлением транспортного шума (ОСШ 15 дБ), чистого речевого сигнала и сигнала с выхода обнаружителя

Значение порога вычисляется в схеме вычисления порога путем анализа первых 150 мс от начала анализа, т.к. на этом интервале речь обычно отсутствует.

Для проведения исследования была выбрана тестовая фраза: «Продолжение отладки устройства». На рис. 12 представлена осциллограмма данной фразы с добавлением транспортного шума (ОСШ 15 дБ), чистого речевого сигнала и сигнала с выхода обнаружителя.

Общее время записи речевого сигнала указанной фразы составило 5 с, а суммарное время активных периодов – 2,64 с, что составляет 52,8 % времени выбранного речевого сигнала.

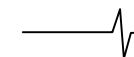
Количество ошибок обнаружения активной речи составило 3,56 %, количество ошибок обнаружения речевых пауз – 1,61 %.

Для проверки восприятия речи на слух после обра-

ботки была произведена запись речевого сигнала с помощью блока аудиозаписи, управляемого командами с обнаружителя пауз. Интервалы, в которых присутствовала речь, оставались без изменения. В интервалы, в которых находились паузы, записывался нуль. Качество полученного речевого сигнала практически не отличалось от исходного. Слова были хорошо различимы, речь легко воспринималась на слух.

### Заключение

Таким образом, можно сделать вывод о том, что разработан новый способ обнаружения пауз в речи, который позволяет существенно повысить точность разделения на активные участки речи и паузы. При этом качество восстанавливаемой речи остается практически на том же уровне.



Предметом последующего исследования является сравнительная оценка количества ошибок и вероятности правильного обнаружения участков активной речи и паузы при различных уровнях ОСШ.

### Литература

1. Kondo A.M. Digital Speech. Coding for Low Bit Rate Communication Systems. – John Wiley & Sons, Ltd. 2004. – 442 p.
2. Шелухин О.И. Цифровая обработка и передача речи / О.И. Шелухин, В.Г. Лукьянцев; Под ред. О.И. Шелухина. – М.: Радио и связь, 2000. – 456 с.: ил.
3. ITU-T (1996) A silence compression scheme for G.729 optimised for terminals conforming to ITU-T V.70, ITU-T Rec. G.729 Annex B.
4. ITU-T (1996) Dual rate speech coder for multimedia communications transmitting at 5.3 and 6.3 kbit/s. Annex A: Silence compression scheme, ITU-T Rec. G.723.1 Annex A.
5. ITU-T (1996) Coding of speech at 8 kbit/s using conjugate-structure algebraic code excited linear prediction (CS-ACELP), ITU-T Rec. G.729.
6. ITU-T (1996) Dual rate speech coder for multimedia communications transmitting at 5.3 and 6.3 kbit/s, ITU-T Rec. G.723.1.
7. ETSI (1998) Digital cellular telecommunications system (phase 2+); Voice activity detector (VAD) for full rate speech traffic channels, GSM 06.32 (ETSI EN 300 965 v7.0.1).
8. ETSI (1999) Digital cellular telecommunications system (phase 2+); Voice activity detector (VAD) for full rate speech traffic channels, GSM 06.42 (draft ETSI EN 300 973 v8.0.0).
9. ETSI (1997) Digital cellular telecommunications system; Voice activity detector (VAD) for enhanced full rate (EFR) speech traffic channels, GSM 06.82 (ETS 300 730), March.
10. ETSI (1998) Digital cellular telecommunications system (phase 2+); Voice activity detector (VAD) for adaptive multi-rate (AMR) speech traffic channels, GSM 06.94 v7.1.1 (ETSI EN 301 708).
11. DeJaco P., Gardner W., and C. Lee (1993) 'QCELP: The North American CDMA digital cellular variable rate speech coding standard', in IEEE Workshop on Speech Coding for Telecom, pp. 5–6.
12. TIA/EIA (1997) Enhanced variable rate codec, speech service option 3 for wideband spread spectrum digital systems, IS-127.
13. TIA/EIA (1998) High rate speech service option 17 for wideband spread spectrum communication systems, IS-733.
14. Пат. 2436173 Российская Федерация, МПК G10L 15/00, G10L 11/02, Способ обнаружения пауз в речевых сигналах и устройство его реализующее / Витязев В.В., Розов В.И., Волченков В.А.; заявитель и патентообладатель Рязанский государственный радиотехнический университет. – № 2010124342/08, заяв. 15.06.10; опубл. 10.12.11, Бюл. 34.

### VOICE ACTIVITY DETECTION METHODS

*Volchenkov V.A., Vityazev V.V*

Problems of accuracy increase in voice activity detection (VAD) are considered. General information about standard VAD methods and performance comparison of VADs are given. A new method of voice activity detection is offered.

### 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'» и Выставках: «ExpoElectronica», «СвязьЭкспокомм», «ЭЛЕКТРОНИКА: компоненты, оборудование, технологии» (г. Москва) и др.

Научно-технический журнал «Цифровая обработка сигналов» включен в Список изданий, рекомендуемый ВАК РФ для публикации результатов научных исследований соискателями ученой степени доктора и кандидата технических наук в области радиотехники, связи, вычислительной техники, электроники, информационно-измерительных и управляющих систем.

#### Планируемые сроки издания отдельных номеров журнала:

№ 2 – июнь 2013 г. Тематический выпуск: «Теория и методы цифровой обработки сигналов» (по материалам международной научно-технической конференции «Цифровая обработка сигналов и ее применение – DSPA' 2013».

№ 3 – сентябрь 2013 г. Тематический выпуск: «Цифровая обработка изображений».

№ 4 – декабрь 2013 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 681.513.6

## ПРИЛОЖЕНИЕ ПРИНЦИПОВ ЦИФРОВОЙ АДАПТИВНОЙ ФИЛЬТРАЦИИ К ЗАДАЧЕ УПРАВЛЕНИЯ ДИНАМИЧЕСКИМ ОБЪЕКТОМ С НЕИЗВЕСТНОЙ МАТЕМАТИЧЕСКОЙ МОДЕЛЬЮ

*Гудкова Н.В., канд. техн. наук, доцент каф. САУ Южного федерального университета,  
e-mail: tala\_gud@rambler.ru*

**Ключевые слова:** динамический объект, адаптивная модель, функциональная обратная связь, адаптивный трансверсальный фильтр, метод наименьших квадратов, нерегулярные колебания, дрейф выходного сигнала.

### Введение

Известно, что большинство традиционных методов синтеза автоматических систем управления базируется на использовании математических моделей управляемых объектов. Однако эти методы применимы далеко не всегда, т.к. на практике получить адекватное математическое описание динамических свойств сложного технологического процесса зачастую трудно, а иногда и невозможно.

В подобных случаях выходом из положения может стать применение адаптивных регуляторов, предназначенных для управления так называемыми неизвестными объектами (НО), то есть объектами с неизвестной математической моделью, иначе, объектами типа «черный ящик» [1–4].

Статья посвящена исследованию одного из способов управления НО, который базируется на принципах прямого и обратного адаптивного моделирования динамического объекта в процессе его функционирования [1]. Предполагается, что объект обладает колебательными свойствами, имеет детерминированные параметры и структуру, а его адаптивные модели реализуются в виде трансверсальных фильтров (АТФ) с весовыми коэффициентами, перестраиваемыми по методу наименьших квадратов (LMS).

### Алгоритм функционирования системы управления

На рис. 1 показана одна из возможных структур адаптивной системы, которая предназначена для подавления колебаний в управляемом объекте и минимизации установившейся ошибки между временными отсчетами задающего (входного) воздействия  $g_k$  и выходного сигнала системы  $x_k$ , где  $k = 0, 1, 2, \dots$

Для управления объектом в системе формируются два одновременно протекающих адаптивных процесса: процесс адаптивной идентификации (прямое моделирование НО) и процесс формирования управляющего воздействия (обратное адаптивное моделирование НО). Строго говоря, эти процессы не являются независимы

*Рассматривается один из способов управления непрерывными колебательными объектами с неизвестной математической моделью, базирующийся на принципах прямого и обратного адаптивного моделирования. Для целей управления в исследуемой системе формируются два одновременно протекающих адаптивных процесса: процесс адаптивной идентификации (прямое моделирование объекта) и процесс формирования управляющего воздействия (обратное моделирование). Адаптивные модели объекта реализуются в виде адаптивных трансверсальных фильтров с весовыми коэффициентами, перестраиваемыми по методу наименьших квадратов в режиме реального времени. В отличие от традиционных замкнутых систем автоматического управления в рассматриваемой адаптивной структуре отсутствует физическая отрицательная обратная связь между выходом и входом. Ее роль играет функциональная обратная связь, замыкающаяся через адаптивный процесс. Показано, что предлагаемая методика синтеза такой системы позволяет минимизировать ошибки управления и демпфировать колебания, а также компенсировать случайный низкочастотный дрейф выходного сигнала в управляемом объекте. Результаты имитационного моделирования системы подтверждают эффективность предложенных решений.*

ми, но при достаточно медленной адаптации можно считать, что они протекают независимо.

**I. Подсистема адаптивной идентификации** НО предназначена для реализации адаптивной модели (АМ), которая перестраивает свои весовые коэффициенты таким образом, чтобы минимизировать среднеквадратическую ошибку идентификации объекта  $СКО_1 = E[\varepsilon_k^2]$ .

В схеме  $\varepsilon_k = x_k - y_k$ , где  $y_k$  – выходной сигнал АМ.

Дискретное уравнение адаптивной модели имеет вид

$$y_k = \sum_{l=0}^L w_{lk}^{AM} u_{k-l}, \quad (1)$$

где  $u_k$  – управляющее воздействие,  $L$  – длина АТФ,

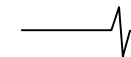
$w_{lk}^{AM}$  – временные отсчеты  $l$ -го весового коэффициента фильтра на  $k$ -й итерации обучения.

Алгоритм LMS<sub>1</sub>, перестраивающий параметры АМ, представляет собой рекуррентное выражение

$$w_{l(k+1)}^{AM} = w_{lk}^{AM} + 2\mu^{AM} u_{k-l} \varepsilon_k, \quad (2)$$

где  $\mu^{AM}$  – параметр (шаг) сходимости алгоритма идентификации.

**II. Подсистема адаптивного обратного моделирования** НО предназначена для формирования управ-



ляющего воздействия  $u_k$ , которое минимизирует среднеквадратическую ошибку управления  $СКО_2 = E[e_k^2]$ , где

$$e_k = x_{\text{э}k} - x_k. \quad (3)$$

Здесь  $x_{\text{э}k}$  – сигнал, который является эталонным для выходного сигнала управляемого объекта. В рассматриваемой системе он представляет собой задержанный на  $m$  тактов сигнал  $g_k$ , т.е.  $x_{\text{э}k} = g_{k-m}$ .

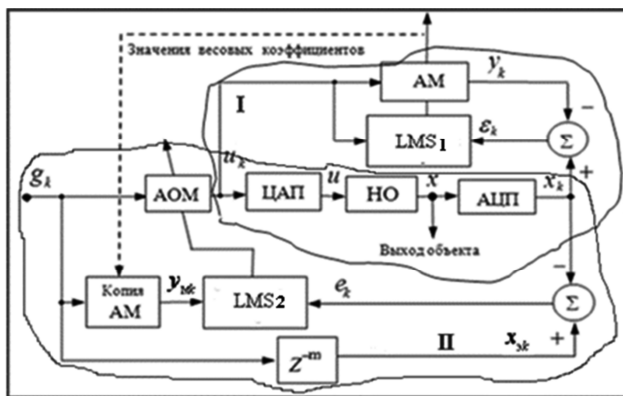


Рис. 1. Структурная схема адаптивной системы управления НО

Роль регулятора, формирующего управляющее воздействие, играет адаптивная обратная модель объекта (АОМ), дискретное уравнение которой имеет вид

$$u_k = \sum_{l=0}^L w_{lk}^p g_{k-l}. \quad (4)$$

Алгоритм LMS<sub>2</sub> по аналогии с (2) перестраивает весовые коэффициенты  $w_{lk}^p$  по формуле

$$w_{l(k+1)}^p = w_{lk}^p + 2\mu^p y_{m(k-l)} e_k, \quad (5)$$

где  $\mu^p$  – шаг сходимости алгоритма управления. Сигнал  $y_{mk}$ , входящий в выражение (5), является выходным сигналом копии АМ управляемого объекта и вычисляется по формуле

$$y_{mk} = \sum_{l=0}^L w_{lk} g_{k-l}. \quad (6)$$

Для подавления низкочастотного дрейфа сигнала  $x_k$ , а также в целях повышения запаса устойчивости и быстродействия системы в АОМ вводится весовой коэффициент смещения  $w'_k$ , адаптивная перестройка которого также осуществляется с помощью алгоритма LMS по формуле

$$w'_{k+1} = w'_k + 2\mu^p e_k. \quad (7)$$

Закон управления (4) в этом случае принимает вид

$$u_k = w'_k + \sum_{l=0}^L w_{lk}^p g_{k-l}. \quad (8)$$

После завершения переходных процессов в системе величина СКО<sub>2</sub> стремится к своему минимальному значению, а отклик объекта  $x_k$  на задающее воздействие

$g_k$  становится приблизительно равным эталонному сигналу  $x_{\text{э}k}$ .

### Методика выбора параметров настройки адаптивных фильтров АМ и АОМ

Обычно качество систем управления динамическими объектами принято оценивать по запасу устойчивости, быстродействию и точности обработки так называемых медленных (по сравнению с переходными процессами) детерминированных задающих воздействий (уставок) вида:

1.  $g(t) = g_0 1(t)$  (управление положением);
2.  $g(t) = \Omega t$  (управление движением с постоянной скоростью);
3.  $g(t) = at^2$  (управление движением с постоянным ускорением);
4.  $g(t) = A_0 \cos(\omega t + \varphi)$  или  $g(t) = A_0 \sin(\omega t + \varphi)$  (гармонические колебания или регулярная качка).

Параметрами настройки адаптивных моделей в рассматриваемой системе являются число коэффициентов АТФ и параметры сходимости алгоритмов LMS  $\mu^{\text{AM}}$  и  $\mu^p$ . Они должны выбираться из условий обеспечения требуемых показателей качества переходных процессов (колебательности и времени регулирования) и допустимых величин ошибок управления в установившихся режимах [5].

В настоящее время в специальной литературе практически отсутствуют рекомендации по выбору параметров настройки адаптивных фильтров, применяемых в системах управления, а приводимые примеры носят частный характер. Это обусловлено большим разнообразием задач, а также отсутствием во многих случаях строгого математического обоснования их решения.

В работах [2, 3] предложена методика синтеза программируемых АТФ, предназначенных для идентификации и управления непрерывными объектами типа «черный ящик» при минимальной априорной информации о входных сигналах и динамике объекта. Эта методика не дает оптимальных решений, но, как показала практика, для большого класса систем она обеспечивает приемлемое качество процессов управления.

Исходными данными для расчета являются:

- 1) время установления  $t_{\text{уст}}$  и, по возможности, вид переходной характеристики неизвестного объекта;
- 2) приблизительный частотный диапазон и предельная мощность входных сигналов адаптивных фильтров.

Поскольку процессы в управляющей системе носят дискретный характер, в первую очередь необходимо выбрать интервал квантования сигналов по времени  $T$ . Для систем автоматического управления обычно он выбирается из соотношения

$$T_{\text{алг}} \leq T \leq (0,01 \div 0,001) t_{\text{уст}}, \quad (9)$$

где  $T_{\text{алг}}$  – время обработки информации в управляющем компьютере и преобразователях ЦАП/АЦП.

Из выражения (9) видно, что нижняя граница для интервала  $T$  определяется быстродействием используемой цифровой аппаратуры, а верхняя – инерционностью объекта управления. Поэтому выбор величины  $T$  обычно является результатом компромиссного решения.

Для выбора параметров настройки адаптивного фильтра используются соотношения вида

$$\mu(L+1) \approx \frac{M}{P} \quad (10)$$

и

$$(L+1) \cong \lceil MT_{\text{ско}} \rceil, \quad (11)$$

где  $P$  – средняя мощность сигнала на входе АТФ;  $M$  ( $0 < M < 1$ ) – допустимое относительное среднее значение СКО, от которого зависит устойчивость, точность и скорость адаптации фильтра;  $T_{\text{ско}}$  – число итераций адаптивного процесса в фильтре.

Анализ показал, что для получения хорошего качества управления в рассматриваемой структуре величины  $T_{\text{ско}}$  АМ и АОМ должны выбираться из условия  $T_{\text{ско}} \leq t_{\text{уст}} / T$  при значениях  $M \approx 0,001 - 0,2$ .

Следует отметить, что зачастую мощность сигналов на входах адаптивных фильтров заранее не известна. В таких случаях для начала в качестве расчетной величины можно принять значение предельно допустимой мощности постоянного входного сигнала. Это приведет к затягиванию процессов адаптации, но повысит запас устойчивости системы управления. В ходе эксплуатации системы параметры настройки могут уточняться.

Радикальным решением проблемы выбора оптимальных значений параметра  $\mu$  для каждой из адаптивных моделей является их расчет в режиме реального времени, процедура которого предложена в [6].

Поясним эту процедуру на примере расчета текущих значений параметра  $\mu$  для подсистемы идентификации неизвестного объекта.

1. Вычисляется текущее среднее значение мощности сигнала на входе АМ по формуле  $P_k = E[u_k^2]$ . Для усреднения можно использовать цифровой адаптивный фильтр низких частот (ФНЧ), реализуемый в виде дискретного уравнения [7]

$$P_k = (1 - \lambda)(u_k^2 - u_{k-1}^2) + (1 - 2\lambda)P_{k-1}, \quad (12)$$

где  $0 < \lambda < 0,5$  – параметр, отвечающий за точность и скорость сходимости этого алгоритма.

2. Далее вычисляется текущее значение  $\mu_k^{\text{AM}}$  по формуле (10), которая в данном случае модифицируется

$$\mu_k^{\text{AM}} \approx \frac{M}{(L+1)P_k}.$$

Аналогичным образом рассчитывается параметр  $\mu_k^{\text{АОМ}}$  для АОМ.

Помимо параметров адаптивных моделей на качество управления влияет также число тактов задержки  $m$  эталонного сигнала  $x_{\text{э}k}$ , которая учитывает инерционные свойства канала управления. Ее выбор носит эмпирический характер. Практика показала, что эта задержка не должна превышать половины длины АОМ.

Необходимо подчеркнуть, что в отличие от традиционной замкнутой системы автоматического управления в рассматриваемой схеме отсутствует физическая отрицательная обратная связь между выходом и входом. Ее роль играет функциональная (алгоритмическая) обратная связь, замыкающаяся через адаптивный процесс.

### Имитационное моделирование адаптивной системы управления на ПК

Ниже приводятся результаты компьютерного моделирования процессов в системе адаптивного управления НО, имитационная модель которого представлена передаточной функцией колебательного звена

$$W(s) \approx \frac{0,8}{(s^2 + 0,4s + 1)}$$

с временем установления переходной характеристики  $t_{\text{уст}} \approx 20$  с.

Выбраны следующие параметры адаптивной системы.

Число тактов задержки эталонного сигнала  $m = 0$ . Интервал квантования сигналов по времени  $T = 0,2$  с. Нормированные параметры типовых входных воздействий:  $g_0 = 1$ ;  $\Omega = 0,008$ ;  $a = 0,000028$ ;  $A_0 = 1$ ;  $\omega = 0,4$ .

Предельные нормированные мощности входного и управляющего воздействий  $E[g_k^2] = E[u_k^2] = 1$ . Длина фильтров, реализующих АМ и АОМ,  $L = 3$ . Допустимые относительные средние значения СКО идентификации и управления  $M_1 = 0,1$  и  $M_2 = 0,02$ , соответственно. Используется алгоритм управления (8). На рис. 2-7 показаны для сравнения некоторые характерные процессы в неуправляемом объекте и в адаптивной системе (в качестве типовых воздействий использовались их дискретные модели).

Полученные данные можно прокомментировать следующим образом.

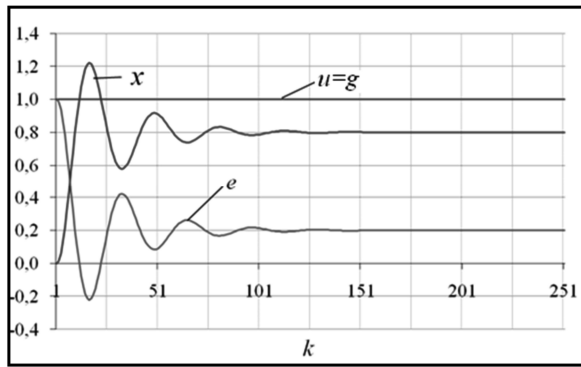
В неуправляемом объекте:

1) переходные характеристики при  $g_k = 1$  имеют колебательный характер, а статическая ошибка управления  $e_{\text{ст}} = g_0 - x_0$  постоянна и не равна нулю (рис. 2а);

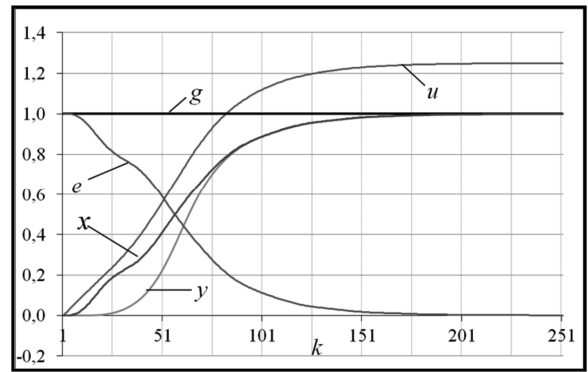
2) при  $g_k = \Omega \cdot kT$  скоростная ошибка  $e_{\text{ск}}$  линейно возрастает во времени с постоянной скоростью, равной величине  $\Omega \cdot e_{\text{ст}}$ . (рис. 3а);

3) при  $g_k = a \cdot (kT)^2$  ошибка по ускорению  $e_{\text{уск}}$  возрастает во времени с постоянным ускорением, равным величине  $a \cdot e_{\text{ст}}$  (рис. 4а);

4) при  $g_k = A_0 \cdot \text{Cos}(\omega kT + \varphi)$  ошибка установившегося режима изменяется по косинусоидальному закону с амплитудой, равной  $e_{\text{ст}}$  неуправляемого объекта (рис. 5 а).

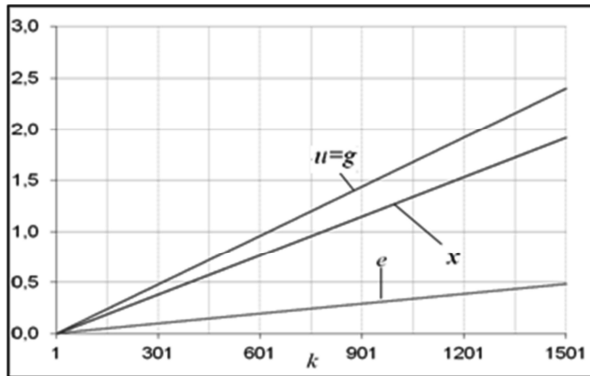


а)

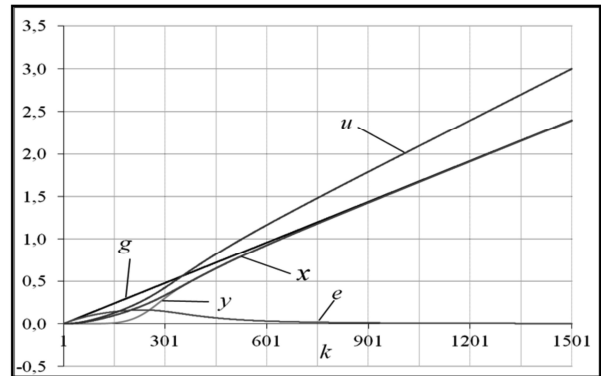


б)

Рис. 2. Результаты компьютерного моделирования процессов в неуправляемом НО (а) и в адаптивной системе (б) при  $g_k = 1$

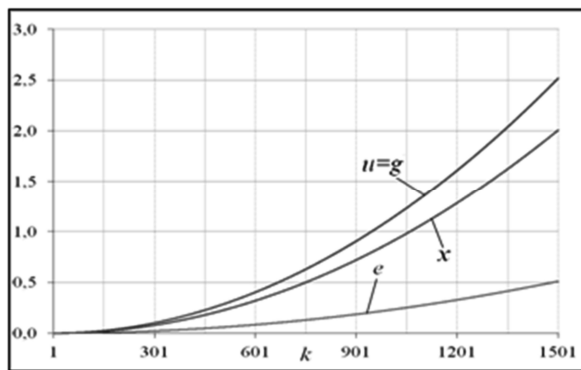


а)

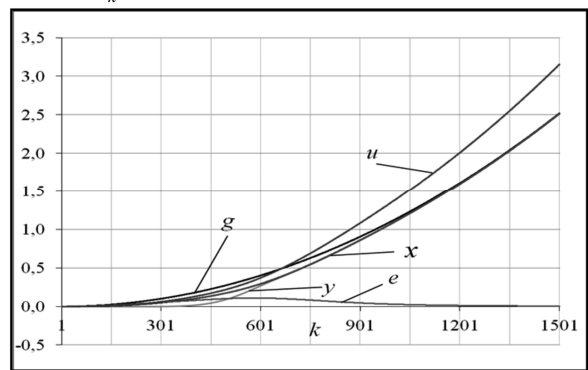


б)

Рис. 3. Результаты компьютерного моделирования процессов в неуправляемом НО (а) и в адаптивной системе (б) при  $g_k = \Omega kT$

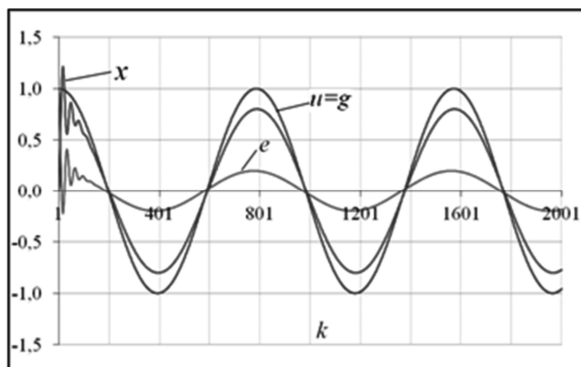


а)

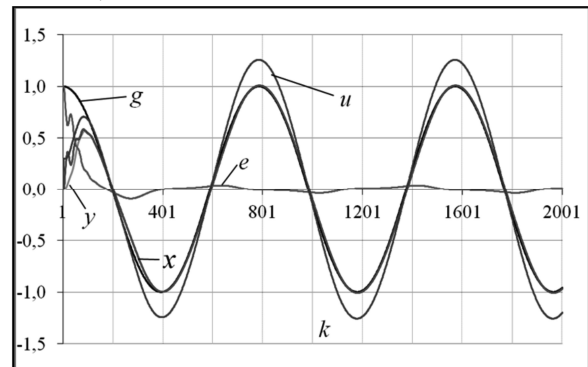


б)

Рис. 4. Результаты компьютерного моделирования процессов в неуправляемом НО (а) и в адаптивной системе (б) при  $g_k = a \cdot (kT)^2$



а)



б)

Рис. 5. Результаты компьютерного моделирования процессов в неуправляемом НО (а) адаптивной системе (б) при  $g_k = A_0 \cdot \cos(\omega kT + \varphi)$

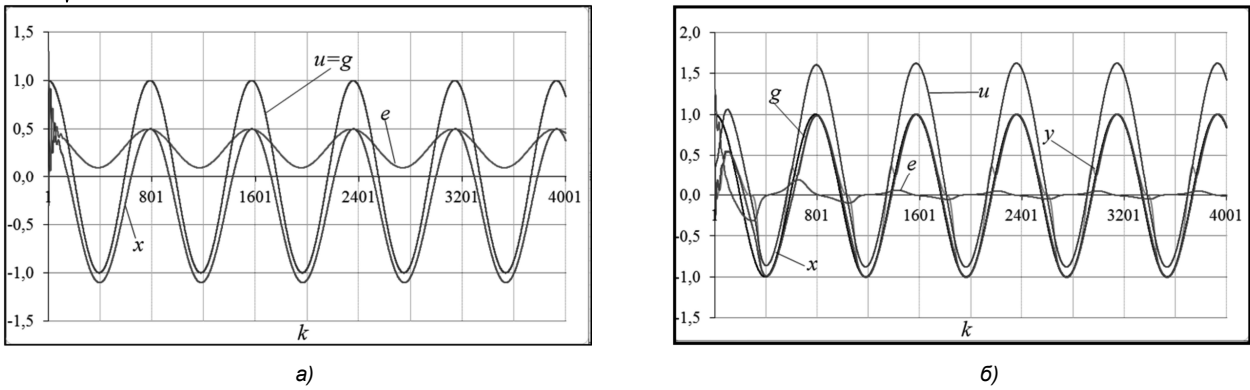


Рис. 6. Результаты компьютерного моделирования процессов в неуправляемом НО (а) и в адаптивной системе (б) при  $g_k = A_0 \cos(\omega kT + \varphi)$  и постоянном дрейфе выходного сигнала  $x_o(t)$

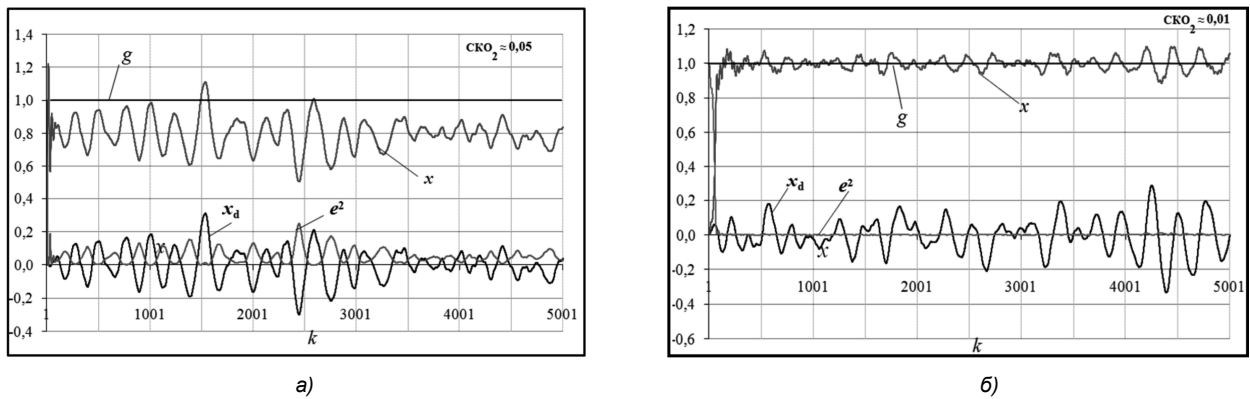


Рис. 7. Результаты компьютерного моделирования процессов в неуправляемом НО (а) и в адаптивной системе (б) при  $g_k = 1$  и дрейфе выходного сигнала  $x_o(t)$  в виде низкочастотных нерегулярных колебаний

В адаптивной системе:

- 1) полностью подавляются собственные колебания объекта, при этом установившиеся ошибки при обработке типовых задающих воздействиях равны нулю (рис. 2б – рис. 5б);
- 2) рис. 6а и 6б иллюстрируют способность адаптивной системы подавлять постоянный дрейф выходного сигнала неуправляемого объекта;
- 3) на рис. 7а и 7б показаны процессы компенсации адаптивной системой случайного дрейфа выходного сигнала в виде низкочастотных нерегулярных колебаний в неуправляемом объекте (их имитационная модель получена автором путем фильтрации «дискретного белого шума» специальным алгоритмом).

### Заключение

Анализ результатов проведенных исследований позволяет сделать следующие выводы:

- отличительной особенностью функционирования рассматриваемой системы является одновременное выполнение адаптивных процедур прямого и обратного моделирования управляемого объекта, предназначенных для формирования управляющих воздействий в режиме реального времени.
- адаптивные модели объекта реализуются в виде адаптивных трансверсальных фильтров, для синтеза которых не требуется его математическое описание.
- имитационное моделирование на ПК показывает, что адаптивная система:

- 1) успешно демпфирует собственные колебания и компенсирует постоянный и низкочастотный дрейф выходного сигнала в виде нерегулярных колебаний;

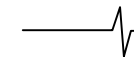
- 2) обладает способностью приводить к нулю установившиеся ошибки управления объектом в режимах регулярной качки, а также в режимах управления по положению, по скорости и ускорению, что дает основание отнести данную структуру к классу астатических систем, по меньшей мере, третьего порядка.

Таким образом, можно считать, что рассмотренный способ управления колебательными объектами может успешно конкурировать с традиционными замкнутыми системами, особенно, при неизвестном или недостаточно адекватном математическом описании объекта и сложности реализации высокого порядка астатизма.

### Литература

1. Уидроу Б., Стирнз С. Адаптивная обработка сигналов / М: Радио и связь, 1989.
2. Гудкова Н.В. Алгоритмы адаптивной идентификации технических объектов. Автоматизация и современные технологии. М: Машиностроение. №8. 2005.
3. Гудкова Н.В. Цифровое управление техническими объектами с применением адаптивного обратного моделирования. Автоматизация и современные технологии. М: Машиностроение. №4, 2006.
4. Widrow B., Walach E. Adaptive Inverse Control. A Signal Processing Approach, Wiley, Hoboken, NJ, 2008.
5. Бесекинский В.А., Попов Е.П. Теория систем автоматического управления. Изд. 4-е, перераб. и доп. СПб, Изд-во «Профессия», 2004.
6. Гудкова Н.В. Адаптивные алгоритмы регулирования уровня дискретных сигналов. Цифровая обработка сигналов. № 2, 2005.





7. Гудкова Н.В. Адаптивные цифровые фильтры в системах промышленной электроники. Автоматизация и современные технологии. М: Машиностроение. №7, 2005.

**APPLICATION THE PRINCIPLE DIGITAL ADAPTIVE FILTERING TO CONTROL DYNAMICAL PLANT WITH UNKNOWN MATHEMATICAL MODEL**

*Gudkova N.*

The paper considers a method of control to the unknown plant based on the direct and inverse adaptive modeling.

Adaptive models are adaptive transversal filters with the weights, calculated by Least Mean Square algorithm. The control system is developed. The minimization of the error between system input and plant output is the target. There are two simultaneous adaptive processes in the system. The direct adaptive identification is used to build the plant adaptive model. The inverse model is a control device for the plant. The physical negative feedback between an output and an input of the adaptive system is absent. However, the adaptive system has a functional feedback. The results of computer simulation of an adaptive control system are presented.

---

**ПОЗДРАВЛЕНИЯ ЮБИЛЯРУ**



*21 февраля заведующему кафедрой ЭВМ,  
доктору технических наук, профессору,  
Заслуженному деятелю науки и техники РФ,  
Почетному гражданину города Рязани*

**ЗЛОБИНУ  
ВЛАДИМИРУ КОНСТАНТИНОВИЧУ**

*исполнилось 75 лет*

Вся деятельность Злобина В.К. тесным образом связана с Рязанским радиотехническим институтом-университетом. Свой трудовой путь, после окончания в 1960 году кафедры ЭВМ по специальности «Математические и счетно-решающие приборы и устройства», он начал на этой же кафедре в должности младшего научного сотрудника. Злобин В.К. более 45 лет проработал в РРТИ-РГРТА-РГРТУ в различных должностях, в том числе в 1983-2007 годах (почти 25 лет) в должности ректора, с 1980 года по настоящее время (более 30 лет) в должности заведующего кафедрой ЭВМ. Одновременно по совместительству и на общественных

началах исполнял обязанности зам. начальника и научного руководителя ОКБ «Спектр», научного руководителя НИИ «Фотон». Под его руководством подготовлено 15 кандидатов и докторов наук. Опубликовано 2 монографии, 4 учебно-методических пособия с грифом УМО, множество научных статей в российских и зарубежных изданиях по цифровой обработке изображений.

За многолетний добросовестный труд Злобин В.К. награжден правительственными наградами, среди которых Орден «Знак Почета». Ему присвоено почетное звание «Заслуженный деятель науки и техники РФ». Успехи в области развития высшего профессионального образования и плодотворная научная деятельность отмечены многими ведомственными наградами, такими как знак Минобразования СССР «За отличные успехи в работе», медаль Роскосмоса «Звезда Циолковского», Золотая медаль им. академика В.Ф. Уткина, медаль им. С. Королева Федерации космонавтики России, медаль им. Альфреда Нобеля Российской академии естествознания. В 2007 году Злобину В.К. присвоено почетное звание «Почетный гражданин г. Рязани».

В 2009 году Злобину В.К. присвоено почетное звание основателя научной школы «Цифровая обработка изображений». За заслуги в области отечественного образования Злобин В.К. награжден нагрудным знаком Российской академии естествознания «Золотая кафедра России».

Желаем Владимиру Константиновичу доброго здоровья и дальнейших творческих успехов.

*Редколлегия журнала «Цифровая обработка сигналов»*

УДК 681.142.2

## ПОВЫШЕНИЕ ВОССТАНАВЛИВАЕМОСТИ КЛАСТЕРНЫХ НЕЙРОПРОЦЕССОРНЫХ СИСТЕМ ОБРАБОТКИ ДАННЫХ

*Григоренко Д.В., директор ЗАО «Рязаньприбор», аспирант кафедры ЭВМ Рязанского государственного радиотехнического университета*

*Ручкин В.Н., д.т.н., профессор кафедры информатики и вычислительной техники Рязанского государственного университета имени С.А. Есенина, e-mail: v.ruchkin@rsu.edu.ru*

**Ключевые слова:** нейропроцессорные системы, теоретико-множественная модель, классы эквивалентности, рефлексивность, транзитивность, кластеризация, восстанавливаемость, безотказность, вероятность безотказной работы.

*Исследуются возможности повышения восстанавливаемости кластерных нейропроцессорных систем конвейерной, векторной, конвейерно-векторной или векторно-конвейерной структур обработки данных на базе современного отечественного микрокомплекта ИМ 640Х.*

### Введение

Двумя основными проблемами систем телекоммуникаций, обработки транзакций и управления базами данных являются обеспечение высокой производительности и достаточно длительного функционирования [1]. Наиболее эффективный способ достижения заданного уровня производительности – применение параллельных масштабируемых архитектур [2, 3]. Для обеспечения длительного функционирования необходимо повышение надежности, готовности и удобства обслуживания, которые взаимосвязаны и применяются совместно. Повышение надежности основано на принципе предотвращения неисправностей посредством уменьшения интенсивности отказов и сбоев, например, высокой и сверхвысокой степени интеграции или архитектурных совершенствованиях аппаратуры. Увеличение уровня готовности предполагает подавление в определенных пределах влияния отказов и сбоев с помощью средств контроля и коррекции ошибок или аппаратную и программную избыточность реализации различных вариантов отказоустойчивых архитектур. Главной задачей создания «систем высокой готовности» (HighAvailabilitySystems) является снижение времени простоя системы, минимизация которой требует различной стратегии и технологии.

Авторами предлагается концептуальная модель проектирования, позволяющая посредством кластеризации реализовать систему как одну из возможных структур и за счет резервирования повысить восстанавливаемость кластерных нейропроцессорных систем.

### Математические методы кластерного анализа

С развитием компьютерных сетей кластерами стали называть два и более компьютеров (серверы или рабочие станции), соединенных в единую систему специальным программным и аппаратным обеспечением [3]. Такие кластеры можно рассматривать как гетерогенную вычислительную систему с распределенной памятью и распределенным управлением. В результате в системе или сети обеспечивается достаточно высокий уровень распараллеливания, надежности, готовности и удобства

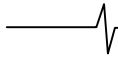
обслуживания при относительно низких затратах.

Одним из строгих определений термина «кластеризации» является определение: «реализация объединения машин, представляющегося единым целым для операционной системы, системного программного обеспечения, прикладных программ и пользователей». Машины, кластеризованные вместе подобным образом, могут при отказе одного процессора очень быстро перераспределить работу на другие процессоры как внутри кластера, так и между кластерами. Таким образом, организованные системы имеют следующие свойства: разделение ресурсов, высокую готовность, высокую пропускную способность, удобство обслуживания системы и расширяемость.

Важный раздел теории распознавания составляют методы кластеризации (или автоматической классификации, таксономии, самообучения, обучения без учителя, группировки), решающие задачи разбиения объектов (при заданных признаковых пространствах или матрицах близостей объектов) на классы эквивалентности [4], причем эквивалентность объектов базируется на мерах близости, сходства и т.п. Из перечисленного набора используемых названий, близких друг другу и воспринимаемых почти как синонимы, будем далее использовать термин «кластер» как синоним термину «класс», но обозначающий множество близких объектов, полученное как результат решения задачи кластерного анализа.

Методы кластерного анализа позволяют решать задачи минимизации числа эталонов, поиска эталонных описаний, выявления структурных свойств классов и многие другие вопросы анализа данных. Принципы, согласно которым объекты объединяются в один кластер, являются обычно «внутренним делом» конкретного алгоритма кластеризации. Пользователь, зная данные принципы, может в определенных пределах интерпретировать результаты каждого конкретного метода.

В отличие от задач распознавания, различные методы кластеризации могут приводить к решениям, имеющим весьма существенные различия. Таким образом, кроме набора разнообразных методов кластеризации,



практический интерес представляет наличие средств автоматической обработки результатов, полученных независимо различными алгоритмами [3-5]. Существует ряд подходов для решения задачи кластерного анализа.

Предварительно следует сделать два замечания. Во-первых, различают задачи кластерного анализа (и соответственно алгоритмы) с заданным (или известным) числом кластеров, а также с не заданным (неизвестным) числом кластеров. В последнем случае оптимальная кластеризация и число кластеров находятся в результате решения единой задачи. Во-вторых, кроме обычной постановки задачи кластеризации, как задачи поиска разбиений, существуют постановки как задачи поиска покрытий и структур на заданном множестве прецедентов.

Далее основная задача кластеризации будет рассматриваться прежде всего как задача поиска разбиений выборки признаковых описаний  $I(S_1), I(S_2), \dots, I(S_m), I(S) = (x_1(S), x_2(S), \dots, x_n(S))$ , заданной числовой таблицей  $T_{nm}$

В «нестрогой» постановке данная задача формулируется как поиск разбиения выборки на группировки (классы, кластеры, таксоны) близких объектов, причем само искомое разбиение находится как решение некоторой оптимизационной задачи, как результат сходимости некоторой итерационной процедуры, как результат применения некоторой детерминированной процедуры и т.п.

В общем случае пусть рассматривается задача кластеризации на  $l$  кластеров. Выборку признаковых описаний объектов будем обозначать как  $X = \{x_1, x_2, \dots, x_m\}$ ,  $x_i = \{x_{i1}, x_{i2}, \dots, x_{in}\}$ .

Разбиением  $K = \{K_1, K_2, \dots, K_l\}$  выборки на  $X = \{x_1, x_2, \dots, x_m\}$  групп является произвольная совокупность непересекающихся подмножеств множества  $X$ , покрывающая все объекты выборки

$$K_i \subseteq X, i = 1, 2, \dots, l,$$

$$\bigcup_{i=1}^l K_i = X, K_i \cap K_j = \emptyset, i \neq j$$

Пусть задан некоторый критерий качества  $F(K)$  разбиения  $K$ . Тогда задача кластеризации будет состоять в нахождении разбиения  $K^*$ , доставляющего экстремум критерию

$$F(K) : F(K^*) = \underset{K \in \{K\}}{extr} F(K)$$

В качестве примеров могут быть использованы критерии [6].

1 Сумма внутриклассовых дисперсий или сумма квадратов ошибок.

$$F(K) = \sum_{j=1}^l \sum_{x_i \in K_j} \rho^2(x_i, y_j)$$

где  $y_j = \frac{1}{n_j} \sum_{x_i \in K_j} x_i$ ,  $n_j = |K_j| \approx$  число объектов в группе  $K_j$ .

Решением задачи кластерного анализа при данном критерии считается такое разбиение  $K^*$ , которое доставляет минимум функционалу  $F(K)$

2. Критерии на базе матриц рассеяния. Матрица рассеяния для группы  $K_j$  определяется как

$$\sum_j = \sum_{x_i \in K_j} (x_i - y_j)(x_i - y_j)^t,$$

а матрица внутригруппового рассеяния как  $\sum = \sum_{i=1}^l \sum_j$  (здесь символ  $t$  означает символ транспонирования).

Известно несколько определений критериев кластеризации на базе матриц внутригруппового рассеяния. Например, выбор в качестве критерия определителя матрицы внутригруппового рассеяния  $F(K) = |\Sigma|$ .

Решение задачи кластерного анализа при данном критерии также находится в результате дискретной минимизации.

Широко известен метод « $k$ -внутригрупповых средних». В этом методе строится последовательность разбиений  $K = \{K_1^i, K_2^i, \dots, K_l^i\}$   $i = 1, 2, \dots$  как результат выполнения следующих однотипных итераций.

Пусть разбиение  $K$  выбрано случайно. Для группы  $K$  находится ее центр  $y_1 = \frac{1}{n_1} \sum_{x_j \in K_1^i} x_j$ .

Далее в группу  $K_l^{i+1}$  зачисляются все элементы выборки, которые ближе к  $y_1$  чем к аналогично полученным  $y_2, y_3, \dots, y_l$ .

Группа  $K_l^{i+2}$  строится аналогично, но относительно множества объектов  $\frac{X}{K_l^{i+1}}$  и т.д.

После вычисления  $K_1^{i+1}, K_2^{i+1}, \dots, K_l^{i+1}$  их центры пересчитываются, и вычислительный процесс повторяется.

Известен метод Фореля, который является представителем подходов, в котором кластеры находятся не в результате оптимизации некоторого критерия, а с помощью итерационных процедур – движения гипершаров фиксированного радиуса в сторону мест «сгущения» объектов [3, 6].

Пусть фиксировано некоторое положительное число  $R$ . Выбирается случайный элемент  $x_i \in X$  и гипершар радиуса  $R$  с центром в  $y_1 = x_i$ :  $R_1 = \{x : \rho(x, y_1) \leq R\}$ .

Полагаем  $K_1^1 = \{x_i : x_i \in X \cap R_1\}$ . Вычисляется  $K_1^2 = \{x_i : x_i \in X \cap R_2\}$ .

Вычисляется центр новой сферы  $y_2 = \frac{1}{|K_1^1|} \sum_{x_j \in K_1^1} x_j$

и группа  $R_2 = \{x : \rho(x, y_2) \leq R\}$ .

Процесс заканчивается вычислением такой группы объектов  $K_1^t = \{x_i : x_i \in X \cap R_t\}$ , для которой  $K_1^t = K_1^{t+1}$ .

**Теоретико-множественная модель**

Для решения поставленной задачи введем понятие равенства кластеров  $CL_l$  и  $CL_k$  обработки информации  $CL_l = CL_k$ , под которым далее понимается равенство длин указанных кластеров и совпадение их с точностью до команды  $MK_i^{(k)} = MK_i^{(l)}$ , т.е.

$$|CL_l| = |CL_k| \quad \forall l, k = 1, N;$$

$$MK_i^{(k)} = MK_i^{(l)}, \quad \forall i = 1, |CL_l|. \quad (1)$$

В работе показывается, что вводимая кластерная структура  $KS_w \in S$  обработки информации есть отношение эквивалентности и удовлетворяет условиям рефлексивности, симметричности и транзитивности.

Действительно, любой произвольный кластер обработки  $CL_l$  информации, удовлетворяющий (1), выполняется параллельно сам себе, т.е.  $CL_l \xrightarrow{KS_w} CL_l$ . Таким образом, справедливо условие рефлексивности любого кластера  $CL_l \in PR^{(j)}$  обработки информации.

Если кластер обработки информации  $CL_l$  равен кластеру  $CL_k$  и, следовательно, параллелен этому кластеру, тогда кластер обработки  $CL_k$  равен кластеру обработки  $CL_l$  и, следовательно, одновременно во времени с ним может выполняться, т.е.

$$CL_l, CL_k \in PR^{(j)} : CL_l \xrightarrow{KS_w} CL_k \Rightarrow CL_k \xrightarrow{KS_w} CL_l. \quad (2)$$

Другими словами выполняется условие симметричности любых двух кластеров  $CL_l, CL_k \in PR^{(j)}$  обработки информации.

Если кластер обработки информации  $CL_l$  равен кластеру обработки  $CL_k$ , а кластер  $CL_k$ , в свою очередь, равен  $CL_q$ , тогда кластер обработки  $CL_l$  равен и, следовательно, параллелен кластеру обработки  $CL_q$ , т.е.

$$CL_l, CL_k, CL_q \in PR^{(j)} EKS_w : CL_l \xrightarrow{KS_w} CL_k, \quad CL_k \xrightarrow{KS_w} CL_q \Rightarrow CL_l \xrightarrow{KS_w} CL_q. \quad (3)$$

Таким образом, справедливо условие транзитивности двух произвольных кластеров  $CL_l, CL_q \in PR^{(j)}$  обработки информации.

Поэтому справедливо утверждение о том, что отношение кластерной структуры обработки  $KS_w$  есть отношение эквивалентности.

Действительно, отношение кластерной структуры  $KS_w$  обработки информации, которое указывает на то, что любые два произвольно взятых кластера  $CL_l$  и  $CL_k$ , удовлетворяющие (1), могут выполняться одновремен-

но на разных процессорных модулях под управлением команд, принадлежащих указанным кластерам

$$CL_l, CL_k \in PR^{(j)} : CL_l \xrightarrow{KS_w} CL_k \quad (4)$$

Таким образом, введенное отношение кластерной структуры обработки информации  $KS_w$  по определению ставит в соответствие некоторой  $j$ -й программе обработки информации  $PR^{(j)}$  совокупность независимых и неравных кластеров  $CL_l$ , число которых равно числу классов эквивалентности  $L$ , а кратность  $(CL_l)^q$  определяется порядком класса эквивалентности  $|a_l|$ :

$$\forall j = 1, N \quad PR^{(j)} \xrightarrow{S_w} \{(CL_l)^q\}, \quad \forall q = 1, |a_l|; \quad \forall l = 1, L \quad (5)$$

Каждый класс эквивалентности имеет своего представителя кластера  $CL_l$ , имеющего число макрокоманд равное порядку кластера  $|CL_l|$ . Выражение (5) является решением задачи кластерного анализа, так как задает искомое разбиение исходной программы  $PR^{(j)}$  на кластеры.

**Кластерное представление нейропроцессорных систем**

В результате возможны следующие варианты решения задачи (5).

1. Число классов эквивалентности равно  $L$ , и порядок каждого класса равен единице  $a_l = 1$ . Тогда  $j$ -ый алгоритм информации представляет собой кортеж из  $L$  кластеров

$$A^{(j)} \xrightarrow{S_w} \{CL_l\} = \langle CL_1, CL_2, \dots, CL_l, \dots, CL_L \rangle, \quad (6)$$

каждый из которых является представителем своего класса эквивалентности и обеспечивает входной информацией последующий  $CL_{l+1}$  кластер. Взяв число процессорных модулей равное  $L$  и назначив каждому  $l$ -му модулю соответствующий ему кластер, получим нейропроцессорную реализацию конвейерного типа на базе нейропроцессора NM 640X. [9], представленную на рис. 1.

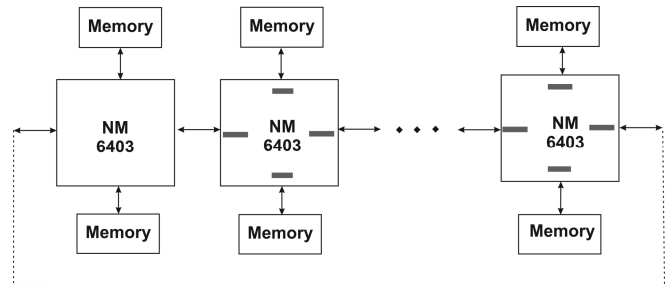
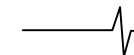


Рис. 1. Кластерная конвейерная структура обработки на базе нейропроцессора NM 640X

Далее определяются основные параметры получаемого конвейера обработки информации: количество модулей обработки  $L$ , объем оперативной памяти (Memory)  $|CL|$ , производительность как время цикла конвейера  $T_0$ , время простоя  $T_{np}$  и др [5-7].

2. Число классов эквивалентности  $L$  равно единице и порядок класса  $|a_l|$  равен  $q$ . Тогда  $j$ -му алгоритму обработки  $A^{(j)}$  информации соответствует набор  $q$  совершенно одинаковых кластеров  $A^{(j)} \xrightarrow{S_k} \{(CL)^q\}$ . При



этом, если выходная информация каждого кластера  $CL_i, \forall i = 1, q$  является входной для каждого последующего  $CL_{i+1}$ , тогда имеем кластерную структуру конвейерного типа с числом кластеров равным  $q$ , каждый из которых функционирует по командам кластера  $CL$  (рис. 1).

Если входная информация требуется одновременно для всех кластеров, тогда, назначая  $q$  процессорных модулей на обработку информации, получаем нейропроцессорную реализацию векторного или параллельного типа  $A^{(j)} \xrightarrow{Sw} \{CL_1\} = \langle CL_1, CL_2, \dots, CL_L \rangle$ , в которой все  $q$  процессорных модулей функционируют по одному и тому же кластеру  $CL_1$  (рис.2).

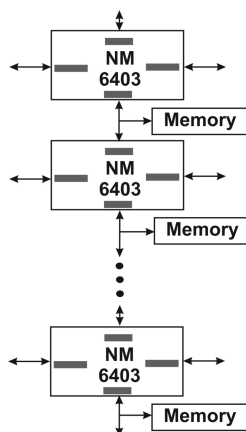


Рис. 2. Кластерная векторная структура обработки

Аналогично определяются основные параметры получаемой кластерной векторной структуры обработки информации: количество модулей обработки –  $L$ , объем оперативной памяти (Memory) –  $|CL|$ , производительность конвейера –  $T_0$ , время простоя –  $T_{np}$  и др.

3. Число классов эквивалентности равно  $L$  и порядок  $p$  каждого  $l$ -го класса равен  $|a_l|$ . Этот случай является общим и предполагает отношения не только между кластерами внутри класса, но и между кластерами различных классов.

Пусть обмен информацией между представителями классов  $CL_l, \forall l = 1, L$  осуществляется последовательно. Если кластеры внутри  $l$ -го класса обмениваются последовательно, тогда, назначая число процессорных модулей  $N_0 = L * \sum_{i=1}^L |a_i|$  получим известную конвейерную структуру. Если для подпрограмм внутри  $l$ -го класса требуется информация одновременно, тогда, назначая число процессорных модулей равное

$N_0 = L * \max |a_l|, \forall l = 1, L$ , получим конвейерно-векторную структуру обработки данных.

Если кластеры внутри  $l$ -го класса обмениваются информацией последовательно, имеем векторно-конвейерную структуру обработки информации. Указанные структуры являются частными случаями более общей – кластерной матричной структурой обработки, приведенной на рис. 3.

Далее для каждой получаемой структуры обработки информации определяются основные параметры: количество модулей обработки  $L$ , объем оперативной памяти (Memory)  $|CL|$ , производительность как время цикла конвейера  $T_0$ , время простоя  $T_{np}$  и др.

Таким образом, введение кластеризации позволило распараллелить программу обработки согласно (5) и в зависимости от исследуемых алгоритмов обработки информации  $PR^{(j)}$  получить кластерные структуры обработки: конвейерную, векторную [1], или новые структуры: конвейерно-векторную, векторно-конвейерную и др. [5, 7]. В результате в нейропроцессорную систему (НПС) вводится аппаратная избыточность в виде отдельных нейропроцессорных модулей, и тем самым появляется возможность повышения безотказности, восстанавливаемости и готовности кластерных НПС.

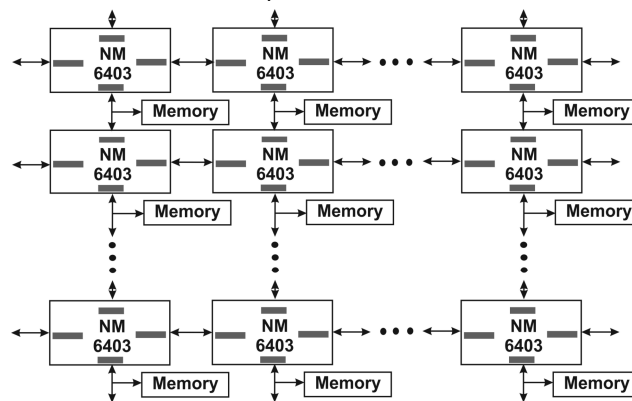


Рис. 3. Кластерная матричная структура обработки

### Восстанавливаемость кластерных нейропроцессорных систем

Как известно [1, 2], под надежностью понимают свойство вычислительной аппаратуры выполнять возложенные на нее функции, сохраняя свои эксплуатационные показатели в течение заданного промежутка времени в пределах, установленных в техническом задании или технических условиях.

Надежность аппаратуры обычно связывается с понятиями работоспособности, безотказности, ремонтпригодности и долговечности.

Под работоспособностью понимается состояние системы, в нашем случае НПС обработки данных [3, 5], при котором она способна выполнять возложенные на нее функции с параметрами, установленными требованиями технической документации.

Способность НПС обработки специальных данных сохранять работоспособность в течение некоторого времени наработки без вынужденных простоев называется безотказностью.

Как правило, на практике большинство НПС эксплуатируются длительное время. При этом они многократно включаются в работу и выключаются. Возможные случаи нарушения состояния работоспособности сопровождаются ремонтом, т.е. восстановлением рабочего состояния. Поэтому надежность аппаратуры наряду с работоспособностью и безотказностью оценивается также ремонтпригодностью, определяемой приспособленностью НПС к предупреждению, обнаружению и устранению от-

казов и неисправностей при проведении технического обслуживания и восстановления (ремонта). Качество и быстрота восстановления аппаратуры определяется не только ремонтпригодностью, но и организацией службы эксплуатации и совершенством автоматизации нейропроцессорной системы обработки специальных данных.

Для нахождения показателей ремонтпригодности (восстановливаемости) пользуются случайной величиной – временем выполнения операций по техническому обслуживанию. Эта величина в зависимости от цели накопления статистических данных может быть временем ремонта (восстановления) аппаратуры НПС после возникновения отказов и неисправностей, временем технической подготовки аппаратуры с учетом и без учета времени восстановления или временем выполнения регламентных работ.

Оперативным показателем ремонтпригодности аппаратуры НПС является вероятность восстановления аппаратуры за заданное время  $p_e(t)$  как вероятность того, что время восстановления  $t_e$  не превысит заданного для этих целей времени  $t$ :  $p_e(t) = P\{t_e \leq t\}$ .

По определению этой вероятности ясно, что она может представлять функцию распределения времени выполнения операций по восстановлению аппаратуры НПС.

Значительная часть электронных схем построена таким образом, что отказ хотя бы одного из элементов НПС ведет к отказу всей НПС в целом. Такое соединение элементов с точки зрения надежности называется последовательным (основным). Если все элементы в последовательном соединении работают независимо, то вероятность безотказной работы схемы НПС из  $N$  элементов за промежуток времени  $t$  находится по формуле

$$p(t) = p_1(t)p_2(t)\dots p_i(t)\dots p_N(t) = \prod_{i=1}^N p_i(t).$$

В ряде случаев электронные схемы устроены так, что отказ одного из элементов не приводит к отказу схемы в целом. В этом случае говорят, что такие схемы имеют функциональное резервирование. Однако в тех случаях, когда не удается обычными методами добиться высокой надежности аппаратуры, приходится прибегать к резервированию, которое может быть осуществлено на стадии конструирования или в процессе эксплуатации аппаратуры путем установки двух, трех однопотипных нейропроцессоров.

Резервирование в «физическом» смысле характеризуется параллельным соединением нейропроцессоров, в котором только отказ всех элементов приводит к отказу соединения в целом. В данном соединении из  $k$  элементов наряду с основным элементом имеется  $(k-1)$  резервных элементов. Каждый из которых может быть включен в работу или отключен при отказе.

В общем случае могут быть созданы резервные цепи, в которых все элементы одновременно участвуют в работе, а отказ одного или нескольких нейропроцессоров не нарушает работоспособности цепи. При этом вероятность отказа параллельного соединения

$$Q(t) = q_1(t)q_2(t)\dots q_i(t)\dots q_k(t) = \prod_{i=1}^k [1 - p_i(t)], \text{ где } q_i(t)$$

– вероятность отказа  $i$ -го нейропроцессора параллельного соединения.

Тогда вероятность безотказной работы параллельного соединения нейропроцессоров равна

$$P(t) = 1 - \prod_{i=1}^k [1 - p_i(t)]. \quad (7)$$

Для исследования наиболее рационального (наилучшего) способа повышения надежности введем и рассмотрим следующие соотношения. Пусть НПС состоит из  $N$  последовательно соединенных элементов и с целью повышения надежности резервируется  $m$  резервными цепями. Резервированная система не обслуживается (отказавшие цепи не ремонтируются). Отношение количества резервных цепей к числу основных называется *кратностью резервирования*. Для случаев резервирования, показанных на рисунках, число резервных цепей  $m$  совпадает с кратностью резервирования. В общем случае кратность резервирования может быть и дробным числом.

Вероятность безотказной работы при общем резервировании находится из условия, что отказ всей системы, включающей одну основную и  $m$  резервных цепей, произойдет после того, как независимо друг от друга откажут все  $(m+1)$  параллельных цепей. Тогда вероятность отказа  $Q_{общ}(t)$  системы равна:  $Q_{общ}(t) =$

$$= Q_1(t)Q_2(t)\dots Q_j(t)\dots Q_{m+1}(t) = \prod_{j=1}^{m+1} Q_j(t), \text{ а вероятность}$$

безотказной работы  $P_{общ}(t)$  системы составляет:

$$P_{общ}(t) = 1 - Q_{общ}(t) = 1 - \prod_{j=1}^{m+1} Q_j(t) = 1 - \prod_{j=1}^{m+1} \left[ 1 - \prod_{i=1}^N p_i(t) \right] \quad (8)$$

где  $Q_j(t)$  – вероятность отказа за время  $t$   $j$ -й резервной цепи;  $p_i(t)$  – вероятность безотказной работы за время  $t$   $i$ -го элемента цепи (основной или резервной).

Если все  $(m+1)$  цепей в параллельном соединении равнонадежны, то вероятность безотказной работы определяется

$$P_{общ}(t) = 1 - \left[ 1 - \prod_{i=1}^N p_i(t) \right]^{m+1} \quad (9)$$

Вероятность безотказной работы при раздельном резервировании определяется из условия, что НПС состоит из  $N$  последовательно соединенных звеньев, а каждое звено из  $(m+1)$  параллельно соединенных элементов, причем отказы элементов в звене – события независимые, и поэтому вероятность определяется

$$P_{разд}(t) = \prod_{i=1}^N p_{i3в}(t) = \prod_{i=1}^N [1 - q_{i3в}(t)] =$$

$$= \prod_{i=1}^N \left[ 1 - \prod_{j=1}^{m+1} q_{ij}(t) \right] = \prod_{i=1}^N \left\{ 1 - \prod_{j=1}^{m+1} [1 - p_{ij}(t)] \right\} \quad (10)$$

где  $p_{i3в}(t)$ ,  $q_{i3в}(t)$  – вероятность безотказной работы и отказа  $i$ -го звена соединения соответственно; а  $p_{ij}(t)$ ,  $q_{ij}(t)$  – вероятность безотказной работы и отказа в  $i$ -м звене  $j$ -го элемента (основного или резервного) соответственно.

Если все  $(m+1)$  элементов в звене равнонадежны, то вероятность раздельного резервирования определяется:

$$P_{разд}(t) = \prod_{i=1}^N \left\{ 1 - [1 - p_i(t)]^{m+1} \right\} \quad (11)$$

Сравнение формул (9) и (11) позволяет установить, что для всех значений  $p_i(t)$ ,  $N$  и  $m$  (исключая тривиальный случай, когда  $p_i(t) = 0$  и  $N = 1$ ) величины  $P_{общ}(t) < P_{разд}(t)$ .

**Пример.** Пусть аппаратура нейропроцессорной системы состоит из десяти нейропроцессоров, каждый из которых имеет вероятность безотказной работы за время  $t_p(t) = 0.8$ . Имеются две резервные цепи. При общем резервировании  $P_{общ}(t) = 1 - (1 - 0.8^{10})^3 = 0.27$ , а при раздельном резервировании  $P_{разд}(t) = [1 - (1 - 0.8^3)^{10}] = 0.92$ . Это наглядно свидетельствует о значительной эффективности раздельного резервирования в сравнении с общим.

## Заключение

Предложенная методика позволяет представить проектируемую систему в виде совокупности кластеров и реализовать ее как одну из полученных структур: конвейерную, векторную, конвейерно- векторную или векторно- конвейерную на базе современного отечественно- го микрокомплекта NM 640X или K1879 BMX.

При общем резервировании отказ любого из элементов рабочей цепи вызывает необходимость включения целиком резервной цепи, в то время как при раздельном резервировании отказ одного из элементов вызывает необходимость включения лишь одного эле-

мента, что приводит к повышению восстанавливаемости кластерных нейропроцессорных систем.

## Литература

1. Сотсков Б.С. Физика отказов и определение интенсивности отказов // О надежности сложных технических систем Под ред. Берга А.И., Бруевича Н.Г., Гнеденко Б.В., Голикевича Т.А. М.: Советское Радио, 1966
2. Нечипоренко В.И. Структурный анализ и методы построения надежных систем. М.: Советское Радио, 1968.
3. Архитектура компьютерных систем и сетей: Учеб. пособие, Т.П. Барановская, В.И. Лойко, М.И. Семенов, А.И. Трубилин; Под.ред. В.И. Лойко. – М.: Финансы и статистика, 2003. – 256 с.: ил.
4. Кузнецов О.П., Адельсон-Вельский Г.М. Дискретная математика для инженеров М.: Энергия, 1980.
5. Злобин В.К., Ручкин В.Н. Нейросети и нейрокомпьютеры. СПб.: БХВ-Петербург, 2011, -256 с.
6. Григоренко Д.В., Ручкин В.Н. Кластеризация систем обработки специальных данных// Методы и средства обработки и хранения информации Рязань: РИО Рязанского государственного радиотехнического университета, 2012.
7. Ручкин В.Н. Проектирование и выбор специализированных средств обработки информации. М.: Московский государственный открытый университет, 1997. - 120 с. : ил.
9. Виксне П.Е., Фомин Д.В., Черников В.М. Однокристальный цифровой нейропроцессор с переменной разрядностью операндов // Известия Вузов, Приборостроение, 1996, т.39, №7

## IMPROVING RECOVERABILITY NEUROPROCESSOR CLUSTER OF DATA PROCESSING SYSTEMS

*Grigorenko D.V., Ruchkin V.N.*

Investigate the possibility of improving recoverability cluster neuroprocessor conveyor systems studied, vector, conveyor-vector or vector-pipelined structures based data set of the modern domestic micro sets NM 640X.

## 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:

- Текст - текстовый редактор Microsoft Word (2003, 2007).
- Таблицы и рисунки должны быть пронумерованы. На все рисунки, таблицы и библиографические данные указываются ссылки в тексте статьи.
- Объем статьи до 12 стр. (ширифт 12). Для заказных обзорных работ объем может быть увеличен до 20 стр.
- Название статьи на русском и английском языках.

### Рукопись статьи сопровождается:

- краткой аннотацией на русском и английском языках;
- номером УДК;
- сведениями об авторах (Ф.И.О., организация, должность, ученая степень, телефоны, элект-ронная почта);
- ключевыми словами;
- актом экспертизы (при наличии в вашей организации экспертной комиссии).

### 2) Требования к иллюстрациям:

- Векторные (схемы, графики) - желательно использование графического редактора Corel DRAW.
- Растровые (фотографии, рисунки) - М 1:1, разрешение не менее 300dpi, формат tiff, jpg.



## Автомобильные усилители мощности:



› Два варианта работы выходного каскада усилителя - «ламповый» и «транзисторный»;

› Цифровые аудио входы — оптический TOSLINK, USB, BlueTooth;

› Аналоговые входы — RCA, возможность

использования укороченного аналогового аудиотракта;

› Обработка входного аудиосигнала встроенным 56-битным DSP в формате 24 бит/96кГц (встроенные АЦП/ЦАП) с передискретизацией входного цифрового сигнала на 96кГц;

› Возможность у пользователя самостоятельно, в визуальной среде программирования, формировать оптимальную обработку звука средствами DSP;

› Дискретная схемотехника усилителя с MOSFET-выходом и использованием транзисторных сборок, экранирующие переборки, блок питания с двойной стабилизацией;

› Мощность 4 канала по 80 Вт(RMS) каждый на нагрузке 4 Ом;

› Диапазон воспроизводимых частот 10Гц - 50 кГц.

## Акустические системы Hi-End класса для домашнего применения:



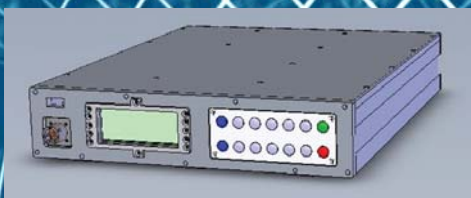
› Активные 2-х полосные напольные и полочные АС, на базе 20см, 16см, 13 см головок фирмы Peerless серии HDS;

› Отдельные усилители, работающие в «ламповом» режиме под НЧ-СЧ и ВЧ-головки (Full Bi-Amp) с коррекцией АЧХ и по-канальной фильтрацией средствами встроенного DSP (24 бит/96кГц);

› Аудио входы/выходы — аналоговый RCA, цифровые оптический TOSLINK и BlueTooth, оба – с передискретизацией входного цифрового потока на 96 кГц;

› возможность настройки под имеющийся инфракрасный ПДУ (кодировки RC-5, NEC).

## Аппаратура передачи специальных данных ПД-441-01:



› Собственная реализация 2-х и 4-х проводного модема по стандартам до V.32bis, с обменом по протоколам V.42, V.42bis, V.8, значительно превосходящая по своим характеристикам аналогичную продукцию фирм Zyxel, 3Com;

› Построение специализированных сетей с использованием интерфейсов Ethernet 10/100, RS-485, G1-И, 3G,

радиостанций и др.;

› Голосовая IP-связь с маскированием речи;

› Работа со специализированными носителями информации.

# ЗАО «НПО «Рязаньприбор» – Мы можем это!

тел. +7-(4912)-29-84-48  
факс +7-(4912)-76-65-74

info@rznprb.com  
www.rznprb.com



# ВЫСОКОПРОИЗВОДИТЕЛЬНЫЕ РЕШЕНИЯ ДЛЯ ЦИФРОВОЙ ОБРАБОТКИ РАДИОСИГНАЛОВ



Цифровой радиоприём  
Цифровой синтез сигналов  
Системные синхронизаторы  
Высокоскоростные АЦП и ЦАП  
Мультипроцессорные модули ЦОС  
Модули обработки сигналов на FPGA



ЗАО "Инструментальные Системы"  
[www.insys.ru](http://www.insys.ru), [info@insys.ru](mailto:info@insys.ru)  
(495)781-2750, 781-2751