ЦИФРОВАЯ ОБРАБОТКА СИГНАЛОВ

Научно-технический журнал № 1/2018

Издается с 1999 года Выходит четыре раза в год

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

ЗАМЕСТИТЕЛИ ГЛАВНОГО РЕДАКТОРА:

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

РЕДАКЦИОННАЯ КОЛЛЕГИЯ:

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

Адрес редакции:

г. Москва, ул. Авиамоторная, д. 8 Научный центр МТУСИ Тел.: (+7) 903-201-53-33 E-mail: rntores@mail.ru vityazev.v.v@rsreu.ru http://www.dspa.ru

Издатель:

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

> Подписной индекс по каталогу ОАО «Роспечать» – 82185

Подписано в печать 29.03.18 г. Формат 60х90/8. Гарнитура «Arial». Печать офсетная.

Бумага офсетная. Печ.л. 6,5. Тираж 500 экз.

Заказ № 1236<mark>.</mark> Отпечатано в ООО НПЦ «Информационные технологии» Рязань, ул. Островского, д. 21/1 тел.: (4912) 98-69-84

Издание зарегистрировано в Министерстве Российской Федерации по делам печати, телерадиовещания и средств массовых коммуникаций. Свидетельство о регистрации ПИ № 77-1488

от 14.01.2000 г.

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

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

B HOMEPE:

m	ехнологий в условиях вызовов цифровой экономики
Б	ыховский М.А.
Γι πe	перфазовая модуляция – оптимальный метод редачи цифровых сообщений
C	оловьева Е.Б.
Ре не	журрентные нейронные сети в качестве моделей линейных динамических систем
П	опов Д.И.
O c	птимизация нерекурсивных режекторных фильтров настичной адаптацией
Па Ч	астарнак В.И., Козинин Е.А., Кубышкин А.П., ереватюк И.В., Ефимов Е.А.
И	нтеграция наземных комплексов приема, обработки распространения информации
on c	авспростратения информации n космических систем «Электро» и «Арктика-М» единой территориально распределенной системой ДЗЗ
П	экровский П.С.
Де сп се	зухканальный универсальный алгоритм детектирования ектрально-эффективных радиосигналов с управляемой язью между синфазной и квадратурной составляющими
Tj	пицин Г.С., Топников А.И., Приоров А.Л.
П	остроение комбинированного показателя качества
pe	чевых сигналов для оценки точности
в	/словиях воздействия аддитивных шумов
Tj	ровский Я.А.
C	равнительные характеристики алгоритмов
de	текции стабильных зрительных
H5	званных попленциалов головного мозга электроэнцефалограмме
A	отемьев В.В., Кашин А.В.
Ре бе	ализация целочисленных цифровых рекурсивных фильтров з умножителей на плис отечественного производства
Ч	ан Ван Нгиа.
Πļ	рименение комплексных фильтров
дл и	я снижения пик-фактора сигналов с OFDM модуляцией их реализация на ПЛИС
Л	е Ван Ки.
Ај не	пгоритмические методы компенсации линейности усилителя мощности
Г	рязных И.В., Решетняк С.А.,
Τį	ретьяков Г.Н., Шпак А.В.
Pe	едукция частоты сигнала при нелинейной фильтрации

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

1

Digital Signal Processing

Science & Technical Magazine Issue 1, 2018 year

Is published quarterly since 1999

THE EDITOR-IN-CHIEF:

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

DEPUTY EDITORS-IN-CHIEF:

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

EDITORIAL BOARD:

Adzhemov A.S., Dr.Sci.(Tech.), Professor Alpatov B.A., Dr.Sci.(Tech.), Professor Bartenev V.G., Dr.Sci.(Tech.), Professor Borisov Y.I., Dr.Sci.(Tech.) Bruchanov Y.A., Dr.Sci.(Tech.), Professor Eremeyev V.V., Dr.Sci.(Tech.), Professor Zaitsev G.V., Dr.Sci.(Tech.) Zubarev R.V., Kirpichnikov A.P., Krivocheev M.I., Dr.Sci.(Tech.), Professor Kuznetsov N.A., Dr.Sci.(Tech.), Academician Mistyukov V.G., Mishenkov S.L., Dr.Sci.(Tech.), Professor Petrovsky A.A., Dr.Sci.(Tech.), Professor Prokhorov Y.N., Dr.Sci.(Tech.), Professor Sannikov V.G., Dr.Sci.(Tech.), Professor Speranskii V.S., Dr.Sci.(Tech.), Professor Shinakov Y.S., Dr.Sci.(Tech.), Professor

Editorial office address:

Aviamotornaya, 8, Moscow, Russia Research Center of MTUCI Phone: (+7) 903-201-53-33 E-mail: rntores@mail.ru, vityazev.v.v@rsreu.ru Web: http://www.dspa.ru

Publisher:

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

Computer makeup:

I.A. Blagodarova

FOUNDERS: InSys

Radio Research and Development Institute

CONTENTS:

Zatsarinny A.A., Gorshenin A.K., Volovich K.I., Kondrashev V.A.
The main trends of the development of information technologies within the challenges of the digital economy
Bykhovskiy M. Hyperphase Modulation – the optimal method of message transmission8
Solovyeva E.B. Recurrent neural networks as behavioral models of nonlinear dynamic systems18
Popov D.I. Optimization of non-recursive rejection filters with partly adaptation28
Pastarnak V.I., Kozinin E.A., Kubishkin A.P., Cherevatyuk I.V., Efimov E.A.
and dissemination of information from space systems «Electro» and «Arctica-M» with Unified territorialy distributed system of earth remote sensing
Pokrovskij P.S. Two-channel universal algorithm to demodulate spectral-efficient radio signals with controlled coupling between in-phase and quadrature components
Tupitsin G.S., Topnikov A.I., Priorov A.L. Development of combined speech quality measure for speaker identification accuracy estimation in additive noise environments
Turovsky Ya.A. Comparative characteristics of the algorithms of detection steady state visually evoked potentials of the brain on an electroencephalogram
Artemiev V.V., Kashin A.V. Implementation of integer recursive digital filters without multipliers on fpga of the Russian production
Tran Van Nghia. Application of the complex filters for papr reduction of ofdm signals and FPGA implementation62
Le Van Ky. Algorithmic methods for compensation of power amplifier nonlinearity71
Gryaznykh I.V., Reshetnyak S.A., Tretyakov G.N., Shpak A.V. Reduction of the frequency of the signal in nonlinear filtering76

Subscription index: 82185 ("Rospechat") ISSN: 1684-2634 9 771684 263005

УДК 004.9

ОСНОВНЫЕ НАПРАВЛЕНИЯ РАЗВИТИЯ ИНФОРМАЦИОННЫХ ТЕХНОЛОГИЙ В УСЛОВИЯХ ВЫЗОВОВ ЦИФРОВОЙ ЭКОНОМИКИ

Зацаринный А.А., д.т.н., заместитель директора Федерального исследовательского центра «Информатика и управление» РАН, e-mail: alex250451@mail.ru;

Горшенин А.К., к.ф-м.н., в.н.с. Федерального исследовательского центра «Информатика и управление» РАН, e-mail: agorshenin@frccsc.ru;

Волович К.И., с.н.с. Федерального исследовательского центра «Информатика и управление» РАН, е-mail: kvolovich@frccsc.ru;

Кондрашев В.А., с.н.с. Федерального исследовательского центра «Информатика и управление» РАН, e-mail: vkondrashev@frccsc.ru.

THE MAIN TRENDS OF THE DEVELOPMENT OF INFORMATION TECHNOLOGIES WITHIN THE CHALLENGES OF THE DIGITAL ECONOMY

Zatsarinny A.A., Gorshenin A.K., Volovich K.I., Kondrashev V.A.

The article considers the main trends of the development of information technologies in the context of the implementation of the Digital Economy program of the Russian Federation. A brief analysis of the current state of the digital economy in developed countries, as well as the starting positions of Russia, are given. Information technologies of the digital economy presented by experts of the Davos forum are considered. The role of scientific organizations in the implementation of the Digital Economy program is shown. The content and the statement of the problem «Digital Science» are presented. Information technologies for scientific research and the experience of the Federal Research Center «Computer Science and Control» of the Russian Academy of Sciences are considered.

Key words: digital economy, information technology, digital platforms, digital transformations, scientific services.

Ключевые слова: цифровая экономика, информационные технологии, цифровые платформы, цифровые трансформации, научные сервисы.

Информационные технологии как основа цифровой экономики

В настоящий момент Российская Федерация решительно движется в направлении внедрения подходов и технологий цифровой экономики. Старт этим процессам был дан Президентом России В.В.Путиным, который в Послании Федеральному собранию 1 декабря 2016 г. [1]

предложил «запустить масштабную системную программу развития экономики нового технологического поколения – цифровой экономики» и обозначил ряд ключевых позиций: фундаментальная наука, исследовательская инфраструктура, ІТ-индустрия, собственные передовые разработки, цифровые технологии и, наконец, программа развития цифровой экономики. Все эти стратегические тренды являются очень актуальными для развития нашей страны, они полностью коррелируют с мировыми тенденциями развития экономики, которые в концентрированном виде представлены в известной книге Клауса Шваба [2]. Вместе с тем, следует отметить, что такая направленность во многом была подготовлена Стратегией научно-технологического развития Российской Федерации (утверждена Указом Президента РФ от 1 декабря 2016 г. № 642) [3].

Этот важнейший документ направлен на качественные изменения в области науки и технологий с целью повышения способности государства отвечать на боль-

Рассмотрены основные направления развития информационных технологий в условиях выполнения программы цифровой экономики Российской Федерации. Приводится краткий анализ состояния цифровой экономики в развитых странах, а также стартовых позиций России. Рассмотрены информационные технологии цифровой экономики, представленные экспертами Давосского форума. Показана роль научных организаций в реализации программы цифровой экономики. Представлено содержание и постановка задачи «Цифровая наука». Рассмотрены информационные технологии в интересах научных исследований на основе опыта Федерального исследовательского центра «Информатика и управление» РАН.

> шие вызовы, представляющие совокупность проблем и угроз такой масштабности и сложности, которые не позволяют их разрешить или устранить только за счет количественного увеличения ресурсов. В Стратегии в качестве важнейшего приоритета развития России на период 10-15 лет определен «переход к передовым цифровым, интеллектуальным производственным технологиям, роботизированным системам, новым материалам и способам конструирования, создание систем обработки больших объемов данных, машинного обучения и искусственного интеллекта». При этом отмечено, что для эффективного ответа на большие вызовы необходима консолидация усилий органов государственной власти, научного, образовательного и предпринимательского сообществ, институтов гражданского общества по созданию благоприятных условий для применения достижений науки и технологий в интересах социально-экономического развития России. И важнейшими инструментами для реализации такого подхода должны стать наука и технологии.

В программе цифровой экономики Российской Федерации [4] важнейшее место занимают вопросы развития и применения наиболее современных информационных технологий в различных отраслях промышленности, здравоохранения, науки, культуры, социальной сферы. Возникает необходимость постановки соответствующих задач и проведения ряда взаимоувязанных фундаментальных междисциплинарных исследований в области информатики на основе математических, информационных, логических, психологических, лингвистических и биологических принципов. Для успешного решения таких задач требуется выполнить опережающие фундаментальные исследования, направленные, прежде всего, на разработку и развитие методов формирования высокоинтеллектуальных цифровых платформ, технологий накопления знаний и повышения уровня компетенции интеллектуальных систем, методов и технологий искусственного интеллекта, а также методов целеполагания при выборе в интеллектуальных системах новых целей поведения. Все это ведет к необходимости:

 перестройки всего сектора исследований, разработок и инноваций;

 концентрации ресурсов на получении новых научных результатов;

- развития национальных центров превосходства;

 – созданию эффективных партнерств с иностранными исследовательскими центрами;

– опережающему увеличению расходов на научноисследовательские и опытно-конструкторские работы.

В результате 2017 год во многом стал годом постановки стратегических задач цифровой экономики. Правительством РФ приняты программа «Цифровая экономика Российской Федерации» (Распоряжение Правительства РФ от 28.07.2017 г. № 1632-р) и система управления реализацией программы «Цифровая экономика Российской Федерации» (Постановление Правительства РФ от 28.08.2017 г. № 1030). Ряд документов принят Евразийским экономическим союзом. Проведено широкое обсуждение проблематики цифровой экономики на целом ряде представительных форумов, конференций, семинаров и совещаний (см., например, [5]).

Документы приняты, однако дискуссии по самым основным понятиям активно продолжаются. В частности, дискуссионной стала трактовка самого понятия «Цифровая экономика» (см., например, [6, 7]). В Указе Президента РФ от 9 мая 2017 г. № 203 «О Стратегии развития информационного общества в Российской Федерации на 2017-2030 гг.» дано определение, которое можно считать официальной позицией руководства страны: «Цифровая экономика – хозяйственная деятельность, в которой ключевым фактором производства являются данные в цифровом виде, обработка больших объемов и использование результатов анализа которых по сравнению с традиционными формами хозяйствования позволяют существенно повысить эффективность различных видов производства, технологий, оборудования, хранения, продажи, доставки товаров и услуг». В этом определении главный акцент на использовании в хозяйственной деятельности данных в цифровой форме.

Вместе с тем цифровая экономика является широким, многоаспектным, понятием, реализация принятой программы затронет практически все сферы деятельности как государства, общества, так и каждого отдельного человека. При этом очевидно, что основным инструментом реализации будут информационные технологии. Таким рбразом, в настоящей работе будем считать, что цифровая экономика – это возможность создания цифровых моделей реального мира экономики, которые на основе новых возможностей измерений позволят обеспечить учет самых разных ресурсов в реальной экономике (материалы, техника, интеллектуальные, человеческие ресурсы, инфраструктурные ресурсы) и процессов, которые происходят с этими ресурсами.

Цифровая экономика, по существу, предполагает решение трех взаимоувязанных задач. Первая состоит в создании единого информационного реестра всех ресурсов в цифровой экономике (материалы, техника, интеллектуальные, человеческие ресурсы, инфраструктурные и другие ресурсы). Заметим, что попытка решения этой задачи предпринималась в России неоднократно разными ведомствами, однако приемлемый результат так до сих пор и не получен. Вторая задача – создание и внедрение технологии учета всех процессов, которые приводят к тем или иным изменениям этих ресурсов. Эта задача не нова, но ее решение должно быть на самом современном научно-технологическом уровне. Третья задача – самая сложная. Единый реестр ресурсов должен быть заполнен и оперативно обновляться актуальными, достоверными и объективными исходными данными.

Только при реализации такого подхода может быть обеспечена эффективность управленческих решений в цифровой экономике на всех уровнях. Естественно, что такой подход приведет к минимизации человеческого фактора и сокращению числа уровней в иерархии системы управления.

Отсюда следует, что информационные технологии приобретают особую значимость в условиях взятого курса на цифровую экономику.

Краткое состояние в мире и стартовые позиции России в цифровой экономике

Лидером цифровой экономики может считаться Великобритания [8], поскольку ее доля в ВВП цифровой экономики наибольшая, она оценивается порядка 18 %. Основные применения – это спутники, хранилища данных, робототехника, большие данные. Департамент науки определяет стратегические приоритеты и ответственность несет за их правильный выбор (но не за исполнение). Этим департаментом были проведены исследования ключевых будущих технологий и инноваций, среди которых выявлен ряд технологий многоцелевого назначения:

- передовые материалы;
- спутники;
- хранилища энергии;
- робототехника и автономные системы;
- агро-наука;

регенеративная медицина;

_/

– большие данные;

- синтетическая биология.

В США создан Институт цифрового производства и инновационного дизайна в Чикаго, который сегодня является одним из крупнейших центров цифровой экономики. Доля в ВВП цифровой экономики – почти 11 %. В Германии родился термин «Индустрия 4.0», которым сегодня оперируют во многих публикациях, на конференциях, и даже в документах. Японский премьерминистр Синдзо Абэ на ежегодной IT-конференции СеВIT в Ганновере в марте 2017 года использовал понятие «Индустрия 5.0».

Япония традиционно активно занимается исследованиями в области робототехники. В ноябре 2017 г. японская компания Nippon Telegraph and Telephone Company (NTT) официально запустила в эксплуатацию прототип квантовой вычислительной системы.

Китай в цифровой экономике обозначил стратегические перспективы до 2050 года. Поставлены очень амбициозные цели – стать ведущей технологической державой в мире. И эти цели подкреплены конкретными нормативными государственными документами. Доля КНР в ВВП цифровой экономики – около 10 %.

Каковы стартовые позиции России? Менее 4 % ВВП в цифровой экономике. По данным Организации экономического сотрудничества и развития (ОЭСР) продукция России в ежегодном мировом обороте на рынке высоких технологий и наукоемкой продукции составляет всего 0,3 % (продукция США – 35 %, Японии – 20 %, Германии – 13 %, Китая – 12 %, Южной Кореи – 5 %). Индекс глобальной конкурентоспособности – 43-е место, индекс глобальной инновационности – 45-е, индекс диверсифицированности экономики – 45-е. При среднемировом числе роботов на 10 тыс. занятых, равном 65, в России всего 2 робота (по данным Международной федерации робототехники IFR), в Южной Корее – 478, Японии – 314, Германии – 292, США – 64, Китае – 36.

По каждому направлению утвержденной программы цифровой экономики Российской федерации (нормативное регулирование, кадры и образование, формирование исследовательских компетенций и технических заделов, информационная инфраструктура и информационная безопасность) определены ответственные, однако среди них нет научных организаций, а также федеральных органов власти, в ведении которых находятся сотни ведущих научных организаций страны (ФАНО, Минобрнауки России), а также ведущие промышленные организации (Минпроторг России). Безусловно, такая ситуация ведет к исключительной ориентации мероприятий на бизнес-сообщество, что далеко не всегда подразумевает рост наукоемкого потенциала.

По оценкам экспертов Давосского форума, внедрение цифровой экономики позволит достичь целого ряда результатов, которые представляются сегодня немыслимыми. Рассмотрим некоторые ожидаемые результаты цифровой экономики к 2025 году (прогнозы экспертов Давосского форума в 2015 году):

 – 10 % людей носят одежду, подключенную к сети Интернет;

- 90 % людей имеют возможность неограниченного

и бесплатного (поддерживаемого рекламой) хранения данных;

- триллион датчиков, подключенных к сети Интернет;

– первый робот-фармацевт в США;

– 10 % очков для чтения подключены к сети Интернет;

 – 80 % людей с цифровым присутствием в сети Интернет;

производство первого автомобиля при помощи 3D печати;

 первое правительство, заменяющее перепись населения источниками больших данных;

 первый имеющийся в продаже имплантируемый мобильный телефон;

 – 5 % потребительских товаров создано с помощью технологии 3D печати;

90 % населения используют смартфоны;

 90 % населения имеет регулярный доступ к сети Интернет;

 – беспилотные автомобили составляют 10 % от общего количества автомобилей на дорогах США;

 – первая пересадка печени, созданной с использованием 3D печати;

 – 30 % портативных аудиторских проверок проводит искусственный интеллект (ИИ);

 правительство впервые собирает налоги при помощи цепочки блоков (технологии блокчейн);

 – более 50 % домашнего интернет-трафика приходится на долю приложений и устройств;

 превышение количества поездок/путешествий на автомобилях для совместного использования над поездками на частных автомобилях;

 – первый город с населением более 50000 без светофоров;

 – 10 % всемирного валового продукта хранится по технологии цепочки блоков (технологии блокчейн);

 – первый ИИ-робот в составе корпоративного совета директоров.

Наука как отрасль цифровой экономики

Достижение планируемых показателей и целей развития российского общества с использованием инструментов цифровой экономики возможно за счет активного участия научных организаций в реализации программы цифровой экономики на основе:

– фундаментальных научных исследований;

 участия в создании исследовательской инфраструктуры;

 исследований в области сквозных информационных технологий;

- участия в проектах различного уровня.

Вместе с тем, сегодня представляется логичной постановка задачи «Цифровая наука», а именно – цифровая экономика с позиций науки или, наоборот, наука в условиях вызовов цифровой экономики. Научные организации, академические институты, научно-исследовательские центры ведущих ВУЗов должны быть вовлечены в максимальное количество проектов цифровой экономики.

Российская наука обладает всеми необходимыми компонентами для рассмотрения в качестве отрасли

цифровой экономики: наличие развитой инфраструктуры, организационных структур, нормативной базы, высокого уровня компетенций и высококвалифицированных научных коллективов. При этом все перечисленные компоненты становятся «цифровыми» и наука, как отрасль экономики, также становится «цифровой».

Другим не менее важным обстоятельством является объективная смена подходов к проведению научных исследований, а именно переход к новой парадигме в научных исследованиях, основанной на анализе накопленных данных в конкретных предметных областях, естественно, в формализованном цифровом виде. Проведение таких исследований становится неотъемлемой частью различных областей науки, экономики, бизнеса на основе инструментария интенсивного использования данных. Часто области научного знания в науках с интенсивным использованием данных называют «Хинформатикой» (например, биоинформатика, геоинформатика, астроинформатика, нейроинформатика, медицинская информатика и др.).

Научные и образовательные организации обладают достаточно широкой, распределенной по территории страны, сетью центров коллективного пользования и уникальных научных установок, обладающих колоссальным спектром научных услуг в различных областях науки. Сегодня формализуется, описывается, структурируется имеющаяся исследовательская инфраструктура в виде центров коллективного пользования в академических институтах, уникальных научных установок, суперкомпьютерных центров и так далее. Необходима систематизация этих ресурсов и повышение эффективности их использования. На повестке дня – объективная потребность преобразования этой сети в современную исследовательскую инфраструктуру, которая предоставляла бы широкий спектр возможностей по научным сервисам не только для научных организаций, но и для внешних пользователей.

Чрезвычайно важно отметить, что в последние годы нарастающим итогом формируется институциональная цифровая среда практически во всех сферах нашей жизни: цифровые корпорации, цифровые университеты, цифровые институты, цифровое общество, цифровой спорт, цифровая культура и т.д.

ФИЦ ИУ РАН – научная организация, которая создана в рамках реализации принципов объединения, интеграции и укрупнения научных центров, на основании приказа Федерального агентства научных организаций России, – позиционирует свои научные исследования в рамках утвержденных направлений применительно к цифровой экономике. Наибольший акцент сделан на методах анализа больших данных, методах искусственного интеллекта, методах формирования цифровых платформ, обеспечения информационной безопасности.

Для эффективных исследований в интересах формирования информационных технологий цифровой экономики в ФИЦ ИУ РАН создается современная цифровая исследовательская платформа [9], представляющая собой совокупность трех компонентов. Первый – центр компетенций, в котором концентрируются знания в конкретной области. Второй – материальная среда, сеть центров обработки данных. Наконец, третий – это огромная совокупность сервисов, которые должны реализовываться на этой цифровой платформе (аналитических, образовательных, библиотечных, вычислительных, аналитических и др.). Эти сервисы должны предоставлять услуги образованию, науке, коммерции, промышленности, государственным структурам.

Современная исследовательская инфраструктура ФИЦ ИУ РАН имеет гибридную вычислительную архитектуру, в которой вычислительные ресурсы предоставляются исследователям в виде традиционных облачных услуг в режимах программных (SaaS), платформенных (PaaS) и инфраструктурных (IaaS) сервисов. Кроме того, отрабатывается технология предоставления исследователям научного сервиса как услуги (RaaS, Research as a Service) [10] в виде предметно-ориентированных программ.

В ближайшие годы планируется создание территориально распределенного кластера высокопроизводительных вычислений для научных исследований с ожидаемой общей производительностью до 650 терафлопс. При этом создаваемые вычислительные мощности активно задействуются в рамках научных ученых ФИЦ ИУ РАН и их коллег. Так, в 2017 году с использованием высокопроизводительных вычислений выполнены исследовательские работы по следующим тематикам:

 молекулярно-динамическое моделирование процесса взаимодействия частиц;

 – квантово-механические расчеты структурных свойств многокомпонентных материалов;

– обратные задачи подводной акустики;

 интеллектуальный поиск и анализ больших массивов текстов;

классификация изображений;

интеллектуальный анализ данных;

– сегментация трехмерных медицинских изображений.

Особую значимость приобрела проблема управления научными сервисами. Научными коллективами ФИЦ ИУ РАН в рамках фундаментальных исследований по заказу ФАНО России обоснованы [9] концептуальные и системотехнические подходы к созданию системы управления научными сервисами (СУС), которые обеспечиваются центрами коллективного пользования (ЦКП) и уникальными научными установками (УНУ) академических институтов.

Научный сервис сегодня определен как совокупность действий и средств для обеспечения процесса по выполнению конкретной работы и выделения для этого оборудования, расходных материалов и ресурсов, а также, самое важное, квалифицированных кадров, способных обеспечить научный сервис (измерение, эксперимент и т.д.). Разработаны модели СУС, в которых учитываются субъекты системы управления в виде администраторов системы, в виде пользователей (заказчиков), персонала ЦКП (УНУ), а также аналитиков, которые получают доступ к обобщенной информации для ее анализа с возможностью формирования отчетов по видам ЦКП, видам услуг, научным направлениям и с различными временными срезами. Эта система реализована в виде действующего макета на базе центра обработки данных ФИЦ ИУ РАН. Также разработан ряд инновационных системотехнических решений по построению управления деятельностью организационных систем, которые могут найти применение при реализации программы цифровой экономики.

Заключение

На основании выше изложенного следует, что:

 современные информационные технологии являются основой построения и развития цифровой экономики;

 приведенные количественные характеристики задела, созданного в наиболее технологически развитых странах мира, констатируют значительное отставание отечественной цифровой экономики от ведущих зарубежных стран;

 необходимо уделять существенное внимание построению отрасли «Цифровая наука» как ключевому драйверу цифровой институциональной среды с развитием современной исследовательской инфраструктуры и решением приоритетных задач в области искусственного интеллекта, роботехники и т.д.

Литература

1. Послание Президента РФ В.В. Путина Федеральному Собранию, 1 декабря 2016 года (доступно по адресу http://kremlin.ru/events/president/news/copy/53379)/

2. Шваб К. Четвертая промышленная революция. – М.: Эксмо, 2016. – 208 с.

3. Стратегия научно-технологического развития Российской Федерации (утверждена Указом Президента РФ №642 от 01.12.2016 г., доступна по адресу http://static. kremlin.ru/media/events/files/ru/uZiATIOJiq5tZsJgqcZLY9Yy L8PWTXQb.pdf).

4. Программа «Цифровая экономика Российской Федерации» (утв. Распоряжением Правительства РФ от 28.07.2017 №1632-р; http://static.government.ru/media/files/ 9gFM4FHj4PsB79I5v7yLVuPgu4bvR7M0.pdf).

5. Аналитика развития, безопасности и сотрудничества: Большая Евразия – 2030. Сборник материалов 4-й Международной конференции 29.11.2017г., Общественная палата РФ. – М.: Когито-Центр, 2017. – 253 с.

6. Иванов В.В., Малинецкий Г.Г. Цифровая экономика: мифы, реальность, перспектива. – М.: Российская академия наук, 2017. – 63 с.

7. Введение в цифровую экономику / А.В. Кешелава В.Г. Буданов, В.Ю. Румянцев и др.; под общ. ред. А.В. Кешелава. – ВНИИГеосистем, 2017. – 28 с.

8. Соколов И.А., Куприяновский В.П., Намиот Д.Е., Дрожжинов В.И., Быков А.Ю., Синягов С.А., Карасев О.И., Добрынин А.П. Государство, инновации, наука и таланты в измерении цифровой экономики (на примере Великобритании) // International Journal of Open Information Technologies, 2017. Vol. 5. №6. С. 33–48.

9. Зацаринный А.А., Горшенин А.К., Волович К.И., Колин К.К., Кондрашев В.А., Степанов П.В. Управление научными сервисами как основа национальной цифровой платформы «Наука и образование» // Стратегические приоритеты, 2017. №2 (13). С. 103–114

10. Волович К.И., Зацаринный А.А., Кондрашев В.А., Шабанов А.П. О некоторых подходах к представлению научных исследований как облачного сервиса // Системы и средства информатики, 2017. Т.27. №1. С. 73–84.



Витязев В.В.



Многоскоростная обработка сигналов – М.: Изд-во «Горячая линия-Телеком», 2017 г. – 336 с.: ил.

Рассмотрена эволюция теории и технологий многоскоростной обработки сигналов в период с начала 70-х гг. прошлого столетия до наших дней с позиции вклада, который внесли в их развитие работы российских ученых и специалистов в области цифровых информационных технологий реального времени. Описаны методы и алгоритмы многоступенчатой и многокаскадной реализаций цифровых узкополосных фильтров и банков цифровых фильтров на основе эффектов прореживания по времени (децимация во временной области) и по частоте (децимация в частотной области).

Приведена методика оптимизации параметров многоступенчатых и многокаскадных структур цифровых полосовых фильтров. Построение оптимальных структур и расчет параметров фильтров частотной селекции иллюстрируется многочисленными примерами.

Для специалистов, научных работников, преподавателей вузов, аспирантов; будет полезна студентам информационных и инфокоммуникационных направлений подготовки.

УДК 621.391.13

ГИПЕРФАЗОВАЯ МОДУЛЯЦИЯ – ОПТИМАЛЬНЫЙ МЕТОД ПЕРЕДАЧИ ЦИФРОВЫХ СООБЩЕНИЙ

Быховский М.А., д.т.н., профессор Московского технического университета связи и информатики, e-mail: bykhmark@gmail.com.

HYPERPHASE MODULATION – THE OPTIMAL METHOD OF MESSAGE TRANSMISSION

Bykhovskiy M.

Author proposes a new method of message transmission by using Hyperphase Modulation. In HPM, all signal points (SP) of the corresponding signal ensemble (SE) are located on the surface of N-dimensional sphere and the distance (d) between the closest SPs

with the increase in N dimension is measured according to the formula $d \sim \sqrt{N}$. Such SE, as was shown by Shannon in 1959, is optimal given large enough value of N and allows to ensure the transmission of signals in the telecommunication channel with a limited frequency band with a set speed of transmission and maximum possible noise immunity.

The article explores dependence of the speed of message transmission with the use of HPM on minimal phase distance between SPs in this multidimensional SE. It is shown that by applying HPM method it is possible to achieve a substantial net energy gain as compared with other two-dimensional SEs, such as QAM. Such gain grows with the increase of N and can be quite significant.

The author develops an effective, from a calculation standpoint, coding algorithm of the transmitted message – by using the transmitted message number (m) to calculate phase coordinates of signals' SPs. The values of those coordinates are then used for formation of transmitted signals with HPM. The major difference of this coding method from those that are used in modern telecommunication systems is that that method, while improving noise immunity, it does not contribute additional (excessive) symbols into a sequence of informational symbols of a transmitted message.

Key words: A new Method for Digital Modulation- Hyperphase Modulation (HPM), the Design of the Ensemble of Signals, Energy and Spectral Efficiency of HPM.

Ключевые слова: гиперфазовая модуляция, конструкция ансамбля сигналов, энергетическая и спектральная эффективность, многомерные поверхностно-сферические коды.

Введение

В фундаментальных работах К. Шеннона [1-2] для пояснения основных функций передачи и приема цифровых сообщений широко использовались геометрические представления. Было показано, что для передачи сообщений по каналу связи необходимо создать оптимальный ансамбль сигналов – AC (кодов), длительность в полосе частот F которых T. Каждому из сигналов, входящих в AC, может быть поставлена в соответствие сигнальная точка (CT) в N – мерном пространстве (N = 2FT, FT – целое число).

Было также установлено, что координаты этих точек должны быть выбраны так, чтобы расстояние между ближайшими *СТ* было не меньше заданного значения *d*.

Представление сигналов в виде CT в *N*-мерном сигнальном пространстве основано на теореме Найквиста-Котельникова-Шеннона, согласно которой в канале связи с полосой частот *F*, любые сигналы могут быть представлены в виде суммы N = 2FT, элементарных составляющих. Безразмерную полосу частот канала связи с заданным параметром *FT* можно считать нормированной длительностью передаваемых сигналов, для геометрического представления которых удобно использовать их разложение по базису ортогональных функций.

Настоящий материал посвящен изложению проблем построения конкретного оптимального ансамбля сигналов (кодов) в многомерном пространстве состоит из трех частей. В данной статье описана процедура построения многомерных поверхностно-сферических кодов, оптимальность которых была доказана Шенноном. При формировании сигналов, принадлежащих этому коду, осуществляется фазовая модуляция N-1 генераторов элементарных гармонических колебаний, входящих в состав модулятора. Поэтому метод передачи сообщений с помощью поверхностно-сферических кодов логично называть гиперфазовой модуляцией (ГПФМ, HPM – Hyperphase Modulation). Разработан алгоритм, ставящий в соответствие номеру m передаваемого сообщения (алгоритм кодирования сообщений) сигналы $S_m(t)$, входящие в АС. Вычислительная сложность этого алгоритма растет лишь линейно с увеличением N.

На практике наиболее часто для этого применяются гармонические функции и при этом, как правило, используют методы временного и частотного уплотнения каналов.

При передаче сообщений по каналу связи с полосой F методом временного уплотнения (TDMA – Time Division Multiple Access) интервал T, делится на подинтервалы, длительность которых равна $\tau_T = T/(FT)$ и при $t \in T_i = [(i-1)\tau_T, i\tau_T,]$ (i = 1...FT) в канале связи передаются элементарные сигналы $V_{(2i-1),m}(t) = x_{m,(2i-1)} \times \sin \omega(t - T_i)$ и $V_{2i,m}(t) = x_{m,2i} \cos \omega(t - T_i)$, за пределами

интервала T равные 0. Параметры $x_{m,(2i-1)}$ и $x_{m,2i}$ определяются передаваемой цифровой информационной последовательностью символов на входе модулятора. Сигнал на входе приемника за интервал T может быть

записан в виде:
$$S_m(t) = \sqrt{2NP_S} \left[\sum_{i=1}^{F_I} V_{(2i-1),m}(t) + V_{2i,m}(t) \right]$$

где $P_{\rm S}$ – средняя мощность одного элементарного сигнала на передающем конце линии связи.

Другим методом передачи сигналов является частотное уплотнение канала связи ортогональными поднесущими (OFDM – Orthogonal Frequency-Division Multiplexing). Длительность передачи каждой поднесущей равна полной длительности сигнала Т, на каждой из них передаются элементарные сигналы $V_{(2i-1),m}(t) =$ $=x_{m,(2i-1)}\sin\omega_i t$ и $V_{2i,m}(t)=x_{m,2i}\cos\omega_i t$, частоты которых выбраны так, чтобы выполнялись условия $\int_{0}^{\infty} V_{i,m}(t) V_{j,m}(t) dt = 0$ при $i \neq j$. Для этого необходимо, чтобы $|\omega_i - \omega_i| = 2\pi |i - j| / T$ при $i \neq j$. В данном случае

сигнал, принятый за время *T*, также может быть представлен выше приведенным соотношением.

Отметим, что оба метода (TDMA и OFDM) с точки зрения геометрической трактовки алгоритмов передачи и приема сигналов полностью эквивалентны: передаче сигналов во временном интервале T_i в системе TDMA соответствует передача таких же сигналов на частоте ω_i . Однако следует отметить, что физические параметры этих сигналов существенно отличаются. Кроме того, как было показано С. Вейнштейном и П. Эбертом [3], используя быстрое преобразование Фурье, эти сигналы могут быть преобразованы из одного вида в другой.

Шенноном было показано, что CT оптимального ансамбля сигналов $S_m(t)(m = 1...M)$ при больших значениях N должны быть расположены на поверхности N-мерной сферы. Отметим, что в частном случае, когда N =2, двумерная сфера представляет собой окружность и в таком случае принято, что для передачи сигналов используется AC с многофазной модуляцией. При любых значениях N для всех входящих в AC поверхностно-сферических сигналов $S_m(t)$, имеющих длительность T, справедливо соотношение

$$\int_{0}^{T} S_m^{-2}\left(t\right) \frac{dt}{T} = N P_S , \qquad (1)$$

где P_s – мощность сигнала, приходящаяся на одно измерение *N*-мерного пространства, в котором размещены *CT*, соответствующие *AC*.

Параметры $x_{m,(2i-1)}$ и $x_{m,2i}$ могут рассматриваться как нормированные евклидовы координаты CT, соответствующие сигналу $S_m(t)$. Из (1) следует, что нормированные координаты всех CT удовлетворяют условию

$$\sum_{i=1}^{FT} \left[x_{m,(2i-1)}^2 + x_{m,2i}^2 \right] = 1 , \qquad (2)$$

т.е. при нормировке все CT лежат на поверхности N -мерной сферы с радиусом $R_{\rm s}=1$.

На рис. 1 показана структура систем цифровой системы связи [1, 2], фундаментальные исследования которой впервые были начаты Шенноном. Эта система на передающей стороне имеет последовательно соединенные кодер, на вход которого поступает $n \approx \log_2 M$ двоичных информационных символов, где M – количество передаваемых сигналов.



Рис. 1. Схема системы связи

Передаваемые блоки информационных символов по сути представляют собой номер m CT, соответствующий сигналу $S_m(t)$, принадлежащему используемому для передачи сообщений AC. По Шеннону в кодере осуществляется формирование N индексов фазовой модуляции, определяющих значения координат CT, соответствующих блоку входных информационных символов. В модуляторе формируется аналоговый сигнал, принадлежащий AC, выбранному для передачи сообщений. Оптимальный AC должен быть сконструирован таким образом, чтобы евклидовы расстояния между двумя ближайшими CT, соответствующими входящими сигналам, были бы равны d. От значения d зависит количество составляющих в AC и, следовательно, скорость передачи сообщений.

На приемной стороне линии связи принятый сигнал поступает на вход демодулятора, основной функцией которого является определение координат *CT*, соответствующей индексам модуляции принятого сигнала. Значения этих координат (индексов модуляции) подаются на вход декодера. В оптимальном демодуляторе по Шеннону осуществляется определение координат *CT*, принадлежащей одному из передаваемых сигналов вNмерном евклидовом сигнальном пространстве, наиболее близко расположенному к *CT*, соответствующий принятому сигналу, искаженному шумом, действующим на входе демодулятора.

Функциональным назначением декодера является формирование блока двоичных информационных символов, определяющих номер m принятого сообщения, используя для этого определенные в демодуляторе координаты *CT*.

Следует отметить, что, согласно Шеннону [1, 2] в оптимальной системе передачи сообщений по существу осуществляются только модуляция передаваемых сигналов (формирование сигналов $S_m(t)$), а помехоустойчивое кодирование, при котором по каналу связи передаются не только информационные, но и избыточные символы, не используется.

Шенноном было установлено, что, применяя оптимальный *AC*, можно достичь предельно высокой помехоустойчивости приема сигналов в том случае, когда скорость передачи сообщений $R = \log_2 M / T$ меньше пропускной способности канала связи:

 $C = F \cdot \log_2(1 + \rho_0) \tag{3}$

где $\rho_0 = P_{S0} / P_n$ – отношение средней мощности принятого сигнала (P_{S0}) к мощности шума (P_n), действующего на входе демодулятора.

При проектировании системы связи полоса частот канала, по которому передаются сообщения, обычно является заданной. Введем важные нормированные параметры, к значениям которых предъявляются требования при проектировании системы связи. Создаваемая система должна обеспечивать заданную удельную ско-

рость передачи сообщений ($R_f = (R/F) = \frac{\log_2 M}{N/2}$), а канал связи должен иметь удельную пропускную способность ($C_f = C/F$). Отметим, что между параметрами ρ_0 и C_f имеется соотношение $\rho_0 = (2^{C_f} - 1)$ и что условие R < C эквивалентно условию $R_f < C_f$. К системе связи предъявляется также требования обеспечения необходимой надежности связи – значения вероятности ошибочного приема сигналов ($P_{er}(\rho_0,FT)$).

Как было показано Шенноном [1-2], при выполнении условия $R_f < C_f$ вероятность ошибочного приема сигналов ($P_{er}(\rho_0, FT)$) может быть сколь угодно малой при увеличении длительности сигналов, входящих в оптимальный AC. Применяя такой AC, можно при $FT \to \infty$ обеспечить предел $P_{er}(\rho_0, FT) \to 0$. В работе Шеннона [2] были получены зависимости $P_{er}(\rho_0, FT)$ от ρ_0 и FT, из которых при $P_{er}(\rho_0, FT) = const$ могут быть определены соотношения между ρ_0 и FT, позволяющие выбрать размерность N-мерного сигнального пространства, в котором размещены CT сигналов оптимального AC.

В [4, 5] разработаны простые методы определения зависимостей $P_{er}(\rho_0,FT)$, относящиеся к классам поверхностно- и объемно-сферических AC (в литературе используются также эквивалентный термин – сигнально-кодовые конструкции), которые приводят к результатам Шеннона [2]. Из [2, 4, 5] следует, что в оптимальном AC параметр d, определяющий расстояние между ближайшими CT в AC,с увеличением N изменяется пропорционально $\sqrt{N}(d \sim \sqrt{N})$.

В работах Шеннона, посвященных помехоустойчивым кодам, было показано, что в N-мерном пространстве возможно построить ансамбль сигналов, в котором евклидово расстояние между любыми двумя сигналами имеет максимально возможное значение, зависящее только от значений FT, ρ_0 и R_f .

Термин «кодирование» по Шеннону существенно отличается от современного понятия в теории кодирования. К поверхностно- и объемно-сферическим ансамблям сигналов, исследованным Шенноном, широко используемое в этой теории понятие «избыточность», вообще не применимо.

Отметим, что в современных системах связи для передачи кодовых комбинаций, сформированных в кодере, широкое применение находят АС с квадратурной амплитудной модуляцией (QAM – Quadrature Amplitude Modulation) и с амплитудно-фазовой модуляцией (APM -Amplitude and Phase Modulation). Находят применение также и другие виды AC, в частности сигналов с квадратурно-фазовой (QPSK – Quadrature phase-shift keying), с частотной модуляцией (FSK - Frequency-shift keying) и др. [6, 7]. Все указанные АС имеют достаточно простую структуру и относятся к классу двумерных. Принадлежащие им сигналы представляют собой простое гармоническое колебание, модулированное по амплитуде, фазе или частоте. Следует указать, что в классе двумерных АС обеспечить между любыми сигналами достаточно большое значение минимального расстояния d, которое необходимо для их демодуляции с малой вероятностью ошибки, можно только за счет увеличения мощности Р_с передаваемого сигнала. Поэтому в современных системах связи получение малой вероятности ошибки при демодуляции сигналов является проблематичным.

Во всех двумерных AC при заданном значении P_s величина d уменьшается с увеличением R_f . Поэтому, используя двумерные AC, невозможно получить высокую помехоустойчивость и скорость передачи сообщений, которую может обеспечить идеальный по Шеннону AC при $d \sim \sqrt{N}$.

Во многих современных системах передачи сообщений кодирование и модуляция осуществляются раздельно. При этом кодер (рис. 1), построенный на основе современной теории кодирования, осуществляет формирование кодовых комбинаций, имеющих длину *n* и содержащих как информационные символы, поступившие на его вход, так и дополнительные избыточные символы, связанные с информационными определенными математическими соотношениями. Включение в каждую кодовую комбинацию избыточных символов увеличивает расстояние между разными кодовыми комбинациями и позволяет исправлять в них ошибки, возникающие из-за действия шумов при приеме сигналов. Однако избыточные символы уменьшают реальную скорость передачи сообщений. Символы, сформированные в кодере, поступают на вход модулятора. В соответствии с блоком двоичных символов длины k < n $(k \in 1, ..10)$ в модуляторе формируется один из 2^k аналоговых сигналов, принадлежащих АС для передачи сообщений в канал связи.

Важным вкладом в теорию связи стали работы Унгербоека [8], разработавшего методы так называемой решетчатой модуляции (Trellis-CodedModulation – TCM), предназначенной для передачи сообщений в каналах связи, ограниченных по полосе частот. Этот вид модуляции предусматривает введение избыточности в передаваемую информационную последовательность символов для формирования *N*-мерных *AC*, элементами которых являются двумерные *M* -позиционные *AC* типа QAM или PAM. Избыточные символы используются та-

ким образом, чтобы из M^N возможных CT, соответствующих всем комбинациям координат СТ для N последовательно передаваемых в канал связи элементарных двумерных сигналов, обеспечивалась бы передача только тех, для которых расстояние между ближайшими СТ превышает заданное расстояние между ближайшими *СТ* в двумерных *АС* значение *d*. Энергетический выигрыш от использования ТСМ вместо QAM без применения помехоустойчивого кодирования при М =128 и 256 достигает примерно 6 дБ [6]. При этом скорость передачи сообщений уменьшается до 2/3 от той, которая могла быть обеспечена при передаче этих же сообщений с помощью QAM без введения избыточных символов. Исследованию для передачи сообщений ТСМ и других видов многомерных АС посвящен также ряд обстоятельных исследований [9-11].

В ряде работ были выполнены многочисленные исследования, связанные с выяснением возможности применения для передачи сообщений многомерных сигналов. В этих исследованиях оценивались зависимости скорости передачи сообщений с помощью многомерных AC от их длительности и от значения минимального расстояния между ними, были предложены алгоритмы построения AC для сравнительно небольших значений N. Результаты таких исследований приведены, например, в [12], где даны обширные ссылки на соответствующую литературу.

Следует отметить, что общих решений задачи построения многомерного *AC*, с помощью которого сообщения передаются с потенциально возможной [2] помехоустойчивостью, в них не предлагалось.

Как видно из рис. 1, на приемной стороне сигналы поступают на вход демодулятора, в котором при традиционном построении системы связи в результате демодуляции сигнала формируются блоки из *n* символов кодовой комбинации, соответствующей принятому сигналу. Часть этих символов может быть принята с ошибкой. Эти кодовые комбинации поступают на вход декодера, в котором осуществляется исправление ошибок, появившихся на выходе демодулятора, и формируется последовательность информационных символов, являющихся копией переданной информации. Для обеспечения высокой помехоустойчивости приема - малой вероятности ошибки в информационной последовательности, применяются достаточно длинные коды (n ≫1). При их декодировании используются весьма сложные алгоритмы. Современная теория кодирования предлагает для применения в телекоммуникационных системах ряд кодов и методов их декодирования. Наиболее развиты и широко применяются линейные коды, алгоритмы кодирования и декодирования которых основаны на теории Галуа, а также коды Галлагера, сверточные и турбо-коды [6, 7, 13].

Следует отметить, что в работах Шеннона не рассматривались вопросы построения конкретных AC в многомерном пространстве. Построение таких AC требует разработки алгоритма определения координат CT, соответствующих сигналам оптимального AC, таким образом, чтобы при заданном значение R_f наименьшее расстояние d между сигналами, входящими в AC, увеличивалась бы с увеличением N по закону $d \sim \sqrt{N}$ – также, как идеальном по Шеннону AC.

В работах Шеннона не затрагивались также вопросы, связанные с разработкой алгоритма нумерации сигналов, входящих в AC. Однако важно обратить внимание на то, что процедуру кодирования для оптимального кода (AC), содержащего $M = 2^{R_f(FT)}$ сигналов, Шеннон трактовал, как операцию присвоения определенного номера *m* каждому из M возможных сигналов, считая, что двоичная запись этого номера представляет собой информационную последовательность знаков, которые должны быть переданы по каналу связи.

Гиперфазовый модулятор и кодер сообщений

Процедуры построения гиперфазового модулятора, в котором определяются координаты всех CT ансамбля сигналов, размещенных на поверхности N-мерной сферы радиуса 1 и определения количества CT в AC можно представить в ниже следующем виде. Известно [14], что нормированные евклидовы координаты любой CT, соответствующей сигналу $S_m(t)$, расположенной на поверхности N-мерной сферы радиуса 1, могут быть выражены в через ее гиперсферические (фазовые) координаты с помощью следующих N соотношений:

$$x_{m1} = \sin \varphi_{m1}, \quad x_{m2} = \cos \varphi_{m1} \sin \varphi_{m2},$$

$$x_{mk} = (\prod_{i=1}^{k-1} \cos \varphi_{mi}) \sin \varphi_{mk}, \dots \qquad (4)$$

$$\dots x_{m(N-1)} = (\prod_{i=1}^{N-2} \cos \varphi_{mi}) \sin \varphi_{m(N-1)}, \dots, x_{mN} =$$

$$= (\prod_{i=1}^{N-2} \cos \varphi_{mi}) \cos \varphi_{m(N-1)}.$$

Для упрощения записи этих формул обозначим: $R_{m1} = 1$, $R_{m2} = \cos(\varphi_{m1})$ и $R_{mk} = (\prod_{i=1}^{k-1} \cos \varphi_{mi})$. Тогда приведенные формулы примут более компактный вид: $x_{m1} = R_1 \sin \varphi_{m1}$, $x_{m2} = R_{m2} \sin \varphi_{m2}$,

$$x_{mk} = R_{mk} \sin \varphi_{mk}, \dots x_{m(N-1)} = R_{m(N-2)} \sin \varphi_{m(N-1)},$$

$$x_{mN} = R_{m(N-2)} \cos \varphi_{m(N-1)}.$$
(5)

Эти соотношения показывают, что при заданном значении P_s каждая CT в N-мерном пространстве определяется N евклидовыми координатами x_{mk} (k = 1...N), зависящими от (N-1) параметров φ_{km} – ее фазовых координат. Эти параметры могут быть использованы для передачи сообщений в качестве информационных. Особенностью представления координат CT, расположенных на поверхности гиперсферы, является их взаимозависимость: для сигнала $S_m(t)$ евклидова координата по оси x_1 зависит только от одной фазовой координата по оси N-мерного пространства – от фазовых координат φ_{m1} , $\varphi_{m2}, \ldots \varphi_{m(k-1)}$. Важно указать диапазон возможных значений параметров φ_{km} [14]:

V-0,5 $\pi \le \varphi_{mk} \le 0,5\pi$, если $1 \le k \le (N-2)$, и $0 \le \varphi_{m(N-1)} \le 2\pi$. (6)

Из (5) следует, что если значения координат x_{mk} CT для $1 \le k \le (N-2)$ заданы, то ее координаты $x_{m(N-1)}$ и x_{mN} лежат на окружности, радиус которой зависит от значений x_{mk} $(k = 1, \ldots (N-2))$ и равен $R_{m(N-2)}$.

Алгоритм синтеза AC, предназначенного для передачи сообщений по каналу связи с удельной скоростью R_f и высокой помехоустойчивостью, состоит в выборе фазовых координат φ_{mk} таким образом, чтобы расстояние между ближайшими CT, размещенными на поверхности N-мерной сферы, было бы равно заданному значению d. Кроме того, количество сигналов в N-мерном AC (и соответствующих им CT) должно быть таким, чтобы этот ансамбль позволял передавать сообщения по каналу связи с заданной удельной скоростью R_f . Для этого количество CT, размещенных на поверхности N-мерной сферы, должно быть равно $M = 2^{R_f(N/2)}$.

При синтезе AC необходимо, чтобы каждый из информационных параметров (фазовых координат CT) мог изменяться дискретно, т.е. $\varphi_{mk} = I_{mk} \Delta \varphi_{mk}$, где I_{mk} – целые числа $(-M_k \leq I_{mk} \leq M_k)$ – индексы модуляции передаваемого сообщения, соответствующие его номеру, однозначно определяющие значение m. Операции вычисления индексов выполняются на передающем конце линии связи в блоке их формирования, который можно назвать кодером. Формирование сигнала $S_m(t)$, зависящего от фазовых координат φ_{mk} (k = 1...(N-1)), осуществляется в гиперфазовом модуляторе.

Процедура формирования сигналов ГПФМ определяется формулами(4) и (5).

Ниже показано, что значения евклидовых координат CT для AC, их количество (M), а также значения $\Delta \varphi_{mk}$ полностью определяются выбором только одного параметра $\Delta \varphi_{\rm l}$, причем значения $\Delta \varphi_{mk}$ и M_{mk} зависят не только от $\Delta \varphi_{\rm l}$, но и, как следует из (4) и (5), от того, какие значения присвоены фазовым координатам φ_{mi} при $i \leq (k-1)$.

Найдем соотношения, определяющие $\Delta \varphi_{mk}$ и M_{mk} , используя для этого условие минимального евклидово расстояния d между любыми двумя соседними CT, соответствующими принадлежащим AC сигналам $S_{m_{\rm I}}(t)$ и

$$d^{2} = \int_{0}^{T} \left[S_{m_{1}}(t) - S_{m_{2}}(t) \right]^{2} \frac{dt}{T} =$$

$$= NP_{S} \left\{ \sum_{i=1}^{FT} \left[x_{m_{1},(2i-1)} - x_{m_{2},(2i-1)} \right]^{2} + \left[x_{m_{1},2i} - x_{m_{2},2i} \right]^{2} \right\}.$$
(7)

Нормируя расстояние d к $\sqrt{NP_S}$ ($d_0 = d / \sqrt{NP_S}$), соотношение (7) можно записать:

$$d_0^2 = \left[\sum_{i=1}^{FT} (x_{m_1,(2i-1)} - x_{m_2,(2i-1)})^2 + (x_{m_1,2i} - x_{m_2,2i})^2\right].$$
 (8)

Значения *k*-х фазовых координат (*k*-х информационных параметров) в рассматриваемом *AC* могут принимать значения, равные $\varphi_{mk} = I_{mk} \Delta \varphi_k$, где I_{mk} – целые положительные или отрицательными числа. Возможный диапазон их изменений будет определен ниже.

Отметим, что минимальное фазовое расстояние между двумя сигналами $S_{m_1}(t)$ и $S_{m_2}(t)$, отличающимися друг от друга только k-й угловой координатой, равно $\Delta \varphi_{mk}$ (при этом $\min \left| I_{m_1k} - I_{m_2k} \right| = 1$). Выполнив в (6) выкладки с учетом (5), найдем нормированное евклидово расстояние между ними:

$$d_0^2 = 4\sin^2(\Delta \varphi_1 / 2) = 4 \left(\prod_{i=1}^{k-1} \cos \varphi_{mi}\right)^2 \sin^2(\Delta \varphi_{mk} / 2)$$
при $k \ge 2$ (9)

Второе соотношение в (9) относится к случаю, когда у сигналов $S_{m_1}(t)$ и $S_{m_2}(t)$ значения первых фазовых координат φ_{m_i} совпадают при $i \leq (k-1)$, и отличаются только значения $\varphi_{m,k}$ и $\varphi_{m,k}$.

Используя уравнения (9), получим все основные параметры AC: значения φ_{mk} всех фазовых координат CT, значения M_{mk} , определяющие возможный диапазон изменений параметра I_{mk} для каждого значения k, количество CT для рассматриваемого AC (M), а также удельную скорость передачи сообщений R_f , которая может быть обеспечена этим AC.

Из (9) следуют соотношения, определяющие зависимость $\Delta \varphi_{\scriptscriptstyle mk}$ и $M_{\scriptscriptstyle mk}$ от $\Delta \varphi_{\rm 1}$:

$$(\prod_{i=1}^{k-1} \cos \varphi_{mi})^2 \sin^2(\Delta \varphi_{mk} / 2) = \sin^2(\Delta \varphi_1 / 2).$$
(10)

Рассмотрим построение AC для случая, когда $N \ge 3$. Формула (10) позволяет последовательно находить значения $\Delta \varphi_{mk}$ в зависимости от $\Delta \varphi_1$, а также значения M_{mk} , определяющие возможный диапазон изменения параметров I_{mk} . Величины $\Delta \varphi_{mk}$ равны:

$$\Delta \varphi_{mk} (\Delta \varphi_1, \overline{I}_{m(k-1)}) =$$

$$= 2 \arcsin[|\sin(\Delta \varphi_1 / 2)| / \prod_{i=1}^{k-1} |\cos(I_{mi} \Delta \varphi_{mi})|], \qquad (11)$$

Отметим, что значения $\Delta \varphi_{mk}$ зависят только от абсолютного значения индекса $|I_{m1}|$. Запись формул более компактна, если указать на зависимость индексов I_{mi} от многих параметров (i = 1...k), в (11) и в последующих формулах строка из k символов I_{mk} записывается в виде $\vec{I}_{mk} = (I_{m1}, I_{m2}, ... I_{mk})$.

Определить из (11) значение $\Delta \varphi_{mk}(\Delta \varphi_1, I_{m(k-1)})$ можно только в том случае, если выполняется условие

S(t)

$$\left|\sin(\Delta \varphi_{1}/2)\right| / \prod_{i=1}^{k-1} \left|\cos(I_{mi}\Delta \varphi_{mi})\right| \le 1.$$
(12)

٨

Следовательно, целые числа $I_{\it mk}$ при любом значении k должны лежать в интервале $-M_k(\Delta \varphi_1, I_{m(k-1)}) \leq$ $\leq I_{mk} \leq M_k(\Delta \varphi_1, \vec{I}_{m(k-1)}),$ $M_1(\Delta \varphi_1, \vec{I}_{m(k-1)}) =$ где $= int \{ \arccos[|\sin(\Delta \varphi_1 / 2)|] / \Delta \varphi_1 \}$ и при $k \ge 2$ $\vec{1}$

$$M_{k}(\Delta \varphi_{1}, I_{m(k-1)}) =$$

$$= \operatorname{int}\left\{\frac{\operatorname{arccos}\left[\left|\sin(\Delta \varphi_{1} / 2)\right| / (\prod_{i=1}^{k-1} \left|\cos(I_{mi}\Delta \varphi_{mi})\right|)\right]}{\Delta \varphi_{mk}}\right\}.$$
(13)

Формулы (11)-(13), где int(x) – целая часть числа x, определяют все фазовые координаты СТ, т.е. их евклидовы координаты. Знание $\Delta \varphi_{mk}(\Delta \varphi_1, I_{m(k-1)})$ $M_{k}(\Delta \varphi_{1}, I_{m(k-1)})$ позволяет составить схему гиперфазового модулятора, указанного на функциональной схеме системы связи на рис. 1.

Формулы (11) и (13) показывают, что каждая фазовая координата может принимать разные значения в зависимости от чисел, входящих в состав строки $I_{m(k-1)} = (I_{m1}, I_{m2} \dots I_{m(k-1)}).$

Важно отметить, что целые числа I_{mk} в интервале $-M_{k}(\Delta \varphi_{1}, I_{m(k-1)}) \leq I_{mk} \leq M_{k}(\Delta \varphi_{1}, I_{m(k-1)}),$ могут рассматриваться как индексы вложенных подмножеств разного порядка, на которые может быть разбито множество всех СТ, размещенных на поверхности N-мерной сферы. Подмножества сигнальных точек 1-го порядка, которым присваиваются индексы |I_{m1}| (для краткости обозначены далее $\prod_{|I_{u_1}|}$), отличаются друг от друга тем, что принадлежащие им СТ имеют разные абсолютные значения 1-й фазовой координаты $\varphi_1(I_{m1}) = I_{m1} \Delta \varphi_1$, зависящей от индекса $I_{\scriptscriptstyle m1}$. Количество таких подмножеств $\Pi {
m M}_{|I_{n}|}$ равно $G_{\rm l}=2M_{\rm l}(\Delta {
m \phi}_{\rm l})+{
m l},$ где $M_{\rm l}(\Delta {
m \phi}_{\rm l})=$ $= int \{ \arccos[\sin(\Delta \varphi_1 / 2)] / \Delta \varphi_1 \}$ – число значений, которые может принимать индекс |I_{m1}|. Каждое из подмножеств 1-го порядка с индексом |I_{m1}| может быть разбито на подмножества 2-го порядка, которые обозначены $\prod_{|I_{u_1}|,|I_{u_2}|}$ и отличаются друг от друга тем, что принадлежащие им СТ имеют разные значения 2-й фазовой координаты $\varphi_2(\vec{I}_{m2}) = I_{m2} \Delta \varphi_2(\Delta \varphi_1, |I_{m1}|)$ при $|I_{m1}| = const.$ Количество таких подмножеств равно $G_2(I_{m1}) =$ $= 2M_2(\Delta \varphi_1, |I_{m1}|) + 1$. В общем случае, когда определены фазовые координаты $CT \ \varphi_1(I_{m1}), \ \varphi_2(\vec{I}_{m2}), \dots, \varphi_{mk}(\vec{I}_{mk}) =$ $= I_{mk} \Delta \varphi_{mk} (\Delta \varphi_1, \vec{I}_{m(k-1)}),$ входящих в одно из подмножеств k -го порядка – $\prod_{|I_m|,...,|I_m|}$, то это подмножество может быть разбито на $G_{(k+1)}(\vec{I}_{mk}) = 2M_{(k+1)}(\Delta \varphi_1, \vec{I}_{mk}) + 1$ подмножеств (k+1)-го порядка.

В каждое из этих подмножеств входят СТ, у которых (k+1) -я фазовая координата определяется формулой $\varphi_{m(k+1)}(\vec{I}_{m(k+1)}) = I_{m(k+1)} \Delta \varphi_{mk}(\Delta \varphi_1, \vec{I}_{mk})$. При этом $-M_{(k+1)}(\Delta \varphi_1, \vec{I}_{mk}) \le I_{m(k+1)} \le M_{(k+1)}(\Delta \varphi_1, \vec{I}_{mk})$ и $|I_{mi}| =$ = const при $i \leq k$.

Отметим важные соотношения между значениями $\Delta \varphi_{mk}(\Delta \varphi_1, I_{mk})$ и $M_k(\Delta \varphi_1, I_{m(k-1)})$, которые следуют из (11) и (13):

$$\Delta \varphi_{mk} (\Delta \varphi_1, I_{mk}) \ge \Delta \varphi_{m(k+1)} (\Delta \varphi_1, I_{m(k+1)})$$

$$\mathsf{M} \ M_{(k+1)} (\Delta \varphi_1, \vec{I}_{mk}) \le M_k (\Delta \varphi_1, \vec{I}_{m(k-1)}) .$$
(14)

Из (14) следует, что с увеличением k увеличиваются значения $\Delta \varphi_{mk} (\Delta \varphi_1, I_{mk})$, а количество подмножеств, на которое может быть разбито подмножество k -го порядка, уменьшается. Максимальное количество таких подмно-

(

жеств
$$\Pi_{M_{|I_{m1}|}}$$
 равно $M_1(\Delta \varphi_1) = \operatorname{int} \left\{ \operatorname{arccos} \frac{\left[\operatorname{sin} \left(\frac{\Delta \varphi_1}{2} \right) \right]}{\Delta \varphi_1} \right\}$

В том случае, если $\Delta \varphi_1 \ll 1$, то $M_1(\Delta \varphi_1) \approx int[\pi / 2\Delta \varphi_1]$.

Сигнальные точки, входящие в состав подмножеств (N-1)-го порядка, у которых (N-1)-я фазовая координата равна $\varphi_{m(N-1)} = I_{m(N-1)} \Delta \varphi_{m(N-1)} (\Delta \varphi_1, I_{m(N-2)})$, при заданных значениях $\vec{I}_{m(N-2)}$ лежат на двумерных окружностях. Так как для таких СТ должно выполняться условие $0 \le \varphi_{m(N-1)} \le 2\pi$, то индексы $I_{m(N-1)}$ должны лежать в интервале

$$0 \le |I_{m(N-1)}| \le M_{(N-1)}(\Delta \varphi_1, \vec{I}_{m(N-2)}) =$$

= int(2\pi / |\Delta \varphi_{m(N-1)}(\Delta \varphi_1, \vec{I}_{m(N-2)})|), (15)

где $M_{(N-1)}(\Delta \varphi_1, I_{m(N-2)})$ – количество *CT*, лежащих на одной окружности.

Формулы (13) и (15) позволяют определить общее количество *CT* для *AC* в *N* -мерном пространстве:

$$M_{sN}(\Delta \varphi_{1}) = \sum_{I_{m1}=-M_{1}(\Delta \varphi_{1})}^{M_{1}(\Delta \varphi_{1})} \dots \sum_{I_{m(N-1)}=-M_{N}(\Delta \varphi_{1}, \vec{I}_{m(N-2)})}^{M_{N}(\Delta \varphi_{1}, \vec{I}_{m(N-2)})} M_{(N-1)}(\Delta \varphi_{1}, \vec{I}_{m(N-2)}),$$
(16)

а также удельную скорость передачи сообщений $R_r = [2\Delta \log_2 M_{sN}(\Delta \varphi_1)] / N.$

Основные характеристики сигналов с гиперфазовой модуляцией

С помощью формул (16) и (17) на рис. 2 построены зависимости удельной скорости передачи сообщений $R_{f}(\Delta \varphi_{1})$ от параметра $\Delta \varphi_{1}$ (град) для исследованного AC, а также для AC, рассмотренного в работе Шеннона [2].

Для многомерных поверхностно-сферических АС зависимости $R_{f}(\Delta \varphi_{1})$ построены при значениях N =3, 4, и 6. Из рис. 2 видно, что кривые $R_{f}(\Delta \varphi_{1})$ от значения N зависят довольно слабои при N > 6 они практически

(17)

совпадают с кривой при N =6. Зависимость $R_f(\Delta \varphi_1)$ от $\Delta \varphi_1$ гораздо более существенна. Для передачи сообщений с высокой удельной скоростью следует выбирать AC, у которых параметр $\Delta \varphi_1$ имеет небольшое значение.

Представляет интерес вопрос о выигрыше в увеличении минимального расстояния *d* между соседними *CT* в *N*-мерном ансамбле сигналов с ГПФМ по отношению к ансамблю сигналов с QAM, который сегодня широко применяются на практике. Это связано с тем, что помехоустойчивость сигналов любого многопозиционного ансамбля зависит от минимального расстояния *d* между его *CT*, при этом, чем больше это расстояние, тем выше помехоустойчивость приема сигналов.



Рис. 2. Зависимости удельной скорости передачи сообщений $R_f(\Delta \varphi_1)$

Такой выигрыш определяется следующей формулой [6] $\mu(R_f N) = 10 \log(d_{AC}^2 / d_{QAM}^2)$ (18)

По формуле (18) на рис. З построены графики зависимости $\mu(R_{\ell}, FT)$ при R_{ℓ} = 2, 4, 6 и 8.





В настоящее время в системах связи для передачи сообщений широко применяются, в частности, двумерные AC, такие, например, как QAM. Квадрат минимального расстояния между этой AC равен $d_{QAM}^2 = 6P_S \div (2^{R_f} - 1)$ [6], где R_f – удельная скорость передачи сообщений в системе связи, в которой применяется AC

с QAM. В исследованном *N*-мерном $AC d_{AC}^2 = = 4NP_S \{ \sin[\Delta \varphi_1(R_{fN})/2]^2 \}$. Квадрат минимального расстояния, как следует из (7), равен d_{AC}^2 и увеличивается с ростом N (в этой формуле зависимость $\Delta \varphi_1(R_f)$ является функцией, обратной $R_f(\Delta \varphi_1)$, представленной на рис. 2).

Из рис. З следует, что определенный (меньше 2 дБ) выигрыш в $\mu(R_f,N)$ от применения многомерного AC может быть получен даже при сравнительно небольших значениях $FT \cong 2$, если $R_f \le 2 \div 6$. При больших значениях $R_f \ge 8$ от применения такого AC заметный выигрыш $\mu(R_f,N)$ может быть получен только в том случае, если $FT \ge 9$. Отметим, что, как известно [6], для двумерных AC сигналы с многофазной модуляцией (частный случай сигналов с ГПФМ) также уступают по своей энергетической эффективности сигналам QAM, если $R_f \ge 3$. Из рис. З также следует, что с увеличением R_{f_f} для получения заметного выигрыша в $\mu(R_f,N)$ сигналов с ГПФМ по сравнению QAM, требуется увеличивать длительность применяемых многомерных сигналов.

Отметим, что при заданном $P_{er}(\rho_0,FT)$ – значении вероятности ошибочного приема сигналов, энергетический выигрыш системы связи с ГПФМ по сравнению с системами, использующими AC с QAM, меньше, чем $\mu(R_{f,N})$, так как на прием сигналов, входящих в оптимальный АС ГПФМ, действуют шумы на каждую из N ортогональных составляющих этих сигналов.

Отметим, что Унгербоеком [6, 8] разработана решетчатая кодированная модуляция, сочетающая многопозиционную модуляцию и сверточное кодирование с избыточными символами. В ней при удельной скорости передачи сообщений $R_f \leq 4$ был получен выигрыш в $\mu(R_{f,}N)$ порядка 6 дБ [6, 8]. Для двумерных сигналов этот эффект определяет предельный энергетический выигрыш системы TCM по сравнению с системой, в которой для передачи сообщений применяются сигналы с QAM без использования помехоустойчивых кодов.

Алгоритм кодирования сообщений

Выше были получены формулы, позволяющие рассчитать количество сигналов $M_{sN}(\Delta \varphi_1)$, входящих в состав AC), а также удельная скорость передачи сообщений в системе связи, использующей ГПФМ модуляцию.

Построенный AC обеспечивает минимальное евклидово расстояние между разными сигналами AC, равное $d \sim \sqrt{N}$, и позволяет передавать по каналу связи различные сообщения, каждому из которых могут быть присвоены номера $m = 0...M_{sN}(\Delta \varphi_1) - 1$. При передаче информационного сообщения, имеющего номер m и представленного в виде информационной двоичной последовательности знаков, эта последовательность поступает на вход системы связи рис. 1.

Для того, чтобы передать такую последовательность по каналу связи с помощью одного из формируемых в модуляторе сигналов $S_m(t)$ с ГПФМ, необходимо преобразовать номер передаваемого сообщения *m* в последовательность его индексов – чисел I_{mk}, с помощью которых могут быть определены все фазовые координаты *СТ*, соответствующей сигналу $S_m(t)$. Алгоритм решения этой задачи (по Шеннону алгоритм кодирования передаваемого сообщения) должен иметь минимально возможную вычислительную сложность. Предлагаемый рекуррентный алгоритм вычисления чисел І_{ти} (кодирования) основан на разбиении всего множества СТ на подмножества 1-го, 2-го, ... (N-2) -го порядка в соответствии со значениями индексов I_{mi} ($i \le (N-2)$), от которых зависят значения фазовых координат (φ_{mi}) сигнальных точек АС. Последний индекс І_{т(N-1)} определяет координату $\varphi_{m(N-1)}$ у CT, которая лежит на соответствующей окружности и зависит от всех значений $\phi_{\scriptscriptstyle mi}$ при $i \le (N-2)$.

Число подмножеств *i*-го порядка равно $G_i(I_{m(i-1)}) =$ $=2M_{i}(\Delta \varphi_{1}, \vec{I}_{m(i-1)})+1$ – количеству значений, которое могут принимать индексы І_{ті} при определенном значении *i*, где $M_i(\Delta \varphi_1, I_{m(i-1)})$ рассчитываются по формуле (13). Структура множества всех СТ для ГПФМ такова, что, например, в каждое из подмножеств 1-го порядка при заданном значении $I_{m1} = const$ входит $G_2(I_{m1}, I_{m2})$ подмножеств 2-го порядка. При этом каждому такому подмножеству относятся СТ, первая фазовая координата которых равна $\varphi_{m1} = I_{m1} \Delta \varphi_1$, а вторая $-\varphi_{m2} =$ $= I_{m_2} \Delta \varphi_{m_2}(I_{m_1})$. Эта координата определяется значением I_{m2} , изменяющееся в интервале $-M_2(I_{m1}) \le I_{m2} \le$ $\leq M_2(I_{m1})$. Подобным же образом в определяются подмножества CT 3-го порядка, координаты φ_{m1} и φ_{m2} которых уже определены (т.е. $I_{m1} = const$ и $I_{m2} = const$). Эти CT входят в $G_3(I_{m1}, I_{m2}, I_{m3})$ подмножеств 3-го порядка и т.д. В подмножества (N-1) -го порядка при $I_{mk} = const$ при $k \le (N-2)$ входят CT, расположенные на окружностях, причем на каждой из них количество СТ равно $M_{(N-1)}(\Delta \varphi_1, I_{m(N-2)}).$

Отметим несколько важных свойств зависимости $M_{(N-1)}(\Delta \phi_1, \vec{I}_{m(N-2)})$ от $I_{mi}.$

1. Значения $M_{(N-1)}(\Delta \varphi_1, \vec{I}_{m(N-2)})$ при $\vec{I}_{m(N-2)} = const$ зависят, как следует из (15), только от абсолютных значений индексов I_{mk} (т.е. от $|I_{mk}|$). При этом индексы $I_{m(N-1)}$ изменяются в пределах $0 \le I_{m(N-1)} \le M_{(N-1)} \times \times (\Delta \varphi_1, \vec{I}_{m(N-2)})$.

2. Суммарное количество СТ, лежащих на всех таких

окружностях, является суммой слагаемых $M_{(N-1)} \times \times (\Delta \varphi_1, \vec{I}_{(N-2)m})$ в (16), имеющих одинаковые значения. Число слагаемых, входящих в эту сумму, равно числу возможных комбинаций знаков + и - у индексов I_{mk_i} (i = 1,...L). Поэтому, если только L значений $|I_{mk_1}|$, $|I_{mk_2}|,...,|I_{mkL}|$ не равны нулю, а все остальные равны нулю, то суммарное количество CT в подмножествах (N-1)-го порядка равно $2^L M_{(N-1)}(\Delta \varphi_1, \vec{I}_{(N-2)m})$.

3. Номер каждой из 2^{L} окружностей может быть присвоен, в принципе, произвольно. Но рациональное правило присвоения номера может быть таким: в строке $\vec{I}_{m(N-2)}$, которая входит в формулу для $M_{(N-1)}(\Delta \varphi_1, \vec{I}_{(N-2)m})$, на позициях, где расположены индексы $I_{mk} \neq 0$, ставиться число $\gamma_{mk} = 0, 5[1 + sign(I_{mk})]$, здесь функция sign(x) = 1 при x > 0 и sign(x) = -1 при x < 0, при этом $\gamma_{mk} = 1$ при $I_{mk} > 0$ и $\gamma_{mk} = 0$ при $I_{mk} < 0$. В результате номер соответствующей окружности оказывается записанным в этой строке в двоичной форме.

4. Величины $M_{(N-1)}(\Delta \varphi_1, I_{m(N-2)})$ имеют одно и тоже значение при любых перестановках индексов I_{mk_1} , I_{mk_2} ,..., I_{mk_k} .

Для определения индексов I_{mk} необходимо знать количество *CT*, содержащихся в подмножествах 1-го, 2-го и т.д. порядков. Формула, позволяющая рассчитать количество *CT* в соответствующем подмножестве k-го порядка в том случае, когда заданы значения индексов I_{mi} ($i \leq (k-1)$):

$$A_{k}(\vec{I}_{m(k-1)}) =$$

$$= \sum_{I_{mk} = -M_{k}(\Delta \varphi_{1})}^{M_{k}(\Delta \varphi_{1})} \cdots \sum_{I_{m(N-1)} = -M_{(N-1)}(\Delta \varphi_{1}, \vec{I}_{m(N-2)})}^{M_{(N-1)}(\Delta \varphi_{1}, \vec{I}_{m(N-2)})} M_{(N-1)}(\Delta \varphi_{1}, \vec{I}_{m(N-2)})).$$
(18)

При изменении знака одного из индексов, входящих в строку $\vec{I}_{m(k-1)}$, количество CT в соответствующем подмножестве не изменяется, так как $A_k(I_{m1},..,|I_{m(k-1)}|) = = (I_{m1},..,-|I_{m(k-1)}|)$. Отсюда следует, что

$$A_{k}(|I_{m1}|,..|I_{mi}|..|I_{m(k-1)}|) = 2A_{k}(I_{m1},..I_{m(k-1)}).$$
(19)

Абсолютные значения индексов $|I_{mi}|$ могут изменяться в следующих пределах: $0 \leq |I_{mi}| \leq M_i (\Delta \varphi_1, \vec{I}_{m(i-1)}).$

Алгоритм присвоения номеров *CT*, расположенным на N-мерной сфере, используя (18) и (19) для расчета $A_k(\vec{I}_{m(k-1)})$, таков:

Вычисляют $A_1(|I_{m1}|)$ – количество CT в подмножествах 1-го порядка при разных значениях $|I_{m1}| = 0, ...M_1(\Delta \varphi_1)$, а также суммарное количество CT подмножеств 1-го порядка для которых $|I_{m1}| \le K_{m1} \le M_1(\Delta \varphi_1)$, равное $Al \equiv (K_m 1) = \sum_{(|I_m||=0)}^{(K_m 1)} A_1(|I_m 1|)$. Отме-

тим, что $A1\Xi(M_1(\Delta \varphi_1)) = \sum_{(l_m 1|=0)}^{(M_1(\Delta \varphi_1)} A_1(|I_{m1}|)) = M_s N(\Delta \varphi_1),$ т.е. суммарное количество CT во всех подмножествах 1-го порядка равно полному количеству сигналов в AC. Поэтому номера всех CT, расположенных на поверхности N-мерной сферы, имеют значения $0 \le l \le [M_{sN}(\Delta \varphi_1) - 1].$ Причем, CT, относящиеся к подмножеству 1-го порядка с индексом, равным $|I_{m1}| = 0$, имеют номера $0 \le l \le \le A_1(0) - 1$, если же $|I_{m1}| = 1, ...M_1(\Delta \varphi_1)$, то CT этих подмножеств они имеют номера $A_1(|I_{m1}| - 1) \le l \le \le A_1(|I_{m1}|) - 1$.

Вычисляют $M_1(\Delta \varphi_1)$ значений $Al \equiv (K_{m1})$ и, выполняя сравнение числа m с числами $Al \equiv (K_{m1})$, определяют к какому из подмножеств CT 1-го порядка относится номер m. Индекс $|I_{m1}|$ определяют таким образом: если $D_1(K_{m1}) = A_{l \equiv}(K_{m1} - 1) \le m \le A_{l \equiv}(K_{m1})$, то $|I_{m1}| = (K_{m1} - 1)$.

Прямое нахождение индекса $|I_{m1}|$ требует выполнения $N_1=M_1(\Delta \varphi_1)$ сравнений. Однако это число можно существенно уменьшить, если воспользоваться алгоритмом решения задач о взвешивании монет [15]. В [15] показано, что для определения индекса $|I_{m1}|$ достаточно выполнить только $N_{01}=\log_2(N_1)$ сравнений.

Вычислив индекс $|I_{m1}|$ и определив подмножество CT 1-го порядка, содержащее $\Delta N_1 = [A_{1\Xi}(K_{m1}) - -A_{1\Xi}(K_{m1} - 1)]$ CT, в котором имеется одна CT, соответствующая номеру m, аналогичным образом вычисляют индекс $|I_{m2}|$ при $|I_{m1}| = const.$

Для этого, найденное подмножество CT 1-го порядка разбивается на $M_2(\Delta \varphi_1, |I_{m1}|)(|I_{m2}|=0, \ldots M_2(\Delta \varphi_1, |I_{m1}|)-1)$ подмножеств 2-го порядка. Вычисление индекса $|I_{m2}|$ осуществляется следующим образом:

Вычисляют $A_2(|I_{m1}|,|I_{m2}|)$ – количество CT в подмножествах 2-го порядка при разных значениях $|I_{m2}|(|I_{m2}|=0,...M_2(\Delta \varphi_1,|I_{m1}|)-1)$, а также суммарное количество CT подмножеств 2-го порядка, для которых $|I_{m2}| \le K_{m2} \le M_2(\Delta \varphi_1,|I_{m1}|)$, равное $A2\Xi(K_m2,|I_m1|) = \sum_{\substack{(I_m2|=0)\\ (I_m2|=0)}}^{(K_m2)} A_2(|I_m1|,|I_m2|)$. Отметим, что $A2\Xi(M_2(\Delta \varphi_1,|I_m1|)) = \sum_{\substack{(I_m2|=0)\\ (I_m2|=0)}}^{(M_2(\Delta \varphi_1,|I_m1|)} A_2(|I_m1|,|I_m2|) = A_1(|I_m1|)$, т.е. суммар-

ное количество *CT* во всех подмножествах 2-го порядка равно полному количеству *CT* в подмножествах 2-го порядка равно полному количеству *CT* в подмножества 2-го порядка, которое соответствует индексу $|I_{m1}|$. Причем, *CT*, относящиеся к подмножеству 2-го порядка с индексом, равным $|I_{m2}| = 0$, *CT* имеют номера $A_{1\Xi}(K_{m1}-1) \le \le l \le A_{1\Xi}(K_{m1}-1) + A_2(0, |I_{m2}|) - 1$, а если $|I_{m1}| = 1, ..., M_2(\Delta \varphi_1, |I_{m1}|)$, то *CT* имеют номера $D_1(K_{m1}) + A_2(|I_{m2}| - -1, |I_{m1}|) \le l \le D_1(K_{m1}) + A_2(|I_{m2}|, |I_{m1}|) - 1$.

Вычисляют $A_1(|I_{m1}|)$ значений $A2\Xi(K_{m2},|I_{m1}|)$ и, выполняя сравнение числа m с числами $A2\Xi(K_{m2},|I_{m1}|)$, определяют к какому подмножеству CT 2-го порядка относится номер m, т.е. определяют индекс $|I_{m2}|$: если $D_2(K_{m2}) = D_1(K_{m1}) + A_{2\Xi}(K_{m2} - 1, |I_{m1}|) \le m \le +D_1(K_{m1}) + A_{2\Xi}(K_{m2}, |I_{m1}|) - 1$, то $|I_{m2}| = (K_{m2} - 1)$.

Прямой способ нахождения $|I_{m2}|$ требует выполнения $N_2 = M_2(\Delta \varphi_1, |I_{m1}|)$ операций сравнения, а при использовании оптимального метода [11] достаточно выполнить только $N_{02} = \log_2(N_2)$ сравнений. Вычислив индекс $|I_{m2}|$ и определив подмножество CT 2-го порядка, содержащее $\Delta N_2 = [A_{2\Xi}(K_{m2}, I_{m1}) - A_{2\Xi}(K_{m2} - 1, I_{m1}))]$ CT, в числе которых имеется CT, соответствующая номеру m, переходим к определению индекса $|I_{m3}|$.

Индекс $|I_{m3}|$ определяют аналогичным образом, учитывая значения уже найденных индексов $|I_{m1}| = const$ и $|I_{m2}| = const$. Для этого, найденное подмножество CT 2-го порядка разбивается на $M_3(\Delta \varphi_1, |I_{m1}|, |I_{m2}|)(|I_{m3}| = 0, ...M_3(\Delta \varphi_1, |I_{m1}|, |I_{m2}|))$ подмножеств 3-го порядка. Продолжая подобные вычисления, последовательно находим индексы $|I_{m1}|, |I_{m2}|, ...$

 $|I_{m(N-2)}|.$

Следует отметить, что, по мере нахождения этих индексов, интервал чисел, в котором может находиться CT, соответствующая номеру передаваемого m сообщения, все время уменьшается. После определения всех индексов $|I_{m1}|, |I_{m2}|, \dots |I_{m(N-2)}|(|I_{m(N-2)}| = (K_{(N-2)} - -1))$ он становится равным $\Delta N_{(N-2)} = [A_{(N-2)\Xi}(K_{(N-2)}, \vec{I}_{m(N-2)}) - A_{(N-2)\Xi}(K_{(N-2)} - 1, \vec{I}_{m(N-2)})]$. При вычислении $|I_{mk}|$ находится также нижняя граница этого интервала, которая вычисляется по рекуррентной формуле $D_{(N-2)} \times D_{(N-2)}$

$$\times (K_{m(N-2)}) = D_{(N-3)}(K_{m(N-3)}) + A_{(N-2)\Xi}(K_{(N-2)} - 1, |I_{m(N-2)}|).$$

Подмножество CT (N-1)-го порядка, номера которых находятся в интервале $[D_{(N-2)}(K_{m(N-2)}), D_{(N-2)}(K_{m(N-2)}) + \Delta N_{(N-2)}]$. Эти CT лежат на окружностях, число которых, как указывалось выше, равно 2^L (число L равно количеству $|I_{mk}| \neq 0$ в найденной последовательности индексов $\vec{I}_{m(N-2)}$), при этом на каждой окружности лежит $M_{(N-1)}(\Delta \varphi_1, \vec{I}_{(N-2)m})$ сигнальных точек. Определим число $m_1 = [m - D_{(N-2)}(K_{m(N-2)})]$, которое является относительным номером CT, соответствующей номеру m. Это число определяет относительное положение этой CT, отсчитываемое от первой CT подмножества (N-1)-го порядка и содержащее $Z_{(N-1)}(\vec{I}_{(N-2)m}) = 2^L M_{(N-1)}(\Delta \varphi_1, \vec{I}_{(N-2)m})$ CT.

Порядок нумерации этих окружностей определяется знаками индексов в $\vec{I}_{(N-2)m}$. Номер окружности, на которой расположена *CT*, соответствующая числу *m*, равен

$$Y(\vec{I}_{(N-2)m}) = \operatorname{int}\left[\frac{m_1}{M_{(N-1)}(\Delta \varphi_1, \vec{I}_{(N-2)m})}\right]$$
. Записав это число

в двоичной форме в виде последовательности длиной (N-2) двоичных символов 0 и 1, можно определить знаки в строке абсолютных значений индексов $I_{(N-2)m}$, руководствуясь сформулированным выше правилом нумерации окружностей. Если на k-ой позиции двоичной последовательности, определяющей номер окружности, стоит символ 0, то индекс $I_{mk} = - |I_{mk}|$, в противном случае $I_{mk} = \mid I_{mk} \mid$. Остаток от деления числа m_1 на $M_{_{(N-1)}}(\Delta arphi_1,ec{I}_{(N-2)m})$ определяет значение индекса $I_{m(N-1)}$: если он равен 0, то *CT* является последней на окружности с номером $Y(\vec{I}_{(N-2)m})$ и $I_{m(N-1)} = M_{(N-1)} \times$ $\times (\Delta \varphi_1, \tilde{I}_{(N-2)m}) - 1$, в противном случае, он равен дробной части результата такого деления, т.е. $I_{m(N-1)} =$ $= m_1 - Y(\vec{I}_{(N-2)m})M_{(N-1)}(\Delta \varphi_1, \vec{I}_{(N-2)m}).$ Описанный алгоритм позволяет найти все индексы I_{mk} с учетом их знаков.

Процедура нахождения значений I_{mk} , соответствующих номеру m сообщения, поступившего на вход системы связи, требует вычисления K_A чисел $A_i(\vec{I}_{mi}) \times$

$$\begin{split} \times (K_{A} &= \sum_{i=1}^{N} M_{i} (\Delta \varphi_{1}, \vec{I}_{m(i-1)}) \leq N[2M_{1} (\Delta \varphi_{1}) + 1]) \quad \text{и выпол-} \\ \text{нения} \quad K_{S} &= N \sum_{i=1}^{N-1} \log_{2} [G_{i} (\vec{I}_{mi})] \leq N \left\{ \log_{2} \left[2M_{1} (\Delta \varphi_{1}) + 1 \right] \right\} \end{split}$$

сравнений этих чисел с числом *m*. Таким образом, сложность описанного алгоритма кодирования сообще-

ний растет линейно с увеличением N.

Заключение

В данной работе представлен метод формирования N-мерного ансамбля сигналов (AC) с гиперфазовой модуляцией, для котором все сигнальные точки (CT) соответствующего AC расположены на поверхности N-мерной сферы. Такой AC, как было показано Шенноном [2], является оптимальным при достаточно больших значениях N. Этот AC позволяет обеспечить передачу сигналов в канале связи с ограниченной полосой частот и заданной удельной скоростью передачи и максимально возможной помехоустойчивостью.

Принципиальное отличие этого метода передачи сообщений от тех, которые используются в современных системах связи, состоит в том, что для повышения помехоустойчивости в нем не предусмотрено внесение в последовательность информационных символов передаваемого сообщения помехоустойчивых кодов, содержащих помимо информационных дополнительные избыточные (проверочные) символы. Внесение избыточных символов снижает скорость передачи сообщений, хотя и позволяет на приемном конце линии связи исправлять ошибки в информационной последовательности, возникшие при демодуляции принятого сигнала.

Описан алгоритм гиперфазовой модуляции – метода формирования сигналов $S_m(t)$ такого ансамбля, а также разработан эффективный в вычислительном отношении метод кодирования сообщений – процедуры вычисления по номеру *m* передаваемого сообщения индексов модуляции I_{mk} , которые определяют фазовые координаты CT, соответствующей сигналу $S_m(t)$. Применение рассмотренного оптимального *N*-мерного *AC* даже при сравнительно небольших значениях *N* реализует выигрыш в величине минимального расстояния между *CT* в оптимальном *AC* по сравнению с широко распространенным *AC* с *QAM*. С увеличением *N* этот выигрыш растет.

Литература

1. Shannon C. Communication in the presence of noise, Proc. IRE, № 1, 1949. (Перевод на русский язык статьи «Связь при наличии шума», опубликованной в книге Шеннон К. Работы по теории информации и кибернетике. М.: Из-во иностранной литературы / под ред. Р.Л. Добрушина и О.Б. Лупанова).

2. Shannon C. Probability of error for optimal codes in Gaussian channel. Bell SystemTechn. J., May, 1959. (Перевод на русский язык статьи «Вероятность ошибки для оптимальных кодов в гауссовском канале», опубликованной в книге Шеннон К. Работы по теории информации и кибернетике. М.: Из-во иностранной литературы. // Под ред. Р.Л. Добрушина и О.Б. Лупанова).

3. Weinstein S.B., Ebert P.M. Data transmission by frequency division multiplexing using the discrete Fourier transform. IEEE Trans. COM-19, № 10, 1971.

4. Быховский М.А. Вероятность ошибки для оптимальных многомерных кодов в гауссовом канале связи и их основные характеристики. Электросвязь № 2, 2016.

 Быховский М.А. Помехоустойчивость приема оптимальных сигналов, расположенных на поверхности N-мер-ного шара.
 Электросвязь № 3, 2016.

 Прокис Дж. Цифровая связь.//Перевод с английского под ред. Д.Д. Кловского.//М.: Советское радио, 2000.

 Дж. Кларк мл., Дж. Кейн. Кодирование с исправлением ошибок в системах цифровой связи. М.: Радио и связь, 1987

8. Ungerboeck G. Trellis-Coded Modulation with Redundant Signal Sets. PartlandII, IEEECommun.Mag., vol. 25, № 2, 1987.

9. Forney G.D., Gallager R.G., Lang G.R., Longstaff F.M., Qureshi S.U. Efficient Modulation for Bandlimited Channels. IEEE J. Se1ectd Areas in Commun. vol. SAC-2, n. 5, № 9, 1984.

10. Forney G.D., Wei L.F. Multidimensional constellations-Part I: Introduction, figures of merit and generalized cross constellations. IEEE I. Select. Areas Commun., vol. 7, № 8, 1989.

11. Forney G.D. Multidimensional constellations-Part II: Voronoi constellations. IEEE I. Select. Areas Commun., vol. 7, № 8, 1989.

12. Бураченко Д.Л., Савищенко Н.В. Геометрические модели сигнально-кодовых конструкций. СПб.: Военная академия связи, 2012.

13. Галлагер Р. Коды с малой плотностью проверок на четность. – М.: Мир, 1966

14. N-мерная сфера. https://en.wikipedia.org/wiki/N-sphere.

15. Альсведе Р., Вегенер И. Задачи поиска. М.: Мир, 1982.

УДК 004.032.26

РЕКУРРЕНТНЫЕ НЕЙРОННЫЕ СЕТИ В КАЧЕСТВЕ МОДЕЛЕЙ НЕЛИНЕЙНЫХ ДИНАМИЧЕСКИХ СИСТЕМ

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@hotbox.ru.

RECURRENT NEURAL NETWORKS AS BEHAVIORAL MODELS OF NONLINEAR DYNAMIC SYSTEMS

Solovyeva E. B.

The classification of recurrent neural networks used as the mathematical models of nonlinear dynamic systems described by the input-output ratio is proposed. The structures, properties, advantages and disadvantages of different types of recurrent networks are considered. The presented analysis is useful for choosing the mathematical model of a nonlinear dynamical system a priori, when it is necessary to evaluate which of the known neural network structures meets the requirements for model characteristics, such as accuracy, computational complexity, robustness, hardware implementation, more than others do.

Key words: neural networks, recurrent, models of nonlinear dynamical systems, classification of neural networks, model characteristics.

Ключевые слова: нейронная сеть, рекуррентная, модель динамической системы, классификация сетей, характеристики модели.

Введение

В последние десятилетия нейронные сети находят применение во многих областях техники, где они используются для решения многочисленных задач, включая идентификацию, моделирование и синтез нелинейных динамических систем (НДС). Указанные задачи можно

сформулировать в аппроксимационной постановке, когда требуется построить математическую модель нелинейного оператора, однозначно отображающего множество входных сигналов во множество выходных сигналов динамической системы [1-14]. Нейронные сети привлекательны в ситуации, когда полиномиальные модели нелинейного оператора с ростом степени дают медленно уменьшающуюся погрешность аппроксимации оператора.

Рекуррентные нейронные сети (recurrent neural network, RNN) [1], [4], [7-14] считаются развитием однонаправленных персептронных сетей (или сетей прямого распространения, feed-forward neural networks) путем добавления обратных связей. В каждом контуре (петле) обратной связи присутствует элемент единичной задержки, благодаря которому поток сигналов может считаться однонаправленным (выходной сигнал предыдущего временного цикла рассматривается как априори заданный, который просто увеличивает размерность входного вектора). С помощью обратной связи можно накапливать информацию и использовать ее при обработке сигналов. Таким образом, в сети реализуется «память», что принципиально меняет характер ее работы и позволяет анализировать любые последовательности данных, в которых важно, в каком порядке идут значения (например, речь, текст, изображение и т.д.). В итоге, порядок следования сигналов играет существенную роль в задаче.

Предложена классификация рекуррентных нейронных сетей, применяемых в качестве математических моделей нелинейных динамических систем, описываемых по соотношению вход-выход. Рассмотрены структуры, свойства, достоинства и недостатки разных видов рекуррентных сетей. Представленный анализ полезен для выбора математической модели нелинейной динамической системы априори, когда следует оценить, какая из известных структур нейронных сетей более других отвечает требованиям, предъявляемым к характеристикам модели, таким как точность, вычислительная сложность, робастность, аппаратная реализация.

Рекуррентная нейронная сеть может содержать меньшее число параметров по сравнению с многослойной персептронной сетью, выполняющей ту же задачу. Однако алгоритм обучения RNN, адаптирующий значения синаптических весов, более сложен вследствие зависимости сигналов в текущий момент времени от их значений в предыдущие моменты и соответственно ввиду более громоздкой формулы для расчета вектора градиента. Кроме того, из-за наличия обратных связей необходимо исследовать вопрос устойчивости сети [1], [4], [7-14].

Несмотря на указанные трудности, за последние годы достигнут значительный успех в применении рекуррентных нейронных сетей для решения задач моделирования, синтеза, распознавания, предсказания, управления и т.д. [3].

В зависимости от расположения обратных связей, влияющих на организацию взаимодействия нейронов, можно выделить два класса RNN [15]:

 – глобально рекуррентные сети, в которых допускаются обратные связи между нейронами одного слоя или разных слоев;

 – локально рекуррентные сети, содержащие обратные связи внутри нейронов. Обратные связи между нейронами одного слоя или разных слоев отсутствуют.

Рассмотрим структуры, модели и свойства сетей, образующих выделенные классы.

Глобально рекуррентные сети

К глобально рекуррентным сетям относятся четыре вида сетей:

 – полносвязная рекуррентная сеть, называемая также рекуррентной сетью реального времени (real time recurrent network, RTRN) [4], [8];

– частично рекуррентные сети, к которым относятся, например, сеть Элмана (Elman network), сеть Джордана (Jordan network), многослойная рекуррентная персептронная сеть (recurrent multi-layer perceptron network, RMLPN) [1], [4], [8-14];

- сети в пространстве состояний [1], [4], [8];

– клеточные нейронные сети (cellular neural networks, CNN) [10], [16-21].

Полносвязная рекуррентная сеть

Наиболее общая структура глобальной рекуррентной нейронной сети, в которой присутствуют любые связи между нейронами, предложена Уильямсом и Зипсером. Такая полносвязная структура часто называется сетью Уильямса–Зипсера (Williams-Zipser network) или рекуррентной сетью реального времени (RTRN), поскольку она предназначена для обработки сигналов в реальном времени [4], [8].

Структура RTRN показана на рис. 1. Обратная связь RTRN включает задержанные выходные сигналы сети.



Рис. 1. Структура RTRN

RTRN описывается следующими уравнениями: $\mathbf{X}_{i}(n) = [u_{1}(n), u_{2}(n), ...,$

$$\begin{split} & u_p(n), 1, y_1(n-1), y_2(n-1), ..., y_N(n-1)]^t = \\ & = [x_1(n), x_2(n), ..., x_{p+N+1}(n)]^t , \\ & v_i(n) = \sum_{l=1}^{p+N+1} w_{i,l} x_l(n) = \mathbf{W}_i^t \mathbf{X}_i(n) , \ i = 1, 2, ..., N , \end{split}$$

 $y_i(n) = F(v_i(n)), i = 1, 2, ..., M,$

где $\mathbf{X}_i(n)$ – вектор-столбец, содержащий для *i*-го нейрона в момент времени *n* входные сигналы, сигнал смещения с единичным значением, сигналы с выводов обратной связи; $\mathbf{W}_i^t = \begin{bmatrix} w_{i,1}, w_{i,2}, ..., w_{i,p+N+1} \end{bmatrix}$ – вектор синалтических весов *i*-го нейрона; *t* – знак транспонирования вектора; *F* – функция активации; *p*, *N*, *M* – количество внешних входов, выводов обратной связи и выходных сигналов сети соответственно.

Сеть, изображенная на рис. 1, содержит M выходов и (N-M) скрытых нейронов.

RTRN может моделировать широкий класс динамических систем [4], [8]. Однако, структура RTRN оказывается слишком сложной для практических применений. Действительно, наличие непосредственной связи между числом нейронов и числом выводов обратной связи не позволяет раздельно регулировать порядок динамики и нелинейные свойства модели. Кроме того, обучение сети сложное, медленно сходящееся к решению, сопровождаемое необходимостью поддержания устойчивости сети [4], [8].

Частично рекуррентные сети

Архитектура частично рекуррентных сетей основана на многослойном персептроне, включающем дополнительный слой, называемый контекстным слоем. Сигналы этого слоя описывают внутреннее состояние сети.

Обратные связи частично рекуррентных сетей более структурированы по сравнению с полносвязными сетями, поэтому они быстрее обучаются и имеют меньшее число проблем потери устойчивости. Тем не менее, число состояний сети по-прежнему зависит от количества скрытых или выходных нейронов, что существенно ограничивает гибкость сети.

Среди частично рекуррентных сетей наиболее популярны три структуры [1], [4], [8], [9]:

- сеть Элмана (Elman network);
- сеть Джордана (Jordan network);

– рекуррентная многослойная персептронная сеть (recurrent multi-layer perceptron network, RMLPN).

Рассмотрим упомянутые структуры и соответствующие им математические модели.

Сеть Элмана

Сеть Элмана, вероятно, является самым известным примером частично рекуррентной нейронной сети. Она часто применяется в системах управления движущимися объектами для обнаружения изменений характеристик сигналов [4], [8].

Структура сети Элмана показана на рис. 2. В структуре присутствуют частичная рекуррентность в виде обратной связи в скрытом слое и однонаправленная связь между скрытым и выходным слоями.

Сеть Элмана описывается системой рекуррентных уравнений:

$$v_j(n+1) = \sum_{i=1}^p w_{ji}^{(1)} u_i(n) + \sum_{i=1}^N w_{ji}^c x_i(n) + b_j^{(1)},$$

$$x_{j}(n+1) = F_{1}(v_{j}(n+1)), \quad j = 1, 2, ..., N,$$
$$y_{j}(n+1) = F_{2}\left(\sum_{i=1}^{N} w_{ji}^{(2)} x_{i}(n+1) + b_{j}^{(2)}\right), \quad j = 1, 2, ..., M,$$

или уравнениями в матричной форме:

 $\mathbf{X}(n+1) = \mathbf{F}_{1}(\mathbf{W}^{(1)}\mathbf{U}(n) + \mathbf{W}^{c}\mathbf{X}(n) + \mathbf{B}^{(1)}),$ $\mathbf{Y}(n+1) = \mathbf{F}_{2}(\mathbf{W}^{(2)}\mathbf{X}(n+1) + \mathbf{B}^{(2)}),$

где $\mathbf{U}(n)$ – вектор внешних входных сигналов в момент времени n; p – число внешних входов сети; $\mathbf{X}(n)$ – вектор сигналов контекстного слоя в момент времени n; N – число сигналов в контекстном слое; $\mathbf{W}^{(1)}$, \mathbf{W}^{C} , $\mathbf{W}^{(2)}$ – матрицы синаптических весов внешних входных сигналов, сигналов контекстного и выходного слоя соответственно; $\mathbf{B}^{(1)}$, $\mathbf{B}^{(2)}$ – векторы весов смещений в нейронах скрытого и выходного слоя соответственно; \mathbf{F}_1 , \mathbf{F}_2 – векторы функций активации в скрытом и выходном слое соответственно; $\mathbf{Y}(n+1)$ – вектор выходных сигналов сети в момент времени (n+1); M – количество выходов сети.



Рис. 2. Структура сети Элмана

Сеть Джордана

Сеть Джордана решает тот же класс задач, что и сеть Элмана, однако обладает лучшими аппроксимирующими и предсказывающими свойствами благодаря более «глубокой» памяти и дополнительному слою нелинейных функций активации. Сеть Джордана успешно применяется для распознавания, различения и классификации временных последовательностей [4], [8].

Структура сети Джордана представлена на рис. 3. Сеть Джордана – это многослойная персептронная сеть с одним скрытым слоем и контурами обратной связи между выходами и дополнительными входами сети, сформированными в контекстном слое. Контекстный слой содержит также повторные контуры обратной связи с коэффициентом ослабления памяти α , $0 \le \alpha \le 1$. В итоге, на выходе контекстного слоя присутствуют как текущие, так и взвешенные предыдущие сигналы обратной связи. Описанное свойство обеспечивает более «глубокую» память сети Джордана по сравнению с сетью Элмана.



Рис. 3. Структура сети Джордана

Сеть Джордана описывается системой рекуррентных уравнений:

$$\begin{split} v_{j}(n+1) &= \sum_{i=1}^{p} w_{ji}^{(1)} u_{i}(n) + \\ &+ \sum_{i=1}^{N} w_{ji}^{c} \left(y_{i}(n) + \alpha y_{i}(n-1) \right) + b_{j}^{(1)} \\ x_{j}(n+1) &= F_{1}(v_{j}(n+1)) , \ j = 1, 2, ..., N , \\ y_{j}(n+1) &= F_{2} \left(\sum_{i=1}^{N} w_{ji}^{(2)} x_{i}(n+1) + b_{j}^{(2)} \right), \ j = 1, 2, ..., M , \end{split}$$

или уравнениями в матричной форме:

$$\mathbf{X}(n+1) = \mathbf{F}_{1}(\mathbf{W}^{(1)}\mathbf{U}(n) + \mathbf{W}^{c}(\mathbf{Y}(n) - \alpha\mathbf{Y}(n-1)) + \mathbf{B}^{(1)})'$$

$$\mathbf{Y}(n+1) = \mathbf{F}_{2}(\mathbf{W}^{(2)}\mathbf{X}(n+1) + \mathbf{B}^{(2)}),$$

где U(n) – вектор внешних входных сигналов в момент времени *n*; *p* – число внешних входов сети; X(n+1) – вектор выходных сигналов скрытого слоя в момент времени (*n*+1); *N* – число сигналов в контекстном слое;

 $\mathbf{W}^{(1)}$, \mathbf{W}^{C} , $\mathbf{W}^{(2)}$ – матрицы синаптических весов внешних входных сигналов, сигналов контекстного и выходного слоя соответственно; $\mathbf{B}^{(1)}$, $\mathbf{B}^{(2)}$ – векторы весов смещений в нейронах скрытого и выходного слоя соот-

ветственно; \mathbf{F}_1 , \mathbf{F}_2 – векторы функций активации в скрытом и выходном слое соответственно; $\mathbf{Y}(n+1)$ – вектор выходных сигналов сети в момент времени (n+1); M – количество выходов сети.

Рекуррентная многослойная персептронная сеть

Рекуррентная многослойная персептронная сеть (RMLPN) построена на основе MLPN (multi-layer perceptron network) путем добавления обратных связей между соседними нейронами внутри скрытых слоев, а также обратных связей нейронов на себя. Нелинейные свойства сети формируются в прямых связях, а динамические свойства – в обратных связях.

Сеть RMLPN успешно применяется для идентификации динамических систем [1], [4], [8], [9].

Блок-схема RMLPN с двумя скрытыми слоями показана на рис. 4.



Рис. 4. Блок-схема RMLPN с двумя скрытыми слоями Модель RMLPN имеет вид системы уравнений:

 $\mathbf{NET}^{(1)}(n+1) = \mathbf{F}_1(\mathbf{W}_a^{(1)}\mathbf{NET}^{(1)}(n), \mathbf{W}_b^{(1)}\mathbf{U}(n)),$

 $NET^{(2)}(n+1) =$

=
$$\mathbf{F}_2(\mathbf{W}_a^{(2)}\mathbf{NET}^{(2)}(n), \mathbf{W}_b^{(2)}\mathbf{NET}^{(1)}(n+1)),$$

 $y(n+1) = \mathbf{F}_3(w_a^{(3)}y(n), \mathbf{W}_b^{(3)}\mathbf{NET}^{(2)}(n+1)),$

где U(n) – вектор входных сигналов в момент времени *n*; **NET**^(*i*)(*n*) – вектор выходных сигналов *i* -го скрытого слоя в момент времени n; $W_a^{(i)}$, *i* = 1, 2 – матрица синаптических весов нейронов, соединенных с узлами обратной связи в *i*-м скрытом слое; $W_b^{(i)}$ – матрица синаптических весов нейронов, соединенных с узлами прямой связи в *i*-м скрытом (*i* = 1, 2) или в выходном (*i* = 3) слое сети; **F**₁, **F**₂, **F**₃ – векторы функций активации двух скрытых слоев и выходного слоя соответственно; *y*(*n*+1) – выходной сигнал сети в момент времени (*n*+1).

Недостатками RMLPN является повышенная сложность, зависящая от количества скрытых нейронов, и высокая длительность процесса обучения сети.

Рекуррентные сети в пространстве состояний

Блок-схема рекуррентной сети в пространстве состояний показана на рис. 5. Выходные сигналы скрытого слоя (сигналы пространства состояний) проходят через блок единичных задержек и поступают на вход сети. Размерность пространства состояний задает порядок сети. Пользователь может выбрать количество нейронов скрытого слоя, формирующих сигналы для контура обратной связи.



Рис. 5. Блок-схема рекуррентной сети в пространстве состояний

Математическая модель сети, блок-схема которой представлена на рис. 5, имеет вид:

 $\mathbf{X}(n+1) = \mathbf{F}(\mathbf{W}_{a}\mathbf{X}(n), \mathbf{W}_{b}\mathbf{U}(n)),$

$$\mathbf{Y}(n+1) = \mathbf{C}\mathbf{X}(n+1)$$

где X(n), U(n), Y(n) – векторы сигналов соответственно скрытого слоя, а также на входе и выходе сети в дискретный момент времени n; W_a , W_b – матрицы синаптических весов нейронов скрытого слоя, которые соединены с узлами соответственно обратной связи и источника входного слоя (предполагается, что внешнее смещение, применяемое к скрытым нейронам, учтено в матрице весов W_b); F – вектор нелинейных функций активации в скрытом слое; С – матрица синаптических весов линейных нейронов выходного слоя, в которой учитывается внешнее смещение, применяемое к выходным нейронам.

Отметим, что сходную с показанной на рис. 5 структуру имеет сеть Элмана за исключением того факта, что выходной слой сети Элмана в общем случае нелинейный.

Математическими моделями рекуррентных сетей в пространстве состояний являются системы нелинейных разностных (с дискретным временем) или дифференциальных (с непрерывным временем) уравнений состояния. Среди моделей данного типа можно выделить модели Хопфилда и Гроссберга, широко применяемые на практике [1], [4], [8].

Рекуррентные сети в пространстве состояний обладают рядом преимуществ [1], [8]:

 – модели данных сетей описывают широкий класс нелинейных динамических систем;

– число состояний (порядок) сети не равно числу скрытых нейронов. Только те скрытые нейроны, выходные сигналы которых поступают в обратную связь, проходят через элементы задержки и приходят на вход, определяют состояния сети. Нейроны выходного слоя не влияют на пространство состояний;

– поскольку сигналы состояния поступают через обратную связь на вход сети, эти сигналы доступны для измерения или определения. Данное свойство полезно для нахождения начальных условий сети в разные моменты времени.

Несмотря на то, что рассматриваемые сети кажутся более перспективными по сравнению с полносвязными и частично рекуррентными сетями, их применение сталкивается со следующими трудностями [1], [8]:

 построение математической модели в пространстве состояний возможно на компактных подмножествах сигналов и на конечных временных интервалах, в результате динамические характеристики сети не всегда описываются в полной мере;

 неточные начальные условия сети замедляют ее обучение, особенно при коротких последовательностях входных и выходных сигналов;

 возможно появление неустойчивости сети в процессе ее обучения.

Отмеченные недостатки ярко проявляются в тех случаях, когда начальные условия не известны и их невозможно оценить.

Клеточные нейронные сети

Клеточная нейронная сеть (cellular neural network, CNN) – это объединение локально связанных клеток (нейронов), где каждая клетка – нелинейная динамическая система со входом, выходом и некоторым состоянием, описываемым динамическим уравнением. Каждая клетка взаимодействует непосредственно с соседними клетками. Взаимодействие клеток, не являющихся соседними, возникает из-за эффекта распространения динамики в сети [10], [16-21].

CNN с локально связанными клетками производит параллельную обработку сигналов. Данное свойство отличает CNN от других сетей и является преимуществом, позволяющим чрезвычайно быстро выполнять сложные преобразования одномерных и многомерных сигналов в телевидении, робототехнике, дистанционном управлении, нейрофизиологии. CNN широко применяются для линейной и нелинейной фильтраций, детектирования различных свойств изображений (очертаний, цвета, глубины, движения и т.д.), генерирования нелинейных колебаний и хаотических процессов, распознавания образов и решения нелинейных частных дифференциальных уравнений. На основе аналоговых и дискретных сетей разработана CNN-технология (высоко интегрированные чипы) и создана сотовая архитектура компьютера [10], [16-21].

Клеточные нейронные сети содержат обратные связи, поэтому они являются рекуррентными сетями. Обратные связи могут приводить к возникновению неустойчивости, которая проявляется в блуждающей смене состояний нейронов, не приводящей к стационарным состояниям. В общем случае ответ на вопрос об устойчивости динамики произвольной системы с обратными связями крайне сложен и до настоящего времени является открытым.

СNN-модели разнообразны. Они различаются параметрическим и динамическим описаниями клеток, а также топологиями сетей [10], [16-21].

Рассмотрим стандартную двумерную (2-dimensional, 2-D) CNN в непрерывном времени. Клетку, расположенную на позиции (i, j) в двумерной области с размером $M \times N$, обозначим C_{ij} . Окрестность клетки N_{ij}^r , где r

 размер окрестности (положительное целое число), формируется согласно выражению

$$N_{ij}^{r} = \{C_{kl} \text{ при max}(|k-i|, |l-j|) \le r,$$

 $1 \le k \le M; 1 \le l \le N \}.$

Например, при r = 1 создается окрестность размером 3×3 , показанная на рис. 6, при r = 2 – окрестность размером 5×5 .

Нелинейный динамический процесс в клетке C_{ij} описывается двумя уравнениями (дифференциальным уравнением 1-го порядка и нелинейным уравнением на выходе клетки):

$$\frac{dx_{ij}(t)}{dt} = -\gamma x_{ij}(t) +
+ \sum_{kl \in N_{ij}^r} \mathbf{A}_{ij,kl} y_{kl}(t) + \sum_{kl \in N_{ij}^r} \mathbf{B}_{ij,kl} u_{kl}(t) + I_{ij}.$$

$$y_{ij}(t) = F(x_{ij}(t)),$$
(2)

где $x_{ij}(t), u_{ij}(t), y_{ij}(t)$ – переменная состояния, входной и выходной сигналы клетки соответственно; γ – константа; I_{ij} – постоянная смещения клетки; $\mathbf{A}_{ij,kl}$ – матрица трафарета обратной связи; $\mathbf{B}_{ij,kl}$ – матрица трафарета прямой связи (входного трафарета или трафарета управления); F – нелинейная функция.



Рис. 6. Клетка C_{ij} и ее окрестность размером 3x3 Структура клетки с моделью (2), изображена на рис. 7, где \int – операция интегрирования.



Рис. 7. Структура клетки CNN в непрерывном времени

Во многих приложениях CNN изотропна, т.е. пространственно инвариантна [10], [16-21]. Изотропная сеть

управляется повторяющимся трафаретом (*A*, *B*, *I*), содержащим фиксированные для всей нейронной сети матрицы $\mathbf{A}_{ij,kl}$, $\mathbf{B}_{ij,kl}$ размером $(2r+1) \times (2r+1)$ и смещение I_{ij} . В случае изотропной CNN слагаемые в модели (2) представляются как

$$\sum_{kl \in N_{ij}^{r}} \mathbf{A}_{ij,kl} y_{kl}(t) = \sum_{k=-r}^{r} \sum_{l=-r}^{r} a_{i-k,j-l} y_{kl}(t),$$
$$\sum_{kl \in N_{ij}^{r}} \mathbf{B}_{ij,kl} u_{kl}(t) = \sum_{k=-r}^{r} \sum_{l=-r}^{r} b_{i-k,j-l} u_{kl}(t),$$
$$I_{ij} = I,$$

где $a_{i-k,j-l}$, $b_{i-k,j-l}$ – элементы матриц $\mathbf{A}_{ij,kl}$, $\mathbf{B}_{ij,kl}$ соответственно.



Рис. 8. Кусочно-линейная функция с насыщением

Функция *F* в выражении (2) часто задается в виде [10], [16-21]:

 – кусочно-линейной функции с насыщением, изображенной на рис. 8 и описываемой уравнением

$$y_{ij}(t) = F(x_{ij}(t)) = \frac{1}{2} \left(\left| x_{ij}(t) + 1 \right| - \left| x_{ij}(t) - 1 \right| \right);$$

- знаковой функции

$$y_{ij}(t) = F\left(x_{ij}(t)\right) = \operatorname{sgn}\left(x_{ij}(t)\right) = \begin{cases} +1 & \text{для } x_{ij}(t) \ge 0, \\ -1 & \text{для } x_{ij}(t) < 0, \end{cases}$$

- ступенчатой функции;

 – функции из класса сигмоидальных, обладающих следующими свойствами:

$$\left| F(x_{ij}(t)) \right| \le \text{Const},$$
$$\frac{dF(x_{ij}(t))}{dx_{ii}(t)} \ge 0.$$

CNN преобразовывает сигналы при ограничении, действующем на начальное состояние клетки: $|x_{ii}(0)| \le 1$,

и при ограничении, действующем на входной сигнал: $|u_{ii}(t)| \leq 1$.

Описание CNN в дискретной временной области формируется из выражения (2) в результате следующих преобразований:

- аппроксимации производной

$$\frac{dx_{ij}(t)}{dt} \approx \frac{x_{ij}(t) - x_{ij}(t - \Delta t)}{\Delta t} = x_{ij}(n) - x_{ij}(n-1),$$

где *n* – нормированное дискретное время, пусть $\Delta t = 1$;

 перехода от дифференциального уравнения к рекурсивному разностному уравнению

$$x_{ij}(n) - x_{ij}(n-1) = -\gamma x_{ij}(n-1) + \sum_{kl \in N_{ij}^r} \mathbf{A}_{ij,kl} y_{kl}(n-1) + \sum_{kl \in N_{ij}^r} \mathbf{B}_{ij,kl} u_{kl}(n-1) + I_{ij}.$$
(3)

В итоге, при $\gamma = 1$ на основе (2) и (3) записывается модель клетки C_{ij} дискретной CNN (discrete-time CNN, DTCNN):

$$\begin{aligned} x_{ij}(n) &= \sum_{kl \in N_{ij}^{r}} \mathbf{A}_{ij,kl} y_{kl}(n-1) + \\ &+ \sum_{kl \in N_{ij}^{r}} \mathbf{B}_{ij,kl} u_{kl}(n-1) + I_{ij}, \\ y_{ii}(n) &= F(x_{ii}(n)) . \end{aligned}$$
(4)

Модель (4) соответствует структуре клетки DTCNN, изображенной на рис. 9, где z^{-1} – элемент задержки.



Рис. 9. Структура клетки DTCNN

При независимом от времени воздействии из выражения (4) формируется модель клетки изотропной DTCNN:

$$x_{ij}(n) = \sum_{kl \in N_{ij}^r} \mathbf{A}_{ij,kl} y_{kl}(n-1) + \sum_{kl \in N_{ij}^r} \mathbf{B}_{ij,kl} u_{kl} + I,$$

 $y_{ij}(n) = F(x_{ij}(n)) .$

Локально рекуррентные сети

Локально рекуррентные сети содержат обратные связи внутри нейронов, при этом отсутствуют обратные связи между нейронами одного слоя и разных слоев.

Среди локально рекуррентных сетей выделяются два вида структур:

 – сети со статическими прямыми связями и, так называемыми, динамическими нейронами [4];

– блочно-ориентированные нейронные сети Винера, Гаммерштейна, Винера–Гаммерштейна и т.д. [12].

Рассмотрим упомянутые структуры и соответствующие им нелинейные модели.

Сети с динамическими нейронами

В локальных рекуррентных сетях нейроны, описываемые динамическими моделями, соединены статическими связями. Различия между динамическими моделями нейронов обусловлены локализацией обратных связей.

Динамическая модель нейрона с обратной связью между выходом и входом

Структура нейрона с локальной обратной связью между выходом и входом, предложенная Гори, показана на рис. 10 [4]. В данной структуре выходной сигнал нейрона обрабатывается фильтром с конечной импульсной характеристикой (КИХ-фильтром), выходные сигналы фильтра суммируются со взвешенными внешними сигналами, поступающими на нейрон, и суммарный сигнал преобразуется в блоке активации.



Рис. 10. Структура нейрона с обратной связью между выходом и входом

Рассматриваемый динамический нейрон описывается моделью:

$$x(n) = \sum_{i=1}^{m} w_i u_i(n) + \sum_{i=1}^{r} d_i y(n-i) ,$$

$$y(n) = F(x(n)) ,$$

где $u_i(n)$, i = 1, 2, ..., m – входные сигналы нейрона; w_i , i = 1, 2, ..., m – синаптические веса входных сигналов; d_i , i = 1, 2, ..., r – синаптические веса сигналов обратной связи; x(n) – сигнал активации; F – нелинейная функция активации; y(n) – выходной сигнал нейрона.

На основе динамических нейронов строится сеть Хопфилда обычно в виде структуры, изображенной на рис. 11 [1], [4], [8], [22]. В структуре задержанные в контурах обратной связи выходные сигналы нейронов собираются на узлах входа сети и направляются к нейронам. Каждый нейрон суммирует взвешенные сигналы, обрабатывает суммарный сигнал $x_i(n+1)$ функцией активации F и в результате формирует выходной сигнал $y_i(n+1)$.

Дискретная модель для сети Хопфилда (рис. 11) имеет вид:

$$x_i(n+1) = \sum_{j=1}^r d_{ij} y_j(n) + b_i,$$

$$y_i(n+1) = F(x_i(n+1)),$$

где $x_i(n+1)$ – сигнал активации *i*-го нейрона в момент времени (n+1); d_{ij} , j = 1, 2, ..., r – синаптические веса входных сигналов *i*-го нейрона; b_i – вес смещения *i*-го нейрона; F – функция активации; $y_i(n+1)$ – выходной сигнал *i*-го нейрона в момент времени (n+1).



Рис. 11. Структура сети Хопфилда

Соответствующая непрерывная модель Хопфилда представляется в виде:

$$\frac{dx_i(t)}{dt} = \sum_{j=1}^r d_{ij} y_j(t) + b_i$$

$$y_i(t) = F(x_i(t)),$$

где *t* – переменная непрерывного времени.

Модель Хопфилда – это система нелинейных разностных или обыкновенных дифференциальных уравнений, которую можно использовать для описания нелинейных динамических устройств в пространстве состояний.

Сеть Хопфилда применяется [1], [4], [8] в качестве

– интерфейса между аналоговыми и цифровыми устройствами, где входные сигналы сети являются аналоговыми, а выходные сигналы – дискретными. Нейронный интерфейс обладает способностью к обучению, поэтому нейронный аналого-цифровой преобразователь адаптируется, чтобы компенсировать начальные несоответствия устройств или долгосрочные дрейфы;

- ассоциативной памяти;

 – аппарата для решения задач комбинаторной оптимизации.

Физичность топологии аналоговой сети Хопфилда обуславливает возможность ее аппаратной реализации на СБИС [22].

Динамическая модель нейрона с обратной связью между активацией и входом

Структура нейрона с локальной обратной связью, между блоком активации и входом, изображена на рис. 12. На вход нейрона поступают внешние сигналы и состояния сигнала активации из обратной связи. Сумма взвешенных входных сигналов преобразуется функцией активации в выходной сигнал нейрона.



Рис. 12. Структура нейрона с обратной связью между активацией и входом

Модель рассматриваемого нейрона была исследована Фраскони и описана следующими уравнениями [4]:

$$x(n) = \sum_{i=1}^{m} w_i u_i(n) + \sum_{i=1}^{r} d_i x(n-i),$$

$$y(n) = F(x(n)),$$
(5)

где $u_i(n)$, i = 1, 2, ..., m – внешние входные сигналы; w_i , i = 1, 2, ..., m – синаптические веса внешних входных сигналов; d_i , i = 1, 2, ..., r – синаптические веса сигналов обратной связи; x(n) – сигнал активации; F – нелинейная функция активации; y(n) – выходной сигнал нейрона.

Суммирование во втором слагаемом правой части (5) можно интерпретировать как обработку сигнала активации фильтром с конечной импульсной характеристикой (КИХ-фильтром).

Динамическая модель нейрона с двумя контурами обратной связи

На основе теории цифровых фильтров строятся более сложные структуры динамического нейрона, один вариант из которых показан на рис. 13. В данной структуре сигнал активации и выходной сигнал нейрона обрабатываются разными КИХ-фильтрами в соответствующих контурах обратной связи.

Нелинейная модель рассматриваемого нейрона имеет вид:

$$x(n) = \sum_{i=1}^{m} w_i u_i(n) + \sum_{i=1}^{r} d_i x(n-i) + \sum_{i=1}^{k} a_i y(n-i) ,$$

$$y(n) = F(x(n)) ,$$

где $u_i(n)$, i = 1, 2, ..., m – внешние входные сигналы

нейрона; w_i , i = 1, 2, ..., m – синаптические веса внешних воздействий; x(n) – сигнал активации; d_i , i = 1, 2, ..., r – синаптические веса состояний сигнала активации в обратной связи; a_i , i = 1, 2, ..., k – синаптические веса состояний выходного сигнала в обратной связи; F – нелинейная функция активации; y(n) – выходной сигнал нейрона.



Рис. 13. Структура нейрона с двумя контурами обратной связи

Блочно-ориентированные сети

Выходной сигнал блочно-ориентированной сети является комбинацией выходных сигналов блоков (модулей) сети. Идентификация структуры, оценка параметров и проверка модели нейронной сети с применением подмоделей отдельных блоков часто более удобны, чем аналогичные операции, выполняемые с полной моделью сети.

Блочно-ориентированная сеть состоит из нелинейных статических и линейных динамических модулей. К простым соединениям таких модулей относятся [12]:

 Структура Винера, изображенная на рис. 14, а, в которой за линейным динамическим модулем следует безынерционная (с нулевой памятью) нелинейность.

Операторная модель Винера записывается в виде

$$y(n) = f(H(q^{-1})[u(n)])$$

где u(n), y(n) – входной и выходной сигналы соответственно; $H(q^{-1})$ – линейный оператор рациональной формы, действующий на сигнал u(n); q^{-1} – оператор временной задержки на один такт; f – нелинейная функция активации.

Линейный оператор описывается формулой

$$H(q^{-1}) = \frac{B(q^{-1})}{A(q^{-1})},$$
(6)

где $A(q^{-1})$, $B(q^{-1})$ – линейные операторы, представленные соответственно в выражениях:

$$A(q^{-1}) = 1 + a_1 q^{-1} + \dots + a_{N_a} q^{-N_a} , \qquad (7)$$

$$B(q^{-1}) = b_1 q^{-1} + \dots + b_{N_b} q^{-N_b} .$$
(8)

 Структура Гаммерштейна, показанная на рис. 14, б, в которой за безынерционной нелинейностью следует линейный динамический модуль.

Операторная модель Гаммерштейна имеет вид

 $y(n) = H(q^{-1}) \left[f\left(u(n)\right) \right].$

Согласно данной модели линейный оператор $H(q^{-1})$ действует на сигнал, являющийся результатом обработки воздействия u(n) нелинейной функцией f.



Рис. 14. Простые блочно-ориентированные структуры:

а – структура Винера, б – структура Гаммерштейна

Структуры Винера и Гаммерштейна объединяются для создания более сложных соединений [12]. Рассмотрим некоторые из таких соединений.

Структура Винера–Гаммерштейна, представленная на рис. 15, а, включает в себя структуру Винера (модули $H_1(q^{-1})$ и f) и структуру Гаммерштейна (модули R и $H_2(q^{-1})$).



Рис. 15. Архитектуры сетей на основе структур Винера и Гаммерштейна: а – структура Винера–Гаммерштейна, б – обобщенная структура Гаммерштейна, в – динамический персептрон

Модель Винера–Гаммерштейна записывается в виде операторного уравнения

$$y(n) = H_2(q^{-1}) \Big[f(H_1(q^{-1})[u(n)]) \Big].$$

Обобщенная структура Гаммерштейна, изображенная на рис. 15, б, описывается операторной моделью:

$$v(n) = \sum_{i=1}^{N} H_i(q^{-1})[u^i(n)].$$

Динамический персептрон, представленный на рис. 15, в, имеет операторную модель:

$$y(n) = f\left(H(q^{-1})\left[\sum_{i=1}^{m} w_i u_i(n)\right]\right),$$

где $H(q^{-1})$ – линейный оператор из выражения (6).

Операторной модели динамического персептрона соответствует система уравнений:

$$\begin{aligned} x(n) &= \sum_{i=1}^{m} w_{i} u_{i}(n) ,\\ v(n) &= \sum_{i=1}^{N_{b}} b_{i} x(n-i) - \sum_{j=1}^{N_{a}} a_{j} v(n-j) \\ y(n) &= f(v(n)) , \end{aligned}$$

где a_i и b_i – параметры линейных операторов $A(q^{-1})$ и $B(q^{-1})$ из выражений (7) и (8) соответственно.

Заключение

Математическое моделирование нелинейных динамических систем рассмотрено как задача описания соотношения вход/выход систем с применением нейронных сетей, в частности, рекуррентных сетей. В зависимости от расположения обратных связей сети разделены на глобально и локально рекуррентные. Внутри указанных классов выделены виды рекуррентных сетей, для которых рассмотрены аппроксимирующие свойства, достоинства и недостатки, области применения.

Приведенный анализ полезен для выбора структуры нейронной сети априори (до ее обучения, т.е. до построения математической модели нелинейной системы). Данный аспект важен, поскольку задача моделирования нелинейных систем в общей постановке сложна, поэтому целесообразно априори выбрать конструктивную модель, обеспечивающую высокую точность аппроксимации соответствия вход/выход нелинейной динамической системы.

Отметим, что развитие RNN связано не только с моделированием нелинейных динамических сетей, но и с моделированием последовательностей, в которых соседние точки зависят друг от друга, и эту связь нельзя игнорировать. Для преобразования последовательностей с внутренними зависимостями разработаны специальные структуры RNN и методы их обучения, применяемые в задачах распознавания, классификации, предсказания и т.д. Указанному направлению развития RNN посвящена литература из области глубокого обучения нейронных сетей [23], [24].

Литература

1. Хайкин С. Нейронные сети: полный курс. – М.: Издательский дом «Вильямс», 2016.

2. Janczak A. Identification of nonlinear systems using neural networks and polynomial models. A Block-Oriented Approach. – Berlin: Springer-Verlag Berlin Heidelberg, 2005.

3. Speech, audio, image and biomedical signal processing using neural networks / Ed.: B. Prasad, S. R. Mahadeva Prasanna. – Berlin: Springer-Verlag Berlin Heidelberg, 2008.

4. Patan K. Artificial neural networks for the modelling and fault diagnosis of technical processes. – Berlin: Springer-Verlag Berlin Heidelberg, 2008.

5. Tang H., Tan K. C., Yi Z. Neural networks: computational models and applications. – Berlin: Springer-Verlag Berlin Heidelberg, 2007.

6. Dreyfus G. Neural networks: methodology and applications. – Berlin: Springer-Verlag Berlin Heidelberg, 2005.

7. Нейронные сети. STATISTICA Neural Networks: Методология и технологии современного анализа данных / Под ред. В. П. Боровикова. – М.: Горячая линия-Телеком, 2008.

8. Осовский С. Нейронные сети для обработки информации. – М.: Финансы и статистика, 2002.

9. Медведев В. С., Потемкин В. Г. Нейронные сети. МАТLAB 6. – М.: ДИАЛОГ-МИФИ, 2002.

10. Bianchini M., Maggini M., Jain L. C. Handbook on neural information processing. – Berlin: Springer-Verlag Berlin Heidelberg 2013.

11. Michel A.N., Liu D. Qualitative analysis and synthesis of recurrent neural networks. – New York: Marcel Dekker, 2002.

12. Mandic D.P., Chambers J.A. Recurrent neural networks for prediction: learning algorithms, architectures and stability. – New York: John Wiley & Sons, Inc., 2001.

13. Соловьева Е.Б. Полиномиальные и нейронные

модели нелинейных дискретных систем. – СПб.: Изд-во СПбГЭТУ «ЛЭТИ», 2014.

14. Bichkov U.A., Inshakov U.M., Solovyeva E.B., Scherbakov S.A. Analysis of mathematical models of continuous and discrete nonlinear systems. – СПб.: Изд-во СПбГЭТУ «ЛЭТИ», 2017.

15. Solovyeva E. Types of recurrent neural networks for non-linear dynamic system modelling // Proceedings of 2017 XX IEEE International Conference on Soft Computing and Measurements (SCM2017). – St. Petersburg: Saint-Petersburg Electrotechnical University «LETI». Russia, St. Petersburg, May 24–26, 2017, P. 1–4.

16. Chaos, CNN, Memristors and beyond. A festschrift for Leon Chua. / Ed.: A. Adamatzky, G. Chen. – World Scientific Publishing Co. Pte. Ltd., 2013.

17. Chen W.K. Feedback, nonlinear and distributed circuits. – New York: Taylor & Francis Group, LLC., 2009.

18. Yalcin M.E., Suykens J.A.K., Vandewalle J.P.L. Cellular neural networks, multi-scroll chaos and synchronization. – Singapore: World Scientific Publishing Co. Pte. Ltd., 2005.

19. Slavova A. Cellular neural networks: dynamics and modelling. – Dordrecht: Springer Science + Business Media, 2003.

20. Dogaru R. Universality and emergent computation in cellular neural networks. – Singapore: World Scientific Publishing Co. Pte. Ltd., 2003.

21. Chua L.O., Roska T. Cellular neural networks and visual computing: foundations and applications. – Cambridge: Cambridge Univ. Press, 2002.

22. Du K.-L., Swamy M.N.S. Neural networks in a softcomputing framework. – London: Springer-Verlag London Ltd, 2006.

23. Гудфеллоу Я., Бенджио И., Курвилль А. Глубокое обучение. – М.: ДМК Пресс, 2017.

24. Николенко С., Кадурин А., Архангельская Е. Глубокое обучение. – СПб.: Питер, 2018.

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.

• Таблицы и рисунки должны быть пронумерованы. На все рисунки, таблицы и библиографические данные указываются ссылки в тексте статьи.

• Объем статьи до 12 стр. (шрифт 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, электронная почта); ключевыми словами на русском и английском языках; актом экспертизы (при наличии в вашей организации экспертной комиссии).

2) Требования к иллюстрациям:

• Векторные (схемы, графики) – желательно использование графических редакторов Adobe Illustrator или Corel DRAW.

Растровые (фотографии, рисунки) – М 1:1, разрешение не менее 300dpi, формат tiff.

УДК 621.391:621.396.96

ОПТИМИЗАЦИЯ НЕРЕКУРСИВНЫХ РЕЖЕКТОРНЫХ ФИЛЬТРОВ С ЧАСТИЧНОЙ АДАПТАЦИЕЙ

Попов Д.И., д.т.н., профессор кафедры радиотехнических систем Рязанского государственного радиотехнического университета, e-mail: adop@mail.ru.

OPTIMIZATION OF NON-RECURSIVE REJECTION FILTERS WITH PARTLY ADAPTATION

Popov D.I.

The optimization method and the principles of the construction of rejector filters (RF) with partial adaptation are considered only to the Doppler phase of the clutter. The proposed method for optimizing RF characteristics with partial adaptation is based on the application of the minimax principle to the efficiency criterion in question and makes it possible in the a priori range of the spectral-correlation clutter parameters to increase the efficiency of its rejection with minimal losses in comparison with the marginal efficiency corresponding to full adaptation. System functions and structural schemes are given for cascade and canonical forms of RF implementation. The analysis of the RF with partial adaptation establishes the relationship between the efficiency of the RF and the volume of the training sample, depending on the order and structure of the RF and the clutter parameters. It is shown that the losses in RF efficiency are insignificantly dependent on the clutter parameters.

Key words: adaptation, Doppler phase, training sample set, optimization, clutter, minimax principle, rejector filters, rejection efficiency.

Ключевые слова: адаптация, доплеровская фаза, обучающая выборка, оптимизация, пассивные помехи, принцип минимакса, режекторные фильтры, эффективность режектирования.

Введение

В когерентно-импульсных радиолокационных системах высокой скважности при обнаружении сигналов движущихся целей на фоне пассивных помех, создаваемых мешающими отражениями от неподвижных или медленно перемещающихся объектов, основной операцией является режектирование спектральных составляющих помехи [1-5]. Априорная неопределенность спектрально-корреляционных характеристик помехи, а также их неоднородность

и нестационарность в зоне обзора существенно затрудняют реализацию предельных возможностей режектирования помехи. Преодоление априорной неопределенности параметров помехи в соответствии с методологией адаптивного байесовского подхода основывается на методах адаптации к неизвестным корреляционным параметрам помехи - аргументу и модулю коэффициентов межпериодной корреляции – путем замены этих параметров их состоятельными оценками [6], что приводит к построению адаптивных алгоритмов и систем обработки [7, 8], в частности, алгоритмов адаптивного режектирования с комплексными весовыми коэффициентами и соответствующих адаптивных режекторных фильтров (АРФ) [9]. Реализация данных АРФ в цифровом виде требует высокого быстродействия выполнения арифметических операций. Избежать указанных трудностей можно путем предварительной компенсации доплеровского сдвига фазы помехи. В работе

Рассмотрены метод оптимизации и принципы построения режекторных фильтров (РФ) с частичной адаптацией – только к доплеровской фазе помехи. Предложенный метод оптимизации характеристик РФ с частичной адаптацией основан на применении принципа минимакса к рассматриваемому критерию эффективности и позволяет в априорном диапазоне изменения спектрально-корреляционных параметров помехи повысить эффективность ее режектирования с минимальными потерями по сравнению с предельной эффективностью, соответствующей полной адаптации. Приведены системные функции и структурные схемы при каскадной и канонической формах реализации РФ. Проведенный анализ РФ с частичной адаптацией устанавливает связь между эффективностью РФ и объемом обучающей выборки в зависимости от порядка и структуры РФ и параметров помехи. Показано, что потери в эффективности

> [10] синтезированы алгоритмы оценивания и предложены принципы построения и структурные схемы автокомпенсаторов доплеровской фазы пассивных помех с прямой и обратной связью. Особенности адаптации к корреляционным свойствам помехи на выходе автокомпенсатора и последующего ее режектирования рассмотрены в работе [11]. Определенное упрощение процедуры адаптации достигается в АРФ каскадного типа [12]. Другим вариантом упрощения процедуры адаптации является переход от комплексных весовых коэффициентов к действительным, что ограничивает область целесообразного применения соответствующих АРФ при ограниченной и сравнительно малой в зависимости от порядка фильтра и ожидаемых параметров помехи величине ее доплеровской скорости [13].

> Необходимым условием эффективного режектирования помехи является адаптация к центральной часто

те спектра, определяемой доплеровским сдвигом частоты или фазы помехи, что реализуется в РФ с частичной адаптацией. Достаточным условием для достижения предельной эффективности режектирования является полная адаптация к спектрально-корреляционным свойствам помехи [7-9], что не всегда оправдано на практике. Ниже рассматривается оптимизация характеристик нерекурсивных режекторных фильтров с частичной адаптацией в априорном диапазоне изменения спектрально-корреляционных параметров помехи.

Постановка задачи

В общем случае адаптация нерекурсивного режекторного фильтра заданного порядка *m* сводится к выбору вектора весовых коэффициентов $\mathbf{G} = \{G_k\}$ ($k = \overline{0, m}$) по критерию максимума эффективности АРФ [9]:

$$\mu_{\max} = \max_{\mathbf{G}} [\mathbf{G}^{\mathsf{T}} \mathbf{G}^{\mathsf{T}} \mathbf{G}^{\mathsf{T}} (\mathbf{R} + \lambda \mathbf{I})\mathbf{G}] =$$
$$= \max_{\mathbf{G}} [(\mathbf{G}^{\mathsf{T}*} \mathbf{R} \mathbf{G} / \mathbf{G}^{\mathsf{T}*} \mathbf{G}) + \lambda]^{-1}, \qquad (1)$$

где символы т и * означают соответственно транспортирование и комплексное сопряжение; **R** – корреляционная матрица помехи, элементы которой $R_{jk} = \rho_{jk} \exp[i(j-k)\varphi]; \varphi$ – доплеровский сдвиг фазы помехи за период повторения; λ – отношение шум–поме-

ха; I – единичная матрица.

Из экстремальных свойств характеристических (собственных) чисел матриц следует [14], что минимумом отношения Рэлея в выражении (1) является наименьшее собственное число α_{\min} матрицы **R**. При этом $\mu_{\max} =$ $= (\alpha_{\min} + \lambda)^{-1}$, а оптимальный вектор **G** не зависит от λ и находится как собственный вектор матрицы **R**, соответствующий α_{\min} , в виде $\{G_k\} = \{g_k \exp(ik\varphi)\}$, где коэффициенты g_k определяются величинами ρ_{jk} в соответствии с алгоритмами работы [9].



Рис. 1. Зависимости выигрышей в эффективности РФ АРФ с полной адаптацией по отношению к классическим фильтрам с биномиальными весовыми коэффициентами $g_k = (-1)^k C_m^k$ реализуют выигрыш $\Delta \mu$ в эффективности, составляющий в зависимости от порядка АРФ от 2 до 6 дБ. Однако, как следует из приведенного выше соотношения для величины $\mu_{\rm max}$, на эффективность АРФ и величину соответствующих выигрышей существенное влияние оказывает наличие собственного шума приемника. Зависимости выигрышей $\Delta \mu$ от нормированной ширины гауссовского спектра помехи β_{Π} , что соответствует коэффициентам корреляции

$$\rho_{jk} = \exp\{-[\pi\beta_{\Pi}(j-k)]^2/2, 8\} = \rho^{(j-k)^2}$$

для величин $\lambda = 10^{-6}$, 10^{-4} , 10^{-2} приведены на рис. 1 сплошными кривыми соответственно 1, 2, 3. Как видим, с увеличением параметра λ происходит уменьшение выигрышей, обусловленных адаптацией коэффициентов g_k . В связи с этим при ограниченном динамическом

диапазоне помехи λ^{-1} представляют интерес режекторные фильтры с частичной адаптацией – только к доплеровскому сдвигу фазы помехи. При этом весовые коэффициенты g_k являются фиксированными. Кроме

того, при произвольной величине λ^{-1} использование РФ с частичной адаптацией является актуальным с точки зрения аппаратурного упрощения, состоящего в исключении измерителей коэффициентов корреляции помехи и вычислителя весовых коэффициентов РФ.

Критерий оптимизации РФ

Синтез РФ с частичной адаптацией сводится к оптимизации характеристик фильтра – вектора весовых коэффициентов или взаимного положения нулей системной функции – в заданном интервале изменения параметров помехи в соответствии с принятым критерием и к выбору его структуры. Для решения задачи оптимизации используем параметрическое описание модуля корреляционной функции помехи, приводящее в общем случае к некоторому вектору параметров $\boldsymbol{\beta} = \{\beta_l\}$, элементами которого могут быть коэффициенты корреляции ρ_{1k} , $k = \overline{2, m+1}$. При описании корреляционной функции в виде взвешенной суммы гауссовской и экспоненциальной функций с весовыми коэффициентами соответственно β_0 и $1 - \beta_0$ элементами вектора $\boldsymbol{\beta}$ являются β_0 и $\rho_{12} = \rho$ [7].

Использование критерия оптимизации, усредненного по пространству параметров вектора $\boldsymbol{\beta}$, предполагает введение меры $P(\boldsymbol{\beta})$, имеющей вероятностный смысл или отражающей степень значимости потерь в эффективности обработки при изменении $\boldsymbol{\beta}$ [6]. С учетом этого вектор $\mathbf{g} = \{g_k\}$ выбирается на основе критерия

$$\mu_{\max} = \max_{\mathbf{g}} \int_{\mathbf{\beta}} \mu(\mathbf{\beta}) dP(\mathbf{\beta}) .$$
⁽²⁾

Техника отыскания решения в соответствии с (2) включает операцию интегрирования обратной корреляционной матрицы помехи по мере $P(\beta)$. В общем случае выполнить аналитическое интегрирование невозможно, что приводит к численным методам интегрирования, ограничивая возможности синтеза по данной методики. И, наконец, введение более или менее произвольной меры $P(\beta)$ затруднительно в условиях априорной неопределенности и, в конечном счете, предполагает ее отсутствие.

Существенно проще в сравнении с критерием (2) определяется минимаксное правило решения [6]. Такое правило соответствует оптимуму только для некоторого значения вектора $\beta = \tilde{\beta}$ и находится в результате применения принципа минимакса к величине относительных потерь в эффективности

 $\delta\mu(\mathbf{\beta},\mathbf{\hat{\beta}}) = \mu(\mathbf{g},\mathbf{\beta}) / \tilde{\mu}(\mathbf{\tilde{g}},\mathbf{\beta}),$

где $\mu(\mathbf{g}, \boldsymbol{\beta})$ – эффективность при оптимизации вектора $\mathbf{g} = \{g_k\}$ по критерию (1) в соответствии с вектором $\boldsymbol{\beta}$; $\tilde{\mu}(\tilde{\mathbf{g}}, \boldsymbol{\beta})$ – эффективность при том же векторе $\boldsymbol{\beta}$, но при оптимизации вектора $\mathbf{g} = \tilde{\mathbf{g}}$ в соответствии с вектором $\tilde{\boldsymbol{\beta}}$.

Условие выбора значения β и соответствующего ему вектора $\tilde{\mathbf{g}}$ на основе принципа минимакса записывается в виде

 $\min_{\tilde{\boldsymbol{\beta}}} \max_{\boldsymbol{\beta}} \delta \mu(\boldsymbol{\beta}, \tilde{\boldsymbol{\beta}}) \leq \max_{\boldsymbol{\beta}} \delta \mu(\boldsymbol{\beta}, \tilde{\boldsymbol{\beta}}) \,. \tag{3}$

При известной форме корреляционной функции помехи ее свойства описываются скалярной величиной, в качестве которой может использоваться коэффициент корреляции ρ или нормированная ширина спектра помехи β_{Π} . В последнем случае для реализации условия (3) необходимы априорные данные об интервале $\Delta\beta_{\Pi}$, одномерный поиск в пределах которого приводит к искомым $\tilde{\beta}_{\Pi}$ и $\tilde{\mathbf{g}} = \{\tilde{g}_k\}$. В частности, для фильтров 2- и 3-го порядков ввиду симметрии весовых коэффициент тов g_k оптимизируется только один коэффициент или взаимное угловое смещение комплексно-сопряженных нулей системной функции в z-плоскости, также характеризуемых одним параметром.

При гауссовской аппроксимации спектра помехи и априорном интервале $\Delta\beta_{\Pi} = 0, 1...0, 3$ на рис. 1 штриховыми кривыми 1, 2, 3 для тех же λ и параметра $\tilde{\beta}_{\Pi}$ соответственно 0,1, 0,125, 0,15 приведены результаты, характеризующие выигрыши в оптимизированном по критерию (3) РФ с частичной адаптацией. Как видим, по сравнению с полной адаптацией сокращается интервал изменения β_{Π} , соответствующий выигрышам. Возникающие при этом потери зависят от параметров β_{Π} и λ , уменьшаясь с ростом λ . В частности, при $\lambda > 10^{-4}$ величина потерь в рассматриваемом интервале $\Delta\beta_{\Pi}$ не превышает 1 дБ. Заметим, что предложенный фильтр проще соответствующего фильтра с полной адаптацией.

Принципы построения и структурные схемы РФ

В конечном счете, в РФ с частичной адаптацией искомый весовой вектор имеет вид $\tilde{\mathbf{G}} = \{\tilde{G}_k\} = \{\tilde{g}_k \exp(ik\hat{\varphi})\}, k = \overline{0, m}, где \hat{\varphi}$ – оценка доплеровского сдвига фазы помехи за период повторения T. Синтез структуры РФ с частичной адаптацией удобно провести на основе системной функции в z-плоскости. Тогда при каскадной и канонической формах реализации фильтра для соответствующих системных функций имеем:

$$H\{z, \exp(i\hat{\varphi})\} = \prod_{k=1}^{m} \{1 - z^{-1} \exp[i(\theta_{0k} + \hat{\varphi})]\} =$$
$$= \sum_{k=0}^{m} \tilde{g}_k z^{-k} \exp(ik\hat{\varphi}) , \qquad (4)$$

где величины θ_{0k} , задающие положение нулей $z_{0k} = \exp(i\theta_{0k})$, могут быть оптимизированы непосредственно по критерию (3) или вычислены по коэффициентам \tilde{g}_k , например,

при
$$m = 2$$
 $\tilde{\theta}_{0(1,2)} = \pm \tilde{\theta}_r = \pm \arccos(-\tilde{g}_1 / 2);$
при $m = 3$ $\theta_{01} = 0$, $\tilde{\theta}_{0(2,3)} = \pm \tilde{\theta}_r = \pm \arccos(|1 + \tilde{g}_1| / 2)).$

Структурная схема РФ 3-го порядка (m = 3) с частичной адаптацией при каскадной форме реализации приведена на рис. 2. Обработка цифровых кодов комплексных отсчетов $U_j = x_j + iy_j$ последовательно осуществляется в трех каскадно включенных череспериодных компенсаторах (ЧПК), содержащих запоминающие на период повторения T устройства ($3Y_T$), комплексные умножители (\dot{x}) и сумматоры (Σ). В блоке измерения (БИ) по входным отсчетам в каждом периоде повторения вычисляется оценка максимального правдоподобия доплеровского сдвига фазы помехи $\hat{\phi}$ в соответствии с алгоритмом работы [10]:

 $\hat{\varphi} = \arg V = \operatorname{arctg}(\operatorname{Im} V/\operatorname{Re} V)$,

где n — объем обучающей выборки, определяемый числом независимых отсчетов со смежных элементов разрешения по дальности, за исключением отсчета с номером l = n/2 + 1.





$$V = \sum_{\substack{l=1\\l \neq n/2+1}}^{n+1} U_{j-1,l}^* U_{jl} = |V| e^{i\hat{\varphi}}$$

Так как адаптивную обработку на основе поступающих оценок для уменьшения ошибок рассогласования следует осуществлять после предварительной задержки исходных данных для среднего элемента разрешения временного строба, образующего обучающую выборку, то целесообразно исключить из обучающей выборки соответствующий этому элементу отсчет U_{il} с

номером l = n/2+1. Тогда в случае сигнала, соизмеримого с помехой, или разрывной помехи при обработке

элемента разрешения, содержащего сигнал, ослабление или подавление сигнала за счет его влияния на используемые оценки не происходит.

Адаптация к доплеровскому сдвигу фазы помехи в РФ на рис. 2 осуществляется соответствующим смещением нулей в каждом ЧПК на угол $\hat{\phi}$ путем комплексного умножения задержанных отсчетов на вычисляемую в косинусно-синусных функциональных преобразователях величину $\exp(i\hat{\phi}) = \cos\hat{\phi} + i\sin\hat{\phi}$. Кроме того, нули второго и третьего звеньев взаимно смещаются на величину $\tilde{\theta}_r$, оптимизируемую в соответствии с условием (3). При оптимизации θ_r в интервале изменения β_{Π} опре-

деляющим является нижнее граничное значение $\beta_{\Pi}^{({\rm H})}$. При этом оптимизированная величина

 $\tilde{\theta}_r = 1, 4\pi \beta_{\Pi}^{(\mathrm{H})}$.

Выигрыш $\Delta \mu$ по сравнению с неоптимизированным

вариантом ($\theta_r = 0$) при $\lambda < 10^{-6}$ достигает 3,6...3,9 дБ, а проигрыш по отношению к РФ с полной адаптацией является минимальным.

Каноническая форма реализации, вытекающая из второй формулы (4), соответствует РФ с комплексными весовыми коэффициентами $\tilde{G}_k = \tilde{g}_k \exp(ik\hat{\phi})$. Более просто РФ реализуется при раздельном решении задачи автокомпенсации доплеровской фазы помехи и последующего режектирования «остановленной» помехи в фильтре с действительными весовыми коэффициентами. Соответствующие автокомпенсаторы с прямой и обратной связью описаны в работе [10]. Автокомпенсаторы первого типа целесообразно использовать в обзорных радиолокаторах, а второго типа – в радиолокаторах сопровождения.

С целью минимизации вычислительных операций режекторный фильтр следует выполнять в виде каскадного соединения звеньев 1- и 2-го порядка для нечетного m и звеньев 2-го порядка для четного m [12]. Весовые коэффициенты звеньев 2-го порядка a_l непосредственно оптимизируются по критерию (3) или определяются по коэффициентам \tilde{g}_k с помощью соответствующих системных функций. В частности, при нечетном m:

$$H(z) = \sum_{k=0}^{m} \tilde{g}_k z^{-k} =$$

= $(1 - z^{-1}) \prod_{l=1}^{(m-1)/2} (1 + \tilde{a}_l z^{-1} + z^{-2}).$

Структурная схема фильтра 3-го порядка (m = 3) в виде каскадного соединения звеньев 1- и 2-го порядка с автокомпенсатором (АК) на входе приведена на рис. 3, Выходные величины автокомпенсатора $U_j e^{-i j \hat{\varphi}} = u_j e^{i [j(\varphi - \hat{\varphi}) + \varphi_{\rm H}]}$ с точностью до погрешностей измерения оценки $\hat{\varphi}$ не содержат доплеровских сдвигов фазы помехи [10], что позволяет осуществлять последующее режектирование помехи фильтром с действительными весовыми коэффициентами. Для звена 2-го порядка на рис. З действительный весовой коэффициент $\tilde{a} = -2\cos\tilde{\Theta}_r = 1 + \tilde{g}_1$. При $\beta_{\Pi}^{(\mathrm{H})} = 0,05$ и 0,1 коэффициент $\tilde{a} = -1,95$ и -1,8 соответственно. Заметим, что в сумматорах и умножителе соответствующие операции осуществляются раздельно для каждой квадратурной проекции (x_j, y_j) . В данном фильтре задачи частичной адаптации и последующего режектирования помехи выполняются при минимальном числе вычислительных операций и соответствующих арифметических блоков.



Рис. 3. Структурная схема РФ с автокомпенсатором

Анализ эффективности РФ

Рассмотрим эффективность РФ с частичной адаптацией в зависимости от объема обучающей выборки и параметров помехи. Используя результаты работы [11], можно показать, что подавление помехи в рассматриваемых (рис. 2, 3) фильтрах определяется выражением $(\sigma_{db}^2 / \sigma^2)_{\pi} =$

$$= \sum_{j,k=0}^{m} \tilde{g}_{j} \tilde{g}_{k} \rho_{jk} \exp[-(j-k)^{2} \sigma_{\hat{\phi}}^{2}/2] + \lambda \sum_{j=0}^{m} \tilde{g}_{j}^{2},$$

The npu $m = 2$ $\tilde{g}_{0} = g_{0} = \tilde{g}_{2} = g_{2} = 1,$
 $\tilde{g}_{1} = -2\cos\tilde{\Theta}_{r} = \tilde{a};$ npu $m = 3$ $\tilde{g}_{0} = g_{0} = -\tilde{g}_{3} = -g_{3} = 1,$

 $ilde{g}_1=- ilde{g}_2=-1-2\cos ilde{\Theta}_r=-1+ ilde{a}$; $\sigma_{\hat{\varphi}}^2$ – дисперсия оценки $\hat{\varphi}$.

Для автокомпенсатора с обратной связью [10]:

$$\sigma_{\hat{\varphi}}^2 = \frac{\eta[(1+\lambda)^2 - \rho^2]}{2(2-\eta)n\rho^2},$$
(5)

где η – весовой коэффициент сглаживающего фильтра в цепи обратной связи автокомпенсатора; n – объем обучающей выборки, определяемый числом независимых отсчетов со смежных элементов разрешения по дальности, усредняемых при вычислении оценки $\hat{\varphi}$.

Дисперсия оценки $\hat{\varphi}$ в аналогичных РФ (рис. 2) и автокомпенсаторе с прямой связью (рис. 3) определяется также выражением (5) при $\eta = 1$ [10].

Адаптация к доплеровской фазе φ приводит к потерям в эффективности, зависящим от величины n. Расчеты при гауссовской функции корреляции помехи показывают, что потери несущественно зависят от величины β_{Π} . При m = 2, n = 8, $\eta = 1$, $\beta_{\Pi} = 0,05...0,2$ и $\lambda \le 10^{-6}$ потери составляют около -0,5 дБ, а при m = 3 аналогичные потери в зависимости от величины β_{Π} составляют -0,7...-0,9 дБ.

Заключение

Таким образом, предложенный метод оптимизации нерекурсивных режекторных фильтров с частичной адаптацией, основанный на применении принципа минимакса к рассматриваемому критерию эффективности, позволяет в заданном интервале изменения параметров помехи повысить эффективность ее режектирования с минимальными потерями по сравнению с предельной эффективностью, соответствующей полной адаптации. Рассмотрены принципы построения и приведены структурные схемы соответствующих РФ, адаптирующихся к доплеровской фазе пассивной помехи. Проведенный анализ РФ с частичной адаптацией устанавливает связь между эффективностью РФ и объемом обучающей выборки в зависимости от порядка и структуры РФ и параметров помехи. Показано, что потери в эффективности РФ несущественно зависят от параметров помехи.

Литература

1. Skolnik M.I. Introduction to Radar System, 3rd ed., New York: McGraw-Hill, 2001. – 862 p.

2. Richards M.A., Scheer J.A., Holm W.A. (Eds.). Principles of Modern Radar: Basic Principles. New York: SciTech Publishing, IET, Edison. 2010. – 924 p.

3. Попов Д.И. Статистическая теория радиотехнических систем: учеб. пособие. Рязань: РГРТУ. 2011. – 80 с.

4. Melvin W. L., Scheer J.A. (Eds.). Principles of Modern Radar: Advanced Techniques. New York: SciTech Publishing, IET, Edison, 2013. – 846 p.



www.bhv.ru

Отдел оптовых поставок E-mail: opt@bhv.spb.su



5. Справочник по радиолокации: в 2 кн. Кн. 1 / под ред. М.И. Сколника; пер. с англ. под ред. В.С. Вербы. М.: Техносфера, 2014. – 672 с.

6. Репин В.Г., Тартаковский Г.П. Статистический синтез при априорной неопределенности и адаптация информационных систем. М.: Сов. радио. 1977. – 432 с.

7. Попов Д.И. Адаптивная обработка сигналов на фоне пассивных помех // Известия вузов. Радиоэлектроника. – 2000. – Т. 43, № 1 (451). – С. 59-68.

8. Попов Д.И. Оптимальная обработка многочастотных сигналов // Известия вузов России. Радиоэлектроника. – 2013. – Вып. 1. – С. 32–39.

9. Попов Д.И. Адаптивные режекторные фильтры с комплексными весовыми коэффициентами // Вестник Концерна ПВО «Алмаз – Антей». – 2015. – № 2 (14). – С. 21-26.

10. Попов Д.И. Автокомпенсация доплеровской фазы пассивных помех // Цифровая обработка сигналов. – 2009. – № 2. – С. 30–33.

11 Попов Д.И. Адаптивное подавление пассивных помех // Цифровая обработка сигналов. – 2014. – № 4. – С. 32-37.

12. Попов Д.И. Адаптивные режекторные фильтры каскадного типа // Цифровая обработка сигналов. – 2016. – № 2. – С. 53-56.

13. Попов Д.И. Адаптивные режекторные фильтры с действительными весовыми коэффициентами // Цифровая обработка сигналов. – 2017. – № 1. – С. 22-26.

14. Гантмахер Ф.Р. Теория матриц. М.: Наука, 1988. – 552 с.

Солонина А.И.

Цифровая обработка сигналов в зеркале MATLAB

РЕЦЕНЗЕНТЫ: Е.Б. Соловьева, д-р техн. наук, завкафедрой теоретических основ электротехники Санкт-Петербургского государственного электротехнического университета «ЛЭТИ»; А.А. Монаков, д-р техн. наук, проф. кафедры радиотехнических систем Санкт-Петербургского государственного университета аэрокосмического приборостроения.

Учебное пособие представляет собой базовый курс по цифровой обработке сигналов с традиционными темами: линейные дискретные системы; эффекты квантования в цифровых системах с фиксированной точкой; дискретные сигналы с описанием в частотной области, включая алгоритмы дискретного и быстрого преобразований Фурье; случайные процессы и их статистические характеристики; методы непараметрического и параметрического спектрального анализа; цифровые фильтры общего и специального назначения; многоскоростные системы и их полифазные структуры; банки фильтров; основы адаптивной обработки сигналов.

Теория ЦОС преподносится с позиции моделирования в MATLAB: теоретические аспекты базовых методов и алгоритмов ЦОС увязываются с их моделированием в MATLAB и иллюстрируются соответствующими примерами.

Книга предназначена для студентов и преподавателей вузов, а также специалистов в области цифровой обработки сигналов.

Солонина Алла Ивановна, кандидат технических наук, доцент кафедры радиосистем и обработки сигналов Санкт-Петербургского государственного университета телекоммуникаций имени проф. М.А. Бонч-Бруевича с многолетним научно-методическим и педагогическим стажем, автор многочисленных публикаций и книг по цифровой обработке сигналов, в том числе «Алгоритмы и процессоры цифровой обработки сигналов», «Основы цифровой обработки сигналов. Курс лекций», «Цифровая обработка сигналов. Моделирование в MATLAB», «Цифровая обработка сигналов. Моделирование в Simulink», «Цифровая обработка сигналов и MATLAB», выпущенных издательством «БХВ-Петербург». УДК 621.396

ИНТЕГРАЦИЯ НАЗЕМНЫХ КОМПЛЕКСОВ ПРИЕМА, ОБРАБОТКИ И РАСПРОСТРАНЕНИЯ ИНФОРМАЦИИ ОТ КОСМИЧЕСКИХ СИСТЕМ «ЭЛЕКТРО» И «АРКТИКА-М» С ЕДИНОЙ ТЕРРИТОРИАЛЬНО РАСПРЕДЕЛЕННОЙ СИСТЕМОЙ ДЗЗ

Пастарнак В.И., к.т.н., старший научный сотрудник, главный конструктор АО «НПО Лавочкина», e-mail: npol@laspace.ru;

Козинин Е.А., заместитель начальника отдела АО «НПО Лавочкина», e-mail: kozinin@laspace.ru; Кубышкин А.П., ведущий инженер-конструктор АО «НПО Лавочкина», e-mail: npol@laspace.ru; Череватюк И.В., инженер-электроник АО «НПО Лавочкина»», e-mail: kozinin@laspace.ru; Ефимов Е.А., инженер-электроник АО «НПО Лавочкина»», e-mail: npol@laspace.ru.

INTEGRATION OF GROUND COMPLEX FOR RECEPTION, PROCESSING AND DISSEMINATION OF INFORMATION FROM SPACE SYSTEMS «ELECTRO» AND «ARCTICA-M» WITH UNIFIED TERRITORIALY DISTRIBUTED SYSTEM OF EARTH REMOTE SENSING

Pastarnak V.I., Kozinin E.A., Kubishkin A.P., Cherevatyuk I.V., Efimov E.A.

Components and purpose of ground segments of space systems «Electro» and «Arctica-M» and their future integration with Unified territorialy distributed system of Earth remote sensing which being created by State space corporation «Roscosmos» are presented in the article.

Key words: Earth remote sensing; ground complex of acquisition, processing and distribution of information; Unified territorially distributed system; geostationary hydrometeorological space system; highly elliptical orbital space system.

Ключевые слова: дистанционное зондирование Земли, наземный комплекс приема, обработки и распространения информации, единая территориально распределённая информационная система, геостационарная гидрометеорологическая космическая система, высокоэллиптическая космическая система.

Введение

В рамках Федеральной космической программы России планируется развитие геостационарной гидрометеорологической космической системы (ГГКС) «Электро-Л». Созданный ранее наземный комплекс приема, обработки и распространения информации от космического аппарата (КА) «Электро-Л» № 1 (НКПОР-Э) требует дальнейшего развития для последующих аппаратов этого типа и прежде всего в части взаимодействия с создаваемой в России Единой территориально-распределенной информационной системы дистанционного зондирования Земли (ЕТРИС ДЗЗ).

Научно-технический задел по теме «Электро» может быть эффективно использован для создания нового класса систем глобального наблюдения Земли типа «Арктика-М», который нацелен на детальное изучение стратегически важных для России приполярных областей, что не возможно выполнить с помощью системы «Электро-Л».

В настоящей работе акцент сделан на особенности построения наземных средств обработки информации с ориентацией на космические системы серии «Электро-Л»

Представлены принципы построения и задачи наземных средств обработки данных от космических систем «Электро-Л» и «Арктика-М». Рассмотрены вопросы интеграции этих наземных средств с Единой территориальной распределенной информационной системой дистанционного зондирования Земли, создаваемой Госкорпорацией «Роскосмос».

> (функционирующих на геостационарных орбитах) и систем «Арктика-М» (работающих на высокоэллиптических орбитах). Большое внимание уделено вопросу использования полученной и обработанной информации в интересах конкретных потребителей путем интеграции в систему ЕТРИС ДЗЗ.

Космическая система «Электро-Л»

Геостационарная гидрометеорологическая космическая система «Электро-Л» предназначена для информационного обеспечения задач оперативной метеорологии, гидрологии, агрометеорологии, мониторинга климата и окружающей среды.

Задачей КА «Электро-Л» является получение многозональных цифровых изображений земного диска на фоне космоса, на основе которых формируются спутниковые гидрометеорологические продукты: глобальные карты облачности и подстилающей поверхности, температурные карты морской поверхности, карты перемещения облачных образований и т.п.

КА «Электро-Л» также решают задачи получения гелиогеофизических данных на высоте орбиты с последующей передачей их в НКПОР-Э, ретрансляции метеоданных с платформ сбора данных и других видов целевой информации.

Возможности бортовой целевой аппаратуры позволяют обеспечить:

 – квазинепрерывное наблюдение всего земного диска (частота получения глобальной информации – 0,5 часа);

 – получение снимков фрагментов облачного покрова и земной поверхности за счёт сканирования земного диска с временным интервалом 15 минут;

 непрерывное получение гелиогеофизической информации на высоте орбиты;

 – ретрансляцию метеорологических данных с платформ сбора данных;

 – ретрансляцию обработанной метеорологической информации и массивов цифровой информации;

– ретрансляцию сигналов от аварийных буев системы КОСПАС-САРСАТ.

Передача информации с КА «Электро-Л» на пункты приема производится:

 – круглосуточно, с интервалом 0,5 часа (15 минут в учащенном режиме) в части информации, сформированной бортовой системой сбора данных (БССД) от многозонального сканирующего устройства (МСУ-ГС), гелиогеофизического аппаратурного комплекса (ГГАК-Э) и бортового комплекса управления;

 – непрерывно, в части гелиогеофизической информации от ГГАК-Э;

 круглосуточно, в соответствии с циклограммой ретрансляции различных видов целевой информации.

Покрытие земной поверхности зонами обзора КА «Электро-Л» представлено на рис. 1.

НКПОР-Э осуществляет прием, обработку, накопление и распространение всех видов целевой информации, передаваемой и ретранслируемой с КА «Электро-Л», планирования и контроля работы целевой аппаратуры и средств приема и обработки данных наземного комплекса, выдачи необходимых данных для управления КА в наземный комплекс управления (НКУ-Э).

НКПОР-Э создан с учетом совместимости по информационным продуктам с космическими аппаратами международной метеорологической спутниковой системы и представляет собой административную территориально-распределенную систему, составные части которой расположены в различных регионах Российской Федерации.

Область применения информационных продуктов:

 – анализ и прогноз погоды в региональном и глобальном масштабах;

 – анализ и прогноз состояния акваторий морей и океанов;

- анализ и прогноз условий для полетов авиации;

 – анализ и прогноз гелиогеофизической обстановки в околоземном космическом пространстве, состояние ионосферы и магнитного поля Земли;

мониторинг климата и глобальных изменений;

- контроль чрезвычайных ситуаций;

– экологический контроль окружающей среды и др.

Схема территориального размещения составных частей НКПОР-Э представлена на рис. 2.

НКПОР-Э решает следующие целевые задачи:

– планирование, составление и коррекцию (при необходимости) программ работы бортовой целевой аппаратуры с учетом имеющегося ресурса, условий функционирования составных частей космического комплекса, состояния средств НКПОР-Э и заявок потребителей с последующей выдачей их в НКУ-Э в установленные сроки;

– прием на частоте 7500 МГц со скоростью 30,72 Мбит/с многоспектральных снимков облачности и подстилающей земной поверхности в пределах всего наблюдаемого диска Земли, получаемых с МСУ-ГС, а также гелиогеофизической информации и оперативноконтрольной информации;

 прием на частоте 1693 МГц со скоростью 5 Кбит/с гелиогеофизических данных, полученных на высоте орбиты, их обработку и выдачу подразделениям Федеральной службы России по гидрометеорологии и мониторингу окружающей среды;

– передачу метеорологической информации с платформ сбора данных непосредственно на КА «Электро-Л», в режиме многостанционного доступа с частотновременным разделением каналов, с периодичностью, привязанной к синоптическим срокам и информации о



Рис. 1. Покрытие земной поверхности зонами обзора КА «Электро-Л»

штормовых предупреждениях в диапазоне частот от 401 до 403 МГц со скоростями передачи 100 или 1200 бит/с и объемом передаваемых данных: для международных каналов 5192 бит в каждом сообщении; для национальных каналов до 15000 бит в каждом сообщении;



Рис. 2. Схема территориального размещения составных частей НКПОР-Э

– прием метеорологической информации, ретранслированной через КА «Электро-Л», одновременно с 300 платформ сбора данных в режиме частотно-временного разделения каналов в диапазоне (1697 ± 1) МГц со скоростями приема 100 или 1200 бит/с и объемом принимаемых данных от 5192 бит до 15000 бит в каждом сообщении;

 выполнение телекоммуникационных функций по распространению, обмену обработанными гидрометеорологическими и гелиогеофизическими данными;

 – прием сигналов от аварийных радиобуев системы КОСПАС-САРСАТ, ретранслированных бортовым радиотехническим комплексом (БРТК) КА «Электро-Л» на частоте 1544,5 МГц со скоростью передачи 400 бит/с, и передачу информации в наземную службу системы КОСПАС-САРСАТ; – анализ и оценку качества информации, принимаемой в сеансе связи;

 – формирование и выдачу данных для обмена информацией с НКУ-Э и составными частями НКПОР-Э;

– управление и координацию работы составных частей НКПОР-Э.

Космическая система «Арктика-М»

Высокоэллиптическая гидрометеорологическая космическая система (ВГКС) «Арктика-М» с орбитальной группировкой в составе двух аппаратов предназначена для информационного обеспечения решения задач оперативной метеорологии, гидрологии, агрометеорологии, мониторинга климата и окружающей среды.

ВГКС «Арктика-М» имеет принципиальные отличия от ГГКС «Электро-Л»:

 движение КА во время съемки осуществляется по высокоэллиптической орбите;

 в ВГКС на рабочих участках попеременно будут функционировать 2 КА;

 обеспечивается непрерывная круглосуточная съемка арктического региона Земли выше 60° с.ш., недоступного для наблюдения с геостационарной орбиты, при зенитных углах съемки не более 70°;

 материалы съемки для тематической обработки в первую очередь требуются российским потребителям.

Космические аппараты «Арктика-М» № 1, 2, входящие в состав орбитальной группировки ВГКС «Арктика-М», обеспечивают выполнение следующих основных целевых задач:

 – многоспектральная съемка облачности и подстилающей земной поверхности в пределах всего наблюдаемого диска Земли с периодичностью 30 минут (штатный режим) и 10-15 минут (учащенный режим);

 передача на пункты приёма информации данных многоспектральной съемки, оперативно-контрольной и гелиогеофизической информации;

 – сбор и передача на пункты приёма информации гелиогеофизических данных;

сбор и передача данных с платформ сбора данных;

– ретрансляция сигналов аварийных буев системы КОСПАС-САРСАТ.

Покрытие земной поверхности зонами обзора КА «Арктика-М» представлено на рис. 3.



Рис. 3. Покрытие земной поверхности зонами обзора КА «Арктика-М»

Прием, обработку данных и комплексное планирование информационного ресурса ВГКС осуществляет НКПОР-АМ, представляющей собой административную территориально-распределенную систему, составные части которой расположены в различных регионах Российской Федерации.

Укрупненный состав ВГКС «Арктика-М» представлен на рис. 4.



Рис. 4. Состав ВГКС «Арктика-М»

НКПОР-АМ, входящий в состав ВГКС «Арктика-М», обеспечивает выполнение следующих основных задач:

 прием и передачу радиосигналов по космическим каналам связи с КА «Арктика-М» во всех рабочих диапазонах;

 – демодуляцию, декодирование, распаковку, первичную и тематическую обработку принимаемой информации;

 планирование и составление программ работы целевой аппаратуры с учетом имеющегося ресурса, функционирования космической системы и заявок потребителей с последующей выдачей программ в наземный комплекс управления в установленные сроки;

 – сопряжение центров приема и обработки данных НКПОР-АМ с наземными каналами обмена данными для передачи информации основным потребителям с целью ее вторичной обработки;

 – оперативный контроль и анализ функционирования бортовой целевой аппаратуры;

 – поддержку функций поиска и спасения системы КОСПАС-САРСАТ.

С учетом особенностей функционирования ВГКС «Арктика-М», при создании наземных средств приема информации необходимо реализовать следующие доработки:

 применение антенных постов, обеспечивающих перенацеливание электрической оси антенны в широком диапазоне углов;

 – учет доплеровского сдвига частот вследствие наличия радиальной скорости движения КА;

 – нестандартный X диапазон частот по линии вниз (7865 МГц вместо 7500 МГц).

При проектировании наземных комплексов обработки информации ВГКС «Арктика-М» требуется создание новых технологий, алгоритмов, методик и программного обеспечения решения целевых задач. Главными проблемами являются организация межпрограммного взаимодействия и формирование изображений уровня обработки 1.5 (геометрическая и радиометрическая нормализация, геодезическая привязка).

Съемка с высокоэллиптической орбиты арктической территории со слабо выраженными береговыми линиями существенно усложняет процесс нормализации и геодезической привязки целевой информации МСУ-ГС. С целью обеспечения надежности доставки снимков МСУ-ГС до потребителей НКПОР-АМ будет иметь сложную распределенную структуру, подразумевающую двойное горячее и холодное резервирование вычислительных комплексов. В этих условиях для выполнения требований циклограммы по времени передачи и обработки целевой информации целесообразно минимизировать количество комплексов программно-технических средств в составе НКПОР-АМ.

Интеграция НКПОР-Э и НКПОР-АМ с ЕТРИС Д33

Для интеграции НКПОР-Э и НКПОР-АМ с ЕТРИС Д33 необходимо в первую очередь:

– внести дополнения в тактико-технические задания по темам «Электро-Л», «Арктика-М»;

 – разработать и утвердить положение по взаимодействию НКПОР-Э, НКПОР-АМ с ЕТРИС ДЗЗ, включающее распределение квот бортовых ресурсов по ретрансляции информации между ведомствами;

 образовать оперативно-техническое руководство из представителей головных предприятий по созданию космических комплексов, эксплуатирующих организаций, разработчиков ЕТРИС ДЗЗ, заинтересованных ведомств (по образцу главной оперативной группы управления).

В целях минимизации временных и финансовых издержек на организацию и создание ЕТРИС ДЗЗ целесообразно максимально использовать уже имеющиеся центры НКПОР министерств, ведомств и организаций, учитывая:

 имеющуюся инфраструктуру (территория, здания и сооружения, коммуникации, средства связи и др.);

– разрешительные документы на осуществление космической связи;

подготовленный квалифицированный персонал;

 налаженное взаимодействие с обеспечивающими организациями (службы контроля радиоэфира, учебные заведения, провайдеры, коммунальные службы, транспортные кампании и др.).

Наиболее предпочтительным вариантом организации взаимодействия является реализация на базе существующего СПО НКПОР-Э и НКПОР-АМ функции обмена информацией с ЕТРИС ДЗЗ по протоколам информационно-логического взаимодействия при минимально необходимой доработке используемых форм обмена и СПО.

Подобный принцип работы заложен при обмене информацией между НКПОР-Э (НКПОР-АМ) и НКУ, а также с потребителями тематических продуктов. Специальное программное обеспечение, реализующее данный способ обмена информацией, функционирует с 2010 года и зарекомендовало себя с положительной стороны.

Взаимодействие ЕТРИС ДЗЗ с НКПОР-Э (НКПОР-
АМ) предлагается организовать следующим образом:

– в ЕТРИС Д33 оператор или программное обеспечение формирует заявку на получение и выдачу информации по проведенным съемкам с КА ГГКС «Электро-Л» и ВГКС «Арктика-М»;

– заявка из ЕТРИС ДЗЗ поступает в НКПОР-Э (НКПОР-АМ) и принимается к реализации;

 – НКПОР-Э (НКПОР-АМ) после приема и предварительной обработки информации формирует файлы и выдает запрашиваемые данные в ЕТРИС Д33;

 – ЕТРИС ДЗЗ производит хранение, дальнейшую обработку и предоставление полученной информации заинтересованным пользователям.

Учитывая опыт эксплуатации СЧ НКПОР-Э при работе с КА «Электро-Л» № 1, 2, целесообразно в случаях, когда прием информации на основные приемные средства НКПОР-Э и НКПОР-АМ не представляется возможным, использовать для этого резервные приемные средства НКПОР-Э и НКПОР-АМ или приемные средства ЕТРИС ДЗЗ (ЦПО и КККИ Госкорпорации «Роскосмос» и им подобные), а затем передавать принятую информацию в ФГБУ «НИЦ «Планета» по наземным каналам связи (рис. 5).



Рис. 5. Схема информационного взаимодействия центра НКПОР с ЕТРИС ДЗЗ

Для обеспечения возможности более частого использования телекоммуникационных каналов БРТК в перерывах между съемками МСУ-ГС целесообразно провести работы по усовершенствованию или замене конструкции системы управления остронаправленных антенн (в настоящее время существует ограничение на перенацеливание не более 2 раз в сутки).

В целях увеличения пропускной способности каналов БРТК целесообразно провести работы по увеличению скорости передачи данных (в т.ч. с использованием более сложных способов модуляции сигналов, а также кодирования передачи данных). Это позволит передавать в НКПОР-Э данные, полученные от МСУ-ГС, без буферизации в БССД, в т.ч. массивы данных, полученные от 4-х проходов сканирования в ИК-каналах.

Также для динамического конфигурирования информационной системы по обработке информации предлагается перейти от модульного построения программного обеспечения к конфигурированию через отдельно работающие библиотеки.

Ещё одним ключевым шагом по сокращению времени обработки и доставки информации потребителю, повышению автоматизации и сокращению затрат на создание и эксплуатационных расходов может быть реализация конфигурации программно-технических средств наземных сегментов космических систем дистанционного зондирования Земли, включающей в себя интеграцию работы СПО на базе единых технических средств и общего программного обеспечения, использование виртуализации и кластеризации, сращивание с приемнопередающими станциями. Частично такая конфигурация была реализована в проекте «Электро-Л», без интеграции с приемно-передающими станциями она реализуется в проекте «Арктика-М» и в окончательном виде предполагается к реализации в проекте «Электро-М» на АО «НПО Лавочкина» совместно со смежными организациями. Это позволит в буквальном смысле устанавливать на каждом объекте эксплуатации одну вычислительную стойку, пару-тройку рабочих мест операторов, необходимое количество модульных приемных и передающих антенных комплексов, тиражировать их на любое количество объектов эксплуатации путем конфигурации необходимых модульных компонентов, при этом решая все задачи НКПОР по взаимодействию со средствами ЕТРИС.

В дальнейшем предполагается переход на свободно распространяемое общее программное обеспечение на базе операционных систем Linux.

Заключение

По рассмотренным исследованиям можно сделать следующие выводы.

1. Системы геостационарного типа («Электро-Л») и системы высокоэллиптического наблюдения Земли («Арктика-М») имеют общие и отличительные принципы построения наземных средств обработки информации от них, которые реализуются в НКПОР-Э и НКПОР-АМ.

2. Рассмотрена важная задача - интеграция НКПОР-Э и НКПОР-АМ с ЕТРИС ДЗЗ, которая позволит всем заинтересованным организациям получить оперативный доступ к целевой информации с КА «Электро-Л» и «Арктика-М» для решения собственных задач.

Литература

1. Норенков И.Т. Телекоммуникационные технологии и сети. М.: Изд. МГТУ им. Баумана, 2000. 248 с.

2. Олифер Н.А., Олифер В.Г. Компьютерные сети. Принципы, технологии, протоколы: Учебник для ВУЗов. 2-е изд. СПб.: Питер, 2004. 864 с.

3. Клейнрок Л. Теория массового обслуживания: Пер. с англ. // Под ред. В.И. Нейман. М.: Машиностроение, 1979. 432 с.

4. Кучерявый Е.А. Управление трафиком и качество обслуживания в сети Интернет // СПб, Наука и Техника. 2004.

УДК 621.391

ДВУХКАНАЛЬНЫЙ УНИВЕРСАЛЬНЫЙ АЛГОРИТМ ДЕТЕКТИРОВАНИЯ СПЕКТРАЛЬНО-ЭФФЕКТИВНЫХ РАДИОСИГНАЛОВ С УПРАВЛЯЕМОЙ СВЯЗЬЮ МЕЖДУ СИНФАЗНОЙ И КВАДРАТУРНОЙ СОСТАВЛЯЮЩИМИ

Покровский П.С., к.т.н., инженер кафедры радиоуправления и связи Рязанского государственного радиотехнического университета, e-mail: paulps@list.ru.

TWO-CHANNEL UNIVERSAL ALGORITHM TO DEMODULATE SPECTRAL-EFFICIENT RADIO SIGNALS WITH CONTROLLED COUPLING BETWEEN IN-PHASE AND QUADRATURE COMPONENTS

Pokrovskij P.S.

The demodulation of spectral-efficient radio signals (T-OQPSK, GMSK, FQPSK, and special synthesized) without processing the coupling between in-phase and quadrature components of complex envelope is described. It made possible the route of metric calculation changed to two-channel or three-channel filtering in Viterbi algorithm. It led to considerable decrease by 4 times required computational resources of filtering block as compared with optimal demodulation algorithm. It is shown the proposed two-channel universal algorithm to demodulate spectral-efficient radio signals has a maximum loss in noise immunity of less than 0,3 dB as compared with optimal demodulation algorithm in case bit error rate equals 10⁻⁴, while it has a gain by 2,6 dB relative to the single-channel algorithm under the same conditions.

Key words: synthesis of demodulation device, universal shaper of radio signals, spectral efficient modulations, FQPSK, GMSK, T-OQPSK.

Ключевые слова: синтез устройства детектирования, универсальный формирова-тель радиосигналов, спектрально-эффективные виды модуляции, FQPSK, GMSK, T-OQPSK.

Введение

Дефицит частотного ресурса в настоящее время вынуждает создавать новые системы передачи информации, способные на вторичной основе работать в уже занятой полосе [1, 2]. Для реализации данного принципа проектируемые системы должны обладать высокими когнитивными способностями, широким функционалом средств радиомониторинга, а также значительной вариативностью параметров используемых радиосигналов [3]. Одним из возможных подходов к созданию подобных приемопередающих устройств является применение универсальных алгоритмов формирования и обработки радиосигналов [4].

Для большинства спектрально-эффективных видов модуляции, описанных в [5], подобные процедуры можно получить с использованием модели радиосигналов с управляемой связью между синфазной и квадратурной составляющими (РУССКС). В частности, в [6] разработана схема универсального формирователя таких радиосигналов, как T-OQPSK, FQPSK и GMSK, относящиеся к спектрально-эффективным [5]. В [7] обосновано оптимальное многоканальное устройство обработки РУССКС, а в [8] – универсальный одноканальный блок детектирования подобных сигналов. При этом главным недостатком первого устройства является высокая вычислительная сложность, экспоненциально возрастаю-

Рассмотрена возможность для приема спектрально-эффективных радиосигналов (T-OQPSK, GMSK, FQPSK и специально синтезированных) не учитывать на приемной стороне связь между синфазной и квадратурной составляющими комплексной огибающей, что позволило свести расчет расстояний в алгоритме Витерби процесса детектирования к двухканальной или трехканальной фильтрации. Это привело к значительному (до четырех раз) сокращению требуемых вычислительных ресурсов блока обработки по сравнению с оптимальным алгоритмом. Показано, что предложенный двухканальный универсальный алгоритм детектирования спектрально-эффективных радиосигналов с управляемой связью между синфазной и квадратурной составляющими характеризуется максимальным проигрышем в помехоустойчивости менее 0,3 дБ оптимальному алгоритму обработки при вероятности битовой ошибки на уровен $P_0 = 10^{-4}$, при этом выигрыш по данному пока-

зателю качества относительно одноканального алгоритма достигал 2,6 дБ при тех же условиях.

> щая с увеличением позиционности обрабатываемых сигналов. Данный недостаток существенно ограничивает возможности по практическому использованию РУССКС в системах с адаптацией к условиям функционирования в реальном масштабе времени.

> Одноканальный алгоритм детектирования позволяет решить данную проблему, однако характеризуется существенным проигрышем в помехоустойчивости, который для отдельных видов модуляции может достигать 2,6 дБ [8]. В этой связи актуальной является разработка алгоритма обработки спектрально-эффективных радиосигналов, обобщающего достоинства этих двух устройств – высокая помехоустойчивость и низкая вы

числительная сложность.

Таким образом, целью данной работы является разуниверсального алгоритма детектирования работка спектрально-эффективных радиосигналов с низкой вычислительной сложностью.

Обоснование алгоритма детектирования

Как известно [9], в случае действия аддитивного «белого» гауссовского шума (АБГШ) оптимальным по критерию максимума правдоподобия алгоритмом детектирования радиосигналов является алгоритм, максимизирующий следующую функцию: $\Lambda(r(t)) -$

$$= C \exp(-\frac{1}{N_0} (\int_0^{T_0} |s(t)|^2 dt - 2 \int_0^{T_0} \operatorname{Re}\{r(t)s^*(t)\}dt)), \qquad (1)$$

где r(t) – комплексная огибающая принятой смеси сигнала и шума, s(t) – комплексная огибающая эталонного сигнала, N_0 – спектральная плотность мощности действующего шума, С - нормировочный множитель, *T*₀ – период анализа принимаемого сигнала.

Введем следующие обозначения $r(t) = I_r(t) + j \cdot Q_r(t)$ и $s(t) = I(t) + j \cdot Q(t)$. При этом спектрально-эффективные радиосигналы, такие как T-OQPSK, GMSK и FQPSK [5], можно представить как РУССКС, синфазная и квадратурная составляющие которых определяются в виде [10]:

$$\begin{split} I(t) &= I_{1}(t) - c_{2}I_{2}(t) - c_{3}I_{3}(t) ,\\ Q(t) &= Q_{1}(t) - c_{2}Q_{2}(t) - c_{3}Q_{3}(t) ,\\ c_{2} &= (1 - A_{1})/2 , \ c_{3} &= (1 - A_{2})/2 ,\\ I_{1}(t) &= \sum_{i=0}^{N-1} a_{i}p_{1}(t - iT) ,\\ Q_{1}(t) &= \sum_{i=0}^{N-1} b_{i}p_{1}(t - iT - T/2) ,\\ I_{2}(t) &= \sum_{i=0}^{N-1} (a_{i-1} + a_{i})p_{3}(t - iT) ,\\ Q_{2}(t) &= \sum_{i=0}^{N-1} (b_{i-1} + b_{i})p_{3}(t - iT - T/2) ,\\ I_{3}(t) &= \sum_{i=0}^{N-1} a_{i} \mid b_{i-1} + b_{i} \mid p_{3}(t - iT - T/2) ,\\ Q_{3}(t) &= \sum_{i=0}^{N-1} b_{i-1} \mid a_{i-1} + a_{i} \mid p_{3}(t - iT) . \end{split}$$

Здесь А, и А, - коэффициенты, задающие соответственно «глубину» межсимвольной связи и связи между синфазной и квадратурной составляющими, Т – длительность информационного символа, а, и b, символы информационных последовательностей $\{a_i\}_0^{N-1}$ и $\{b_i\}_0^{N-1}$, соответственно передаваемых по каналу синфазной и квадратурной составляющих. Элементарные импульсы $p_1(t)$ и $p_2(t)$ описываются выражениями [10]:

$$p_{1}(t) = \operatorname{rect}(t / (2T)) \cdot (1 - \cos(\pi t / T)) / 2,$$

$$p_{3}(t) = \operatorname{rect}(t / T) \cdot (1 - \cos(2\pi t / T)) / 2,$$

rge

 $\operatorname{rect}(u) = \begin{cases} 1 & u \in [0,1); \\ 0 & u \notin [0,1). \end{cases}$

 N_{-1}

С учетом длительности и свойств симметрии импульса $p_1(t)$ составляющие $I_1(t)$ и $Q_1(t)$ можно представить в следующем виде:

$$I_{1}(t) = \sum_{i=0}^{N-1} (a_{i-1} + (a_{i} - a_{i-1})p_{1}(t - iT)) \cdot \operatorname{rect}(t/T - i),$$

$$Q_{1}(t) = \sum_{i=0}^{N-1} (b_{i-1} + (b_{i} - b_{i-1})p_{1}(t - iT - T/2)) \cdot \operatorname{rect}(t/T - i - 1/2).$$

Интеграл корреляции из формулы (1) можно представить в виде

$$\int_{0}^{T_{0}} \operatorname{Re}\{r(t)s^{*}(t)\}dt = \int_{0}^{T_{0}} I_{r}(t)I(t)dt + \int_{0}^{T_{0}} Q_{r}(t)Q(t)dt .$$
 (2)

Первое слагаемое в формуле (2) для рассматриваемого класса радиосигналов равно:

$$\int_{0}^{I_{0}} I_{r}(t)I(t)dt = \int_{0}^{I_{0}} I_{r}(t)I_{1}(t)dt - -c_{2}\int_{0}^{I_{0}} I_{r}(t)I_{2}(t)dt - c_{3}\int_{0}^{I_{0}} I_{r}(t)I_{3}(t)dt.$$
(3)

Можно показать, что составляющие данного выражения описываются формулами:

$$\int_{0}^{T_{0}} I_{r}(t)I_{1}(t)dt = \sum_{i=0}^{L-1} (a_{i-1}I_{rM}[i] + (a_{i} - a_{i-1})I_{r1}[i]), \qquad (4)$$

$$\int_{0}^{T_{0}} I_{r}(t) I_{2}(t) dt = \sum_{i=0}^{L-1} (a_{i-1} + a_{i}) I_{r3}[i],$$
(5)

$$\int_{0}^{T_{0}} I_{r}(t) I_{3}(t) dt = \sum_{i=0}^{L-1} a_{i} | b_{i-1} + b_{i} | I_{r3}[i+1/2],$$
(6)

где L – количество информационных символов, попавших в интервал анализа T_0 .

Здесь использованы следующие обозначения:

$$I_{rM}[i] = \int_0^T I_r(t+iT)dt$$
, $I_{r1}[i] = \int_0^T I_r(t+iT)p_1(t)dt$,
 $I_{r3}[i] = \int_0^T I_r(t+iT)p_3(t)dt$.

С учетом выражений (4)...(6) формулу (3) можно представить в виде:

$$\int_{0}^{T_{0}} I_{r}(t)I(t)dt = \sum_{i=0}^{L-1} (a_{i-1}I_{rM}[i] + (a_{i} - a_{i-1})I_{r1}[i]) - c_{2}\sum_{i=0}^{L-1} (a_{i-1} + a_{i})I_{r3}[i] - c_{3}\sum_{i=0}^{L-1} a_{i} \mid b_{i-1} + b_{i} \mid I_{r3}[i+1/2].$$

Аналогично для второго слагаемого формулы (2) в рамках рассматриваемого класса радиосигналов получаем выражение:

$$\begin{split} &\int_{0}^{L_{0}} Q_{r}(t)Q(t)dt = \\ &= \sum_{i=0}^{L-1} (b_{i-1}Q_{rM}[i+1/2] + (b_{i}-b_{i-1})Q_{r1}[i+1/2]) - \\ &- c_{2}\sum_{i=0}^{L-1} (b_{i-1}+b_{i})Q_{r3}[i+1/2] - c_{3}\sum_{i=0}^{L-1} b_{i-1} \mid a_{i-1}+a_{i} \mid Q_{r3}[i]. \\ & \text{Здесь использованы следующие обозначения:} \end{split}$$

$$Q_{rM}[i+1/2] = \int_0^T Q_r(t+iT+T/2)dt ,$$

$$Q_{r1}[i+1/2] = \int_0^T Q_r(t+iT+T/2)p_1(t)dt ,$$

$$Q_{r_3}[i+1/2] = \int_0^T Q_r(t+iT+T/2)p_3(t)dt \, .$$

Энергия комплексной огибающей эталонного сигнала определяется формулой:

$$\int_{0}^{T_{0}} |s(t)|^{2} dt = \int_{0}^{T_{0}} I(t)I(t)dt + \int_{0}^{T_{0}} Q(t)Q(t)dt$$

Представим синфазную и квадратурную составляющие комплексной огибающей эталонного сигнала соответственно в виде выражений:

$$I(t) = \sum_{i=0}^{N-1} x_i(t-iT) - c_3 \sum_{i=0}^{N-1} u_i p_3(t-iT-T/2),$$

$$Q(t) = \sum_{i=0}^{N-1} y_i(t-iT-T/2) - c_3 \sum_{i=0}^{N-1} v_i p_3(t-iT).$$

Здесь

$$u_{i} = a_{i} | b_{i-1} + b_{i} |, v_{i} = b_{i-1} | a_{i-1} + a_{i} |,$$

$$x_{i}(t) = (a_{i-1} + (a_{i} - a_{i-1})p_{1}(t) - c_{2}(a_{i-1} + a_{i})p_{3}(t)) \cdot \operatorname{rect}(t/T),$$

$$y_{i}(t) = (b_{i-1} + (b_{i} - b_{i-1})p_{1}(t) - c_{2}(b_{i-1} + b_{i})p_{3}(t)) \cdot \operatorname{rect}(t/T).$$

С учетом введенных обозначений энергия комплексной огибающей эталонного сигнала равна

$$\int_{0}^{T_{0}} |s(t)|^{2} dt = \sum_{i=0}^{L-1} \int_{0}^{T} x_{i}^{2}(t) dt + \sum_{i=0}^{L-1} \int_{0}^{T} y_{i}^{2}(t) dt + c_{3}^{2} \sum_{i=0}^{L-1} \int_{0}^{T} (u_{i} p_{3}(t))^{2} dt + c_{3}^{2} \sum_{i=0}^{L-1} \int_{0}^{T} (v_{i} p_{3}(t))^{2} dt - c_{3} \sum_{i=0}^{L-1} \int_{0}^{T} x_{i}(t) (u_{i-1} p_{3}(t+T/2) + u_{i} p_{3}(t-T/2)) dt - c_{3} \sum_{i=0}^{L-1} u_{i} \int_{0}^{T} p_{3}(t) (x_{i+1}(t-T/2) + x_{i}(t+T/2)) dt - c_{3} \sum_{i=0}^{L-1} \int_{0}^{T} y_{i}(t) (v_{i+1} p_{3}(t-T/2) + v_{i} p_{3}(t+T/2)) dt - c_{3} \sum_{i=0}^{L-1} \int_{0}^{T} y_{i}(t) (v_{i+1} p_{3}(t-T/2) + v_{i} p_{3}(t+T/2)) dt - c_{3} \sum_{i=0}^{L-1} v_{i} \int_{0}^{T} p_{3}(t) (y_{i-1}(t+T/2) + v_{i} (t-T/2)) dt - c_{3} \sum_{i=0}^{L-1} v_{i} \int_{0}^{T} p_{3}(t) (y_{i-1}(t+T/2) + v_{i} (t-T/2)) dt - c_{3} \sum_{i=0}^{L-1} v_{i} \int_{0}^{T} p_{3}(t) (y_{i-1}(t+T/2) + v_{i} (t-T/2)) dt - c_{3} \sum_{i=0}^{L-1} v_{i} \int_{0}^{T} p_{3}(t) (y_{i-1}(t+T/2) + v_{i} (t-T/2)) dt - c_{3} \sum_{i=0}^{L-1} v_{i} \int_{0}^{T} p_{3}(t) (y_{i-1}(t+T/2) + v_{i} (t-T/2)) dt - c_{3} \sum_{i=0}^{L-1} v_{i} \int_{0}^{T} p_{3}(t) (y_{i-1}(t+T/2) + v_{i} (t-T/2)) dt - c_{3} \sum_{i=0}^{L-1} v_{i} \int_{0}^{T} p_{3}(t) (y_{i-1}(t+T/2) + v_{i} (t-T/2)) dt - c_{3} \sum_{i=0}^{L-1} v_{i} \int_{0}^{T} p_{3}(t) (y_{i-1}(t+T/2) + v_{i} (t-T/2)) dt - c_{3} \sum_{i=0}^{L-1} v_{i} \int_{0}^{T} p_{3}(t) (y_{i-1}(t+T/2) + v_{i} (t-T/2)) dt - c_{3} \sum_{i=0}^{L-1} v_{i} \int_{0}^{T} p_{3}(t) (y_{i-1}(t+T/2) + v_{i} (t-T/2)) dt - c_{3} \sum_{i=0}^{L-1} v_{i} \int_{0}^{T} p_{3}(t) (y_{i-1}(t+T/2) + v_{i} (t-T/2)) dt - c_{3} \sum_{i=0}^{L-1} v_{i} \int_{0}^{T} p_{3}(t) (y_{i-1}(t+T/2) + v_{i} (t-T/2)) dt - c_{3} \sum_{i=0}^{L-1} v_{i} \int_{0}^{T} p_{3}(t) (y_{i-1}(t+T/2) + v_{i} (t-T/2)) dt - c_{3} \sum_{i=0}^{L-1} v_{i} \int_{0}^{T} p_{3}(t) (y_{i-1}(t+T/2) + v_{i} (t-T/2)) dt - c_{3} \sum_{i=0}^{L-1} v_{i} \int_{0}^{T} p_{3}(t) (y_{i-1}(t+T/2) + v_{i} (t-T/2)) dt - c_{3} \sum_{i=0}^{L-1} v_{i} \int_{0}^{T} p_{3}(t) (y_{i-1}(t+T/2) + v_{i} (t-T/2)) dt - c_{3} \sum_{i=0}^{L-1} v_{i} \int_{0}^{T} v_{i} (t-T/2) dt - c_{3} \sum_{i=0}^{L-1} v_{i} (t-T/2) dt - c_{3} \sum_{i=0}^{L-1} v_{i} (t-T/2) dt - c_{$$

С учетом монотонности экспоненциальной функции задача максимизации функции $\Lambda(r(t))$ эквивалентна задаче минимизации функции вида:

$$\rho[L] = \int_0^{T_0} |s(t)|^2 dt - 2 \int_0^{T_0} I_r(t) I(t) dt - 2 \int_0^{T_0} Q_r(t) Q(t) dt .$$

Анализ данного выражения с учетом слагаемых показывает, что при расчете метрик на интервале времени [iT, (i+1)T) учет межканальной связи требует наличия сведений об информационных символах $a_{i-1}, a_i, a_{i+1}, b_{i-2}, b_{i-1}, b_i$ и существенно затрудняет обработку на приемной стороне.

В этой связи предлагается рассмотреть использование квазиоптимального универсального алгоритма детектирования спектрально-эффективных радиосигналов, где в качестве эталонного сигнала применяется отличающийся от передаваемого отсутствием управляемой связи между синфазной и квадратурной составляющими, т.е. значение коэффициента $c_3 = 0$.

В этом случае выражение для $\rho[L]$ упрощается до следующего вида:

$$\rho[L] = \sum_{i=0}^{L-1} \int_0^T x_i^2(t) dt + \sum_{i=0}^{L-1} \int_0^T y_i^2(t) dt -$$

$$-2\sum_{i=0}^{L-1} (a_{i-1}I_{rM}[i] + (a_i - a_{i-1})I_{r1}[i]) +$$

$$+2c_2\sum_{i=0}^{L-1} (a_{i-1} + a_i)I_{r3}[i] -$$

$$-2\sum_{i=0}^{L-1} (b_{i-1}Q_{rM}[i+1/2] + (b_i - b_{i-1})Q_{r1}[i+1/2]) +$$

$$+2c_2\sum_{i=0}^{L-1} (b_{i-1} + b_i)Q_{r3}[i+1/2].$$

При практической реализации вычисление на приемной стороне значения $\rho[L]$ целесообразно вести итеративно согласно формуле

$$\rho[L] = \sum_{i=0}^{L-1} \Delta \rho_i .$$
Здесь
(7)

$$\Delta \rho_{k} = \int_{0}^{T} x_{k}^{2}(t) dt + \int_{0}^{T} y_{k}^{2}(t) dt - 2(a_{k-1}I_{rM}[k] + (a_{k} - a_{k-1})I_{r1}[k]) + 2c_{2}(a_{k-1} + a_{k})I_{r3}[k] - (b_{k-1}Q_{rM}[k+1/2] + (b_{k} - b_{k-1})Q_{r1}[k+1/2]) + (b_{k} - b_{k-1})Q_{r1}[k+1/2] + (b_{k} - b_{k-1})Q_{r3}[k+1/2].$$

Можно показать, что

$$\int_{0}^{T} x_{i}^{2}(t)dt = 3(a_{i} - a_{i-1})^{2}T / 8 + 3c_{2}^{2}(a_{i} + a_{i-1})^{2}T / 8 + a_{i-1}a_{i}T - c_{2}a_{i}(a_{i-1} + a_{i})T - c_{2}(a_{i}^{2} - a_{i-1}^{2})T / 2,$$

$$\int_{0}^{T} y_{i}^{2}(t)dt = 3(b_{i} - b_{i-1})^{2}T / 8 + 3c_{2}^{2}(b_{i} + b_{i-1})^{2}T / 8 + b_{i-1}b_{i}T - c_{2}b_{i}(a_{i-1} + b_{i})T - c_{2}(b_{i}^{2} - b_{i-1}^{2})T / 2.$$

В случае четырех позиционных радиосигналов, к которым, в частности, относятся T-OQPSK, FQPSK и GMSK сигналы, $a_k = \{\pm 1\}$. Тогда данные выражения приводятся к следующему виду:

$$\int_{0}^{T} x_{i}^{2}(t)dt = T(2 + (1 + 3c_{2}^{2} - 4c_{2})(a_{i}a_{i-1} + 1)) / 4,$$

$$\int_{0}^{T} y_{i}^{2}(t)dt = T(2 + (1 + 3c_{2}^{2} - 4c_{2})(b_{i}b_{i-1} + 1)) / 4.$$

Дополнительно введем обозначения

$$I_{rf1}[k] = I_{rM}[k] - I_{r1}[k] - c_2 I_{r3}[k],$$

$$I_{rf2}[k] = I_{r1}[k] - c_2 I_{r3}[k],$$

$$Q_{rf1}[k + \frac{1}{2}] = Q_{rM}[k + \frac{1}{2}] - Q_{r1}[k + \frac{1}{2}] - c_2 Q_{r3}[k + \frac{1}{2}],$$

$$Q_{rf2}[k + \frac{1}{2}] = Q_{r1}[k + \frac{1}{2}] - c_2 Q_{r3}[k + \frac{1}{2}],$$
(9)

С учетом сделанного допущения и введенных обозначений значение $\Delta \rho_k$ вычисляется согласно формулы:

$$\Delta \rho_{k} = T + T(1 + 3c_{2}^{2} - 4c_{2})(b_{k-1}b_{k} + a_{k-1}a_{k} + 2)) \div$$

$$\div 4 - 2(a_{k-1}I_{rf1}[k] + a_{k}I_{rf2}[k] +$$

$$+b_{k-1}Q_{rf1}[k + \frac{1}{2}] + b_{k}Q_{rf2}[k + \frac{1}{2}]).$$
(10)

Формула (7) в совокупности с выражением (10) позволяет реализовать квазиоптимальную обработку спектрально-эффективных радиосигналов на основе алгоритма Витерби. При этом используется автомат с четырьмя начальными состояниями, определяемыми значениями символов a_{k-1} и b_{k-1} . Из каждого состояния возможно четыре перехода согласно значениям симво-

лов a_{i} и b_{i} . Таким образом, в данной структуре в отличие от алгоритма обработки, описанного в [7], сокращено число состояний, что снижает вычислительную сложность при практической реализации. Основное снижение требуемых аппаратных ресурсов достигается за счет реализации процедуры принятия решения на основе результатов фильтрации синфазной и квадратурной составляющих (согласно формуле (10)) вместо прямого вычисления значений расстояний между принятой смесью и каждой из возможных реализаций сигнала. Данный подход позволяет вести детектирование по двухканальной схеме обработки синфазной и квадратурной составляющих принятого радиосигнала в отличие от алгоритма, описанного в [7], число каналов которого составляет не менее 8-и и кроме того экспоненциально увеличивается с повышением позиционности принимаемого радиосигнала. Таким образом, предложенный алгоритм детектирования спектрально-эффективных радиосигналов по сравнению с оптимальным позволяет сократить до четырех раз вычислительные затраты.

Вопросы практической реализации

Целесообразно рассмотреть две реализации процедуры фильтрации синфазной и квадратурной составляющих. Первый вариант подразумевает получение значений $I_{rf1}[k]$ и $I_{rf2}[k]$, $Q_{rf1}[k + \frac{1}{2}]$ и $Q_{rf2}[k + \frac{1}{2}]$ согласно формул (8) и (9). При этом в каждом квадратурном канале используется три фильтра – Φ_M , Φ_1 и Φ_3 , импульсные характеристики которых равны соответственно:

$$h_M(t) = \operatorname{rect}(t/T), \ h_1(t) = p_1(T-t) \cdot \operatorname{rect}(t/T),$$

 $h_2(t) = p_2(T-t).$

Второй вариант реализации подразумевает получение значений $I_{rf1}[k]$ и $I_{rf2}[k]$, $Q_{rf1}[k + \frac{1}{2}]$ и $Q_{rf2}[k + \frac{1}{2}]$ непосредственно из сигналов $I_r(t)$ и $Q_r(t)$ при помощи фильтров Φ_{f1} и Φ_{f2} с импульсными характеристиками вида:

$$h_{f1}(t) = p_1(t) \operatorname{rect}(t/T) - c_2 p_3(t) ,$$

$$h_{f2}(t) = (1 - p_1(t)) \operatorname{rect}(t/T) - c_2 p_3(t) .$$

Функциональные схемы, реализующие эти два подхода к реализации предложенного алгоритма обработки спектрально-эффективных радиосигналов с управляемой связью между синфазной и квадратурной составляющими, показаны на рис. 1 и на рис. 2.

Здесь ОГ – опорный генератор, ЛЗ – линия задержки на величину *T* / 2, ключевые схемы работают под управлением непоказанной системы восстановления тактовой частоты, генерирующей импульсы с периодом *T*.

Преимущество первого подхода заключается в возможности в ходе сеанса связи динамически изменять параметры полезного радиосигнала. Вторая схема лишена данного свойства, но позволяет сократить требуемые на выполнение операции фильтрации аппаратные ресурсы.

Проведем сравнение по числу требуемых вычислительных операций предложенного алгоритма и оптимального алгоритма, описанного в [7]. Оптимальное устройство детектирования согласно работе [7] при анализе каждого символьного интервала должно производить расчет расстояний между принятой смесью с 32-я



Рис. 1 – Вариант схемы предложенного квазиоптимального алгоритма приема спектрально-эффективных радиосигналов



Рис. 2 – Вариант схемы предложенного квазиоптимального алгоритма приема спектрально-эффективных радиосигналов

Таблица	1.	Количество	вычислительных	затрат
---------	----	------------	----------------	--------

Вид операции	Оптимальный алго-	Трехканальная реализация	Двухканальная реализа-
	ритм		ция
Умножений	$16 \times q$	$4 \times q + 2$	$4 \times q$
Сложений	$16 \times (q-1) + 64$	$4 \times (q-1) + 2 \times q + 6 + 32$	$4 \times (q-1) + 32$
Суммарно	$32 \times q + 48$	$10 \times q + 34$	$8 \times q + 28$

возможными реализациями опорного сигнала. При этом с учетом свойств элементарных импульсов $p_1(t)$ и $p_3(t)$, используемых для формирования РУССКС, можно показать, что в процессе детектирования для каждой квадратурной компоненты обрабатываемого сигнала необходимо вычислить всего по 8 коэффициентов корреляций.

При рассматриваемых условиях для реализации оптимального детектора РУССКС требуется выполнить $(q \times 8 \times 2)$ операций умножения с накоплением, где $q = T \cdot f_s$, f_s – частота дискретизации принятой смеси. Далее для получения искомых расстояний требуется к коэффициентам корреляции прибавить значения, учитывающие энергию представленных выше функций, т.е. требуется еще 32 операции сложения (8 исходных состояний, из каждого которого возможно 4 перехода [7]). Для определения «выживших» путей алгоритму Витерби необходимо произвести учет предыстории, что сопряжено с дополнительным выполнением 32-х операций сложения.

Трехканальная реализация предложенного алгоритма детектирования РУССКС требует на выполнение фильтрации в блоках Φ_1 и Φ_3 суммарно ($q \times 2 \times 2$) операций умножения с накоплением и q×2 операций сложения в блоках Ф_м. Кроме того, для формирования отсчетов $I_{rf1}[k]$ и $I_{rf2}[k]$, $Q_{rf1}[k+\frac{1}{2}]$ и $Q_{rf2}[k+\frac{1}{2}]$ необходимо выполнить еще 6 операций сложения и 2 операции умножения (см. рис. 1). Двухканальная реализация предложенного алгоритма детектирования РУССКС требует на выполнение процедур фильтрации $(q \times 2 \times 2)$ операций умножения с накоплением (см. рис. 2). Для получения искомых расстояний, на основе результатов фильтрации для учета энергии возможных переходов вне зависимости от рассматриваемой схемы, необходимо дополнительно произвести 16 операций сложения (4 исходных состояния, из каждого которого возможно 4 перехода). Для определения «выживших» путей алгоритму Витерби требуется произвести еще 16 операций сложения.

Общие количества вычислительных операций, выполняемых за каждый символьный интервал при оптимальном приеме РУССКС, а также при трехканальной и двухканальной реализациях предложенного алгоритма детектирования, сведены в табл. 1.

Анализ данных табл. 1 показывает, что минимальный выигрыш в вычислительных затратах предложенного алгоритма детектирования РУССКС (*q* = 2) по сравнению с оптимальным составляет для трехканальной схемы более 2 раз, а для двухканальной – более 2,5 раз. По мере увеличения величины *q* для показанной на рис. 1 схемы данный выигрыш стремится к 3,2 раз, а для схемы, приведенной на рис. 2, – к 4 раз.

Исследование характеристик предложенного алгоритма

Для исследования помехоустойчивости предлагаемого квазиоптимального алгоритма проведено компьютерное моделирование, в рамках которого производен подсчет частоты возникновения битовых ошибок. При этом длина окна алгоритма Витерби равнялась 10, а статистика набиралась до тех пор, пока не будет накоплено 100 битовых ошибок. В качестве тестовых радиосигналов использовались сигналы с GMSK, FQPSK и T-OQPSK, а также специально синтезированные в [11] РУССКС по критерию минимума пик-фактора (П_{тий}) и по

критерию минимума эффективной полосы частот (β_{\min}). Результаты моделирования сведены в табл. 2 в виде проигрышей δg_S , δg_K и δg_{FQPSK} в помехоустойчивости (при $P_0 = 10^{-4}$) оптимальному алгоритму [7] соответственно предложенного устройства, устройства, обоснованного в [8], и FQPSK-детектора, описанного в [12].

№	Вид сигнала	$\delta g_{\scriptscriptstyle S}$, дБ	$\delta g_{\scriptscriptstyle K}$, дБ	$\delta g_{\rm FQPSK},$ дБ
1	GMSK BT=0,3	0,0	0,2	0,5
2	GMSK BT=0,4	0,0	0,2	0,6
3	GMSK BT=0,5	0,0	0,0	0,7
4	FQPSK	0,3	0,9	0,8
5	T-OQPSK	0,0	1,3	0,8
6	Π_{\min}	0,2	0,5	0,5
7	eta_{\min}	0,0	2,6	1,4

Таблица 2. Проигрыш в помехоустойчивости

Кроме того, на рис.3, а...3, г для примера приведены кривые помехоустойчивости предложенного (кривые 1), описанного в [8] (кривые 2) и оптимального (кривые 3) детекторов [7] РУССКС.

Анализ представленных на рис. З зависимостей и данных табл. 2 показывает, что предложенный метод обработки выигрывает (до 1,3 дБ в случае Т-OQPSK-сигнала и до 2,6 дБ в случае синтезированных по критерию минимума полосы частот радиосигналов при $P_0=10^{-4}$) у одноканального метода, описанного в [8]. При этом проигрыш оптимальному методу приема зависит от глубины управляемой связи между синфазной и квадратурной составляющими полезного радиосигнала, но на рассматриваемой выборке не превышает 0,3 дБ в случае с FQPSK-сигналом при $P_0=10^{-4}$.



Рис.3. Характеристики помехоустойчивости синтезированного алгоритма обработки спектрально-эффективных радиосигналов: FQPSK (под а), GMSKBT = 0,25 (под б), GMSKBT = 0,5 (под в), T-OQPSK (под г)

Заключение

В настоящей статье рассмотрена возможность для приема спектрально-эффективных радиосигналов (T-OQPSK, GMSK, FQPSK и специально синтезированных в [11] РУССКС) использовать на приемной стороне эталонный сигнал, отличающийся от полезного отсутствием связи между синфазной и квадратурной составляющими комплексной огибающей. Показано, что в подобном случае расчет расстояний в алгоритме Витерби процесса детектирования сводится к процедуре двухканальной или трехканальной фильтрации, что позволяет по сравнению с оптимальной схемой до 4-х или до 3,2 раз, соответственно, сократить требуемые вычислительные ресурсы блока обработки. При этом структура и сложность реализации узла расчета расстояний не зависят от позиционности принимаемых радиосигналов. Методом компьютерного моделирования показано, что предложенный двухканальный универсальный алгоритм детектирования спектрально-эффективных радиосигналов характеризуется незначительным проигрышем оптимальному алгоритму обработки, который на множестве рассматриваемых радиосигналов составил менее 0,3 дБ при вероятности битовой ошибки на уровне

 $P_0 = 10^{-4}$. При этом выигрыш по данному показателю качества относительно одноканального алгоритма достигал 2,6 дБ при тех же условиях.

Таким образом, на основе представления в виде РУССКС спектрально-эффективных радиосигналов, таких как T-OQPSK, GMSK и FQPSK, обоснован двухканальный универсальный алгоритм детектирования, который целесообразно использовать наряду с универсальным формирователем, описанным в [6], для реализации перспективных когнитивных систем передачи информации.

Работа выполнена при поддержке гранта Президента РФ № МК-36538.2018.9.

Литература

1. Mitola J. Cognitive radio for flexible mobile multimedia communications // Mobile Multimedia Communications. 1999. (MoMuC'99) 1999 IEEE International Workshop on. IEEE. 1999. pp. 3-10.

2. Sun H. et al. Wideband spectrum sensing for cognitive radio networks: a survey // IEEE Wireless Communications. 2013. Vol. 20. no. 2. pp. 74-81.

3. Wygliski A.M., Nekovee M. HouY.Th. Cognitive Radio Communications and Networks. Principles and Practice. – London: Elsevier, 2010 – 714 p.

4. Кириллов С.Н., Бердников В.М., Покровский П.С, Семин Д.С. Проблемно-ориентированные платформы для реализации универсальных, адаптивных, структурно-защищенных радиосистем передачи информации // Радиотехника, № 5, 2015. С. 6-12.

5. Simon M.K. Bandwidth-Efficient Digital Modulation with Application to Deep-Space Communications. Jet Propulsion Laboratory. California Institute of Technology. URL: https://descanso.jpl.nasa.gov/monograph/series3/complete1.pdf

6. Кириллов С.Н., Покровский П.С. Программноуправ-ляемый формирователь радиосигналов с нелинейными видами модуляции // Нелинейный мир, № 3, 2013. С. 150-157.

7. Покровский П.С. Процедура детектирования радиосигналов с управляемой связью между квадратурными составляющими // Вестник Рязанского государственного радиотехнического университета. 2013. № 3 (45). С. 110-113. 8. Покровский П.С. Синтез квазиоптимального алгоритма детектирования спектрально-эффективных радиосигналов // Материалы XXII ВНТК «Новые информационные технологии в научных исследованиях (НИТ-2017)». Рязань: РГРТУ, 2017. С. 113-114.

9. Прокис Дж. Цифровая связь. Пер с англ. / Под ред. Д.Д. Кловского. – М.: Радио и связь, 2000 – 800 с.

10. Кириллов С.Н., Покровский П.С. Двухкритериальный синтез шестнадцатипозиционных радиосигналов с управляемой связью между синфазной и квадратурной составляющими // Успехи современной радиоэлектроники. №6. 2014. С. 18-25.

11. Покровский П.С. Процедура синтеза радиосигналов с управляемой связью между квадратурными составляющими по двум показателям качества // Вестник Рязанского государственного радиотехнического университета. № 2 (выпуск 52). 2015. С.49-55.

12. Феер К. Беспроводная цифровая связь. Методы модуляции и расширения спектра. – М.: Радио и связь, 2000. – 520 с.



ичеговал овга ИГНАЛОВ Афанасьев А.А.

НОВЫЕ КНИГИ

Брюханов Ю.А.

Цифровые цепи и сигналы / Учебное пособие – М.: Изд-во «Горячая линия-Телеком», 3-изд. перераб. и доп., 2017 г. – 160 с.: ил.

Приведен математический аппарат для анализа сигналов и цепей дискретного времени. Изложены спектральная теория периодических и непериодических цифровых сигналов, теория цепей дискретного времени. Подробно рассмотрены частотные свойства и временные характеристики базовых нерекурсивных и рекурсивных линейных цепей (цифровых фильтров) первого и второго порядков. Изложены методы изменения частоты дискретизации цифровых сигналов: децимация и интерполяция, а также способы их реализации. Уделено внимание теории случайных сигналов и процессов дискретного времени, протекающих в цифровых цепях. Рассмотрены эффекты квантования в цифровых сигналах и цепях.

Для студентов вузов, обучающихся по направлению подготовки «Радиофизика», «Радиотехника» и «Информационные технологии и системы связи».

Афанасьев А.А., Рыболовлев А.А., Рыжков А.П.

Учебное пособие для вузов – М.: Изд-во «Горячая линия-Телеком», 2017 г. – 356 с.: ил.

Изложены базовые теоретические и практические вопросы в области анализа, синтеза и экспериментального исследования систем цифровой обработки сигналов (ЦОС). Приведены основные методы расчета характеристик устройств ЦОС, показаны особенности их анализа в различных условиях функционирования, даны основные подходы к синтезу систем ЦОС, рассмотрены основные происходящие в них физические процессы. Изложенные теоретические сведения сопровождаются практическими примерами решения прикладных задач.

Для студентов вузов, обучающихся по укрупненной группе направлений подготовки 11.00.00 – «Электроника, радиотехника и системы связи».

УДК 004.934

ПОСТРОЕНИЕ КОМБИНИРОВАННОГО ПОКАЗАТЕЛЯ КАЧЕСТВА РЕЧЕВЫХ СИГНАЛОВ ДЛЯ ОЦЕНКИ ТОЧНОСТИ ИДЕНТИФИКАЦИИ ДИКТОРОВ В УСЛОВИЯХ ВОЗДЕЙСТВИЯ АДДИТИВНЫХ ШУМОВ

Тупицин Г.С., к.т.н., сотрудник Ярославского государственного университета им. П.Г. Демидова, e-mail: genichyar@genichyar.com;

Топников А.И., к.т.н., доцент Ярославского государственного университета им. П.Г. Демидова, e-mail: topartgroup@gmail.com;

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

DEVELOPMENT OF COMBINED SPEECH QUALITY MEASURE FOR SPEAKER IDENTIFICATION ACCURACY ESTIMATION IN ADDITIVE NOISE ENVIRONMENTS

Tupitsin G.S., Topnikov A.I., Priorov A.L.

The traditional approach of using the real speaker identification system and an extensive database of test speech signals requires large computational resources. The paper is dedicated to the development of an alternative way to the speaker identification accuracy estimation in additive noise environments. In this paper an indirect speaker identification accuracy estimation technique using some objective speech quality measures and combined measure was proposed. It was shown the proposed combined speech quality measure could be used in the problem of speaker identification accuracy estimation in additive noise environments.

Key words: additive noise, speaker identification accuracy, denoising, speech quality.

Ключевые слова: аддитивный шум, точность идентификации дикторов, шумоподавление, качество речи.

Введение

В последнее десятилетие наблюдается значительный рост интереса к алгоритмам и методам распознавания диктора, что связано с ком-

мерческой значимостью систем биометрии. В идеализированных условиях показатели работы таких систем близки к предельно достижимым. Однако специфика использования систем идентификации по голосу вносит свои коррективы. И наиболее значимые задачи, стоящие перед исследователями в данной области, – это повышение устойчивости систем идентификации к воздействию аддитивных акустических шумов [1, 2] и к так называемым «спуфинг атакам» [3, 4]. Для решения первой задачи часто применяют шумоподавление [1, 2, 5– 8]. При подборе параметров алгоритмов предобработки для системы идентификации диктора естественным показателем качества работы системы является точность идентификации диктора (ТИД), которая вычисляется следующим образом:

$$I = \frac{a}{\Sigma} ,$$

где I – точность идентификации дикторов, a – число правильно идентифицированных тестовых сигналов, Σ – общее число тестовых сигналов. Такой подход выглядит наиболее естественным, но имеет один существенный недостаток – трудоемкость, которая выражается в высоких вычислительных затратах и необходимости многократного запуска алгоритма идентифика

Работа посвящена вопросу косвенной оценки точности идентификации дикторов в условиях воздействия аддитивных акустических шумов. Исследуется оценка на основе отдельных объективных показателей качества. Предлагается комбинированный показатель, основанный на линейной комбинации показателей качества и позволяющий повысить точность косвенной оценки точности идентификации дикторов.

> ции. Кроме того, этот метод требует знания априорной информации для подсчета правильно идентифицированных сигналов. Поэтому поиск альтернативных методик оценки ТИД является актуальным.

> Таким образом, целью исследования является создание методики оценки точности идентификации диктора, позволяющей упростить процесс подбора параметров алгоритмов шумоподавления, используемых в системах распознавания диктора.

Система идентификации диктора и базы речевых сигналов

Для проведения исследований выбрана система идентификации диктора, использующая мел-частотные кепстральные коэффциенты (МЧКК, Mel-Frequency Cepstral Coefficients – MFCC) в качестве признаков и модели гауссовых смесей (МГС, Gaussian Mixture Models – GMM) для реализации классификатора. Для расчета моделей используется универсальная фоновая модель (УФМ, Universal Background Model – UBM), которая обучается на фонограммах большого числа дикторов, записанных в различных акустических условиях [2]. Для ускорения вычислений используется диагональная аппроксимация матрицы ковариаций. Универсальная фоновая модель преобразуется в модель гауссовых смесей обучаемого диктора посредством МАР-адаптации (Maximum a Posteriori adaptation). На этапе классификации рассчитываются значения функции правдоподобия для гипотезы, что входной сигнал принадлежит одной из обученных моделей диктора. Модель, для которой значение функции правдоподобия максимально, принимается за искомую. Все вычисления производились в пакете Matlab.

Для проведения исследований выбрано две базы речевых сигналов. Универсальные фоновые модели строятся отдельно для каждой базы, что позволяет уменьшить число компонент гауссовых смесей при моделировании диктора по сравнению со случаем, когда используется единая УФМ для обеих баз.

> Таблица 1. Характеристики баз речевых сигналов и соответствующих УФМ

Параметр	РУС-31-5	АНГЛ-20-5
Число дикторов (муж- чин/женщин)	31 (31/0)	20 (12/8)
Длительность обучающего сигнала, с	90	72–106
Длительность тестового сиг- нала, с	3	2–3
Число тестовых сигналов	5	5
Число дикторов, используе- мых для получения УФМ (мужчин/женщин)	132 (132/0)	16 (8/8)
Длительность записи голоса каждого диктора, используе- мого для получения УФМ, с	10	35–52

Первая база составлена из речевых сигналов на русском языке. Стиль произношения – нейтральный. Вторая база речевых сигналов на английском языке составлена на основе записей речевого корпуса СНАINS школы компьютерных наук и информатики университетского колледжа Дублина [9]. Данный корпус содержит записи 36 дикторов в различных стилях произношения, из которых в работе используется лишь стиль одиночного чтения (маркируется в корпусе как «Solo reading»). Характеристики баз речевых сигналов «РУС-31-5» и «АНГЛ-20-5», а также соответствующих УФМ, приведены в табл. 1.

Для зашумления речевых сигналов используется аддитивный белый гауссовский шум (АБГШ), генерируемый при помощи встроенной функции пакета Matlab, а также записи реальных шумов из библиотеки Noisex-92: шум в салоне автомобиля (Vehicle Interior Noise – VIN) и шум большого количества одновременно разговаривающих людей (Speech Bubble – SB) [10].

Исследование показателей качества

Основой для разработки методики косвенной оценки ТИД служит предположение, что чем выше качество входных речевых сигналов, тем выше точность идентификации дикторов. Поэтому на первом шаге исследования предлагается оценить корреляцию ТИД со значениями широко используемых объективных показателей качества: PESQ, отношение сигнал/шум (ОСШ), сегментное ОСШ (СегОСШ), WSS, LLR [11–13].

Для определения массива значений показателя качества каждый тестовый сигнал из базы незашумленных сигналов и базы обработанных сигналов попарно подаются на вход блока оценки показателя качества. Результаты усредняются для дальнейшей обработки. Блоксхема данного алгоритма приведена на рис. 1.

Кроме того, предлагается показатель качества на основе признаков речевого сигнала, используемых в исследуемой системе идентификации диктора [14]. В данном случае это МЧКК. Параметры алгоритма извлечения МЧКК выбираются те же, что и для системы идентификации диктора, используемой в исследовании.



Рис. 1. Блок-схема алгоритма оценки усредненного показатели качества для базы тестовых сигналов

В качестве меры близости между МЧКК исследуемого сигнала и МЧКК незашумленного сигнала используется среднее расстояние, которое может быть вычислено по формуле:

$$\mathcal{Q}^{MYKK} = \frac{1}{W} \sum_{w=1}^{W} d(\vec{x}_w^R, \vec{x}_w^A),$$

где \vec{x}_w^R – вектор МЧКК исследуемого сигнала для окна

w, \vec{x}_w^A – вектор МЧКК незашумленного сигнала для окна

w, $d(\vec{x}_w^R, \vec{x}_w^A)$ – расстояние между векторами \vec{x}_w^R и \vec{x}_w^A .

Предлагается использовать следующие методы вычисления расстояния:

Евклидово расстояние – далее «МЧКК-Е»

$$d^{E}(\vec{x}_{w}^{R},\vec{x}_{w}^{A}) = \|\vec{x}_{w}^{R}-\vec{x}_{w}^{A}\|,$$

L1-норма (расстояние городских кварталов, манхэттэнское расстояние) – далее «МЧКК-L1»

$$d^{L1}(\vec{x}_{w}^{R},\vec{x}_{w}^{A}) = \|\vec{x}_{w}^{R} - \vec{x}_{w}^{A}\|_{1},$$

Расстояние Махаланобиса – далее «МЧКК-М»

$$d^{M}(\vec{x}_{w}^{R},\vec{x}_{w}^{A}) = \sqrt{(\vec{x}_{w}^{R}-\vec{x}_{w}^{A})^{T}(\Sigma^{R})^{-1}(\vec{x}_{w}^{R}-\vec{x}_{w}^{A})},$$

где Σ^R – матрица ковариации набора векторов $X^R = \{\vec{x}_1^R, \vec{x}_2^R, ..., \vec{x}_W^R\}$.

При расчете расстояния Махаланобиса используется матрица ковариации набора векторов МЧКК именно зашумленного сигнала, а не эталонного. Это связано с тем, что при использовании матрицы ковариации набора векторов МЧКК эталонного сигнала получить высокой тесноты статистической связи с ТИД не удалось.

Для определения возможности использовать показатели качества речи для оценки точности работы системы идентификации диктора предлагается рассчитать линейные коэффициенты корреляции (коэффициенты корреляции Пирсона) между выборками значений ТИД и соответствующими им значениями показателей качества. Выборки для фиксированного типа шума и базы речевых сигналов можно определить по алгоритму, блок-схема которого представлена на рис. 2.

Каждому значению ТИД сопоставляются усредненные по всем тестовым сигналам значения показателей качества речи. В работе используются две базы речевых сигналов, поэтому выборки для них рассчитываются независимо, а затем объединяются в одну. Для каждого типа шума исследование проводится отдельно.

Линейный коэффициент корреляции между *n*-ым показателем качества речи и ТИД можно оценить по следующей формуле:



Рис. 2. Получение выборок значений ТИД и показателей качества речи

$$r_{n} = \frac{\sum_{t=1}^{T} (Q_{t}^{n} - \overline{Q}^{n})(I_{t} - \overline{I})}{\sqrt{\sum_{t=1}^{T} (Q_{t}^{n} - \overline{Q}^{n})^{2} \sum_{t=1}^{T} (I_{t} - \overline{I})^{2}}},$$

$$\overline{Q}^{n} = \frac{1}{T} \sum_{t=1}^{T} Q_{t}^{n}, \ \overline{I} = \frac{1}{T} \sum_{t=1}^{T} I_{t},$$

где $Q^n = \{Q_1^n, Q_2^n, \dots, Q_T^n\}$ – выборка значений *n*-го показателя качества речи; $I = \{I_1, I_2, \dots, I_T\}$ – выборка значений ТИД; T – объем каждой выборки, который равен произведению числа используемых баз речевых сигналов, значений ОСШ, применяемых алгоритмов предобработки.

Для исследования использовались следующие типы шумов: АБГШ, SB, VIN. Значение ОСШ изменялось от 6 до 15 дБ с шагом 1 дБ. Выбранные алгоритмы предобработки приведены в табл. 2.

N⁰	Описание
1	алгоритм, для которого выходной сигнал равен вход- ному
2	алгоритм на основе ФКС спектрального вычитания [15]
3	алгоритм на основе подхода прямого принятия решения [16] ($\alpha = 0,999$) и ФКС Винера [17]
4	алгоритм на основе подхода прямого принятия решения [16] ($\alpha = 0,9999$) и ФКС Винера [17]
5	алгоритм на основе подхода прямого принятия решения [16] ($\alpha = 0.98$) и ФКС Винера [17]
6	алгоритм на основе подхода прямого принятия реше- ния [16] ($\alpha = 0,98$) и ФКС минимальной средне- квадратичной ошибки кратковременного амплитуд- ного спектра [16]
7	алгоритм на основе подхода прямого принятия реше- ния [16] (α = 0,98) и ФКС минимальной средне- квадратичной ошибки логарифма кратковременного амплитудного спектра [18]
8	двухступенчатый алгоритм шумоподавления [19] ($\alpha = 0.98$) на основе ФКС Винера [17]
9	двухступенчатый алгоритм шумоподавления [19] ($\alpha = 0.98$) на основе ФКС минимальной среднеквад- ратичной ошибки кратковременного амплитудного спектра [16]
10	двухступенчатый алгоритм шумоподавления [19] ($\alpha = 0,98$) на основе ФКС минимальной среднеквад- ратичной ошибки логарифма кратковременного ам- плитудного спектра [18]

Таблица 2. Используемые алгоритмы предобработки

Алгоритмы № 1 и № 2 представляют случаи отсутствия подавления шума и слабого подавления шума соответственно. Алгоритмы № 3 и № 4 представляют случаи избыточного подавления шума, когда помимо точек частотно-временного представления сигнала, содержащих преимущественно шум, ошибочно подавляются точки, содержащие преимущественно речь.

Измерения производились для баз речевых сигналов «РУС-31-5» и «АНГЛ-20-5». Таким образом, при фиксированном типе шума для каждого показателя качества речи и ТИД получена выборка из 200 измерений. Результаты исследования приведены в табл. 3. Значения линейного коэффициента корреляции округлены до сотых. Кроме того, для каждого показателя качества речи рассчитано среднее значение линейного коэффициента корреляции среди используемых типов шума. Наибольшее по модулю значение для определенного типа шума и для среднего выделено жирным шрифтом.

Таблица 3. Значения линейного коэффициента корреляции

Показатель качества	АБГШ	SB	VIN	Среднее
PESQ	0,88	0,72	0,60	0,73
ОСШ	0,87	0,66	0,64	0,72
СегОСШ	0,72	0,80	0,76	0,76
WSS	-0,36	-0,15	-0,60	-0,37
LLR	-0,89	-0,47	-0,29	-0,55
МЧКК-Е	-0,87	-0,50	-0,87	-0,75
МЧКК-L1	-0,78	-0,32	-0,87	-0,66
МЧКК-М	-0,45	-0,56	-0,89	-0,64

Для качественной оценки тесноты статистической связи использована шкала Чеддока. Анализируя результаты, можно отметить, что только СегОСШ имеет высокую тесноту статистической связи с ТИД одновременно для всех используемых типов шума.

В ходе исследования было обнаружено, что значения некоторых показателей качества речи при использовании алгоритмов № 3 и № 4 являются экстремально большими, либо малыми, что снижает значение линейного коэффициента корреляции между данными показателями качества речи и ТИД. В табл. 4 приведены значения линейных коэффициентов корреляции для случая, когда алгоритмы № 3 и № 4 не применялись.

Таблица 4. Значения линейного коэффициента корреляции без использования алгоритмов предобработки № 3 и № 4

Показатель качества	АБГШ	SB	VIN	Среднее
PESQ	0,83	0,65	0,66	0,71
ОСШ	0,87	0,75	0,73	0,79
СегОСШ	0,70	0,78	0,77	0,75
WSS	-0,08	0,16	-0,71	-0,21
LLR	-0,92	-0,27	-0,55	-0,58
МЧКК-Е	-0,87	-0,33	-0,92	-0,71
МЧКК-L1	-0,72	-0,11	-0,91	-0,58
МЧКК-М	-0,85	-0,91	-0,92	-0,89

Можно отметить, что в этом случае показатели качества PESQ, СегОСШ, МЧКК-М обладают высокой теснотой статистической связи с ТИД для всех исследуемых типов шумов. Для МЧКК-М и шума SB и VIN эта связь весьма высокая. Остальные показатели качества не имеют тесноты статистической связи выше заметной для всех типов шумов.

Таким образом, наибольшую тесноту статистической связи с ТИД без использования алгоритмов, избыточно подавляющих шум, имеет предлагаемый показатель качества речи МЧКК-М. При использовании полного набора алгоритмов предобработки наибольшую тесноту статистической связи с ТИД имеет показатель качества речи СегОСШ. Однако значение линейного коэффициента корреляции в обоих случаях может оказаться недостаточно большим для подбора точных параметров алгоритмов шумоподавления. Для повышения точности оценки ТИД предлагается использовать линейную комбинацию значений показателей качества PESQ, ОСШ, СегОСШ, WSS, LRR, МЧКК-Е, МЧКК-L1, МЧКК-М. Похожий подход используется в [20] для создания показателя качества зашумленной речи. Для выявления весов отдельных показателей качества в составе комбинированного используется линейная регрессия. Для обозначения нового комбинированного показателя качества будем использовать название «альтернативная оценка точности идентификации дикторов» (АОТИД). Более качественного результата удается достичь, если весовые коэффициенты для показателей качества находить для каждого шума в отдельности. В качестве примера приведем формулу вычисления АОТИД для АБГШ:

	-0,3522 -	$ ^T $	$\left[Q^{PESQ} \right]$	
	0,0300		Q^{OCIII}	
	-0,0161		Q ^{СегОСШ}	
	-0,0106		Q^{WSS}	
$Q^{AOTU A - A B \Gamma I I I} =$	-0,1503		Q^{LLR}	
	-0,3364		$Q^{MYKK-\Im}$	
	0,1613		$Q^{MYKK-L1}$	
	0,0023		Q^{MYKK-M}	
	0,2241		1	

Для оценки применимости предложенного подхода рассчитаны линейные коэффициенты корреляции между АОТИД и ТИД. Для трех шумов, используемых в исследовании, они составляют 0,96 – для АБГШ, 0,94 – для SB и 0,97 – для VIN.

Диаграммы рассеяния для новых показателей качества и соответствующих типов шума приведены на рис. 3. Анализируя взаимосвязь ТИД и АОТИД, можно отметить, что для АБГШ и шума SB она близка к линейной. Однако для шума VIN при АОТИД менее 0,43 и более 1 имеются более значительные неточности в оценке ТИД.

Отметим, что выбор именно линейной комбинации показателей качества обусловлен необходимостью выполнения принципа суперпозиции. В данном исследовании АОТИД определяется взвешенной суммой усредненных значений показателей качества среди всех тестовых сигналов в базе. На практике АОТИД будет применяться по-иному: сначала будет производиться расчет значений АОТИД для каждого тестового сигнала, а усредняться будут полученные значения АОТИД.

Заключение

Таким образом, в ходе проведенных исследований установлено, что оценка ТИД на основе значения отдельного показателя качества возможна, но не обладает достаточной для решения практических задач достоверностью. Для совершенствования этого подхода предложено использовать линейную комбинацию показателей качества. Разработанная методика, получившая сокращенное название АОТИД, позволяет получить косвенные оценки ТИД без непосредственного участия системы идентификации диктора. Значения линейного коэффициента корреляции между значениями АОТИД и ТИД в исследование составляют 0,94-0,97 в зависимости от типа шума.



Рис. 3. Диаграммы рассеяния значений АОТИД относительно значений ТИД

Литература

1. Tupitsin G., Topnikov A., Priorov A. Two-step noise reduction based on soft mask for robust speaker identification // Proceedings of the 18th Conference of Open Innovations Association FRUCT and Seminar on Information Security and Protection of Information Technolog, FRUCT-ISPIT 2016. 2016. P. 351–356.

2. Тупицин Г.С. Предобработка речевых сигналов в системах автоматической идентификации диктора: диссертация ... кандидата технических наук: 05.12.04. Владимир: Владимирский государственный университет имени Александра Григорьевича и Николая Григорьевича Столетовых, 2015.

3. Щемелинин В.Л., Симончик К.К. Исследование устойчивости голосовой верификации к атакам, использующим систему синтеза // Известия высших учебных заведений. Приборостроение. 2014. Т. 57, No 2. С. 84–88.

4. Shchemelinin V., Kozlov A., Lavrentyeva G., Novo-

selov S., Simonchik K. Vulnerability of Voice Verification System with STC Anti-spoofing Detector to Different Methods of Spoofing Attacks // Speech and Computer. SPECOM 2015. Lecture Notes in Computer Science. 2015. Vol. 9319. P. 480–486.

5. Тупицин Г.С., Топников А.И., Приоров А.Л. Модификация двухступенчатого алгоритма шумоподавления для улучшения качества идентификации диктора в условиях шумов // Информационные системы и технологии. 2015. Т. 6, № 92. С. 39-47.

6. Тупицин Г.С., Кравцов С.А., Топников А.И., Приоров А.Л. Модификация алгоритма оценки бинарной маски в задаче подавления шума для системы идентификации диктора // Проектирование и технология электронных средств. 2015. № 3. С. 32–37.

7. Тупицин Г.С., Топников А.И., Приоров А.Л. Предобработка зашумленных речевых сигналов с помощью бинарных масок в задаче идентификации диктора // Наукоемкие технологии. 2015. Т. 16, № 11. С. 56–61.

8. Тупицин Г.С., Топников А.И., Приоров А.Л. Методика оценки мягкой маски для задачи предобработки зашумленных речевых сигналов в системах идентификации диктора // Успехи современной радиоэлектроники. 2016. № 6. С. 73–80.

9. Cummins F., Grimaldi M., Leonard T., Simko J. The CHAINS Speech Corpus: CHAracterizing INdividual Speakers // Proc of SPECOM. 2006. P. 1–6.

10. Varga A., Steeneken H.J.M. Assessment for automatic speech recognition: II. NOISEX-92: A database and an experiment to study the effect of additive noise on speech recognition systems // Speech Communication. 1993. V. 12, No 3. P. 247–251.

11. International Telecommunication Union. P.862: Perceptual evaluation of speech quality (PESQ), an objective method for end-to-end speech quality assessment of narrowband telephone networks and speech codecs / International Telecommunication Union // ITU-T Recommendation. 2001.

12. Klatt D. Prediction of perceived phonetic distance from critical-band spectra: A first step // ICASSP '82. IEEE International Conference on Acoustics, Speech, and Signal Processing. – Institute of Electrical and Electronics Engineers. 1982. V. 7, P. 1278–1281.

13. Crochiere R., Tribolet J., Rabiner L. An interpretation of the log likelihood ratio as a measure of waveform coder performance // IEEE Transactions on Acoustics, Speech, and Signal Processing. 1980. V. 28, No. 3. P. 318–323.

14. Тупицин Г.С. Использование расстояния между мел-частотными кепстральными коэффициентами для оценки точности идентификации дикторов // Доклады 18-й международной научно-технической конференции «Проблемы передачи и обработки информации в сетях и системах телекоммуникаций». Рязань, 2015. С. 98–99.

15. Boll S. Suppression of acoustic noise in speech using spectral subtraction // IEEE Transactions on Acoustics, Speech, and Signal Processing. 1979. V. 27, No 2. P. 113–120.

16. Ephraim Y., Malah D. Speech enhancement using a minimum-mean square error short-time spectral amplitude estimator // IEEE Transactions on Acoustics, Speech, and Signal Processing. 1984. V. 32, No 6. P. 1109–1121.

17. Lim J., Oppenheim A. Enhancement and bandwidth compression of noisy speech // Proceedings of the IEEE. 1979. V. 67, No 12. P. 1586–1604.

18. Ephraim Y., Malah D. Speech enhancement using a minimum mean-square error log-spectral amplitude estimator // IEEE Transactions on Acoustics, Speech, and Signal Processing. 1985. V. 33, No 2. P. 443–445.

19. Plapous C., Marro C., Mauuary L., Scalart P. A twostep noise reduction technique // IEEE International Conference on Acoustics, Speech, and Signal Processing. 2004. V. 1, P. 289-292.

20. Hu Y., Loizou P. Evaluation of objective quality measures for speech enhancement // IEEE Transactions on Speech and Audio Processing. V. 16, Is. 1. 2008. V. 16. Is. 1. P. 229–238.

новые книги



Горшков Ю.Г.

Обработка речевых и акустических биомедицинских сигналов на основе вейвлетов. – М.: Изд-во «Радиотехника», 2017 г. – 240 с.: ил.

Изложены новые перспективные методы речевых и акустических биомедицинских сигналов на основе вейвлетов, аппаратно-программные средства для решения прикладных задач криминалистического исследования фонограмм, оценки эмоционального состояния человека, шумоочистки и засекречивания речи, а также построения биомедицинских систем диагностики заболеваний сердца и легких. Основное внимание уделено разработке вейвлет-технологий высокоточного анализа речевых и акустических биомедицинских сигналов. Впервые предложены решения многоуровневого вейвлет-преобразования, обеспечивающие визуализацию нестационарных сигналов с повышенным частотно-временным разрешением. Представлена система телемедицины нового поколения с засекречиванием акустических биомедицинских сигналов.

Для ученых и инженеров, работающих над созданием средств обработки биомедицинских сигналов на новых принципах. Может быть полезна преподавателям, аспирантам и студентам технических вузов.

УДК 004.5; 004.94

СРАВНИТЕЛЬНЫЕ ХАРАКТЕРИСТИКИ АЛГОРИТМОВ ДЕТЕКЦИИ СТАБИЛЬНЫХ ЗРИТЕЛЬНЫХ ВЫЗВАННЫХ ПОТЕНЦИАЛОВ ГОЛОВНОГО МОЗГА НА ЭЛЕКТРОЭНЦЕФАЛОГРАММЕ

Туровский Я.А., к.м.н., доцент кафедры цифровых технологий, руководитель лаборатории информационных технологий в медицине факультета компьютерных наук Воронежского государственного университета e-mail: yaroslav_turovsk@mail.ru.

COMPARATIVE CHARACTERISTICS OF THE ALGORITHMS OF DETECTION STEADY STATE VISUALLY EVOKED POTENTIALS OF THE BRAIN ON AN ELECTROENCEPHALOGRAM

Turovsky Ya.A.

The results of comparison of algorithms for analysis of evoked potentials of the brain used in the design of brain-computer interfaces are presented. For carrying out computational experiments, the Fourier transform algorithm was used; the Multivariate Synchronization Index (MSI) algorithm in various modifications: analysis of the original signal, analysis of the accumulated event relation potential and the spectrum of the accumulated event relation potential; linear correlation with analysis from the reference sample of the signal received from the «user», different variants of wavelet filtration. In this case, model signals were used, representing the addition of white noise and a harmonic oscillation simulating a steady state visually evoked potentials. The most stable results (with a number of errors <10 %) for the 3-second analysis period were demonstrated by the MSI of the original signal, the MSI of the Fourier transform results, MSI where the standard was the result of wavelet coherent accumulation, the linear correlation coefficient and MSI where the standard was reconstructed after the wavelet- transformation of the ERP.

Key words: brain computer interface, simulation, SSVEP.

Ключевые слова: интерфейс мозг компьютер, моделирование, SSVEP.

Введение

Одним из направлений развития цифровой обработки сигналов в нейронауках является создание новых и совершенствование уже существующих методов обработки сигналов головного мозга, полученных различными методами. Традиционно, в силу достаточно широкой распространенности и высокой информативности, значительный интерес представляют методы, использующие алгоритмы обработки электроэнцефалограмм (ЭЭГ) – записей электрической активности мозга. Наряду с клиническими решениями, направленными во многом на оценку и прогноз судорожных состояний пациента, фундаментальных исследований, разработка и исследование новых алгоритмов направлены на системы интерфейсов мозг-компьютер

(нейрокомпьютерные интерфейсы). В них сигналы, зарегистрированные с головного мозга, трансформируются в команды, замещающие внешние устройства ввода: клавиатуру, джойстик или «мышь». При этом выделяют два основных направления разработки интерфейсов мозг-компьютер: создание синхронных и асинхронных интерфейсов. Синхронные интерфейсы используют внешнюю стимуляцию пользователя, например, в виде вспышек света или подсвечивания тех или иных пиктограмм на экране монитора. В ответ на такие стимулы в структуре ЭЭГ появляются так называемые «вызван-

Представлены результаты сравнения алгоритмов анализа вызванных потенциалов головного мозга, используемых при конструировании интерфейсов мозг-компьютер. Для проведения вычислительных экспериментов были взяты алгоритм преобразования Фурье; алгоритм Multivariate Synchronization Index (MSI) в различных модификациях: анализ исходного сигнала, анализ накопленного вызванного потенциала и спектра накопленного вызванного потенциала; линейная корреляция с анализом по эталонному образцу сигнала полученного с «пользователя», различные варианты вейвлетфильтрации. При этом применялись модельные сигналы, представляющие собой сложение белого шума и гармонического колебания, имитирующего стабильный зрительный вызванный потенциал. Наиболее устойчивые результаты (с числом ошибок < 10 %) для 3-х секундной эпохи анализа продемонстрировали MSI исходного сигнала, MSI результатов преобразования Фурье, MSI где эталоном выступал результат фильтрации вейвлетом когерентного накопления, коэффициент линейной корреляции и MSI где эталоном был восстановленный после вейвлет-преобразования ВП.

> ные» (связанные с событиями) потенциалы (ВП), как правило, относящиеся к зрительным, либо когнитивным ВП. Их амплитудно-частотные параметры могут использоваться при формировании команд для выполнения тех или иных действий.

> Существует значительное количество алгоритмов выделения и обработки вызванных потенциалов, однако, в большинстве случаев, сравнение проводится исключительно на уровне уже готовых интерфейсов, на эффективность работы которых может повлиять широкий спектр условий, далеко не всегда связанных с устойчи

востью, ресурсоёмкостью или точностью работы алгоритма. К таковым условиям может быть отнесено состояние пользователя (активность, утомление, предшествующий приём пищи), качество датчиков (например, в случае ЭЭГ), их расположение относительно тех или иных отделов мозга, аппаратные характеристики устройства регистрации мозговой активности и т.д. Таким образом, актуальным представляется сравнение именно самих алгоритмов обработки электроэнцефалограмм, содержащих в своей структуре вызванные потенциалы.

В качестве вызванных потенциалов в системах синхронных интерфейсов человек-компьютер наибольшее распространение получили устойчивые зрительные вызванные потенциалы (SSVEP) и потенциал (или компонент) Р300 [1-3]. Ряд других ЭЭГ-феноменов, связанных с теми или иными внешними или внутренними событиями и вызванными потенциалами, хотя и представляет интерес для применения в качестве сигналов, используемых для создания команд, тем не менее, существенно уступают двум указанным выше феноменам по их чувствительности к характеристикам состояния человека. Из двух, указанных выше феноменов, потенциал SSVEP рассматривается в качестве основного для применения в наиболее быстрых интерфейсах.

Целью настоящей работы является проведение вычислительных экспериментов для получения сравнительных характеристик эффективности различных алгоритмов и их программных реализаций при выделении потенциала SSVEP из электроэнцефалограмм в условиях зашумленности сигнала.

Материалы и методы

Как известно [4-6], потенциал SSVEP имеет вид периодического колебания с частотой, равной частоте проводимой стимуляции, а форма сигнала напоминает гармоническое колебание. Поэтому в качестве модели вызванного потенциала была использована синусоида, аддитивно дополненная белым шумом, который имитировал не только сигнал ЭЭГ, но и артефакты, связанные с движением электродов по коже и с электрической активностью перикраниальных мышц [7, 8]. Учитывалось, что амплитуда потенциала SSVEP может колебаться от значений 0,5 мкВ до ≈10 мкВ, а фоновый сигнал ЭЭГ в этих же областях достигает величин 80-100 мкВ [9]. В ходе проведенного моделирования отношение сигнал/шум менялось в пределах от полного отсутствия шума (начальный вариант) до величины 1/300. Рассматривалась модель, при которой следовало выбрать одну из шести возможных частот в диапазоне от 7,5 Гц до 20 Гц, как диапазон частот ЭЭГ, где максимальным образов выражена амплитуда SSVEP [10]. Для выделения вызванных потенциалов, если этого требовали условия численного эксперимента, использовался стандартный подход, основанный на когерентном накоплении сигнала, при этом время регистрации сигнала для всех рассмотренных случаев составляло 1 с или 4 с, а время накопленного вызванного потенциала составляло 0,5 с или 3 с. Анализировались результаты 100 циклов распознавания сигнала.

В качестве алгоритмов многоканальной детекции потенциала SSVEP были выбраны следующие.

1. Преобразование Фурье накопленного, в рамках выделения вызванного потенциала, сигнала в одинаковом для всех частот временном окне. При этом рассчитывалась спектральная плотность мощности (СПМ), как на соответствующей частоте SSVEP, так и в ее окрестностях, определяемых как среднее расстояние между частотами, на которых возможно появление искомого сигнала. По результатам оценки СПМ решение получалось тремя вариантами: выбором наибольшего значения СПМ; выбором самой большой суммы СПМ, получаемой с разных каналов; «голосованием» каналов, когда каждый из каналов ЭЭГ определял свою частоту, а полученные результаты оценки суммировались. Таким образом, выбиралась та частота, которая была выбрана наибольшим числом каналов.

2. Преобразование Фурье, выполняемое, в отличие от предыдущего случая, не в постоянном временном окне, а в окне, содержащем одинаковое число осцилляций каждой из искомых частот. В остальном же выполнялась та же последовательность действий, что и в пункте 1.

3. Алгоритм Multivariate Synchronization Index (MSI) [11, 12], являющийся, по сути, дальнейшим развитием методов канонического корреляционного анализа. Согласно [11, 12], оценивались значения, получаемые при использовании алгоритма MSI для нативного модельного сигнала в сравнении с эталонным сигналом, в качестве которого выступали гармоники искомых частот. При этом аналогично подходам, использованным при преобразовании Фурье, при поиске искомой частоты применялся выбор наибольшего значения индекса синхронизации, а также выбор самой большой суммы индексов синхронизации с разных каналов и «голосование» каналов.

4. Алгоритм MSI, примененный не к нативному сигналу, а к накопленному вызванному потенциалу. В этом случае при сокращении эпохи анализа (что негативно сказывается на распознании сигнала), за счёт когерентного накопления, соотношение сигнал/шум улучшается, что, теоретически, может привести к более удачным результатам распознавания. Выбор искомой частоты по результатам этого преобразования осуществлялся так же, как в пп. 1, 2.

5. Алгоритм MSI, примененный к спектру Фурье накопленного сигнала только в исследуемом частотном диапазоне. Данное ограничение вызвано необходимостью исключить влияние тех спектральных компонент, которые находятся за пределами исследуемого частотного диапазона, но, согласно алгоритму расчета, могут существенно повлиять на результат выбора частоты. Выбор нужной частоты по результатам преобразования осуществлялся так же, как и в пп. 1, 3.

6. Расчет коэффициента линейной корреляции между накопленным вызванным потенциалом и образцом, полученным в ходе обучения интерфейса на конкретном пользователе или генераторе тестового сигнала.

7. Расчет по алгоритму MSI с использованием в качестве эталона образца сигнала, полученного путем пространственного (с разных каналов) когерентного накопле-

8. Расчет по алгоритму MSI с использованием в качестве эталона образца сигнала ВП, полученного путем фильтрации исходного сигнала, заведомо содержащего искомый вызванный потенциал. Данный подход рассматривает возможности использования разных фильтров полученных, например, на основе генетических алгоритмов [13, 14]. При этом, созданные таким образом фильтры, равно как и фильтры, полученные аналитически, могут быть специализированы на выделении тех или иных особенностей исследуемого сигнала. С учетом того, что в модели использовалась синусоидальная функция, как имитатор SSVEP, в качестве фильтра был выбран вейвлет Морлет, представляющий собой амплитудно-модулированную гармоническую функцию, частота которой рассчитывалась исходя из частоты «стимуляции», равной частоте SSVEP. При заданной частоте дискретизации модельного сигнала в 5 кГц, разрешение по частоте в исследуемой области было выше 0.08 Гц для верхней границы исследуемого диапазона.

9. Расчет по алгоритму MSI с использованием в качестве эталона образца сигнала ВП, полученного путем прямого и обратного вейвлет-преобразования накопленного ВП, заведомо содержащего искомый вызванный потенциал. При этом восстановление осуществлялось не по всей матрице коэффициентов вейвлет преобразования. Согласно [15] определялись частотный диапазон восстановления сигнала как разность между двумя соседними локальными минимумами матрицы квадратов коэффициентов вейвлет-преобразования, между которыми располагалась анализируемая частота. Таким образом, анализировался алгоритм восстановления сигнала на основе его оригинальной структуры цепочек локальных максимумов и минимумов матрицы квадратов коэффициентов вейвлет-преобразования [15]. Такой подход позволяет восстанавливать фрагменты сигнала исходя из его существующих частотновременных особенностей, а не из априорных предположений об их существовании.

Общая схема проводимых экспериментов представлена на рис 1. При оценке числа правильных ответов учитывался случайный процент правильных детекций и доверительный интервал их доли.

Результаты и обсуждение

Результаты проведенных вычислительных экспериментов представлены на рис. 2. На рис. 2 использованы следующие обозначения для применяемых алгоритмов: «MSI» – обработка нативного сигнала с использованием алгоритма MSI (п. 3); «MSI накопл.» – применение алгоритма MSI для накопленного вызванного потенциала (п. 4); «MSI-FFT» – применение MSI для анализа спектра Фурье (п. 5); «FFT» – анализ на основе спектра Фурье эпох анализа равной длинны (п. 1); «FFT-C» - применение спектра Фурье, рассчитанного по одинаковому числу осцилляций искомых частот (п. 2); «Линейная корреляция» - расчет на основе результатов корреляции сигнала с образцом, полученным с «пользователя» (генератора тестового сигнала) в ходе обучения интерфейса (п. 6); MSI (фильтр) – эталон для данного алгоритма получен путём фильтрации накопленного сигнала вейвлетом Morlet, центральная частота которого совпадает с искомой в сигнале (п. 8); MSI (по образцу) - (п. 7); MSI (восстановленный) - эталоном служит сигнал, восстановленный на основе подходов по анализу цепочек локальных максимумов и минимумов (п. 9) [15, 16, 17]. В целом методы продемонстрировали высокий уровень распознания полезного сигнала до отношения сигнал/шум 1 к 40, что с учётом амплитуды SSVEP даст от 20 до ~ 400 мкВ «шума» – т.е. фоновой ЭЭГ. С учетом



Рис.1. Схема обработки модельного сигнала для определения эффективности алгоритмов при детекции ВП в ЭЭГ. Примечание – в выносках указаны эталонные сигналы для MS



Рис. 2. Результаты вычислительных экспериментов. Представлена зависимость процента ошибочных классификаций от параметра «шум N», который указывает, что амплитуда шума в N раз превышает амплитуду полезного сигнала. Амплитуда полезного сигнала принята за единицу

того, что амплитуда фоновой ЭЭГ в затылочных областях существенно снижается при умственной нагрузке, имеющуюся точность распознания следует признать вполне достаточной для устойчивого детектирования искомого потенциала. А при наличии депрессии α-ритма при фотостимуляции это дает высокое качество распознавания полезного сигнала. При соотношении сигнал/шум 1/100 и больше распознавание хотя и возможно для ряда алгоритмов, но в целом алгоритмы показывают точность, не позволяющую использовать их для управления дроном, самоходным шасси, набором текста. Иными словами, затруднительно применять те функции, ради которых и создаётся интерфейс мозгкомпьютер. При отношении сигнал/шум, равном 1/300, оказывается, практически, невозможным обеспечить функционирование интерфейса мозг-компьютер с приемлемой скоростью, несмотря даже на когерентное накопление вызванного потенциала.

При уровне сигнал/шум 1/100 наилучшим образом зарекомендовали себя следующие алгоритмы (в скобках указаны эпохи анализа): MSI (3 сек) исходного сигнала, MSI(3 сек) результатов преобразования Фурье, MSI(3 сек) где эталоном выступал результат фильтрации вейвлетом когерентного накопления, коэффициент линейной корреляции (3 сек) и MSI (3 сек), где эталоном был восстановленный после вейвлет-преобразования ВП. При этом преимущества по сравнению с другими исследованными алгоритмами имеют два алгоритма за счёт их существенно меньшей ресурсоемкости - это алгоритм MSI и алгоритм линейной корреляции, являющийся, по сути, свёрткой накопленного вызванного потенциала только по одной функции. Несколько проигрывает им в ресурсоёмкости алгоритм предварительной фильтрации когерентного накопления вейвлетами в известном частотном диапазоне (являющийся, по сути, вариантом линейной корреляции). Однако данная проблема ресурсоёмкости легко решается, если учесть, что генерация эталона осуществляется однократно после обучение программно- аппаратной части интерфейса

мозг-компьютер и непосредственно перед работой пользователя. При уменьшении эпохи анализа до 1 с, практически, все рассмотренные алгоритмы по-прежнему с достаточно высокой точностью выделяют нужные компоненты. Исключение составляет лишь алгоритм MSI, анализирующий накопленный сигнал.

Заключение

В работе представлены результаты сравнения алгоритмов анализа вызванных потенциалов головного мозга, используемых для конструирования интерфейсов мозг-компьютер. Для проведения вычислительных экспериментов был взят алгоритм преобразования Фурье и алгоритм MSI, являющийся развитием методов канонического корреляционного анализа в различных модификациях: анализ исходного сигнала; анализ накопленного вызванного потенциала; анализ спектра накопленного вызванного потенциала. Для исследования применялись модельные сигналы, представляющие собой сложение белого шума и гармоники, имитирующей стабильный зрительный вызванный потенциал. Отношение сигнал/шум менялось в пределах от полного отсутствия шума до 1/300. Полученные результаты показали, что при отношении сигнал/шум от полного отсутствия шума и до 1/40 ни один из рассмотренных алгоритмов не имеет преимущества по точности распознавания вызванного потенциала. При отношения сигнал/шум, равном 1/300, ни один из рассмотренных алгоритмов не обеспечивает необходимую точность распознавания вызванного потенциала. В промежуточном диапазоне таких отношений наиболее устойчивые результаты для 3-х секундной эпохи анализа продемонстрировали MSI исходного сигнала, MSI результатов преобразования Фурье, MSI, где эталоном выступал результат фильтрации вейвлетом когерентного накопления, коэффициент линейной корреляции и MSI, где эталоном был восстановленный после вейвлет-преобразования ВП. Они же оказались наилучшими для коротких 1-секундных отрезков, что позволяет рассматривать эти алгоритмы в качестве перспективных для создания высокоскоростных синхронных интерфейсов мозг-компьютер.

Работа выполнена при финансовой поддержке РФФИ (грант 16-29-08342-офи_м).

Автор выражает признательность С.В. Борзунову, А.А. Вахтину, М.А. Галыгиной С.Д. Кургалину, А.К. Толкачеву.

Литература

1. Zhu D., Bieger J., Molina G., Aarts R.M. A Survey of Stimulation Methods Used in SSVEP-Based BCIs // Computational Intelligence and Neuroscience. – Hindawi Publishing Corporation, 2010. – Article ID 702357.

2. Гнездицкий В.В. Вызванные потенциалы мозга в клинической практике. – М.: МЕДПресс-информ, 2003. – 264 с.

3. Farwell L.A., Donchin E. Talking off the top of your head: towards mental prosthesis utilizing event-related brain potentials // Electroenceph. Clin. Neurophysiol. – 1988. – V. 70. – P. 510-523.

4. Wang Yu., Wang Yi., Cheng Ch., Jung T. Developing Stimulus Presentation on Mobile Devices for a Truly Portable SSVEP-based BCI. http://ieeexplore.ieee.org/document/ 6610738/?section=abstract.

5. Volosyak I. SSVEP based Bremen-BCI boosting information transfer rates // Journ. Neural Eng. – 2011. – V. 8(3). – P. 036020. DOI: 10.1088/1741-2560/8/3/036020.

6. Resalat S.N., Setarehdan S.K. An Improved SSVEP Based BCI System Using Frequency Domain Feature Classification // American Journal of Biomedical Engineering. – 2013. – V. 3(1). – P. 1-8. DOI: 10.5923/j.ajbe.20130301.01.

7. Бойцова Ю.А., Данько С.Г., Медведев С.В. Динамика мощности ЭЭГ в бета- и гамма-диапазонах в условиях нормальной и ослабленной электромиограммы лицевых мышц // Физиология человека. 2016. Т. 42. № 6. С. 5-17.

8. Данько С.Г., Грачёва Л.В., Бойцова Ю.А., Соловьёва М.Л. Электромиограмма перикраниальных мышц в частотных диапазонах бета и гамма при сравнении эмоционально и когнитивно различных состояний // Физиология человека. 2014. Т. 40. № 2. С. 5.

9. Зенков Л.Р. Клиническая электроэнцефалография

(с элементами эпилептологии). – М.: МЕДПресс-информ, 2011.– 356 с.

10. High frequency SSVEPs for BCI applications G. Garcia // http://hmi.ewi.utwente.nl/chi2008/chi2008_files/ garcia.pdf.

11. Multivariate synchronization index for frequency recognition of SSVEP-based brain-computer interface / Ya. Zhang, P. Xu, K. Cheng, D. Yao // Journal of Neuroscience Methods. – 2014. – 221. – P. 32– 40.

12. Comparison of the influence of stimuli color on Steady-State Visual Evoked Potentials / R. Tello, S. Müller, A. Ferreira, T. Freire // Bastos Res. Biomed. Eng. – 2015. – V. 31(3). – P.218-231.

13. Решение задачи автоматической классификации сигналов методом генетического алгоритма Белобродский В.А., Кургалин С.Д., Туровский Я.А. В сборнике: Физико-математическое моделирование систем материалы XIII Международного семинара. 2014. С. 125-132.

14. Белобродский, В.А. Разработка генетического алгоритма для конструирования цифровых фильтров, классифицирующих биомедицинские сигналы, и его апробация на сигналах с известными параметрами / В.А. Белобродский, С.Д. Кургалин, Я.А. Туровский и др. // Биомедицинская радиоэлектроника. – 2015. – № 2. – С. 56-64.

15. Исследование вызванных потенциалов головного мозга на основе адаптивного варианта обратного вейвлет-преобразования Туровский Я.А., Кургалин С.Д., Вахтин А.А., Борзунов С.В., Белобродский В.А. Биофизи-ка. 2015. Т. 60. № 3. С. 547-554.

16. Туровский Я.А. Обработка сигнала электроэннцефалограммы на основе анализа частотных зависимостей и вейвлет-преобразования / Я.А.Туровский С.Д. Кургалин, А.А. Вахтин // Биомедицинская радиоэлектроника. – 2012. – №2. – С.39-45

17. Туровский Я.А. Анализ электроэнцефалограмм на основе исследования изменяющейся во времени структуры локальных максимумов матрицы вейвлет-коэффициентов / Я.А. Туровский, С.Д. Кургалин, А.В. Максимов А.Г. Семёнов // Вестник Воронежского государственного университета. Серия: системный анализ и информационные технологии. – 2012. – №2. – С. 69-74.

Уважаемые коллеги!

Для тех, кто не успел оформить подписку на первое полугодие 2018 года через АО «Роспечать», сохраняется возможность приобретения журналов непосредственно в редакции по адресу: г. Москва, ул. Авиамоторная, дом 8, Научный Центр МТУСИ, ком. 612. Российское научно-техническое общество радиотехники, электроники и связи им. А.С. Попова, метро «Авиамоторная», или оформить Заказ в соответствии с требованиями, выставленными на сайте журнала: www.dspa.ru.

Справки по телефону: (+7 903)201-53-33 (Самсонов Геннадий Андреевич).

E-mail: rntores@mail.ru

УДК 621.396

РЕАЛИЗАЦИЯ ЦЕЛОЧИСЛЕННЫХ ЦИФРОВЫХ РЕКУРСИВНЫХ ФИЛЬТРОВ БЕЗ УМНОЖИТЕЛЕЙ НА ПЛИС ОТЕЧЕСТВЕННОГО ПРОИЗВОДСТВА

Артемьев В.В., аспирант Филиала РФЯЦ-ВНИИЭФ "Научно-исследовательского института измерительных систем им. Ю.Е. Седакова», e-mail: zzzrf413@bk.ru;

Кашин А.В., д.т.н., зам. главного конструктора, начальник научно-исследовательского отдела Филиала РФЯЦ-ВНИИЭФ «Научно-исследовательского института измерительных систем им. Ю.Е. Седакова», e-mail: aKashin@niiis.nnov.ru.

IMPLEMENTATION OF INTEGER RECURSIVE DIGITAL FILTERS WITHOUT MULTIPLIERS ON FPGA OF THE RUSSIAN PRODUCTION

Artemiev V.V., Kashin A.V.

The article is devoted to the synthesis of recursive digital integer filter with the possibility of its implementation in FPGA. The article consists of six parts. The first part is an introduction. The second part provides an overview of the classical approaches of solving the problem of designing multiplierless digital filters. In the third part of the article proposes a method-integer nonlinear programming, allowing to solve the task on the specified set of parameters. Part four describes the hardware implementation on FPGA recursive digital filters without multiplier. In part five given the volume of the matrix space necessary for the implementation of digital filters synthesized on not equidistant set of parameters. The sixth part shows the performance of digital filters. The seventh part is the conclusion.

Key words: digital integer filter, implementation in FPGA, method-integer nonlinear programming.

Ключевые слова: Цифровой целочисленный фильтр, без умножителей, программируемая логическая интегральная схема.

Введение

Цифровой фильтр (ЦФ) является основным устройством во многих системах цифровой обработки сигналов (ЦОС). При проектировании ЦФ для

систем, реализуемых на программируемых логических интегральных схемах (ПЛИС), заказных или полузаказных СБИС, ставятся задачи получения требуемых селективных характеристик ЦФ при минимальном количестве ресурсов занимаемых на кристалле, например, таких как логические элементы. Особо остро вопрос количества необходимых ресурсов стоит для отечественных микросхем базовых матричных кристаллов (БМК), на которых осуществляется изготовление конкретных БИС, и ПЛИС, например, таких как ПЛИС 5578ТС024 (АЕЯР.431260.978ТУ) и 5576ХС7Т (АЕНВ.431260.059ТУ) производства «АО ВЗПП-С» (г. Воронеж), имеющих невысокие показатели количества логических элементов и быстродействия по сравнению с импортными аналогами. Поэтому неоптимальное решение задачи проектирование ЦФ приводит к неоправданному снижению быстродействия, препятствует размещению системы ЦОС на одном или малом числе кристаллов, повышает стоимость изделия и увеличивает потребляемую мощность.

Сложность реализации цифровых фильтров определяется умножителями [1]. На практике при построении различных систем ЦОС широко применяются ЦФ с постоянными коэффициентами и, как правило, высокоскоростные специализированные системы ЦОС оперируют с фиксированной, а не с плавающей точкой. Умножение на постоянный коэффициент можно заменить параллельными операциями сдвига и суммирова-

Рассматриваются вопросы реализации цифрового целочисленного рекурсивного фильтра на ПЛИС без умножителей. Приводятся значения необходимого матричного пространства для ПЛИС 5578TC024 (АЕЯР.431260.978TУ) и 5576XC7T (АЕНВ.431260.059TУ) отечественного производства, а также тактовые частоты, характеризующие быстродействие цифрового фильтра.

> ния/вычитания, что значительно упростит реализацию функции и повысит её быстродействие. Параллельный сдвиг не требует аппаратных и временных затрат и выполняется простым рассогласованием разрядных шин. Цифровые фильтры, в которых умножения на коэффициенты заменены операциями сдвиг/суммирование, называют цифровыми фильтрами без умножителей.

> Целью данной работы является реализация на ПЛИС рекурсивного ЦФ, синтезированного на определённом множестве целочисленных коэффициентов, позволяющем оптимизировать объем матричного пространства ПЛИС и повысить её быстродействие.

Методы проектирования цифровых фильтров без умножителей

Постоянные коэффициенты ЦФ, используемые в операциях умножения, при реализации на ПЛИС, могут быть представлены в виде степени двойки [2]:

$$C = \sum_{i=0}^{R-1} k_i \cdot 2^i ,$$
 (1)

где $k_i = \{0,1\}$ код представления числа. При проектировании ЦФ в базисе ПЛИС целесообразно использовать представление [2], в котором коэффициенты (1) – это сумма и/или разность чисел, равных степени двойки, где $k_i = \{0,1,-1\}$. Такое представление называется знакоразрядном кодом (ЗРК). Для любого числа существует бесконечное множество знако-разрядных представлений

этого числа. Знако-разрядное представление, имеющее минимальное количество 1 и -1, называется каноническим знако-разрядным кодом (КЗРК). Преимуществом при представлении числа в виде КЗРК является меньшее количество сумматоров, необходимых для реализации умножителя, что дает выигрыш при их аппаратной реализации.

Уменьшить количество сдвиговых сумматоров, используемых для реализации умножения, можно также выделением общих подвыражений [2, 13]. В этом случаем постоянный коэффициент представляется в виде

$$C = \prod_{j=1}^{n} \sum_{i=0}^{K-1} k'_{ji} \cdot 2^{i} , \qquad (2)$$

Операции умножения в (2) выполняются сдвигом.

Поскольку требованиям, предъявляемыми к цифровому фильтру по его селективным параметрам, могут удовлетворять множество решений, то возникает задача поиска решения с минимальным общим числом сумматоров, заменяющих умножители на определенное число.

Большинство работ [3-11] по проектированию цифровых рекурсивных фильтров с минимальным количеством ресурсов необходимых для их реализации на ПЛИС основывается на использовании при построении каскада из двух секций все пропускающих фильтров, подключенных в параллель. Как правило, можно выделить два этапа:

1. Поиск хороших начальных точек в области исходных параметров.

 Локальный поиск решений с квантованными коэффициентами в окрестности этих точек.

В работе [7] для поиска оптимальных коэффициентов фильтра предлагается применять алгоритм MM-MFA, основанный на глобальном поисковом алгоритме имитации отжига. Так же для поиска коэффициентов применяются генетические алгоритмы [12]. Однако эти поисковые алгоритмы не гарантируют нахождение наилучшего результата.

Наиболее близка к тематике статьи работа [3], в которой рассматривается вопрос синтеза ЦФ с применением билинейного преобразования аналогово прототипа эллиптических рекурсивных фильтров максимальной добротности (Elliptic filters with maximal Q-factor (EMQF)). В работе для построения ЦФ определяется некоторое множество чисел, оптимальных при реализации умножения с помощью операций сдвига и суммирования, однако алгоритм выборки не гарантирует точную выборку из этого множества, и билинейное преобразование не позволяет осуществлять поиск решения задачи проектирования ЦФ по совокупности характеристик фильтра, например, с учетом ФЧХ и/или ГВЗ.

Метод целочисленного нелинейного программирования

Многофункциональный синтез целочисленных цифровых фильтров по совокупности требуемых функциональных характеристик может быть осуществлен с помощью метода целочисленного нелинейного программирования (ЦНП). Принципиальное отличие ЦНП-синтеза заключается в том, что в данном случае для удовлетворения требуемого функционирования целочисленного ЦФ на стадии его проектирования не используются никакие искусственные приёмы и подходы классических методов [14-16] (аналоговые прототипы, методы аппроксимации, быстрой свёртки и др.), приводящие к существенным ограничениям в выполнении функциональных требований и систематическим ошибкам, а осуществляется прямой поиск требуемых целочисленных коэффициентов фильтра прямо по его математическому определению (модели) [17-19]. Критерием поиска является соответствие совокупного текущего функционирования фильтра его требуемому функционированию.

В качестве базовой структуры ЦФ синтезируемого методом ЦНП используется форма каскадного соединения целочисленных БИХ-звеньев 2-го порядка с передаточной функцией:

$$H(z) = \prod_{i=1}^{m} \frac{b_{0i} + b_{1i}z^{-1} + b_{2i}z^{-2}}{a_{0i} + a_{1i}z^{-1} + a_{2i}z^{-2}} , \qquad (3)$$

где m – число звеньев ЦФ, комплексная переменная

$$z=e^{j\omega}$$
, а $\omega=rac{2\pi f}{F_{\mathcal{A}}}$ — цифровая частота. Из соотноше-

ния (3) получается разностное уравнение для одного звена рекурсивного целочисленного фильтра:

$$y_n = (b_0 x_n + b_1 x_{n-1} + b_2 x_{n-2} - a_1 y_{n-1} - a_2 y_{n-2}) / a_0.$$
 (4)

Операция деления на целочисленный нормирующий коэффициент *a*₀ реализуется при помощи операции побитового сдвига, при условии принадлежности каждого *i*-го нормирующего коэффициента биномиальному целочисленному ряду:

$$a_{0i} \in \{2^q\}, \ q = 0, R-1 \ i = 1, m$$
, (5)

где *R* – разрядность цифровой платформы.

$$F(IX) = \beta_1 f_{AYX}(IX) + \beta_2 f_{\varphi YX}(IX)$$
(6)

Частная целевая функция $f_{A'lX}(IX)$ при этом определяется среднеквадратичной ошибкой о выполнения требований к АЧХ фильтра, а функция $f_{\Phi'lX}(IX)$ – максимальным отклонением текущей фазы фильтра от требуемой линейной ФЧХ фильтра.

Относительно целевой функции (6), задача дискретного целочисленного программирования для синтеза на не эквидистантном множестве *R*-разрядного рекурсивного фильтра в форме каскадного соединения *m*-звеньев второго порядка записывается так:

$$F^{o}(IX^{o}) = \min F(IX), \quad IX \in I^{5m}$$
(7)

$$a_{di} \in I_s \quad d = \overline{1,2} \quad i = \overline{1,m} \tag{8}$$

$$b_{di} \in I_s$$
 $d=0,2$ $i=1,m$

$$a_{0i} \in \{2^q\} \quad q = 0, R - 1 \tag{9}$$

$$\left|Z_{pi}\right| < 1 \quad i=1,m \tag{10}$$

Общая экстремальная задача дискретного синтеза (7) записана относительно целочисленного пространства I^{5m} параметров (коэффициентов фильтра), размерностью 5m. Ограничение (8) задает одно из заранее определённых неэквидистантных множеств чисел Is с малым количеством сумматоров в представлении, из которого осуществляется выборка целочисленных коэффициентов. На рис. 1-3 показаны распределения для множеств параметров I_0 , I_1 , I_2 . Например множество I_0 состоит из чисел равных степени двойки в отрезке от - 2^{R-1} до 2^{R-1} , а I_I из чисел с одним и менее количеством сумматором в представлении. Соотношение (9) определяет принадлежность коэффициентов а_{0i} биномиальному ряду. Функциональные ограничения (10) контролируют условие устойчивости рекурсивного фильтра по всем полюсам коэффициента передачи.



Рис.1. Распределение множества I_0



Рис. 2. Распределение множества I₁



Рис. 3. Распределение множества I₂

Реализация на ПЛИС

Так как отечественные микросхемы ПЛИС по архитектуре соответствуют ранним семействам ПЛИС фирмы Altera, которые не имеют аппаратных умножителей, то с целью получения кросс платформенного HDL описания программирование осуществляется на языке VHDL под заданную структуру фильтра, соответствующую разностному уравнению (4) без явной привязки к аппаратным блокам «умножение с накоплением» и использования IP ядер. Предлагается универсальное VHDL-описание структуры ЦЦФ на ПЛИС.

При аппаратной реализации на ПЛИС используется пять регистров для хранения входных значений x_n , x_{n-1} , x_{n-2} и выходных вычисленных значений y_{n-1} , y_{n-2} . Для одновременного (в параллель) выполнения операций умножения на коэффициенты фильтра применяются пять схем, реализующих умножитель с помощью сдвигов и сумматоров. Получившиеся значения складываются с помощью двух сумматоров, установленных в параллель и двух последовательно. Разрядность сумматоров согласована с выходной разрядностью схем умножителей. Сдвиг осуществляется путем выборки соответствующих цепей разрядов из выходной шины данных сумматора.

На рис. 4-6 приведены синтезированные по VHDL описанию схемы (RTL Schematic) звена рекурсивного целочисленного фильтра для множеств параметров I_0, I_1, I_2 .

Тактовая частота звена фильтра определяется исходя из задержки вычисления значения звена (4), которая, в свою очередь, зависит от времени прохождения сигнала от регистра к регистру по самой длинной цепи. Важно отметить, что при реализации на ПЛИС отрицательные числа представляются в дополнительном коде, что приводит к необходимости использование дополнительного сумматора для перевода чисел в отрицательные.



Рис. 4. Синтезированная схема (RTL Schematic) рекурсивного звена ЦЦФ на множестве параметров I_0



Рис. 5. Синтезированная схема (RTL Schematic) рекурсивного звена ЦЦФ на множестве параметров I₁



Рис. 6. Синтезированная схема (RTL Schematic) рекурсивного звена ЦЦФ на множестве параметров *I*₂

Ресурсы ПЛИС

плис Рассмотрим отечественную микросхему 5578ТС024 АЕЯР.431260.978ТУ. Оценка количества задействованного матричного пространства в ней проведена на основе прямого аналога – EP2C8F256I8 фирмы Altera, используемого при подготовке конфигурационной программы ПЛИС. Компоновка и трассировка осуществляется с опциями по умолчанию в САПР Altera Quartus II 11.1. Для сравнения объема матричного пространства ПЛИС использовались ЦФ полученные с помощью Matlab. При генерировании HDL-описания ЦФ в Matlab имеется возможность указать как реализовывать умножитель: либо с использованием полноценного умножителя, либо с применением встроенного алгоритма получения чисел в КЗРК представлении (CSD) или алгоритма получения КЗРК представления числа с выделением общих подвыражений (Factored-CSD). Разрядность данных, поступающих на фильтр, определена в 16 бит.

Ресурсы, необходимые для реализации шестнадцати разрядного цифрового фильтра из 10 звеньев в ПЛИС 5578TC024, в зависимости от используемого при проектировании множества приведены в табл.1.

По анализу табл. 1 можно сделать вывод, что цифровые фильтры без умножителей, спроектированные методом ЦНП на неэквидистантых множествах I_0 , I_1 , I_2 показывают более высокий результат по экономии объема матричного пространства ПЛИС. Более наглядно результаты приведеные в табл. 1 можно проиллюстрировать с помощъю гистограммы показаной на рис. 7. На гистограмме для значения количества аппаратных умножителей используется маштабирующий коэффициент 100.



Рис. 7. Ресурсы ПЛИС 5578ТС024

Как видно из рис. 7, цифровой фильтр, полученный на неэквидистантном множестве I_1 , сопоставим по количеству требуемых для реализации логических элементов с ЦФ спроектировынным на полноценном множестве, однако не требует аппаратных умножителей 9х9. ЦФ синтезированный на моножетве I_2 занимает на 40 % меньше ресурсов, чем аналогичный полученый с помощъю Matlab. Важно также отметить, что количество аппаратных умножителей необходимое для реализации ЦФ, синтезированного на полном множестве, в ПЛИС 5578TC024 больше числа встроенных умножителей самой ПЛИС.

ПЛИС 5576ХС7Т АЕНВ.431260.059ТУ имеет меньшее количество ресурсов по сравнению с 5578ТС024 и в ней отсутствуют аппаратные умножители. Однако основной особенностью применения данной ПЛИС является её высокая стойкость к внешним воздействующим факторам. Разрядность данных, поступающих на фильтр, была взята 12 бит. Ресурсы, необходимые для реализации одного звена цифрового фильтра в ПЛИС 5576ХС7Т АЕНВ.431260.059ТУ, в зависимости от используемого при проектировании множества приведены в табл. 2.

Таблица 🛛	1. Ресурсы	плис	5578TC024
-----------	------------	------	-----------

Ресурс ПЛИС 5578ТС024	Total logic elements		Total combinational functions		Dedicated logic registers		Embedded Multiplier 9-bit elements	
Всего в ПЛИС 5578ТС024	8256		8256		8256		36	
ЦФ Matlab, аппаратные умножители	1909	23 %	1574	19 %	648	8 %	64	178 %
ЦФ Matlab, умножители на LUT	4719	57 %	4572	55 %	648	8 %	0	0 %
ЦФ с CSD Matlab	4745	57 %	4560	55 %	646	8 %	0	0 %
ЦФ с Factored-CSD Matlab	4588	56 %	4404	53 %	646	8 %	0	0 %
Полное множество, аппаратные умножители	1820	22 %	1153	14 %	827	10 %	90	250 %
Полное множество, умножители на LUT	5039	61 %	5011	61 %	827	10 %	0	0 %
I_0	1287	16 %	1109	13 %	777	9 %	0	0 %
I_1	2097	25 %	1901	23 %	827	10 %	0	0 %
I_2	2884	35 %	2723	33 %	811	10 %	0	0 %

Ресурс ПЛИС 5576ХС7Т	Logic cells
Всего в ПЛИС 5576ХС7Т	1728
ЦФ синтезированный в Matlab	968 (56%)
ЦФ синтезированный с CSD в Matlab	996 (57 %)
ЦФ синтезированный с Factored-CSD в Matlab	928 (53 %)
Полное множество	711 (41%)
I_0	189 (10%)
I ₁	306 (17%)
I_2	428 (24%)

Таблица 2. Ресурсы ПЛИС 5576ХС7Т

По анализу табл. 2 можно так же сделать вывод, что цифровые фильтры без умножителей, спроектированные методом ЦНП на неэквидистантых множествах I_0 , I_1 , I_2 показывают более высокий результат по экономии объема матричного пространства ПЛИС. Более наглядно результаты приведеные в табл. 2 также можно проиллюстрировать с помощъю гистограммы показаной на рис. 8.



Рис. 8. Ресурсы ПЛИС 5576ХС7Т

Как видно из рис. 8 выигрыш по количеству требуемых ресурсов для реализации ЦФ при использовании неэквидистантных множествств I_0 , I_1 , I_2 более существенен чем на рис. 7. ЦФ синтезированный на моножетве I_2 занимает на 50 % меньше ресурсов, чем аналогичный полученый с помощъю Matlab.

Оценка количества используемого матричного пространства, показанная в табл. 2 и на рис. 8, отечественной микросхемы ПЛИС 5576ХС7Т АЕНВ.431260.059ТУ приведена на основе прямого аналога EPF10K30EFC256-3, применяемого при подготовке конфигурационной программы ПЛИС. Компоновка и трассировка осуществляется с опциями по умолчанию в САПР Max II + Plus. Также важно отметить, что VHDL-описание цифровых фильтров, синтезированных с помощью Matlab, не может быть непосредственно применено в проекте САПР Max II + Plus, т.к. в описании используется библиотека IEEE.numeric std.ALL, не поддерживаемая данным САПР. Требуется доработка VHDL-описания цифровых фильтров, синтезированных с помощью Matlab при применении в ПЛИС 5576ХС7Т АЕНВ.431260.059ТУ и CAПPe Max II + Plus.

Быстродействие

Одной из ключевых характеристик ЦФ является его быстродействие – это количество обработанных отчетов фильтром за единицу времени. Данная характеристика определяет ширину полосы цифрового фильтра для систем, работающих в режиме реального времени, где необходимо проводить обработку данных за ограниченный временной интервал. Тактовые частоты работы ЦФ в базисе ПЛИС, одновременно являющиеся частотами дискретизации для систем, работающих в реальном времени, представлены на рис. 9-10.



Рис. 9. Быстродействие ЦФ в ПЛИС 5578ТС024



Рис. 10. Быстродействие ЦФ (одно звено) в ПЛИС 5576ХС7Т

Как видно из рис. 9-10, цифровые фильтры, спроектированные на неэквидистантых множествах параметров I_0 , I_1 , I_2 , показывают более высокий результат по быстродействию по сравнению с фильтрами, спроектированными в Matlab и на полноценном множестве.

Заключение

Разработано универсальное VHDL-описание рекурсивного цифрового фильтра без умножителей, синтезированного методом ЦНП для отечественных ПЛИС/БМК. Цифровые фильтры, спроектированные на неэквидистантых множествах параметров I_0 , I_1 , I_2 , обеспечивают экономию объема матричного пространства и быстродействие на всех представленных ПЛИС по сравнению с цифровыми фильтрами, полученными классическим подходом с применением билинейного преобразования.

Литература

1. Мингазин А.Т. Синтез цифровых фильтров для высокоскоростных систем на кристалле. // «Цифровая обработка сигналов». 2004. №2 с. 14 – 32.

2. Koren I. Computer Arithmetic Algorithms (Second ed.), A.K. Peters, Ltd. (Ed.), 2002.

3. Milic L.D., Lutovac M.D. Design of multiplierless elliptic IIR filters with a small quantization error. // IEEE Trans. Signal Proc. 1999. Vol. 47. № 2. P. 469–479.

4. Lutovac M.D., Milic L.D. Approximate linear phase multiplierless IIR halfband filter. // IEEE Trans. Signal Proc. Lett. 2000. Vol. 7. № 3. P. 52–53.

5. Lutovac M. D. and Milic' Lj. D. «Design of multiplierless

elliptic IIR halfband filters and Hilbert transformers», in Proc. EUSIPCO, '98 Rodos, Greece, Sept. 1998, pp. 291–294.

6. Yli-Kaakinen J., Saramaki T. An algorithm for the design of multiplierless approximately linear-phase lattice wave digital filters. // ISCAS. 2000. May. P. 77–80.

7. Persson P., Nordebo S., Claesson I. A multimode mean field annealing technique to design recursive digital filters. // IEEE Trans. Circuits and Syst.: II. 2001.

8. Yli-Kaakinen J., Saramaki T. A systematic algorithm for the design of multiplierless lattice wave digital filters. // ISCCSP. 2004. Mar. P. 393–396.

9. Milic L. D., Lutovac M. D. Efficient algorithm for the design of high-speed elliptic IIR filters. // Int. J. Electron. Commun. (AEU). 2003. Vol. 57. № 4. P. 255–262.

10. Мингазин А.Т. Синтез цифровых фильтров на основе фазовых цепей с конечной длиной слова коэффициентов. // II Международная конференция «Цифро-вая обработка сигналов и ее применения» (DSPA). 1999. Т. 1. Сентябрь. С. 112–116.

11. Мингазин А.Т. Синтез полуполосных цифровых фильтров без умножителей на основе фазовых цепей. // VI Международная конференция «Цифровая обработка сигналов и ее применения» (DSPA). 2004. Т. 1. Март-Апрель. С. 39-41.

12. Плотников П.В. Повышение эффективности реали-

зации цифровых фильтров в ПЛИС. // Проблемы разработки перспективных микроэлектронных систем – 2006. Сборник научных трудов / под общ. ред. А.Л. Стемпковского. М.:ИППМ РАН, 2006, с.333-338.

13. Алёшин Д.В. Алгоритм синтеза целочисленных умножителей для цифровых КИХ-фильтров. 9-я международная конференция «Цифровая обработка сигналов и её применение». DSPA-2007. с. 96 – 98.

14. Айфичер Э., Джервис Б. Цифровая обработка сигналов: практический подход, 2-е издание. Пер. с англ. М.: Издательский дом «Вильямс», 2004.

15. Антонью А. Цифровые фильтры: анализ и проектирование. М.: Радио и Связь, 1983.

16. Баскаков С.И. Радиотехнические цепи и сигналы. М.: Высшая школа, 2005.

17. Бугров В.Н. Проектирование цифровых фильтров методами целочисленного нелинейного программирования // Вестник ННГУ, 2009, № 6. С. 61 – 70.

18. Артемьев В.В., Бугров В.Н. Синтез цифровых рекурсивных фильтров с линейной фазой // Компоненты и технологии, 2013, № 7. С. 60 – 62.

19. Артемьев В.В., Бугров В.Н., Пройдаков В., Шкелев Е.И., Целочисленные цифровые фильтры – эффективное решение для 8-битовых цифровых платформ // Компоненты и технологии, 2013, № 10. С. 104 – 110.

Уважаемые коллеги!

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

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

Научно-технический журнал «Цифровая обработка сигналов» включен в Перечень изданий, рекомендуемый ВАК РФ для публикации результатов научных исследований соискателями ученой степени доктора и кандидата технических наук в области радиотехники, связи, вычислительной техники, электроники, приборостроения, информационных технологий, информационно-измерительных и управляющих систем. По предварительным итогам за 2015 год по рейтингу Science Index базы РИНЦ (3,394) журнал «Цифровая обработка сигналов» занимает 344-ю позицию из почти 3000 представленных изданий. Импакт-фактор журнала за 5-летний период цитируемости составил 0,535!

Планируемые сроки издания отдельных номеров журнала:

– № 2 июнь 2018 г. Тематический выпуск по материалам 20-й Международной научно-технической конференции «Цифровая обработка сигналов и ее применение-DSPA».

- № 3 сентябрь 2018 г. Тематический выпуск: «Цифровая обработка изображений».

– № 4 декабрь 2018 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 621.396

ПРИМЕНЕНИЕ КОМПЛЕКСНЫХ ФИЛЬТРОВ ДЛЯ СНИЖЕНИЯ ПИК-ФАКТОРА СИГНАЛОВ С OFDM МОДУЛЯЦИЕЙ И ИХ РЕАЛИЗАЦИЯ НА ПЛИС

Чан Ван Нгиа, аспирант кафедры мультимедийных технологий и телекоммуникаций МФТИ, e-mail: nghiamosmipt@gmail.com.

APPLICATION OF THE COMPLEX FILTERS FOR PAPR REDUCTION OF OFDM SIGNALS AND FPGA IMPLEMENTATION

Tran Van Nghia.

In this paper we analyze the reconfigurable complex filtering technique based on the Tone Reservation (TR) and data set in the time-domain and in the frequency-domain to reduce the Peak-to-Average Power Ratio (PAPR) of OFDM signals. Hardware architecture of these filters is proposed and mapped onto FPGA. The simulation results of the proposed methods and experimental results on FPGA for 16-QAM and 64-QAM modulation are presented. These results show that PAPR of OFDM signals is significantly reduced (about 6.6 dB) and OFDM signals achieve the desired threshold after 2-4 iterations.

Key words: Peak-to-Average Power Ratio (PAPR), OFDM modulation, Tone Reservation (TR), clipping and filtering, FPGA, complex filter.

Ключевые слова: пик-фактор, OFDM модуляция, способ резервных поднесущих, ограничение и фильтрация, ПЛИС, комплексный фильтр.

Введение

Благодаря высокой спектральной эффективности при передаче в каналах с многолучевым распространением и высокой скорости передачи технология мультиплексирования с ортогональным частотным разделением каналов OFDM (Orthogonal Frequency-Division Multiplexing) широко используется для решения задач

передачи данных в различных стандартах беспроводных сетей нового типа.

Основным недостатком систем с OFDM модуляцией является высокий пик-фактор (PAPR) передаваемых сигналов вследствие возможного синфазного суммирования большого количества модулированных поднесущих. Пикфактор характеризуется отношением пиковой мощности сигнала к его средней мощности. Из-за этого радиосигнал требует использования широкого линейного динамического диапазона устройств УМ (усилителя мощности), что приводит к увеличению стоимости аппаратуры и к большому энергетическому потреблению. Для УМ с ограниченной линейной областью радиосигнал попадает в его область насыщения, что приводит к нелинейным искажениям (внутриполосному и внеполосному). Поэтому эффективное решение для снижения PAPR и простоты аппаратной реализации является важным вопросом при практическом применении OFDM технологии.

До сих пор было разработано достаточно большое количество разнообразных алгоритмов борьбы с пикфактором, которые разделяются на две части: искажающие и неискажающие способы снижения пик-фактора [1]. Широко известный искажающий способ – метод ограничения амплитуды, и ограничения и фильтрации. Известные неискажающие способы – селективное отоб-

Проведен анализ методов реконфигурируемых комплексных фильтров на базе резервных поднесущих и поднесущих данных во временной и частотной областях и рассмотрено их применение для уменьшения пик-фактора сигналов с OFDM модуляцией. Проанализированы и запрограммированы аппаратные структуры этих фильтров на ПЛИС. Представлены результаты моделирования предлагаемых способов и экспериментальные результаты на ПЛИС для типов модуляции 16-QAM и 64-QAM. Показано, что порекомендованные методы позволяют значительно снизить пикфактор (примерно на 6,6 дБ) и OFDM сигналы достигают желаемого уровня ограничения после 2-4 итераций.

> ражение SLM (Selective Mapping), частичная последовательность передачи PTS (Partial Transmit Sequence), способ резервных поднесущих TR (Tone Reservation), инжектирование тона TI (Tone Injection) и расширение активного созвездия ACE (Active Constellation Extension).

> Ограничение амплитуды и его производные типы [1], [2], [3] являются самыми простыми методами для уменьшения PAPR, которые ограничивают сигналы на заданном пороговом значении. Однако они вызывают внутриполосное и внеполосное искажения. Внеполосное излучение снижает спектральную эффективность и приводит к увеличению энергетических потерь, но оно может быть уменьшено с помощью фильтрации и при использовании повторной операции ограничения и фильтрации. После фильтрации появляются новые пики сигнала, превышающий уровень ограничения. Тем не менее, внутриполосное искажение не может быть уменьшено этим фильтром. Внутриполосное искажение нару-шает ортогональность поднесущих, что приводит к ухуд-шению показателей ошибки (вероятность ошибки BER, коэффициент ошибок модуляции MER), и помехоустойчивость приема снижается из-за влияния искажений. В некоторых случаях невозможно демодулировать сигналы на приёмной стороне. Чем больше ограничение амплитуды, тем выше внутриполосное искажение.

Были предложены способы перестановки сигналов, включающие PTS [4] и SLM [5]. Эти подходы уменьшают PAPR OFDM сигнала путем понижения вероятности возникновения синфазных модулированных поднесущих. Они не вносят внутриполосное и внеполосное искажения в передаваемые сигналы. Однако при этом требуется большое количество вычислений ОДПФ (обратное дискретное преобразование Фурье) для нахождения оптимальных фазовых последовательностей. Поэтому требуются такие ресурсы системы, как повышенная вычислительная производительность, большой объем памяти и высокая скорость обработки сигналов. Другими недостатками этих методов является то, что дополнительная информация должна быть передана от передатчика к приемнику для восстановления исходных данных, и некорректная полученная дополнительная информация в приемнике приводит к ошибкам.

Два метода борьбы с пик-фактором ACE и TR предусмотрены в стандарте системы DVB-T2 [6], [7] и в других OFDM системах, так как они не дают искажений и не требуют дополнительной информации. Метод АСЕ и его производный тип [8], [9] уменьшают PAPR путем увеличения расстояния некоторых крайних точек созвездия от центра без изменения их минимального расстояния. Однако, АСЕ не может быть использован при повороте созвездия, а также неэффективен при модуляции QAM высокого порядка. Метод TR был предложен Tellado [10] на основе импульсного ядра, генерируемого зарезервированными несущими. При этом OFDM сигналы подвергают специальной итеративной обработке, которая каждую итерацию выполняет поиск наибольшего пика OFDM сигнала для формирования сигнала «анти-пик». В общем случае чем больше число итераций, тем более низкое значение PAPR OFDM сигнала будет получено. Это увеличивает задержку обработки, из-за которой появляются «узкие места» обработки данных, что не удовлетворяет реализации аппаратных средств. Меньше 10 итераций рекомендуется в [6].

В данной статье проанализированы два реконфигурируемых комплексных фильтра для подавления пикфактора. Комплексный фильтр реконфигурирован для выделения резервных поднесущих и поднесущих данных. Конфигурируемый по резервным несущим фильтр итеративно используется и существенно подавляет пикфактор OFDM сигнала. Этот фильтр не вносит внутриполосное и внеполосное искажения в передаваемые сигналы. Он использует лишь резервные поднесущие для формирования сигнала «анти-пик». В последней итерации фильтр реконфигурируется для выделения поднесущих данных. В этой итерации амплитуда пикфактора малая. При этом малый шум ограничения вносится в передаваемые сигналы так, что пики OFDM сигналов быстро доводятся до желаемого уровня ограничения, и вторичные пики не появляются. Предлагаемый фильтр может подавить все пики в одной итерации, и повторный процесс фильтрации уменьшает вторичные пики. Этот фильтр может быть эффективно реализован на ПЛИС с использованием ДПФ/ОДПФ [11], [12] в частотном пространстве или с помощью фильтра с конечной импульсной характеристикой (FIR-фильтр) [13] во

временной области. Кроме того, архитектура реконфигурируемого фильтра не отличается от обычного комплексного фильтра, только его коэффициенты (или индексы поднесущих) обновляются, когда это необходимо. Реализация предлагаемого фильтра на ПЛИС затрачивает те же аппаратные ресурсы, что и реализация фильтра в способе ограничения и фильтрации, причем не требуется изменения структуры демодуляции OFDMсигнала.

Анализ предлагаемых алгоритмов и аппаратных архитектур Теоретические основы обработки OFDM сигналов для уменьшения пик-фактора

ОFDM модуляция использует N близко расположенных ортогональных поднесущих, в том числе часть N_r резервных поднесущих $\mathfrak{R} = \{r_0, r_1, ..., r_{N_r-1}\}$, называемых частью TR или \mathfrak{R} , зарезервированных с целью уменьшения пик-фактора. N_C из $(N\!-\!N_r)$ остальных поднесущих, называемых частью \mathfrak{R}^c , используются для передачи данных. ОFDM сигналы получаются с использованием ОДПФ:

$$x_n + c_n = \frac{1}{\sqrt{N}} \sum_{k=0}^{N-1} (X_k + C_k) e^{j2\pi kn/N} , \qquad (1)$$

где X_k и C_k – поднесущие данных и резервные поднесущие соответственно, которые определяются в непересекающихся позициях поднесущих, т.е. $X_k = 0$, $k \in \Re$ и $C_k = 0$, $k \notin \Re$; k – индекс отсчетов в частотной области или индекс поднесущей; c – вектор для уменьшения пикового значения сигнала; x – вектор сигнала данных во временной области; n – индекс отсчетов во временной области.

Суть алгоритма резервных несущих заключается в определении значения резервных поднесущих или генерации вектора с так, чтобы удалить пиковые значения сигнала. Предлагаемые методы основаны на применении к OFDM сигналам оператора ограничения и на реконфигурируемом фильтре. Оператор ограничения определяется следующим образом:

$$\overline{f}_n(x_n) = \begin{cases} x_n, & |x_n| \le A \\ Ae^{j\theta_n}, & |x_n| > A \end{cases},$$
(2)

где A – желаемый уровень отсечения; $x_n = |x_n| e^{j\theta_n}$; θ_n – фаза n-го отсчета входного сигнала.

Шум ограничения представляет собой разность отсчетов OFDM сигнала и их ограниченной версии. Он имеет вид серии импульсов и может быть получен как

$$f_n(x_n) = x_n - \overline{f_n}(x_n) \,. \tag{3}$$

Реконфигурируемый фильтр, обеспечивающий выделение резервных поднесущих, характеризуется следующей частотной характеристикой:

$$H_{k} = \begin{cases} 1, & k \in \Re\\ 0, & \text{иначе} \end{cases}.$$
(4)

Импульсная характеристика этого фильтра получается при применении ОДПФ:

$$h_{n} = IDFT(H) =$$

$$= \frac{1}{\sqrt{N}} \sum_{k=0}^{N-1} H_{k} e^{j2\pi kn/N} = \frac{1}{\sqrt{N}} \sum_{k \in \Re} e^{j2\pi kn/N}.$$
(5)

Этот фильтр может быть рассмотрен в качестве FIRфильтра *N*-го порядка, связь между входным и выходным сигналами которого описывается разностным уравнением (дискретной свёрткой). При этом, для того, чтобы получить вектор для уменьшения пик-фактора, шум ограничения подается на вход фильтра:

$$c_n = \sum_{k=0}^{N-1} h_k f_{n-k} = h_n * f_n , \qquad (6)$$

где (*) – оператор свертки.

٨

Предположим, что существуют P пиков, превышающих пороговое значение A. Они появляются в позициях $n_1, n_2, ..., n_P$. Индексный набор пиков обозначается S_P . Сначала рассмотрим частный случай, когда ограничен только один пик в положении $n_i \in S_P$. В этом случае, шум ограничения может быть выражен следующим образом:

$$f_n = f_{n_i} \delta(n - n_i) , \qquad (7)$$

где $\delta(n)$ – функция дельта Кронекера.

Z-преобразование уравнений (6) и (7) дает величину коррекции в *Z*-области для удаления отсчета пика на позиции *n_i*:

$$C(z) = H(z)F(z) = f_{n_i} z^{-n_i} H(z).$$
(8)

Из уравнения (8) видно, что коррекция представляет собой импульсную характеристику фильтра, циклически сдвинутую на n_i и масштабированную по величине шума ограничения на позиции n_i . После циклического сдвига первичный пик (h_0) импульсной характеристики сдвинут на n_i . Кроме того, из уравнения (5) h_0 обладает самым большим значением, а другие отсчеты ($h_1 \div h_{N-1}$) имеют значение, значительно меньше, чем h_0 , то есть выход фильтра близок к дискретному импульсу и позиция его пика совпадает с позицией пика шума ограничения. Чтобы подавить рассматриваемый пик, амплитуда пика на выходе фильтра должна приближаться к f_{n_i} . Поэто-

му импульсная характеристика фильтра должна выбираться так, чтобы *h*₀ = 1, или:

$$\frac{1}{N_r} \sum_{k \in \Re} e^{j2\pi kn/N} = \frac{\sqrt{N}}{N_r} IDFT(H) = \frac{\sqrt{N}}{N_r} \frac{h_n}{0 \le n \le N-1},$$
(9)

где $\frac{\sqrt{N}}{N_r}$ называют коэффициентом масштаба.

В общем случае шум ограничения во временной области и *Z*-преобразование коррекции представлены следующим образом:

$$f_n = \sum_{n_i \in S_P} f_{n_i} \delta(n - n_i) , \qquad (10)$$

$$C(z) = \frac{\sqrt{N}}{N_r} H(z) F(z) = \frac{\sqrt{N}}{N_r} \sum_{n_i \in S_P} f_{n_i} z^{-n_i} H(z) .$$
(11)

В частотной области шум ограничения и выходной сигнал фильтра могут быть выражены как:

$$F_k = DFT(f_n) = \sum_{n=0}^{N-1} f_n e^{-j2\pi nk/N} = \sum_{n \in S_P} f_n e^{-j2\pi nk/N} , \quad (12)$$

$$C_k = \frac{\sqrt{N}}{N_r} F_k H_k = \frac{\sqrt{N}}{N_r} F_k .$$
(13)

Из уравнения (13) видно, что сохраняются только дискретные частотные составляющие F_k в позициях, соответствующих индексам зарезервированных поднесущих, а остальные составляющие обнуляются. Коррекция c, полученная на выходе фильтра, может значительно уменьшить пики в соответствии с индексами набора S_P . Однако отсчеты коррекции c_n в других позициях не равны нулю, это приводит к тому, что пики, превышающие уровень ограничения, могут появиться в любой из N позиций отсчетов OFDM сигнала. Таким образом, генерирование сигнала для уменьшения пикфактора может повторяться несколько итераций или дальше выполняться фильтром, реконфигурируемым для выделения поднесущих данных.

Авторы в [14] предлагали итеративную оценку и аннулирование шума ограничения. Итеративный процесс оценки может восстанавливать и ликвидировать шум ограничения от приемного сигнала. На практике модуляторы, такие как возбудители стандарта DVB-T, DVB-T2, DAB, требуют MER примерно 43 дБ. Следовательно, в последней итерации производится реконфигурация фильтра по поднесущим данных. Аналогично, частотная и импульсная характеристики фильтра в этой конфигурации приведены ниже:

$$H_k^c = \begin{cases} 1, & k \in \mathfrak{R}^c \\ 0, & \text{иначе} \end{cases},$$
(14)

$$h_{n}^{c}_{0 \le n \le N-1} = \frac{1}{\sqrt{N}} \sum_{k \in \Re^{c}} e^{j2\pi kn/N} = IDFT(H^{c}).$$
(15)

При этом малый шум ограничения вносится в передаваемые сигналы. Искажение будет уменьшаться с ростом количества итераций.

Описание предлагаемых алгоритмов

Из приведенного выше анализа два предлагаемых алгоритма могут быть описаны в алгоритме 1 и алгоритме 2 соответственно.

Два предлагаемых алгоритма отличаются конфигурацией фильтра в шагах 4 и 7. В алгоритме 1 комплексный фильтр, спроектированный на основе пары ДПФ/ОДПФ [12], предназначен для генерирования сигнала коррекции. Вычислительная сложность этого фильтра эквивалентна сложности пары ДПФ/ОДПФ. Для алгоритма 2 использован комплексный FIR-фильтр. Сложность этого алгоритма зависит от выбранной архитектуры FIR-фильтра [13]. ДПФ и ОДПФ могут быть эффективно реализованы на ПЛИС с помощью IP-ядра [11], интегрированного в средах разработки (ISE, Vivado, System Generator for DSP), в то время как комплексный FIR-фильтр недоступен в этих средах. Хотя можно увидеть, что оба алгоритма дают те же самые результаты на Matlab.

Можно заметить, что предлагаемый реконфигурируемый фильтр проектируется на основе простой структурной схемы способа ограничения и фильтрации [2] и отличается от этой схемы в его реконфигурации для выделения резервных поднесущих и поднесущих данных. Реконфигурация фильтра используется для достижения обоих способов ограничения и фильтрации и TR.

Предлагаемое устройство на ПЛИС

Устройство для реализации предлагаемых алгоритмов на ПЛИС реализовано с использованием 16-битных чисел с фиксированной точкой. Структурные схемы алгоритма 1 и алгоритма 2 изображены на рис. 1 а и рис. 1 б соответственно. Структурная схема на рис. 1 а состоит из двух пар ключей и девяти основных блоков: ограничитель уровня сигнала, блок ДПФ, блок выделения необходимых поднесущих (ТR или данных), блок ОДПФ, блок масштабирования, блок буфера *x*, блок обновления с и два блока сложения. Ограничитель уровня сигнала (см. рис. 2) используется для вычисления шума ограничения при конфигурации фильтра, обеспечивающей выделение резервных поднесущих, или для ограничения входного сигнала при другой конфигурации фильтра. Он содержит несколькие арифметических компонент: блоки CORDIC, блок вычитания, блок сравнения и блоки умножения. Конвейерная архитектура используется в проектировании блоков CORDIC для ускорения работы устройства. Входной сигнал x_{ex} ограничителя представляет собой оригинальный OFDM-сигнал x в первой итерации или OFDM-сигнал $x_{ex} = |x|_{ex} e^{j\theta} = x_{ex} I + jx_{ex} Q$, включающий вещественную $x_{ex} I$ и мнимую $x_{ex} Q$ части, подается на вход блока CORDIC Artan, первый выход которого $|x|_{ex} -$ амплитуда входного сигнала – соединен с первым

Алгоритм 1: комплексный фильтр на основе ДПФ/ОДПФ

Инициализация: Выбирают желаемое значение ограничения *A*, часть TR, и максимальное количество итераций *MaxIter*. Итерационный процесс: нижеследующие процедуры повторяются для каждого OFDM символа. Для *m*-го символа:

1. Исходные значения: $x = x^{(m)}$; $c^{(0)} = \{c_n^{(0)} = 0 | 0 \le n \le N - 1\}$.

2. *і* начинается с 1.

3. Вычисляют $f_n(x+c^{(i-1)})$ с использованием (3).

4. Рассчитывают коррекцию $c^{(i)}$ следующим образом:

а) Преобразование f_n в частотную область для получения F_k с использованием (12).

b) Масштабирование составляющих F_k с индексами, совпадающими с индексами резервных поднесущих и обнуление остальных составляющих для получения C_k с использованием (13).

с) Преобразование C_k во временную область с использованием ОДПФ и обновлять вектор $c^{(i)}$: $c^{(i)} = c^{(i-1)} - IDFT(C_k)$.

5. *i* = *i* + 1. Если *i* < *MaxIter*, выполняют возврат к шагу 3. В противном случае выполняют переход к шагу 6.

6. Вычисляют $\overline{f}_n(x+c^{(MaxIter-1)})$ с использованием (2).

7. Преобразуют $\overline{f}_n(x + c^{(MaxIter-1)})$ в частотную область. Выделяют поднесущие данных. Преобразуют их во временную область для получения \overline{x} .

8. Переданный сигнал $x' = \overline{x} + c^{(MaxIter-1)}$

Алгоритм 2: комплексный FIR-фильтр

Инициализация:

• • • • •

4. Вычисляют коррекцию $c = \{c_n | 0 \le n \le N-1\}$ с использованием (6). Обновляют вектор $c^{(i)}$: $c^{(i)} = c^{(i-1)} - \frac{\sqrt{N}}{N_r}c^{(i)}$

7. Отсчеты сигнала $\overline{f}_n(x+c^{(MaxIter-1)})$ поступают на вход фильтра с импульсной характеристикой (15) для получения \overline{x} . 8. Переданный сигнал $x' = \overline{x} + c^{(MaxIter-1)}$.





Рис. 1. Блок-схема высокого уровня предлагаемого устройства на ПЛИС



Рис. 2. Блок-схема блока ограничителя уровня сигнала

входом блока сравнения, первым входом первого ключа и первым входом блока вычитания, а второй выход θ – фаза входного сигнала - соединен с входом блока CORDIC sin-cos. Сигнал порога ограничения А поступает на второй вход блока сравнения и второго ключа. Выход блока сравнения равен нулю при $|x|_{ex} \leq A$, в противном случае он равен единице. Этот выход управляет подключением двух входов первого ключа к выходу $|x|_{op}$ (|x|_{огр} – амплитуда ограниченного входного сигнала), который поступает на второй вход блока вычитания и второго ключа. Выход блока вычитания, представляющий собой амплитуду ограниченного шумового сигнала |x|_{шум}, подается на первый вход второго ключа. Второй ключ подключает |x|огр или |x|шум к выходу в соответствующей конфигурации фильтра, соединенному с первым входом блоков умножения, на второй из которых поступают соответственно сигналы $cos\theta$ и $sin\theta$, генерируемые блоком CORDIC sin-cos. Выход блоков умножения подается на входы блока ДПФ (см. рис. 1 а).

Блок ДПФ преобразует вещественную и мнимую части входа в частотную область. Его выходы (вещественная и мнимая части) поступают на входы блока 3 для выделения необходимых поднесущих (резервных поднесущих или поднесущих данных) при соответствующей конфигурации фильтра. Этот блок включает в себя два модуля памяти и два логических умножителя. Модули памяти хранят позиции резервных поднесущих и поднесущих данных в виде логического значения ('0' или '1'), поступающего на вход логических умножителей для выбора необходимых поднесущих на выходе ДПФ. Это приводит к уменьшению объема встроенной памяти на ПЛИС и к гибкости реконфигурации фильтра. Блок ОДПФ преобразует составляющие, выделенные блоком 3, во временную область.

При конфигурации фильтра для выделения резервных поднесущих сигнал после блока масштабирования поступает на блок 7 для обновления сигнала «антипик», обновленный сигнал «антипик» $c^{(i)}$ на выходе которого передается на блок сложения 8, суммирующий $c^{(i)}$ и оригинальный ОFDM-сигнал x, выход которого подается на вход ограничителя уровня сигнала, чтобы начать следующий шаг итерации. При другой конфигурации фильтра этот сигнал, представляющий собой ограниченный OFDM-сигнал \overline{x} , подается на блок сложения 9, на другие входы которого поступает также обновленный сигнал «анти-пик» с. Блок сложения 9 генерирует переданный сигнал \overline{x} и обновленного сигнала «анти-пик» c.

Проектирование комплексного FIR-фильтра (см. рис. 16) выполняется с использованием двух FIR-фильтров с реальными коэффициентами.

Комплексное число X можно представить как пару вещественной X_I и мнимой X_Q частей, $X = X_I + jX_Q$. Произведение комплексных чисел может быть вычислено как:

$$Y_{I} + jY_{Q} = (X_{I} + jX_{Q})(W_{I} + jW_{Q}) =$$

= $(X_{I}W_{I} - X_{O}W_{O}) + j(X_{I}W_{O} + X_{O}W_{I}).$ (16)

Из уравнения (16) видно, что комплексный фильтр может быть реализован с использованием четырех фильтров с реальными коэффициентами как: два фильтра с коэффициентами W_I и два фильтра с коэффициентами

ентами W_Q используются для фильтрации вещественной X_I и мнимой X_Q частей сигнала. Кроме того, FIR-фильтр может быть конфигурирован для параллельной обработки нескольких сигналов [13]. Поэтому нужны только два FIR-фильтра с коэффициентами W_I и W_Q ,

каждый из которых имеет два входа X_I и X_Q .

Кроме того, из уравнений (6), (9) и (15) видно, что *n*-й выходной отсчет связан с *N* упорядоченных входных отсчетов, циклически сдвинутых на (*n*+1). Например, 0-й выходной отсчет *y*(0) связан с входным вектором {*x*(1), *x*(2),..., *x*(*N*-1), *x*(0)} (*N* входных отсчетов сдвинуты на одну позицию). Аналогично, 1-й отсчет *y*(1) связан с входным вектором {*x*(2), *x*(3),..., *x*(*N*-1), *x*(0), *x*(1)}. В общем случае *n*-й отсчет *y*(*n*) связан с входным вектором {*x*(*n*+1), *x*(*n*+2), ..., *x*(*N*-1), *x*(0), *x*(1), ..., *x*(*n*)}. Поэтому кроме двух FIR-фильтров буферы данных FIFO (first in first out) используются для циклического сдвига входных отсчетов.





На рис. З показан принцип работы алгоритма 2. Действующий сигнал *valid_i* – вход с активным верхним уровнем – активен, когда символ поступает для обработки. Сигнал *data_i* – это сигнал на выходе ограничителя уровня сигнала (см. рис. 2), который включает две части: вещественную и мнимую. Когда *valid_i* активен, этот сигнал поступает на входы фильтров, а также записывается в FIFO для выполнения циклического сдвига. В противном случае, когда *valid_i* неактивен, такой сигнал считывается из FIFO и поступает на вход фильтров. Сигнал *rdy_o* указывает на то, что сигналы на выходах фильтров готовы. Число выходных отсчетов в два раза больше, чем количество входных отсчетов из-за обработки циклического сдвига. Тем не менее, лишь его вторая половина является правильным выходом. Другие блоки на рис. 1 б функционируют как соответствующие блоки на рис. 1 а.

FDATool представляет собой средство для определения цифровых фильтров с графическим интерфейсом пользователя. Блок Xilinx FDATool пакета System Generator for DSP используется для передачи коэффициентов в FIR-фильтры. Импульсная характеристика непосредственно импортирована в FIR-фильтр с помощью функций MATLAB. Амплитудно-частотная характеристика (AЧX) приведена на рис. 4, на котором содержатся линии с равными амплитудами на нормированных позициях, соответствующих индексам резервных поднесущих или поднесущих данных в соответствующей конфигурации фильтра.



Экспериментальные результаты Результаты моделирования в среде Matlab

Эффективность предлагаемых алгоритмов была оценена в среде Matlab с помощью функции выборочного вероятностного распределения (CCDF – complementary cumulative distribution function) и уровня ошибок модуляции MER, где CCDF представляет собой вероятность того, что PAPR превышает заданный порог PAPR0. Были использованы нормированные символы модуляции 16-QAM и 64-QAM в качестве входа системы OFDM, количество поднесущих N = 1024, доля случайно выбранных резервных поднесущих – 5% от общего количества активных несущих, амплитуда ограничения $A \approx 6,62$ дБ.

Таблица	1:	Параметры	uр	результаты	моделирован	ия
---------	----	-----------	----	------------	-------------	----

Параметры системы	Значение параметров				
Тип модуляции	16-QAM	64-QAM			
Количество поднесущих N	1024	1024			
Количество поднесущих наборов ПиП ^с	999	999			
Количество резервных поднесущих	50	50			
Количество OFDM символов	10000	10000			
Количество итераций	3	4			
Уровень ограничения А (дБ)	6,66	6,62			
Результат исследования					
Исходный PAPR	13,35	13,24			
PAPR после обработки	6,71	6,63			
Выигрыш (дБ)	6,64	6,61			
MER (дБ)	50,45	56,76			
Увеличение средней мощности (дБ)	0,072	0,082			

Параметры системы, используемые для моделирования, и результаты исследования представлены в табл. 1. На рис. 5 показаны результаты снижения PAPR с помощью предлагаемых алгоритмов при различном количестве итераций на примере символов модуляции 64-QAM. После четырех итераций PAPR снижается с 13,24 дБ до 6,63 дБ (т.е. абсолютный выигрыш составляет 6,61 дБ) при *MER* = 56,76 дБ и вероятности CCDF 10⁻⁴. Аналогично, для символов модуляции 16-QAM абсолютный выигрыш в снижения пик-фактора составляет 6,64 дБ после трех итераций при *MER* = 50,45 дБ (см. табл. 1).



Puc. 5. PAPR CCDF предлагаемых алгоритмов для модуляции 64-QAM





На рис. 6 в качестве примера проиллюстрированы результаты обработки снижения пик-фактора OFDM сигнала для модуляции 64-QAM с помощью предлагаемых алгоритмов. Качество сигнала с точки зрения MER составляет 56,76 дБ и 50,45 дБ для символов модуляции 64-QAM и 16-QAM, соответственно. Увеличение средней мощности переданного сигнала незначительно (меньше 0,1 дБ). Можно дополнительно уменьшить PAPR при снижении амплитуды ограничения *А* и росте количества итераций.

Результаты реализации на ПЛИС

Предлагаемые способы были реализованы с помощью аппаратных средств и протестированы на чипе XC7A200T-1SBG484C на плате Nexys Video Artix-7 FPGA. Блок-схема устройства на пакете System Generator показана на рис. 7.

Блок 1 используется для генерации псевдослучайных данных и отображения сигнала на нормированные значения созвездия (16-QAM или 64-QAM). Для уменьшения потребления ресурсов ПЛИС псевдослучайные данные генерируются регистром сдвига с линейной обратной связью с помощью любого порождающего полинома. Блок ОДПФ собирает NC ячеек модулированных данных и Nr зарезервированных ячеек, которые обнуляются, в массив активных ячеек OFDM-символа и преобразует их в временную область для формирования оригинального OFDM-сигнала. Блок 4 (предлагаемый способ) выполняет уменьшение пик-фактора сигнала. Блок 5 вычисляет амплитуду оригинального OFDM-сигнала x'_4 .

Результаты снижения PAPR были оценены с использованием Xilinx ChipScope Pro Analyzer (внутрисхемный логический анализатор – ВСЛА), специально разрабатываемого для целей тестовых выводов и отладки на печатной плате. ВСЛА позволяет наблюдать за любым нужным внутренним сигналом. Амплитуды OFDM сигналов до и после снижения PAPR захватываются и фиксируются во внутренних блоках памяти на ПЛИС, потом ВСЛА просматривает захваченные данные на персональном компьютере.

Экспериментальные результаты на пакете System Generator и с помощью анализатора ВСЛА представлены на рис. 8 и рис. 9 соответственно. Амплитуды представлены 18-битным беззнаковым числом с фиксированной точкой, включая 15 наименьших значащих битов на дробной части. На рис. 8 показано, что амплитуда оригинального сигнала составляет примерно 4.5, амплитуда сигнала после уменьшения PAPR составляет примерно 2.1. Таким образом, отношение пиковой амплитуды OFDM сигнала до и после обработки будет более двух, т.е. пиковая мощность снижается больше, чем на 6 дБ.



Рис. 7. Экспериментальная блок-схема алгоритмов на System Generator

Захваченные данные с ВСЛА (см. рис. 9) показывают, что амплитуда OFDM сигналов до (красный) и после (голубой) обработки PAPR составляют соответственно 130000 и 60000, (в ВСЛА 18-битное число данных представлено в виде целого числа). Следовательно, экспериментальные результаты на ПЛИС дают такой же выигрыш, как и результаты на пакете System Generator и в среде Matlab.





Рис. 9. Экспериментальные результаты с использованием ВСЛА

Алгоритм 1 хорошо подходит для длинных OFDM символов за счет длины ДПФ до 64К и его вычислительной эффективности [11]. Затрата ресурсов его реализации на ПЛИС по параметрам в табл. 1 представлена в табл. 2. Видно, что устройство алгоритма 1 потребляет малые ресурсы ПЛИС, в том числе максимальное отношение потребления ресурсов составляет 10 % DSP48E1 (встроенные умножители).

Устройство алгоритма 2 требует наличия двух бу-

феров с объемом *N*х16 бит для циклического сдвига входных отсчетов и четырех модулей памяти для хранения вещественной и мнимой частей коэффициентов фильтра. Объем каждого модуля памяти составляет *N*х16 бит. Кроме того, для реконфигурации FIR-фильтра требуется, как правило, процессорное ядро (MicroBlaze, PowerPC, ARM SoC core, PicoBlaze). Более того, IP-ядро FIR-фильтра поддерживает до 2048 (2K) коэффициентов [13], что не подходит для длинных OFDM символов.

При реализации параллельной архитектуры FIRфильтра по параметрам в табл. 1 требуемое число DSP48E1 превышает возможное количество на чипе XC7A200T. Табл. 3 показывает потребление ресурсов устройства алгоритма 2 при длине OFDM символов N = 256 и других параметрах, как в табл. 1. Тем не менее, задержка обработки алгоритма 2 (примерно в два раза больше длины OFDM символа) меньше, чем задержка алгоритма 1 (примерно в четыре раза больше длины OFDM символа), так как алгоритм 2 обновляет коррекцию только во временной области.

Заключение

В данной статье было предложено два реконфигурируемых комплексных фильтра для снижения пикфактора OFDM сигнала. Фильтр, конфигурируемый для выделения резервных поднесущих, существенно подавляет пик-фактор сигнала. Эта конфигурация фильтра не вносит искажения в передаваемые сигналы. В конфигурации фильтра для выделения поднесущих данных малый шум ограничения вносится в передаваемые сигналы так, что пики OFDM сигналов быстро доводятся до желаемого уровня ограничения, при этом вторичные пики не появляются. Искажение будет уменьшаться с ростом количества итераций.

Средняя мощность переданного сигнала увеличивается из-за конфигурации фильтра по резервным поднесущим, но её увеличение незначительно (примерно 0,1 дБ).

Таблица 2: Пот	ребление ресурсое	в комплексного филь	ьтра на основе I	ΠΦ/ΟΔΠΦ

Device Resources	Available	Used	Utilization
Slice LUTs	134600	6772	5,03 %
Slice Registers	269200	14029	5,21 %
Block RAM Tile	365	9,5	2.60 %
DSP48E1s	740	74	10 %
Maximum Frequency	158,730 MHz		
Output latency (clock cycles)	4276		

Таблица 3: Потребление ресурсов комплексного FIR-фильтра

Device Resources	Available	Used	Utilization
Slice LUTs	134600	20301	15,08 %
Slice Registers	269200	21177	7,87 %
Block RAM Tile	365	2	0,55 %
DSP48E1s	740	520	70,27 %
Maximum Frequency	127,877 MHz		
Output latency (clock cycles)	562		

Реализация комплексного фильтра на базе ДПФ/ОДПФ на ПЛИС хорошо подходит для длинных ОFDM символов и дает высокую скорость обработки, низкое потребление ресурсов ПЛИС и гибкость реконфигурации фильтра. Комплексный FIR-фильтр затрачивает большое количество ресурсов, особенно DSP48E1, но дает низкую задержку обработки.

Оба алгоритма дают одинаковые результаты с точки зрения производительности уменьшения PAPR. В будущем предполагается реализовывать эти алгоритмы в структуре возбудителей стандарта DVB-T/T2 и комбинировать их с другими известными способами, чтобы увеличить эффективность снижения PAPR.

Литература

1. Han S.H. and Lee J.H. An overview of peak-toaverage power ratio reduction techniques for multicarrier transmission // IEEE Wireless Communications, vol. 12, no. 2, pp. 56–65, April 2005.

2. Armstrong J. New OFDM peak-to-average power reduction scheme // Proc. IEEE, VTC2001 Spring, Rhodes, Greece, pp. 756–760, Aug. 2002.

3. Shang-Kang Deng, Mao-Chao Lin, Recursive Clipping and Filtering With Bounded Distortion for PAPR Reduction // IEEE Transactions on Communications, vol. 55, no. 1, pp. 227–230, Jan. 2007.

3. Muller S.H. and Huber J.B. OFDM with reduced peak to average power ratio by optimum combination of partial transmit sequences // Electronics Letters, vol. 33, no. 5, pp. 368-369 February 1997.

4. Bauml R.W., Fischer R.F.H. and Huber J.B. Reducing the peak-to-average power ratio of multicarrier modulation

by selected mapping // IEEE Electronics Letters, vol. 32, no. 22, pp. 2056–2057, Sep. 1996.

5. EN 302 755 V1.4.1, Digital video broadcasting (DVB); Frame structure channel coding and modulation for a second generation digital terrestrial television broadcasting system // European Standard, July 2015.

6. Дворкович В.П., Дворкович А.В. Цифровые видеоинформационные системы (теория и практика) // Москва: Техносфера, 2012. – 1008 с.

7. Krongold B.S. and Jones D.L. PAR reduction in OFDM via active constellation extension // IEEE Trans. Broadcast., vol. 49, no. 3, pp. 258–268, Sep. 2003.

8. Bae K., Andrews J.G. and Powers E.J. Adaptive active constellation extension algorithm for peak-to average ratio reduction in OFDM // IEEE Commun. Lett., vol. 14, no. 1, pp. 39–41, Jan. 2010.

9. Tellado J. Peak to average power reduction for multicarrier modulation // Ph.D. dissertation, Stanford Univ., Stanford, CA, 2000.

10. Xilinx, pg109, Fast Fourier Transform v9.0 // Logi-CORE IP Product Guide, 97p., Nov. 2015.

11. Dinechin F., Takeugming H. and Tanguy J.-M. A 128-tap complex FIR filter Processing 20 giga-samples/s in a single FPGA // 2010 Conference Record of the Forty Fourth Asilomar Conference on Signals, Systems and Computers (ASILOMAR), Pacific Groove, CA, USA, pp. 841–844, Nov. 7–10, 2010.

12. Xilinx, pg149, FIR Compiler v7.2 // LogiCORE IP Product Guide, 131p, Nov. 2015.

13. Chen H. and Haimovish M. Iterative Estimation and Cancellation of Clipping Noise for OFDM Signals // IEEE Commun. Lett., vol. 7, no. 7, pp. 305–307, July 2003.



НОВЫЕ КНИГИ

Литвин С.М., Попов О.Б., Чернышева Т.В.

Аудиопроцессорная обработка сигналов звукового вещания – М.: Изд-во «Горячая линия-Телеком», 2017 г. – 222 с.: ил.

Звуковой сигнал, передаваемый по каналам звукового вещания, был и будет востребован в общем ряду с большим количеством альтернативных источников информации. Рассмотрены основные характеристики сигналов и каналов звукового вещания, базовые процедуры, используемые при обработке звукового вещательного сигнала (ЗВС); основные алгоритмы обработки сигнала в канале передачи. Проведен анализ и оценка искажений сигнала на всех этапах его передачи – от первичного цифрового преобразования до аудиопроцессорной обработки. Особое внимание уделено алгоритмам обработки и проблеме объективной оценки качества ЗВС в системах, не нормируемых в рамках современного метрологического обеспечения.

Для студентов, обучающихся по направлениям подготовки бакалавров и магистров 11.03.02 и 11.04.02 – «Инфокоммуникационные технологии и системы связи», 11.03.01 и 11.04.01 – «Радиотехника», 10.05.02 – «Информационная безопасность телекоммуникационных систем».

621.391

АЛГОРИТМИЧЕСКИЕ МЕТОДЫ КОМПЕНСАЦИИ НЕЛИНЕЙНОСТИ УСИЛИТЕЛЯ МОЩНОСТИ

Ле Ван Ки, аспирант кафедры мультимедийных технологий и телекоммуникаций МФТИ, e-mail: levanky@phystech.edu.

ALGORITHMIC METHODS FOR COMPENSATION OF POWER AMPLIFIER NONLINEARITY

Le Van Ky.

Presently there are many effective methods for nonlinearity compensation of digital radio signal power amplifier. In this paper we analyze several well-known algorithms and propose a new variant of nonlinearity compensation of power amplifier. Computer simulation results confirms the effectiveness of the proposed method.

Key words: adaptive predistortion, nonlinear distortion, Wiener model, Hammerstein model, singular value decomposition.

Ключевые слова: адаптивное предыскажение, нелинейное искажение, модель Винера, модель Хаммерштейна, сингулярное разложение.

Введение

Большинство цифровых радиосистем передачи информации содержат многочастотный групповой сигнал, весьма чувствительный к не-

линейным искажениям. Это, в свою очередь, предъявляет высокие требования к линейности аналогового тракта, и в частности, к линейности усилителя мощности (УМ) передатчика.

Для моделирования нелинейности УМ требуется учет эффекта памяти. Самый общий способ введения эффекта памяти – это использования ряда Вольтерра (Volterra), который рассматривается при проектировании предыскажения данных [1, 2]. Однако большое количество коэффициентов ряда Вольтерра делает его непривлекательным для практического применения. Поэтому обычно реализуется несколько частных случаев ряда Вольтерра, включающих модели Винера (Wiener) (рис. 1), Хаммерштейна (Hammerstein) (рис. 2), а также модель Винера-Хаммерштейна (рис. 3) [3].





Рис. 2. Модель Хаммерштейна



Рис. 3. Модель Винера-Хаммерштейна

Существует достаточно большое число алгоритмических эффективных способов компенсации нелинейности характеристик усилителей мощности цифровых радиосигналов. В статье проанализированы несколько известных алгоритмов и предложен новый вариант компенсации нелинейности усилителя мощности. Приведены результаты компьютерного моделирования, подтверждающие эффективность предложенных решений.

> Модель Винера включается в себя два последовательно соединенных блока – постоянный во времени линейный блок LTI (linear time-invariant) и нелинейный блок без памяти NL (memoryless nonlinearity). Математически эта модель описывается следующим образом:

$$u(n) = \sum_{l=0}^{k-1} a_l z(n-l), \ y(n) = \sum_{\substack{k=1\\k \text{ odd}}}^{k} b_k u(n) |u(n)|^{k-1},$$
$$y(n) = \sum_{\substack{k=1\\k \text{ odd}}}^{k} b_k \left[\sum_{l=0}^{k-1} a_l z(n-l) \right] \left| \sum_{l=0}^{k-1} a_l z(n-l) \right|^{k-1},$$

где *К* – максимальная степень нелинейности, *L* – максимальная задержка, *а* и *с* – коэффициенты линейного и нелинейного преобразований блоков соответственно.

Как и модель Винера, модель Хаммерштейна также включается в себя два блока, но соединенных в обратной последовательности:

$$v(n) = \sum_{\substack{k=1\\k \text{ odd}}}^{K} b_k z(n) |z(n)|^{k-1}, \ y(n) = \sum_{l=0}^{L-1} c_l v(n-l)$$
$$y(n) = \sum_{\substack{l=0\\k \text{ odd}}}^{L-1} c_l \sum_{\substack{k=1\\k \text{ odd}}}^{K} b_k z(n-l) |z(n-l)|^{k-1}.$$

Рассмотрим особенность модель Хаммерштейна. Обозначим $C = [c_0, c_1, \dots, c_{L-1}]^T$, $B = [b_1, \dots, b_K]^T$. Модель Хаммерштейна можно записать следующим образом:

$$y(n) = \sum_{l=0}^{L-1} \sum_{\substack{k=1\\k \text{ odd}}}^{K} c_l b_k z(n-l) \left| z(n-l) \right|^{k-1},$$

Если обозначить $a_{lk} = c_l b_k$, то для матрицу A, сотоящую из элементов a_{lk} , можно выразить как $A = CB^T$ и rank(A) = 1. Таким образом, модель Хаммерштейна является частным случаем модели полиномиальной памяти.

Математическое представление модели полиномиальной памяти выглядит следующим образом:

$$y(n) = \sum_{l=0}^{L-1} \sum_{k=1}^{K} a_{lk} z(n-l) |z(n-l)|^{k-1}$$

Модель Винера-Хаммерштейна является комбинацией моделей Винера и Хаммерштейна. Ее математическое представление выглядит следующим образом:

$$u(n) = \sum_{l=0}^{L_a-1} a_l z(n-l), \quad v(n) = \sum_{\substack{k=1\\k \text{ odd}}}^{K} b_k u(n) |u(n)|^{k-1},$$

$$y(n) = \sum_{l=0}^{L_c-1} c_l v(n-l),$$

$$y(n) = \sum_{\substack{l=0\\k \text{ odd}}}^{L_c-1} c_{l_1} \sum_{\substack{k=1\\k \text{ odd}}}^{K} b_k \left[\sum_{\substack{l=0\\l=0}}^{L-1} a_{l_2} z(n-l_2-l_1) \right] \left| \sum_{\substack{l=0\\l=0}}^{L_a-1} a_{l_2} z(n-l_2-l_1) \right|^{k-1},$$

где *a*, *c* и *b* – коэффициенты линейных блоков и нелинейного блока соответственно.

Методы компенсации нелинейности усилителя мощности

Для решения такого типа задач существует два эффективных популярных метода оценки на основе итерационных алгоритмов Ньютона (Newton) и Нарендры-Галмана (Narendra-Gallman) [4]. Основным недостатком этих алгоритмов являются их чувствительность к начальному предположению и то, что они могут сходиться лишь к локальному минимуму. Чтобы избежать этого, Бай (Bai) [5] использовал комбинацию двух методов – наименьших квадратов (least mean square, LMS) и сингулярного разложения (singular value decomposition, SVD), которые могут привести к глобальному минимуму.

Ниже предложена схема предыскажения (рис. 4) и проведено моделирование трех алгоритмов устранения влияния нелинейности УМ: метода LMS/SVD, метода полиномиальной памяти и комбинационного метода. В качестве блока УМ используются модели Хаммерштейна, Винера и Винера-Хаммерштейна.



В процессе обработки сигнала применяется сингулярное разложение (SVD) прямоугольной вещественной или комплексной матрицы М порядка $m \times n$, имеющее широкое применение при решении многих прикладных задач:

$$M = U\Sigma V^*$$

где \sum – матрица размера $m \times n$, у которой элементы, лежащие на главной диагонали – сингулярные числа матрицы M, а все элементы, не лежащие на главной диагонали, равны нулю; матрицы U (порядка m) и V (порядка n) – унитарные матрицы, состоящие из левых и правых сингулярных векторов матрицы M соответственно (V^* – сопряжённо-транспонированная матрица V).

Сингулярное разложение может быть использовано для нахождения псевдообратных матриц, которые применяются, в частности, в методе наименьших квадратов.

Если $M = V \Sigma U^*$, то псевдообратная к ней матрица M^+ находится по формуле:

$$M^+ = V\Sigma^+ U^*,$$

где Σ^+ – матрица, псевдообратная к матрице Σ , получаемая из неё заменой каждого ненулевого элемента σ на диагонали на обратный к нему $1/\sigma$ с последующим транспонированием самой матрицы.

На рис. 4 *y*(*n*) и *z*(*n*) – соответственно вход и выход блока предыскажения в обучающем режиме. Схема предыскажения с моделью Хаммерштейна подробно представлена на рис. 5.



Рис. 5. Схема предыскажения с моделью Хаммерштейна Методы LMS/SVD и полиномиальной памяти описаны во многих статьях, например, в [5, 6].

Комбинационный метод включает в себя 2 метода: метод наименьших квадратов и итерационный метод Ньютона. Его математическое описание выглядит следующим образом.

$$\hat{z}(n) = \sum_{k=0}^{K-1} \sum_{p=1}^{P} a_{pk} y(n-k) |y(n-k)|^{p-1} + \sum_{q=2}^{Q} b_{q} y(n) \left[\sum_{l=0}^{L-1} c_{l} |y(n-l)| \right]^{q-1},$$

где *P*, *Q* – максимальные степени нелинейности; *K*, *L* – максимальные задержки; *a*, *b* и *c* – коэффициенты каждого блока (приведены выше); *z* и *y* – выход и вход блока предыскажения соответственно.

Погрешность между результатом модуля обучения $\hat{z}(n)$ и выходом блока предыскажения z(n) оценивается так:

$$e(n) = z(n) - \hat{z}(n), \ J = \sum_{n=1}^{N} |e(n)|^2 = \sum_{n=1}^{N} e(n)e^*(n)$$
где e^* – это сопряжённо-транспонированная матрица *e*, *N* – количество отсчетов.

Оптимальные значения a_{pk} , b_q , c_l можно найти с помощью их производных.

$$\begin{split} \frac{\delta J}{\delta a_{kp}^*} &= -\sum_{n=1}^{N} e(n) u_{kp}^*(n) = 0, \ \frac{\delta J}{\delta b_q^*} = -\sum_{n=1}^{N} e(n) v_q^*(n) = 0, \\ \frac{\delta J}{\delta c_l} &= -\sum_{n=1}^{N} 2 \operatorname{Re} \Big[e(n) s_l^*(n) \Big] = 0, \\ \text{rge } u_{kp}(n) &= y(n-k) |y(n-k)|^{p-1}, \\ v_q(n) &= y(n) \bigg(\sum_{l=0}^{L-1} c_l |y(n-l)| \bigg)^{q-1}, \\ s_l(n) &= \sum_{q=2}^{Q} b_q y(q-1) \bigg(\sum_{l=0}^{L-1} c_l |y(n-l_1)| \bigg)^{q-2} |y(n-l)|. \end{split}$$

С помощью метода наименьших квадратов можно найти параметры a_{nk}, b_a .

$$\begin{split} & \left(\left[U V^{(i)} \right]^{H} \left[U V^{(i)} \right] \right) \begin{bmatrix} a \\ b \end{bmatrix} = \left[U V^{(i)} \right]^{H} z, \\ & \mathsf{F} \mathsf{De} \ U = \left[u_{10}, \dots, u_{K0}, \dots, u_{1q}, \dots, u_{KQ} \right], \ V^{(i)} = \left[v_{2}^{(i)}, \dots, v_{Q}^{(i)} \right], \\ & a = \left[a_{10}, \dots, a_{K0}, \dots, a_{1q}, \dots, a_{KQ} \right]^{T}, \ b = \left[b_{2}, \dots, b_{Q} \right]^{T}, \\ & z = \left[z(0), \dots, z(N-1) \right]^{T}, \ u_{kq} = \left[u_{kq}(0), \dots, u_{kq}(N-1) \right]^{T}, \\ & v_{q}^{(i)} = \left[v_{q}^{(i)}(0), \dots, v_{q}^{(i)}(N-1) \right]^{T}. \end{split}$$

Результат решения методом наименьших квадратов: $\begin{bmatrix} \hat{a}^{(i+1)} \\ \hat{b}^{(i+1)} \end{bmatrix} = \left([UV^{(i)}]^H [UV^{(i)}] \right)^{-1} [UV^{(i)}]^H z.$

Используя итерационный метод Ньютона по данным $\hat{a}^{(i+1)}, \hat{b}^{(i+1)},$ необходимо обновить значения $c^{(i+1)}$: $c^{(i+1)} = c^{(i)} - [\nabla_c^2 J(c^{(i)})]^{-1} \nabla_c J(c^{(i)}),$ $\left[\nabla_c J(c^{(i)})\right]_{ll} = \frac{\delta J}{\delta c_l}\Big|_{c^{(i)}}, \quad \left[\nabla_c^2 J(c^{(i)})\right]_{lm} = \frac{\delta^2 J}{\delta c_l \delta c_m}\Big|_{c^{(i)}}.$ $\frac{\delta^2 J}{\delta c_l \delta c_m} = \sum_{n=1}^N 2 \operatorname{Re}\left[-e(n)z_{lm}^*(n) + s_m(n)s_l^*(n)\right],$ $z_{lm}(n) = \frac{\delta s_l(n)}{\delta c_m} =$ $= \sum_{q=2}^O b_q y(q-1)(q-2)\left(\sum_{l_l=0}^{L^{-1}} c_{l_l} |y(n-l_l|)\right)^{q-3} |y(n-l||y(n-m|).$ $\nabla_c J(c^{(i)}) = -2\operatorname{Re}\left\{[S^{(i)}]^H E^{(i)}\right\}, \nabla_c^2 J(c^{(i)}) =$ $= 2\operatorname{Re}\left\{[S^{(i)}]^H S^i - [Z^{(i)}]^H E^{(i)}\right\},$ где $S^{(i)} = \left[s_0^{(i)}, \dots, s_{L^{-1}}^{(i)}\right], \quad Z^{(i)} = \left[\begin{array}{c} z_{10}^{(i)} \dots z_{1,L^{-1}}^{(i)} \\ \vdots \\ \vdots \\ \end{array}\right],$

$$E^{(i)} = \begin{bmatrix} e^{i} \dots 0 \\ 0 \dots 0 \\ \vdots \\ \vdots \\ 0 \dots e^{i} \end{bmatrix}.$$

Итерация методом Ньютона:

$$\hat{c}^{(i+1)} = \hat{c}^{(i)} + [Re\{[S^{(i)}]^H S^i - [Z^{(i)}]^H E^i\}]^{-1} Re\{[S^{(i)}]^H E^{(i)}\}.$$

Результат программной реализации

В эксперименте для моделей УМ Хаммерштейна и Винера-Хаммерштейна при использовании компенсации нелинейности методами LMS/SVD и полиномиального разложения было показано, что полиномов, содержащих только четные степени нелинейного блока, недостаточно для компенсации нелинейности характеристики УМ, в отличие от полиномов, содержащих только нечетные степени. В то же время использование полиномов, содержащих все натуральные степени, не дает существенного улучшения компенсации нелинейности характеристики УМ по сравнению с использованием полиномов, содержащих только нечетные степени, но приводит к увеличению времени вычислений. Поэтому в данной работе были использованы модели предыскажения, построенные с помощью полиномов, содержащих только нечетные степени.

Методы LMS/SVD и полиномиального разложения при использовании модели УМ Винера и комбинационный метод при использовании всех видов моделей УМ, приведенных выше, содержащих все натуральные степени, дают существенное улучшение компенсации нелинейности характеристики УМ по сравнению с использованием полиномов, содержащих только нечетные степени. Поэтому данные методы рассмотрены в двух вариантах. Результат компьютерного моделирования показан в табл. 1.

Амплитудная характеристика УМ и спектры сигналов с применением комбинационного метода и LMS/SVD, когда модель УМ является моделью Винера-Хаммерштейна, представлены на рис. 6 и 7. На рис. 6 размытая часть характеризует искажения сигнала при отсутствии коррекции; коррекция приводит к линеаризации характеристики усилителя (линия черного цвета); прямая темная линия иллюстрирует линейную характеристику.

Из рис. 7 видно, что комбинационный метод обеспечивает более гладкий спектр (правое изображение), чем метод LMS/SVD (левое изображение), и практически полностью компенсирует нелинейность характеристики УМ. Спектр черного цвета изображает исходный сигнал, спектры ярких и темных цветов характеризуют сигналы без предыскажения и с применением методов LMS/SVD или комбинационного метода соответственно.

На рис. 8 показаны сигнальные созвездия на выходе УМ без компенсации нелинейности (слева) и с применением комбинационного метода (справа). Сигнальные созвездия с применением метода LMS/SVD (слева) и метода полиномиальной памяти (справа) изображены на рис. 9.

			Т	аблица 1. Ошибк	и модуляции MER I	после усилит	еля мощности		
Модель	Ошибки моду- ляции MER после УМ (дБ) без предкор- рекции	MER после предкоррекции различными методами (дБ)							
		Metog LMS/SVD		Метод полиномиальной памяти		Комбинационный метод			
		нечетные степени	натуральные степени	нечетные степени	натуральные степени	нечетные степени	натуральные степени		
модель Винера	27	50,8	70,8	52,2	71.3	54,9	71,4		
модель Хаммерштейна	27	37,2		40,3		42	44,4		
модель Винера- Хаммерштейна	27	47,8		48,3		50,9	55,1		





0.9







Рис. 8. Сигнальные созвездия на выходе УМ без компенсации нелинейности (а) и с применением комбинационного метода (б)



Рис. 9. Сигнальное созвездие на выходе УМ

с применением методов LMS/SVD (а) и полиномиальной памяти (б)

Таблица 2. Ошибки модуляции MER после усилителя мощности при большей нелинейности

	Ошибки моду-	МЕR после предкоррекции различными методами (дБ)							
Модель	ляции MER после УМ (дБ) без предкор-	Метод Ll	MS/SVD	Метод полиномиальной памяти		Комбинационный метод			
	рекции	нечетные степени	натуральные степени	нечетные степени	натуральные степени	нечетные степени	натуральные степени		
модель Винера	21	Не работает	54,8	45,6	57,7	46,5	57,8		
модель Хаммер- штейна	21	Не работает	35,3	36,2	36,5	37,5	45,2		
модель Винера- Хаммерштейна	21	Не работает	41,4	41	42	43,4	48		

В эксперименте был также проанализирован вариант со значительной нелинейностью (малыми значениями ошибок модуляции) на выходе после УМ. Легко видеть, что комбинационный метод устойчив и лучше работает по сравнению с двумя другими методами. Результат компьютерного моделирования показан в табл. 2.

Заключение

По результатам проведенного анализа известных методов и алгоритмов компенсации нелинейности усилителя мощности был предложен новый комбинационный метод компенсации на основе методов Ньютона и наименьших квадратов.

По результатам компьютерного моделирования следует вывод о том, что модель предыскажения полиномиальной памяти наиболее простая. Комбинационный метод, хотя сложен в моделировании, но более эффективен по сравнению с другими рассмотренными вариантами.

Литература

1. Eun C. and Powers E.J., «A predistorter design for a memory-less nonlinearity preceded by a dynamic linear system», in Proc. IEEE Global Telecommun. Conf., vol. 1, pp. 152–156, Nov. 1995.

2. Eun C. and Powers E.J., «A new Volterra predistorter based on the indirect learning architecture», IEEE Trans. Signal Processing, vol. 45, pp. 223–227, Jan. 1997.

3. Kang H.W., Cho Y.S. and Youn D.H. «On compensating nolinear distortions of an OFDM system using ecient adaptive predistorter», IEEE Trans. Commun., vol. 47, pp. 522–526, Apr. 1999.

4. Eskinat E., Johnson S.H. and Luyben W.L. «Use of Hammerstein models in identification of nonlinear systems», AIChE J., vol. 37, pp. 255–267, Feb. 1991.

5. Bai E.W. «An optimal two stage identification algorithm for Hammerstein-Wiener nonlinear systems», in Proc. American Contr. Conf., pp. 2756–2760, June 1998.

6. Ding L., Zhou G.T., Morgan D.R., Ma Z., Kenney J.S., Kim J. and Giardina C.R. «Memory polynomial predistorter based on the indirect learning architecture», in Proc. IEEE Global Telecommun. Conf., pp. 967–971, Nov. 2002. УДК 537.86:519.22

РЕДУКЦИЯ ЧАСТОТЫ СИГНАЛА ПРИ НЕЛИНЕЙНОЙ ФИЛЬТРАЦИИ

Грязных И.В., к.т.н., профессор кафедры телекоммуникационных систем (ТС) Московского технологического университета (МТУ) (МИРЭА), e-mail: martenovsk@mail.ru;

Решетняк С.А., д.ф.-м.н., профессор кафедры ТС МТУ (МИРЭА), e-mail: reshets@bk.ru; Третьяков Г.Н., к.т.н., профессор кафедры ТС МТУ (МИРЭА), e-mail: gennady.tretyakov@rambler.ru; Шпак А.В., д.т.н., профессор кафедры ТС МТУ (МИРЭА), e-mail: alexandr.shpak@yndex.ru.

REDUCTION OF THE FREQUENCY OF THE SIGNAL IN NONLINEAR FILTERING

Gryaznykh I.V., Reshetnyak S.A., Tretyakov G.N., Shpak A.V.

The active filter containing a nonlinear element with two tunnel diodes is analyzed. The filter circuit is a series oscillating circuit with an active nonlinear element connected to the capacity in parallel. Forced nonlinear oscillations are investigated by numerical simulation. A resonance was found between the input signal and the process leading to self-oscillations. The values of the amplitudes of the input signals and the spectral region where forced oscillations occur are found. The filter parameters are also set, at which it is possible to obtain output harmonics with frequencies lower than the input frequency.

Key words: frequency of signal, reduction, nonlinear filtering, oscillations.

Ключевые слова: частота сигнала, редукция, нелинейная фильтрация, резонанс, автоколебания.

Введение

Фильтрация сигналов нелинейными радиотехническими системами, как правило, описывается дифференциальными уравнениями с различными типами нелинейности. В настоящее время аналитические методы получения их решений далеки от своего завершения. Основным инструментом исследования этих процессов является численный эксперимент. В дан-

ной работе был развит метод Эйлера в численном анализе указанных уравнений на примере нелинейного фильтра, в котором экспериментальным путем были обнаружены [1-3] вызывающие внимание эффекты стохастического резонанса и стохастической фильтрации (СФ) сигналов. В результате СФ выходное отношение сигнал/шум превышает аналогичное отношение на входе. Известные в литературе методы получения решений указанных уравнений с автоматическим выбором шага по времени требуют значительных затрат по времени и не вполне приемлемы для получения решений в случае большого числа реализаций сигнала и шума на входе фильтра. В отличии от [1-3] в настоящей работе основное внимание уделяется вопросу нелинейной фильтрации гармонических сигналов.

Схема фильтра представляет собой последовательный колебательный контур с параллельно подсоединенным к емкости активным нелинейным элементом. На входе фильтра действует гармонический сигнал, а реакция снимается в виде напряжения u на емкости. Как было ранее показано [1-3], в рабочей области изменения u вольт-амперная характеристика нелинейного элемента имеет вид $I(u) = -Au + Bu^3$, где коэффициенты ряда по степеням u определяются видом N-

Проанализирован активный фильтр, содержащий нелинейный элемент с двумя туннельными диодами. Схема фильтра представляет собой последовательный колебательный контур с параллельно подсоединенным к емкости активным нелинейным элементом. Путем численного моделирования исследованы вынужденные нелинейные колебания. Было обнаружено существование резонанса между сигналом на входе фильтра и процессом, приводящим к автоколебаниям. Найдены значения амплитуд входных сигналов и области спектра, где возникают вынужденные колебания. Установлены также параметры фильтра, при которых возможно получение гармоник выходного сигнала с частотами, меньшими частоты входного сигнала

> образной вольт-амперной характеристики отдельного туннельного диода. Если напряжение отсчитывать в единицах $(RB)^{-1/2}$, а время в единицах RC, где R и C – величины активного сопротивления и емкости, то безразмерное напряжение x на емкости подчиняется следующему дифференциальному уравнению

$$\mu \frac{d^2 x}{dt^2} + \gamma(x) \frac{dx}{dt} + W'(x) = e(t), \tag{1}$$

где $e(t) = h\cos(\omega_s t)$, h и ω_s – амплитуда и частота сигнала на входе фильтра, $\mu = L/(R^2C)$ – параметр инерционности нелинейного фильтра (этот термин используется из-за аналогии уравнений, описывающих нелинейные движения частицы с массой μ и реакции фильтра), L – величина индуктивности в контуре, $\gamma(x) = 1 + \mu(-AR + 3x^2)$ – коэффициент затухания, $W'(x) = -\alpha x + x^3$ – производная потенциальной функции $W(x) = -\alpha x^2/2 + x^4/4$, $\alpha = AR - 1$.

Для численного решения уравнение (1) представляется системой двух уравнений первого порядка:

$$\frac{dx}{dt} = v , \ \frac{dv}{dt} = \frac{1}{\mu} \Big[-\gamma(x)v - W'(x) + e(t) \Big] .$$
(2)

После выбора шага по времени Δt из (2) следует

$$\begin{aligned} x_{n+1} &= x_n + \int_{t_n}^{t_{n+1}} v(t) dt , \\ v_{n+1} &= v_n + (AR - \frac{1}{\mu})(x_{n+1} - x_n) - x_{n+1}^3 + \\ &+ x_n^3 - \frac{1}{\mu} \int_{t_n}^{t_{n+1}} [W'(x) - e(t)] dt, \end{aligned}$$

где x_n и v_n – значения искомых величин в точках отсчетов по времени $t_n = n \Delta t$.

Оценивая интегралы методом трапеций, приходим к следующим рекуррентным соотношениям:

$$\begin{aligned} x_{n+1} &= x_n + \frac{\Delta t}{2} (v_{n+1} + v_n), \\ v_{n+1} &= v_n + (AR - \frac{1}{\mu})(x_{n+1} - x_n) - x_{n+1}^3 + \\ &+ x_n^3 - \frac{\Delta t}{2\mu} \left[W'(x_{n+1}) + W'(x_n) - e_{n+1} - e_n \right]. \end{aligned}$$
(3)

С уменьшением шага ∆*t* точность формул (3) повышается.

С учетом известной зависимости $W'(x) = -\alpha x + x^3$ из формул (3) нетрудно получить кубическое уравнение относительно x_{n+1} , свободные члены которого содержат только x_n и v_n на предыдущем шаге решения. Так как это уравнение всегда имеет только один вещественный корень, то формулы (3) позволяют при заданных начальных условиях находить все остальные значения x_n и v_n .

В ходе решения было выбрано число точек отсчетов по времени $N = 2^{15}$ с выбранным шагом $\Delta t = 0,05$, что соответствует длительности исследуемых процессов $T = N\Delta t$. Полученные в результате решения представляются зависимостями от номера *n* шага по времени, так как дискретное значение времени $t_n = n\Delta t$. Спектр процесса на выходе фильтра является дискретным. Номер гармоники *m* в спектре определяет её частоту $m\Delta f$, $\Delta f = T^{-1}$ – разность между частотами соседних гармоник. Спектры полученных ниже выходных процессов, представляются в виде зависимостей от номера гармоники *т* в спектре. Для частоты сигнала применяется специальный номер *s*. Как было показано в [4], при $\mu > 1$ из-за флуктуаций начальных условий в данном активном фильтре развиваются автоколебания, номер основной гармоники которых обозначается s_a .

В работе для критической потенциальной функции $(\alpha = 0)$ и нулевых начальных условиях были исследованы вынужденные нелинейные колебания, под которыми понимается наличие в спектре выходного процесса основной гармоники с частотой f_s входного сигнала и амплитудой A_s , превышающей значения амплитуд других гармоник. С точки зрения исследования взаимодействия сигнала и шума наибольший интерес представляют вынужденные колебания с устойчивым пре-

дельным циклом на фазовой плоскости.

На рис. 1 представлены АЧХ нелинейного фильтра – зависимости амплитуды основной гармоники A_s от номера S для трех значений h, удовлетворяющих условию (4).



Рис. 1. Зависимость A_s от номера s или частоты входного сигнала при $\mu = 2$. Нижняя, средняя и верхняя кривые соответствуют значениям $h = 0, 2; 0, 4 \ u \ 0, 6$ При этом было выбрано умеренное значение параметра инерционности $\mu = 2$. Ему согласно [4] соответствовала частота автоколебаний с номером $s_a = 115$.

Из полученных на рис. 1 кривых следует, что наибольшие значения A_s достигаются в области, где частота сигнала близка к частоте автоколебаний ($s \cong s_a$). В этом смысле они указывают на существование резонанса между сигналом на входе фильтра и процессом, приводящим к автоколебаниям. Амплитуды основной гармоники A_s в области высоких частот входного сигнала резко падают.

В работе также были исследованы вынужденные колебания при большом значении параметра инерционности. Для выбранного значения $\mu = 10$ выходной процесс описывается «жестким» дифференциальным уравнением. При этом плавные изменения решений во времени сопровождаются резкими скачками. На рис. 2 приведены АЧХ нелинейного фильтра для $\mu = 10$ и

трех значений *h*, удовлетворяющих условию (4).



Рис. 2. Зависимость A_s от номера s или частоты входного сигнала при $\mu = 10$. Нижняя, средняя и верхняя кривые соответствуют значениям h = 0,3; 0,5 и 0,8

Для $\mu = 10$ частота основной гармоники автоколебаний имеет номер $s_a = 51$. Из рис. 2 следует, что максимумы кривых АЧХ расположены в области частот входного сигнала с номерами $s \cong s_a$. Они также указывают на существование резонанса между сигналом на входе фильтра и процессом, приводящим к автоколебаниям. С ростом *h* коэффициент усиления основной гармоники A_s в спектре выходного сигнала падает.

Вынужденные колебания наблюдаются, если амплитуда h входного сигнала имеет достаточно большую величину. Для её оценки амплитуда h сравнивалась с силой $W'(x_0)$ в уравнении (1). Если в качестве характерного значения x_0 выбрать величину, при которой коэффициент затухания обращается в нуль, то приходим к условию на амплитуду входного сигнала $h \ge x_0^3$ или

$$h \ge \left[(1 - 1/\mu)/3 \right]^{3/2}$$
 (4)

Расчеты показали, что при выполнении условия (4) в области низких частот, выше которых происходит резкое падение A_s от s, действительно имеют место вынужденные колебания, период которых определяется частотой f_s входного сигнала, а на фазовой плоскости им соответствуют устойчивые предельные циклы. Значения h существенно меньшие правой части (4), приводят к развитию автоколебаний и воспринимаются системой как ненулевые начальные условия.

Исследование выходного процесса представляет собой многопараметрическую задачу, так как необходимо проанализировать нелинейные колебания в зависимости от амплитуды и частоты входного сигнала, а также от величины параметра инерционности μ фильтра. Из всего множества полученных решений от квазигармонических до квазислучайных в данном случае интерес представляют решения, с помощью которых можно получать основные гармоники с частотами, меньше частоты входного сигнала.

Так для $\mu = 2$, частоты входного сигнала с номером

s = 300 и амплитудой h = 0,5 на рис. З изображены соответственно спектр выходного процесса, его зависимость от времени и отвечающий ему предельный цикл на фазовой плоскости.

Из рис. З а видно, что основная гармоника с максимальной амплитудой в спектре реакции фильтра имеет частоту $f_s/3$. Это означает, что с помощью рассматриваемого нелинейного фильтра можно получать на его выходе сигналы с частотами меньшими частоты сигнала – f_s . Процесс деление частоты на три связан с возникновением решения уравнения (1), которое представлено на рис. З б и содержит биения в течение периода их колебаний. Биения сопровождается «петлями» на фазовом портрете предельного цикла (см. рис. 3 с).

Проведенные исследования показали также, что выходные сигналы, с частотами меньшими f_s , отражают сложное взаимодействие входного сигнала с процессом, приводящим к возникновению автоколебаний. В

этом можно убедиться, если исследовать зависимость частоты основной гармоники выходного сигнала от параметра инерционности при заданных величинах амплитуды и частоты входного сигнала. Такая зависимость была получена и представлена на рис. 4 для h = 0,5 и s = 300 (номер, соответствующий частоте входного сигнала).







Рис. 4. Зависимость номера *s_m* частоты основной гармоники с максимальной амплитудой в спектре выходного сигнала от параметра инерционности *µ*

при
$$h = 0,5$$
 и $s = 300$

Как следует из рис. 4, частоту основной гармоники выходного сигнала можно изменять в широких пределах путем вариации параметра инерционности фильтра μ .

При этом выходной сигнал может не иметь устойчивого предельного цикла, как это показано на рис. 5. Его спектр дан на рис. 6.



Таким образом, проведенное исследование позволяет сделать следующие выводы. Вынужденные нелинейные колебания с периодом входного сигнала и устойчивым предельным циклом на фазовой плоскости возникают при достаточно высоких его амплитудах в низкочастотной области спектра. Ширина этой области зависит от АЧХ основной гармоники и от параметра инерционности фильтра. Наибольшие значения амплитуды наблюдаются при частоте входного сигнала близкой к частоте автоколебаний. Входные сигналы с высокими частотами приводят к развитию автоколебаний. В промежуточной области частот существуют нелинейные колебания, у которых частота основной гармоники меньше частоты входного сигнала и непрерывным образом зависит от параметра инерционности фильтра.



Литература

1. Домбровский А.Н., Решетняк С.А. Стохастический резонанс и фильтрация сигналов в нелинейной электрической системе второго порядка. Радиотехника, 2007, № 9, с. 19-25.

2. Домбровский А.Н., Решетняк С.А. Исследование стохастического резонанса в электрических схемах с несколькими туннельными диодами. Известия вузов. Радиофизика, т. LI, № 9, 2008, с.1-11.

3. Домбровский А.Н., Решетняк С.А. О стохастической фильтрации сигналов в нелинейных электрических системах. Радиотехника и электроника, 2009, т. 54, № 11, с. 1369-1371.

4. Решетняк С.А., Третьяков Г.Н. Об автоколебаниях в нелинейном активном фильтре второго порядка. Цифровая обработка сигналов, 2016, № 2, с. 64-67.



Сборники Трудов конференций (доклады) имеют международную (ISBN) и российскую (УДК ББК) регистрацию. Сборники направляются в 16 крупнейших библиотек Российской Федерации. Работы, опубликованные в материалах международных и общероссийских конференций, засчитываются ВАК РФ при защите диссертаций (Постановление Правительства РФ № 475 от 20 июня 2011 г.). Материалы сборника «ДОКЛАДЫ» размещаются в РИНЦ.

ПРИГЛАШАЕМ ВАС ПРИНЯТЬ УЧАСТИЕ В РАБОТЕ КОНФЕРЕНЦИИ

ОРГАНИЗАТОРЫ:

- Российское научно-техническое общество радиотехники, электроники и связи им. А.С. Попова
- Институт радиотехники и электроники
- им. В.А. Котельникова РАН
- Российская секция IEEE
 - Министерство образования и науки РФ
- Федеральное агентство по промышленности РФ
- ОАО «ГСКБ Концерн ПВО «АЛМАЗ-АНТЕЙ»
- ОАО «Концерн радиостроения «Вега»
- ФГУП «НИИР»
- Балтийский федеральный университет имени И. Канта
- Владимирский государственный университет
- Московский авиационный институт
- Московский энергетический институт

- Московский технический университет связи и информатики
 - Вневедомственный экспертный совет по проблемам воздушно-космической обороны
 - ОАО «Конструкторское бюро 1»

ПРИ УЧАСТИИ:

- Московский институт электронной техники
- Московский государственный технический университет им. Н.Э. Баумана
- Нижегородский госуниверситет им. Н.И. Лобачевского
- Нижегородский технический госуниверситет
- им. Р.Е. Алексеева
- Рязанский государственный радиотехнический университет
- Санкт-Петербургский государственный университет телекоммуникаций
- Санкт-Петербургский государственный электротехнический университет – («ЛЭТИ»)
- Ульяновский государственный технический университет
- Ярославский государственный университет

ПРОГРАММНЫЙ КОМИТЕТ:

Председатель: академик РАН Гуляев Юрий Васильевич

Заместитель Председателя: член-корр. РАН Черепенин Владимир Алексеевич

Члены комитета: академик РАН Бугаев А.С., академик РАН Кузнецов Н.А. академик РАН Чаплыгин Ю.А.

МЕЖДУНАРОДНЫЙ ОРГАНИЗАЦИОННЫЙ КОМИТЕТ:

Председатель: член-корр. РАН Зубарев Юрий Борисович

Заместитель Председателя: член-корр. РАН Черепенин Владимир Алексеевич

Сопредседатели: проф. Аджемов А.С., д.т.н. Ашурбейли И.Р., академик РАН Сигов А.С.,

Члены оргкомитета: доц. Алёшин В.С., проф. Бартенев В.Г., д.т.н. Борисов В.П., д.т.н. Бутенко В.В.,

член-корр. РАН Верба В.С., проф. Витязев В.В., проф. Воробиенко П.П. (Украина), проф. Дроздов Б.В.,

- проф. Кривошеев М.И., проф. Петровский А.А. (Беларусь), проф. Портнов Э.Л., проф. Постников И.И.,
 - проф. Потапов В.Т., проф. Пшеничников А.П., доц. Самсонов Г.А., д.э.н. Сеилов Ш.Ж. (Казахстан),

проф. Скородумов А.И., проф. Сперанский В.С., проф. Степанов С.Н., проф. Хорев А.А., проф. Шинаков Ю.С.

СЕКЦИИ:

1. Радиофизика, распространение радиоволн. 2. Антенны, СВЧ-устройства и технологии. 3. Радиотехника. 4. Радиолокация, в т.ч. зондирование Земли, радионавигация. 5. Системы, сети и устройства телекоммуникаций, в т.ч. космические, лазерные, акустические, волоконно-оптические и иные. 6. Информационные и интернет-технологии, в т.ч. кодирование информации, методы и алгоритмы анализа текстов, устной речи, изображений, искусственный интеллект, применение бионических принципов, методов и моделей. 7. Радиосвязь и . видеоинформатика. 8. Защита информации и информационная безопасность.