Цифровая Обработка Сигналов №1/2020

ЦИФРОВАЯ ОБРАБОТКА СИГНАЛОВ

Научно-технический журнал № 1/2020

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., член-корр. РАН
А.В. Дворкович, д.т.н., профессор В.И. Джиган, д.т.н., проф. В.В. Еремеев, д.т.н. Г.В. Зайцев, Р.В. Зубарев, А.П. Кирпичников, д.т.н., акад. РАН Н.А. Кузнецов, В.Г. Мистюков, д.т.н., проф. С.Л. Мишенков, д.т.н., проф. Ю.Н. Прохоров, д.т.н. А.Л. Приоров, д.т.н., проф. В.Г. Санников,

к.т.н., проф. В.С. Сперанский, д.т.н., проф. Ю.С. Шинаков

Адрес редакции:

r. 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 10.06.20 г. Формат 60х90/8.

Гарнитура «Arial». Печать офсетная. Бумага офсетная. Печ.л. 6,5. Тираж 250 экз.

Заказ № 2463. 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:

АО «Инструментальные системы» ФГУП «НИИ радио»

B HOMEPE:

Тихонов В.Ю., Шинаков Ю.С., Тимошенко А.С., Бахтин А.А. Цифровая обработка нелинейных искажений сигнала в программно-определяемой радиосистеме (SDR)
Кузьмин Е.В. Исследование эффективности беспороговой процедуры поиска псевдослучайного сигнала при ограничении разрядности входных наблюдений9
Киселева Т. П.
Использование последовательностей Задова-Чу для синхронизации по корреляционной кривой циклического префикса OFDM-символов LTE технологии13
Быховский М.А.
– DVB-S2.
определяющего параметры современных систем
спутниковой связи
Годиниа В.С. Пополи изин С.В. Штойн Б.М.
Тооына В.С., Перелыгин С.В., шшейн В.М.
инновационныи метоо измерения импульснои характеристики и нелинейных искажений
Бартенев В.Г.
Коррелятор с адаптивным порогом
Коррелятор с адаптивным порогом 30 Чучупал В.Я. Акустическое и языковое моделирование в сквозных системах распознавания речи 33 Туроеский Я.А., Борзунов С.В., Белобродский В.А. 33 Применение подходов, связанных с минимизацией дисперсии, 8 в задачах классификации многомерных данных 43 Котцов В.А., Котцов П.В. 43 Оперативное определение изменений наблюдаемой сцены 47 Хафизов Р.Г. Обеспечение разрешенного образа при инверсной фильтрации сигналов в условиях неопределенности 50 Памяти Виктора Павловича Дворковича 55

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

Digital Signal Processing

Science & Technical Magazine Issue 1, 2020 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 Dvorkovich A.V. Dr. Sci (Tech), Professor Corresponding Member of RAS Djigan V.I., Dr. Sci (Tech), Professor Eremeyev V.V., Dr.Sci.(Tech.), Professor Zaitsev G.V., Dr.Sci.(Tech.) Zubarev R.V., Kirpichnikov A.P., Kuznetsov N.A., Dr.Sci.(Tech.), Academician Mistyukov V.G., Mishenkov S.L., Dr.Sci.(Tech.), Professor Priorov A.L., Dr. Sci (Tech) 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:

L	
	Tikhonov V.Y., Shinakov Y.S., Tymoshenko A.S., Bakhtin A.A. Digital processing of non-linear signal distortion in a software-defined radio (SDR)
	Kuzmin E.V. Efficiency of the non-threshold spread spectrum signal searching procedure in case of quantization of the incoming observations9
	<i>Kiseleva T.P.</i> Using Zadov-Chu sequences for synchronization along the correlation curve of the cyclic prefix OFDM-symbols LTE technology
	Bykhovskiy M.A. Analysis of the international standard DVB-S2, defining parameters modern satellite communications systems18
	Godina V.S., Perelygin S.V., Stein B.M. Novel approach in measuring impulse response and nonlinear distortion26
	Bartenev V.G. The Correlator with Adaptive Threshold
	Chuchupal V.J. Acoustic and language modeling in end-to-end speech recognition systems
	<i>Turovskyi Ya.A., Borzunov S.V., Belobrodskyi V.A.</i> Dispersion minimization approaches in biomedical multivariable data qualification43
	Kottsov V.A., Kottsov P.V. Rapid detection of changes in the observed scene by logical subtraction of digital images47
	<i>Khafizov R.G.</i> Providing a resolved image with inverse filtering of signals in conditions of uncertainty50

Subscription index: 82185 ("Rospechat") ISSN: 1684-2634 9 771684 263005 УДК 621.391

ЦИФРОВАЯ ОБРАБОТКА НЕЛИНЕЙНЫХ ИСКАЖЕНИЙ СИГНАЛА В ПРОГРАММНО-ОПРЕДЕЛЯЕМОЙ РАДИОСИСТЕМЕ (SDR)

Тихонов В.Ю., аспирант Московского технического университета связи и информатики (МТУСИ), e-mail: sl-tx@yandex.ru;

Шинаков Ю.С., д.т.н., профессор Московского технического университета связи и информатики (МТУСИ); Тимошенко А.С., к.т.н., доцент Московского института электронной техники (МИЭТ), e-mail: timoshenko@edu.miet.ru;

Бахтин А.А., к.т.н., доцент Московского института электронной техники (МИЭТ), e-mail: bah@miee.ru.

DIGITAL PROCESSING OF NON-LINEAR SIGNAL DISTORTION IN A SOFTWARE-DEFINED RADIO (SDR)

Tikhonov V.Y., Shinakov Y.S., Tymoshenko A.S., Bakhtin A.A.

The model of digital predistortion device for compensation of nonlinear signal distortions arising in the power amplifier of the transmission system with the use of OFDM technology is considered. The model is implemented in the form of a software and hardware module and is interfaced with laboratory equipment NI USRP 2943R, containing the receiving and transmitting parts of the transmission system. The results of experimental measurements of nonlinear distortions of OFDM signals in a power amplifier without and using the proposed digital pre-distortion device are presented, confirming a significant expansion of the linear part of the dynamic range of the power amplifier.

Key words: nonlinear inertial device, nonlinear signal distortion, digital predistortion, OFDM technology.

Ключевые слова: нелинейное инерционное устройство, нелинейные искажения сигнала, цифровые предыскажения, технология OFDM.

Введение

Искажения сигналов, возникающие в усилителях мощности (УМ) беспроводных систем передачи, являются препятствием для увеличения качества, скорости передачи данных и мобильности. Широко используемым способом борьбы с такими искажениями является их компенсация, осуществляемая путем введения предыскажений сигналов. Принцип предыскаже-

ний заключается в том, что специально вводится искажение входного сигнала исследуемого УМ дополниустройством тельным устройством, называемым предыскажений (УПрИ), характеристики которого в некотором смысле являются обратными характеристикам усилителя. Это УПрИ вводится перед УМ. Получающаяся цепочка из двух последовательно включенных элементов может обеспечить снижение нежелательных искажений полезного сигнала. Способы предыскажений сигналов представляются важными и во многих других приложениях, например, в телевидении, радиовещании и других радиотехнических системах. Для краткости будем их называть способами или методами линеаризации. Среди них наиболее эффективными являются цифровые методы линеаризации [1].

В течение последних 15 лет проводятся интенсивные исследования различных способов введения предыскажений сигнала для разных моделей УМ без памяти. Нелинейность УМ в этом случае представляется функцией, связывающей мгновенные значения вещественных огибающих радиосигналов на входе и на выходе; это так называемая нелинейность АМ-АМ. В

Рассматривается модель цифрового устройства предыскажения для компенсации нелинейных искажений сигнала, возникающих в усилителе мощности системы передачи с технологией OFDM. Модель реализована в виде программно-аппаратного модуля и сопряжена с лабораторным оборудованием NI USRP 2943R, содержащим приемную и передающую части системы передачи. Приведены результаты экспериментальных измерений нелинейных искажений сигналов OFDM в усилителе мощности без и с применением предлагаемого цифрового устройства предыскажения, подтверждающие значительное расширение линейной части динамического диапазона усилителя мощности.

> более сложном случае приходится рассматривать взаимосвязь и текущего значения фазы радиосигнала на выходе УМ с мгновенным значением огибающей сигнала на входе; обычно в этом случае говорят о нелинейности АМ-ФМ.

> Для УМ без памяти функция предыскажения может быть реализована также как нелинейная функция с последующей полиномиальной аппроксимацией при построении компенсатора [2, 3].

> В последнее время были предложены различные решения для устройств, которые являются как нелинейными, так и инерционными, т.е. обладают памятью. В таких устройствах искажения сигнала имеют более сложный характер [4, 5]. Большинство предлагаемых в этих случаях моделей для УПрИ основаны на полиномиальных моделях с памятью [6], моделях Вольтерра [7] или Винера-Гаммерштейна [8]. Например, последнюю можно представить следующим образом:

$$y_{MWHM}(n) = \tag{1}$$

$$=\sum_{m_1=0}^{P_2-1}\sum_{k=0}^{P_2-1}c_{m_1}b_k\sum_{m_2=0}^{P_1-1}a_{m_2}x(n-m_1-m_2)\left|\sum_{m_2=0}^{P_1-1}a_{m_2}x(n-m_1-m_2)\right|^k.$$

Здесь x(n) – отсчет мгновенного значения комплексной огибающей узкополосного сигнала на входе УМ, y(n) – значение комплексной огибающей сигнала на выходе УМ; все остальные величины являются параметрами модели, значения которых должны быть определены на этапе идентификации модели на основе выборочных значений отсчетов сигналов на входе и выходе УМ. Такая постановка задачи возможна в том случае, когда есть реальное устройство и доступны измерениям необходимые значения сигналов на входе и выходе.



Рис. 1. Структурная схема цифровой системы предыскажений

На рис. 1 представлена предлагаемая структурная схема внедрения в реальное оборудование передатчика системы предыскажений, к которой здесь следует отнести все блоки, охваченные пунктирной кривой. Блок «Устройство предыскажений» обеспечивает коррекцию как нелинейных, так и линейных (частотно зависимых, инерционных) искажений. Блок «Идентификация модели УМ» представляет собой параметрическую нелинейную инерционную модель (1) УМ, значения параметров которой определяются адаптивно в процессе нормального функционирования системы передачи.

Основную задачу системы предыскажений можно сформулировать следующим образом. Устройство предыскажений должно выполнять операции над отсчетами x(n) комплексной огибающей сигнала на входе системы, которые можно рассматривать в некотором смысле как обратные операциям, выполняемым в УМ. Чтобы конкретизировать эти операции в блоке «Идентификация модели УМ» осуществляется контроль ошибки $e^{2}(n) = |k \cdot x(n) - y(n)|^{2}$, значение которой в процессе идентификации сводится к минимально возможному значению путем подбора значений коэффициента *k* и некоторых параметров модели (1), которые на рис. 1 представлены обозначениями $k_1(n)$, $k_2(n)$, $k_3(n)$. Поскольку $k \cdot x(n)$ представляет собой усиленный неискаженный сигнал, то минимизация значения этой ошибки свидетельствует о такой подстройке «Устройства предыскажений», при которой преобразования сигнала в этом устройстве действительно обратны преобразованиям в УМ. При этом структуры модели УМ и устройства предыскажений могут выбираться на основе разных априорных сведениях об УМ.

В данной работе в качестве модели УМ выбрана модификация модели Винера-Гаммерштейна: типовое радиотехническое звено, содержащее последовательное соединение фильтра 1, нелинейного безынерционного преобразователя и фильтра 2. Оба фильтра имеют выбранные структуры с конечным числом параметров, а нелинейный преобразователь описывается степенным полиномом. Идентификация такой модели сводится к оценке значений всех параметров на основе наблюдаемых реализаций процессов y(n) на выходе УМ и x(n)на входе «Устройства предыскажений». Более детальное описание алгоритмов функционирования этих устройств можно найти в [9].

Программно-определяемая радиосистема (англ. Software-defined radio, SDR) – радиотелекоммуникационная система, которая может быть настроена на произвольную полосу частот и принимать различные виды модулированного сигнала, состоящая из программируемого оборудования с программным управлением.

В настоящее время система SDR представляет большой интерес как в теоретической, так и в практической сфере: SDR выполняет значительную часть цифровой обработки сигналов на обычном персональном компьютере или на ПЛИС. цель такой схемы – радиоприемник и/или радиопередатчик произвольных радиосистем, изменяемый путем программной переконфигурации.

Программное радио применяется в военной сфере и сфере беспроводных услуг, так как позволяет обслуживать большое количество абонентов одновременно. В SDR-оборудовании форма модулированного радиосигнала задается в программном обеспечении (ПО). Формируется цифровой сигнал, который затем с помощью широкополосного ЦАП преобразуется в аналоговый на промежуточной частоте (ПЧ).

В приемнике все происходит в обратном порядке. Широкополосный АЦП преобразует в цифровой вид множество узкополосных сигналов, попадающих во входной тракт приемника. В соответствии с встроенным ПО приемник извлекает, преобразует вниз и демодулирует сигналы каждого канала, т.е. технология SDR позволяет изменять эксплуатационные параметры радиооборудования на уровне ПО.

Технология SDR включает в себя комбинацию методов, затрагивающих аппаратную и программную части. Аппаратная часть включает многодиапазонные антенны и радиочастотные преобразователи; широкополосные ЦАП и АЦП; а обработка сигналов ПЧ демодулированных сигналов и результирующего цифрового потока происходит с помощью программируемых процессоров общего назначения.

Достоинства SDR-приемника:

– не требует настройки;

 низкая чувствительность к температуре и разбросу параметров компонентов;

 простота реализации перестраиваемых фильтров с подавлением более 80 дБ;

 высокая точность и широкий диапазон перестройки фазы и частоты гетеродина.

«РЧ-секция, ПЧ-секция и секция базовой станции. РЧ-секция включает только аналоговые аппаратные модули (в то время как две другие целиком содержат цифровую аппаратуру) и отвечает за передачу/прием радиосигнала. ПЧ-секция отвечает за цифро-аналоговое преобразование и модуляцию/демодуляцию сигнала».

Функция программной части SDR сводится к распре-

делению ресурсов аппаратных средств для их использования различными приложениями связи и трансляции протокола второго уровня вышестоящим протоколам (WAP, TCP/IP).

Экспериментальная установка

В процессе проведения эксперимента¹ измерялись АМ-АМ характеристика УМ, входящего в оборудование USRP-2943R, на частотах в диапазоне 3400-3600 МГц [10, 11]. Структурная схема эксперимента представлена на рис. 2. Схема состоит из имитатора комплексной огибаюшей сигнала OFDM. представляющего собой программную часть передатчика, блока передатчика USRP (аппаратная часть эксперимента), который формирует высокочастотный радиосигнал, аттенюатора, приемника USRP (аппаратная часть эксперимента) и приемника сигнала OFDM (программная часть приемника). Передатчик USRP обеспечивает возможность установки значений частоты несущего колебания, выбора режима работы и значения коэффициента усиления УМ. Выход передатчика механически соединен со входом приемника через аттенюатор на 30 дБ. Приемник USRP работает в линейном режиме, не вносит искажений и обеспечивает необходимый уровень радиосигнала на частоте 3,6 ГГц на входе программного модуля приемника.

Имитатор сигнала OFDM формирует данные для 172-х поднесущих и блок из 28 OFDM символов данных. Каждый OFDM символ на интервале циклического префикса имеет N_{cp} = 32 отсчета комплексной огибающей радиосигнала и N_{ff} = 256 отсчетов огибающей на интервале быстрого дискретного преобразования Фурье.

На рис. 3 приведены результаты измерения характеристик АМ-АМ исследуемого передатчика USRP, по-

¹ Экспериментальные исследования выполнялись в лаборатории кафедры «Телекоммуникационные системы» университета МИЭТ.

лученные при разных коэффициентах усиления УМ передатчика и постоянном среднем значении огибающей сигнала на его входе. Каждая точка на этих рисунках представляет собой графическое представление результата одновременных измерений мгновенных значений мощности сигнала на входе УМ (ось абсцисс) и на его выходе (ось ординат). Для этого входом УМ является выход имитатора сигнала OFDM, а выходом – вход программного модуля приемника на рис. 2.



Рис. 2. Структурная схема исследуемой системы

Анализ этих рисунков позволяет сформулировать следующие выводы: 1) при коэффициенте усиления, не превышающем 24 дБ, усилитель можно считать линейным и безынерционным по отношению к огибающей сигнала на входе, поскольку все одновременные отсчеты мгновенных значений огибающих сигналов на входе и выходе практически располагаются на прямой линии; 2) с увеличением коэффициента усиления (увеличение крутизны характеристики АМ-АМ) до значений 26-28 дБ теряется взаимно однозначное соответствие между мгновенными значениями огибающих сигнала на входе и выходе усилителя – одному и тому значению огибающей на входе могут соответствовать разные значения огибающей на выходе усилителя (точки отсчетов рассеиваются вдоль оси ординат); этот эффект свидетельствует о проявлении инерционных свойств усилителя даже относительно огибающей сигнала на входе; 3) при коэффициенте усиления 30 дБ характеристика АМ-АМ свидетельствует о появлении эффекта насыщения, что сопро-



Рис. 3. Характеристики АМ-АМ усилителя NI USRP-2943R при коэффициентах усиления: а) 22 дб; б) 24 дб; в) 26дб; г) 28 дб; д) 30 дб

вождается нелинейными искажениями огибающей сигнала; рассеивание точек отсчетов вдоль оси ординат при появлении эффекта насыщения увеличивается.

Оба указанных эффекта – нелинейность характеристики АМ-АМ и рассеивание точек отсчетов вдоль оси ординат – свидетельствуют о том, что усилитель следует рассматривать как нелинейное инерционное устройство. Для сигналов OFDM эти эффекты приводят к повороту сигнального созвездия и рассеиванию точек сигнального созвездия относительно их номинальных положений. Оба этих следствия являются нежелательными и являются причинами, которые ограничивают либо динамический диапазон входного сигнала усилителя, либо его коэффициент усиления. В обоих случаях ограничивается энергетическая эффективность системы передачи.

Эффективным средством ослабления указанных эффектов представляется введение предыскажений сигнала на входе усилителя. При этом термином «предыскажения» здесь следует понимать устранение как нелинейности характеристики АМ-АМ усилителя мощности, так и рассеивание точек отсчетов вдоль оси ординат на рис. 3. Это означает, что система предыскажений должна вносить как нелинейные, так и инерционные эффекты, обратные аналогичным эффектам УМ. В данной работе предлагается система предыскажений, состоящая из двух блоков (рис. 1). В блоке «Идентификация модели УМ» оцениваются значения необходимых параметров и характеристик выбранной модели реального УМ NI USRP-2943R, которые затем используются в «Устройстве предыскажений» для формирования нужных характеристик этого устройства. В качестве модели УМ используется типовое радиотехническое звено, состоящее из последовательно соединенных фильтра 1, нелинейного безынерционного устройства и фильтра 2. Основная задача идентификации при этом сводится к выбору характеристик и оцениванию значений параметров каждого из элементов этого типового звена, которые затем используются для настройки устройства предыс-кажений.

Система цифровых предыскажений

На рис. 1 представлена структурная схема исследуемой системы, в которой основным исследуемым элементом является усилитель УМ, а остальные блоки представляют собой устройство цифровых предыскажений (УПрИ), к которой отнесены блок идентификации модели усилителя и блок формирования характеристик устройства предыскажения. Алгоритмы идентификации и формирования, представленные в этом разделе, построены на основе работы [9] и адаптированы для работы с оборудованием USRP. При реализации этих алгоритмов оказалось более удобным два функциональных блока УПрИ объединить и представить их в виде последовательного соединения 3-х блоков, изображенных на рис. 4. Рис. 1 удобен для описания режима функционирования УПрИ, а рис. 4 – для описания вычислительных алгоритмов этой системы.

Структура УПрИ на рис. 4 состоит из блока компенсации эффектов памяти после нелинейного преобразования, блока компенсации нелинейных преобразований и блока компенсации эффектов памяти до нелинейного преобразования.

Для описания алгоритмов функционирования этой УПрИ введем следующие обозначения:

 $k_i(n)$ – векторный параметр размерности p_i соответствующего блока, i = 1, 2, 3;

 g_i — множитель шага сходимости алгоритма соответствующего блока, i = 1, 2, 3;

$$\overrightarrow{f_1(x(n), p_1)} = (x(n) \ x(n-1) \ \cdots \ x(n-p_1)),$$
(2)

$$\overline{f_2(x(n), p_2)} = (x^1(n) \ x^2(n) \ \cdots \ x^{p_2}(n)), \tag{3}$$

$$f_3(x(n), p_3) = (x(n) \ x(n-1) \ \dots \ x(n-p_3)) \ . \tag{4}$$

До начала вычислений выполняется инициализация, в результате которой переменным величинам каждого блока присваиваются начальные значения:

$$k_1(0) = k_2(0) = k_3(0) = (1 \ 0 \ 0 \ \cdots \ 0)^T$$

Далее рекурсивно для *n* = 0,1,2,... выполняются следующие вычисления:

$$x_2(n) = \overrightarrow{f_1(x(n), p_1)} \cdot \overrightarrow{k_1(n)}, y_2(n) = \overrightarrow{f_1(y(n), p_1)} \cdot \overrightarrow{k_1(n)}, \quad (5)$$

 – значений на выходе блока компенсации эффектов памяти после нелинейного преобразования;

$$x_{3}(n) = \overline{f_{2}(x_{2}(n), p_{2})} \cdot \overline{k_{2}(n)}, y_{3}(n) = \overline{f_{2}(y_{2}(n), p_{2})} \cdot \overline{k_{2}(n)}.$$
 (6)

 – значений на выходе блока компенсации нелинейных преобразований;

$$\tilde{x}(n) = \overline{f_3(x_3(n), p_3)} \cdot \overline{k_3(n)}$$
(7)



Рис. 4. Структура устройства предыскажений

- значений на выходе УПрИ.

В каждом блоке УПрИ значения векторов параметров в (2), (3) и (4) вычисляются рекуррентно в соответствии со следующими алгоритмами:

а) для блока компенсации эффектов памяти после нелинейного преобразования:

 $err_1(n) = f_1(f_2(f_3(xy_2(n), p_2, p_3)))$ $\times \overline{k_3(n-1)}, p_2)\overline{k_2(n-1)}, p_1)\overline{k_1(n-1)} -f_1(f_2(f_3(yy_2(n), p_2, p_3)\overline{k_3(n-1)}, p_2) \times$ $\times \overline{k_2(n-1)}, p_1) \overline{k_1(n-1)}.$

- поправка для *n*-го такта,

 $\overline{k_1(n)} = \overline{k_1(n-1)} + err_1(n)g_1\overline{f_1(y(n), p_1)} - HOBOE$ (TEKYщее) значение вектора весовых коэффициентов;

б) для блока компенсации нелинейных преобразований:

 $err_{2}(n) = f_{2}(x(n), p_{2})\overline{k_{2}(n-1)} - f_{2}(y(n), p_{2})\overline{k_{2}(n-1)} -$

поправка для *п*-го такта,

$$\overline{k_2(n)} = \overline{k_2(n-1)} + err_2(n)g_2\overline{f_2(y(n),p_2)}$$
 – новое (те-
кущее) значение вектора весовых коэффициентов;

в) для блока компенсации эффектов памяти до нелинейного преобразования:

$$err_3(n) = f_3(x(n), p_3)\overline{k_3(n-1)} - f_3(y(n), p_3)\overline{k_3(n-1)}$$
 – поправка для *n*-го такта;

 $\overline{k_3(n)} = \overline{k_3(n-1)} + err_3(n)g_3\overline{f_3(y(n), p_3)}$ – HOBOE (TEKYщее) значение вектора весовых коэффициентов.

Вышеперечисленные блоки для поиска новых значений векторов коэффициентов используют метод наименьших квадратов. Отличительной особенностью разработанного алгоритма является определение поправки к блоку компенсации эффектов памяти после нелинейного преобразования.

При реализации на лабораторном оборудовании использовалось программное обеспечение NI LabView Communications 2.0. На рис. 5, 6 и 7 изображена реализация блоков устройства предыскажений в программноопределяемой радиосистеме NI USRP 2943R.



Рис. 5. Реализация блока компенсации эффектов памяти после нелинейного преобразования



Рис. 6. Реализация блока компенсации нелинейных преобразований



Рис. 7. Реализация блока компенсации эффектов памяти до нелинейного преобразования

Результаты моделирования

Все вычисления в этом эксперименте (регистрация отсчетов комплексных огибающих, формирование начальных значений, матричные вычисления) выполнялись в ПЛИС оборудования USRP-2943R. Аппаратно решена задача синхронизации системы передачи и устройства предыскажений, так как при моделировании УПрИ для реального УМ задача моделирования системы синхронизации не ставилась. При синхронизации сигнала приходилось учитывать время, затрачиваемое на обновление коэффициентов устройства предыскажений.

На рис. 8 приведены экспериментально полученные зависимости АМ-АМ и АМ-ФМ для оборудования USRP: синее изображение – характеристики без предискажений; черное изображение – эти же характеристики при установившемся режиме функционирования системы предыскажений.



б) АМ-ФМ

Рис. 8. Сравнение характеристик при введении УПрИ

Анализ этих и других аналогичных характеристик позволяет сделать следующие выводы: используемая модель нелинейного УМ с памятью в виде типового радиотехнического звена и предлагаемый адаптивный (рекуррентный) алгоритм ее идентификации обеспечивают возможность построения цифрового предыскажающего устройства, которое обеспечивает существенное уменьшение как линейных, так и нелинейных искажений узкополосного сигнала, сформированного на основе технологии OFDM. Что же касается численных показателей улучшения характеристик усилителя мощности, то здесь можно привести следующие оценки: предлагаемое устройство предыскажений для данного усилителя обеспечивает расширение динамического диапазона этого оборудования на 6 дб, что в свою очередь повышает энергетическую эффективность системы передачи на 4 дб.

Заключение

Модель Винера-Гамерштейна можно рекомендовать как относительно простую модель для описания нелинейных устройств с памятью. Для такого звена в данной работе предложены алгоритмы оценивания параметров этой модели по реальным значениям сигналов на входе и выходе усилителя мощности. Предложенные алгоритмы являются рекуррентными и обеспечивают решения задач идентификации в нормальном режиме функционирования линеаризуемого устройства. Предлагаемая система предыскажений может быть реализована на современных ПЛИС. Введение предыскажающего устройства для линеаризации характеристик усилителей мощности существенно расширяет динамический диапазон для мгновенным значениям огибающей сигнала OFDM; в представленном здесь эксперименте – примерно в 1,5 раза.

Литература

1. H. Alasady, R. Boutros and M. Ibnkahla. «Comparison between digital and analog predistortion for satellite communications», CCECE 2003 – Canadian Conference on Electrical and Computer Engineering. Toward a Caring and Humane Technology (Cat. No.03CH37436), Montreal, Quebec, Canada, 2003, pp. 183-186 vol.1.

2. R. Marsalek, P. Jardin, and G. Baudoin. «From post-distortion to predistortion for power amplifiers linearizationй», IEEE Commun. Lett., vol. 7, no. 7, pp. 308-310, Jul. 2003.

3. A.N. D'Andrea, V. Lottici, and R. Reggiannini.«RF power amplifier linearization through amplitude and phase predistortion», IEEE Trans. Commun., vol. 44, no. 11, pp. 1477-1484, Nov. 1996.

4. D. Lei, R. Raich, and G. T. Zhou. «A Hammerstein predistortion linearization design based on the indirect learning architecture», in Proc. Int. IEEE ICASSP, May 2002, vol. 3, pp. III-2689-III-2692.

5. G. Baudoin, P. Jardin, and R. Marsalek. «Power amplifier linearisation using predistortion with memory», in Proc. 13th Int. Czech-Slovak Scientific Conf. RADIOELEKTRONIKA, Brno, Czech Republic, May 6-7, 2003, pp. 193-196.

6. L. Ding, G.T. Zhou, D.R. Morgan, Z. Ma, J.S. Kenney, J. Kim, and C.R. Giardina. «A robust digital baseband predistorter constructed using memory polynomials», IEEE Trans. Commun., vol. 52, no. 1, pp. 159-165, Jan. 2004.

7. A. Zhu and T.J. Brasil. «An adaptive Volterra predistorter for the linearization of RF high power amplifiers», in Proc. Conf. IEEE MTT, 2002, pp. 461-464.

8. Y. Ding, H. Ohmori, and A. Sano. «Adaptive predistortion for high power amplifier with linear dynamics» in Proc. IEEE Int. MidWest Symp. Circuits and Syst., Hiroshima, Japan, Jul. 2004, pp. 121-124.

9. V.Y. Tikhonov and Y.S. Shinakov. «COMPENSATION OF NONLINEAR DISTORTION IN INERTIAL DEVICES», 2018Systems of Signal Synchronization, Generating and Processing in Telecommunications (SYNCHROINFO), Minsk, 2018, pp. 1-4.

10. A.G. Timoshenko, N.K. Osipenko, A.A. Bakhtin and E.A. Volkova, «5G Communication Systems Signal Processing PAPR Reduction Technique», 2018 Systems of Signal Synchronization, Generating and Processing in Telecommunications (SYNCHROINFO), Minsk, 2018, pp. 1-4.

11. NI USRP-2943 Software Defined Radio Reconfigurable Device, https://www.ni.com/pdf/manuals/374193d.pdf

УДК 621.391.822.3

ИССЛЕДОВАНИЕ ЭФФЕКТИВНОСТИ БЕСПОРОГОВОЙ ПРОЦЕДУРЫ ПОИСКА ПСЕВДОСЛУЧАЙНОГО СИГНАЛА ПРИ ОГРАНИЧЕНИИ РАЗРЯДНОСТИ ВХОДНЫХ НАБЛЮДЕНИЙ

Кузьмин Е.В., к.т.н., доцент кафедры радиотехники ФГАОУ ВО «Сибирский федеральный университет», e-mail: EKuzmin@sfu-kras.ru ; kuzminev@mail.ru.

EFFICIENCY OF THE NON-THRESHOLD SPREAD SPECTRUM SIGNAL SEARCHING PROCEDURE IN CASE OF QUANTIZATION OF THE INCOMING OBSERVATIONS

Kuzmin E.V.

The characteristics of the non-threshold spread spectrum signal searching by the delay procedure in case of quantization effect are studied. The results of the statistical modeling are presented: curves of correct searching probability vs. the reception conditions for various versions of the bit-width of the analog-to-digital conversion. The paper presents estimation of possible losses in noise immunity of the signal searching procedure due to the quantization effect.

Key words: n-bit quantization, quantization noise, spread spectrum signal searching, correct searching probability, cross-correlation function.

Ключевые слова: многоуровневое квантование, шум квантования, поиск псевдослучайного сигнала, вероятность правильного поиска, взаимная корреляционная функция.

Введение

Неотъемлемым этапом обработки псевдослучайных сигналов является поиск по времени запаздывания, в результате которого формиру-

ется грубая оценка задержки сигнала, необходимая для дальнейшей синхронизации приёмника [1, стр. 165]; [2, стр. 268]; [3, стр. 118, стр. 495]. Одним из факторов, способных снижать помехоустойчивость обработки сигналов, является ограниченная разрядность наблюдений аддитивной смеси, вызванная аналого-цифровым преобразованием (АЦП) [3, стр. 490]; [4, стр. 383]. Исследованию влияния данного фактора на характеристики процедуры беспорогового поиска псевдослучайного сигнала и посвящена данная статья.

Цель статьи: исследование влияния ограниченной разрядности наблюдений аддитивной смеси на вероятность правильного исхода беспорогового поиска псевдослучайного сигнала по задержке и определение потерь в помехоустойчивости процедуры поиска.

Модель наблюдений

 $v_{\rm x}(t) = A_{\rm x}v_{\rm x}(t) + v_{\rm x}(t),$

Пусть на входе приёмника действует аддитивная смесь псевдослучайного фазоманипулированного сигнала $y_s(t)$ [5, стр. 50] и белого гауссовского шума $y_n(t)$:

$$y_{s}(t) = a(t)\sin(2\pi f_{s}t - \varphi_{0}), \qquad (1)$$

$$a(t) = \sum_{l=1}^{L} a_{l} \operatorname{rect}(t - (l-1)T_{a}), a_{l} = \{\pm 1\}_{l=1}^{L},$$

где $A_s = \sqrt{2}P_s$ – амплитуда сигнала, P_s – его мощность, f_s – центральная частота спектра сигнала, ϕ_0 –

Исследуются характеристики беспороговой процедуры поиска псевдослучайного сигнала по задержке при учёте эффекта квантования по уровню. Методом статистического моделирования получены семейства зависимостей вероятности правильного выполнения поиска от условий приёма для различных вариантов разрядности аналого-цифрового преобразования. Дана оценка возможных потерь в помехоустойчивости процедуры поиска сигнала, обусловленных эффектом квантования по уровню.

> начальная фаза; a(t) – видеосигнал, соответствующий бинарной псевдослучайной последовательности a_l длины L, rect(t) – парциальный прямоугольный импульс с единичной амплитудой и длительностью T_a . Для реализации цифровой обработки аддитивной смеси $y_{\Sigma}(t)$ необходимо её аналого-цифровое преобразование, предполагающее как минимум две основных операции – дискретизацию по времени с шагом T и квантование по уровню [6, стр. 39], что даёт отсчёты

$$y_{\Sigma_{KB}}(kT) = U\left(\left| \alpha + \frac{m}{2} + \frac{y_{\Sigma}(kT)}{U} \right| - \frac{m}{2} \right),$$

$$U = \frac{D}{m}, \quad m = 2^{q_{aun}},$$

$$y_{\Sigma}(kT) = A_s y_s(kT) + y_n(kT, \sigma_n^2), \quad k = 1, 2, ..., N,$$
(2)

где k – номер выборки; N – количество выборок;] · [– операция выделения целой части числа; U – шаг квантования по уровню; D – диапазон изменения значений смеси (1); m – число уровней квантования; $q_{\rm aun}$ – разрядность АЦП; α – параметр, учитывающий способ квантования: $\alpha = 0,5$ соответствует квантованию с округлением, а $\alpha = 0$ соответствует квантованию с усечением; $y_n(kT, \sigma_n^2)$ – отсчёты дискретизированного шума с дисперсией $\sigma_n^2 = N_0/2T$, N_0 – спектральная плотность мощности.

Процедура поиска сигнала

Практика функционирования многих систем такова, что на начальном этапе приёма сигнал $y_s(t-\tau)$ имеет неизвестную случайную задержку т, оценку которой, как уже было сказано во введении, необходимо получить с грубой точностью - то есть необходимо решить задачу поиска сигнала по задержке, что реализуется на основе оптимального правила распознавания (различения) сигналов [1, стр. 166]. Распознавание сигналов основано на параллельном либо последовательном вычислении взаимных корреляционных функций (ВКФ) аддитивной смеси (1) и копий полезного сигнала, отличающихся вносимой задержкой. По наибольшему значению ВКФ находится оценка задержки $\hat{\tau}$ сигнала, содержащегося в смеси (1) [1, стр. 166]. Используя известный подход к вычислению ВКФ [7], можно записать выражение для выходного эффекта беспороговой процедуры поиска:

$$Z = \left| O Д \Pi \Phi \left\{ \mathcal{I} \Pi \Phi \left\{ \dot{\mathbf{Y}}_{\Sigma} \right\} \cdot \mathcal{I} \Pi \Phi^* \left\{ \dot{\mathbf{S}}_{0} \right\} \right\} \right|, \tag{3}$$

где ОДПФ $\{\cdot\}$ и ДПФ $\{\cdot\}$ – операторы обратного и прямого дискретного преобразования Фурье соответственно; $|\cdot|$ – операция вычисления модуля; звёздочка означает комплексное сопряжение; $\dot{\mathbf{Y}}_{\Sigma}$ и $\dot{\mathbf{S}}_{0}$ – комплексные векторы, содержащие N отсчётов, получаемые соответственно путём квадратурного гетеродинирования отсчётов входных наблюдений $y_{\Sigma}(kT)$ и дискретизированного опорного сигнала $s_{0}(kT)$ (являющегося несмещённой копией полезного сигнала $y_{s}(kT)$), с последующим объединением квадратур с использованием мнимой единицы $j = \sqrt{-1}$

$$\dot{\mathbf{Y}}_{\Sigma} = \{ y_{\Sigma}(kT) \cos(2\pi f_s kT) + j y_{\Sigma}(kT) \sin(2\pi f_s kT) \}_{k=1}^{N}, \\ \dot{\mathbf{S}}_{0} = \{ s_{0}(kT) \cos(2\pi f_s kT) + j s_{0}(kT) \sin(2\pi f_s kT) \}_{k=1}^{N}$$
(4)

Сравнение значений выходного эффекта (3) с пороговым уровнем не используется, поскольку факт присутствия сигнала полагается известным, что характерно для систем с непрерывным сигналом, присутствующим в течение длительного интервала наблюдения. Ожидаемые потери в помехоустойчивости процедуры поиска γ , появляющиеся при учёте влияния АЦП – учёте эффекта квантования по уровню смеси (1), объясняются некоторым ухудшением условий приёма из-за присутствия шума квантования. Нетрудно показать, что при многоуровневом квантовании, потери в помехоустойчивости, обусловленные шумом квантования, составят

$$\gamma = \left(1 + \frac{\sigma_{n_{\rm KB}}^2}{\sigma_n^2}\right)^{-1} = \left(1 + \frac{U^2/12}{\sigma_n^2}\right)^{-1} = \left(1 + \frac{D^2}{12 \cdot 2^{2q_{\rm max}} \sigma_n^2}\right)^{-1},$$

$$10 \lg(\gamma) = -10 \lg\left(1 + \frac{D^2}{12 \cdot 2^{2q_{\rm max}} \sigma_n^2}\right),$$
 (5)

где $\sigma_{n_{\text{KB}}}^2$ – дисперсия шума квантования, равная $U^2/12$ [6, стр. 41]; [8, стр. 380].

Допуская возможным корректный априорный выбор диапазона D, проведём исследование влияния разрядности АЦП $q_{\rm aun}$ на качество беспороговой процедуры

поиска псевдослучайного сигнала $y_s(t-\tau)$ по задержке, реализуемой на основе (3) и (4), с учётом предварительной обработки смеси (1) согласно (2). Для оценки качества процедуры поиска сигнала выбрана вероятность правильного выполнения поиска P, которая оценивалась методом статистического моделирования с применением ранее разработанной методики [9].

Результаты исследования

На рис. 1-4 представлены семейства зависимостей вероятности правильного выполнения поиска Р от энергетического потенциала - отношения мощности сигнала плотности спектральной мошности шума К $q_{\rm arr} = 10 \log(P_{\rm s}/N_0)$, характеризующего условия приёма. Непронумерованными сплошными линиями на рис. 1-4 показаны кривые вероятностей, полученных при подстановке в (4) наблюдений $y_{s}(kT)$ – то есть при отсутствии квантования по уровню. На рис. 1 пронумерованные пунктирные кривые соответствуют различным вариантам ограничения разрядности наблюдений. Кривая 1 соответствует знаковому ограничению наблюдений $sign(y_{\Sigma}(kT))$, где $sign(\cdot)$ – функция знака; кривая 2 соответствует квантованной смеси $y_{\Sigma_{\rm KB}}(kT)$ при разрядности АЦП, равной $q_{\text{aum}} = 2$; кривая 3 – $q_{\text{aum}} = 3$; кривая 4 получена при $q_{\text{ашп}} = 4$. Кривые вероятностей, полученные при $4 < q_{\text{ани}} \le 12$, совпадают с кривой, рассчитанной для случая отсутствии квантования по уровню и на рисунке не показаны. Кривые 2-4 на рис. 1 получены при $\alpha = 0, 5$, то есть при использовании квантования по уровню с округлением.



Рис. 1. Зависимости вероятности правильного выполнения поиска от энергетического потенциала

На рис. 2-4 кривые вероятностей пронумерованы парами, в каждой из которых сплошная линия означает использование квантования по уровню с округлением ($\alpha = 0, 5$), а пунктир – квантование по уровню с усечением ($\alpha = 0$). Нумерация кривых на рис. 2-4 имеет следующий смысл: кривые 1 соответствует квантованной смеси $y_{\Sigma \text{ кв}}(kT)$ при разрядности АЦП, равной $q_{\text{ацп}} = 2$; кривые 2 – соответствует $q_{\text{ацп}} = 3$; кривые 3 получены при $q_{\text{ацп}} = 4$.

Все кривые на рис. 1 и рис. 2 получены при числе статистических испытаний 10^6 , амплитуде полезного

сигнала на входе процедуры поиска $A_s \ll \sigma_n$, значениях $(Tf_s)^{-1} = 4$, N = 4096, L = 511. Шаг изменения отношения $q_{3\pi}$ равен 0,1 дБ. Диапазон $D = 6\sigma_n$ выбран таким образом, чтобы максимальные шумовые выбросы оказывались в пределах «раскрыва» квантователя. Семейства вероятностей правильного выполнения поиска, показанные на рис. 3 и рис. 4, получены при числе статистических испытаний 10^7 , а другие условия статистического эксперимента аналогичны условиям, при которых получен рис. 2. Рис. 4 является «продолжением» рис. 3.



Рис. 2. Зависимости вероятности правильного выполнения поиска от энергетического потенциала



Рис. 3. Зависимости вероятности правильного выполнения поиска от энергетического потенциала



Рис. 4. Зависимости вероятности правильного выполнения поиска от энергетического потенциала

Обсуждение результатов

Как видно из рис. 1, для принятых условий статистического эксперимента ограничение разрядности (2) наблюдений аддитивной смеси (1) оказывает заметное влияние на выбранный показатель качества – на вероятность правильного выполнения поиска сигнала по задержке на основе (3), (4). Полученные зависимости позволяют определить значения энергетического потенциала Q_{ал}, при которых обеспечивается заданный уровень качества процедуры поиска - вероятности правильного выполнения поиска Р, и, таким образом оценить необходимые параметры для (2). В табл. 1 внесены требуемые значения энергетического потенциала Q_{an} , при которых обеспечивается вероятность правильного выполнения поиска P=0,9 и P=0,99 для различных вариантов разрядности АЦП, в том числе и для случая отсутствия АЦП - то есть без ограничения (б / о) разрядности. Как видно из данных таблицы 1, для фиксированных уровней Р максимальные потери в помехоустойчивости процедуры поиска псевдослучайного сигнала по задержке ожидаемо соответствуют случаю знакового ограничения аддитивной смеси $sign(y_{x}(kT))$ и составляют 2 дБ, что совпадает с аналитической оценкой (1,96 дБ) данной в [10, стр. 101]. При разрядности АЦП равной $q_{\text{ann}} = 2$ потери в помехоустойчивости не превышают 0,9 дБ; при разрядности АЦП $q_{aun} = 3$ потери составляют 0,3 дБ. Дальнейшее увеличение разрядности АЦП приводит к исчезающе малым потерям, исчисляемым значениями, меньшими 0,1 дБ. Указанные экспериментально полученные значения потерь в помехоустойчивости процедуры беспорогового поиска для $q_{\rm ann} \ge 2$ хорошо совпадают (расхождение менее 0,1 дБ) со значениями $10 \lg(\gamma)$, полученными согласно (5), а также не противоречат оценкам общего характера [3, стр. 490]; [4, стр. 383]; [10, стр. 101]; [11, стр. 252].

Как видно из рис. 2-4, способ квантования по уровню наблюдений аддитивной смеси (1) оказывает незначительное влияние на вероятность правильного выполнения поиска сигнала по задержке. При разрядности АЦП $q_{\rm aum} \ge 4$, кривые вероятностей, полученные для равных значений $q_{\rm aum}$, но разных α (0 и 0,5), становятся практически неразличимыми. Автор выражает благодарность А.С. Глинченко за полезную консультацию по модификации модели квантователя [6, стр. 39] к виду (2).

Заключение

Исследована эффективность процедуры беспорогового поиска псевдослучайного сигнала по задержке при учёте эффекта квантования по уровню наблюдений аддитивной смеси. Получены семейства зависимостей вероятности P правильного выполнения поиска псевдослучайного сигнала по задержке от значения энергетического потенциала $q_{\rm sn}$ для различных вариантов разрядности АЦП $q_{\rm aun}$. Результаты проведённого статистического моделирования позволяют сделать следующие выводы. Ограничение разрядности входных наблюдений, учитываемое квантованием по уровню аддитивной смеси, оказывает заметное влияние на ход зависимостей вероятности $P(q_{\rm sn})$. При этом характеристики $P(q_{\rm sn})$, полученные для квантования по уровню с округление и усечением, отличаются несущественно.

Установлено, что потери в помехоустойчивости процедуры поиска псевдослучайного сигнала по задержке составляют 2 дБ для случая знакового ограничения аддитивной смеси; 0,9 дБ при разрядности АЦП $q_{aun} = 2$, и менее 0,5 дБ при $q_{aun} \ge 3$. Экспериментально полученные значения потерь в помехоустойчивости процедуры поиска хорошо совпадают с аналитической оценкой (5) и согласуются с известными оценками общего характера.

Таблица 1 – Требуемые значения $\,Q_{_{
m 9n}}$

для достижения вероятности P = 0,9; 0,99

при различных вариантах $\, q_{
m aun} \,$

$q_{ m aun}$	$Q_{_{ m 90}},\;$ дБГц		
	P = 0,9	P = 0,99	
sign	43,4	45,1	
2	42,3	44	
3	41,7	43,4	
4	41,5	43,2	
5-12; б / o	41,4	43,1	

Литература

1. Поиск, обнаружение и измерение параметров сигналов в радионавигационных системах / В.П. Ипатов, Ю.М. Казаринов, Ю.А. Коломенский, Ю.Д. Ульяницкий; под ред. Ю.М. Казаринова. – М.: Сов. радио, 1975. – 296 с.

2. Варакин Л.Е. Системы связи с шумоподобными сигналами / Л.Е. Варакин. – М.: Радио и связь, 1985. – 384 с.

3. ГЛОНАСС. Принципы построения и функциониро-

вания / под ред. А.И. Перова, В.Н. Харисова. – 4-е изд., перераб. и доп. – М.: Радиотехника, 2010. – 800 с.

4. Springer Handbook of Global Navigation Satellite Systems / Eds. Peter J.G. Teunissen, Oliver Montenbruck. – Springer International Publishing AG, 2017. – 1327 c.

5. Тузов Г.И. Статистическая теория приёма сложных сигналов / Г.И. Тузов. – М.: Сов. радио, 1977. – 400 с.

6. Глинченко А.С. Цифровая обработка сигналов: учеб. пособие. 2-е изд., перераб. и доп. / А.С. Глинченко. Красноярск: ИПЦ КГТУ, 2005. – 482 с.

7. J.N. Daigle and N. Xiang. A specialized fast crosscorrelation for acoustical measurements using coded sequences / J. Acoust. Soc. Am., Vol. 119, No. 1, January 2006. P. 330-335.

8. Сергиенко А.Б. Цифровая обработка сигналов / А.Б. Сергиенко – СПб.: Питер, 2003. – 604 с.

9. Кузьмин Е.В. Повышение вероятности правильного поиска шумоподобного сигнала по времени запаздывания на фоне тональной помехи / Е.В. Кузьмин, Ф.Г. Зограф // Успехи современной радиоэлектроники. – 2016. – №11. – С. 137-140.

10. Borio D. A statistical theory for GNSS signal acquisition. PhD thesis / D. Borio. – Politecnico di Torino, Marzo, 2008. – 291 c.

11. Understanding GPS: principles and applications / Eds. Elliott Kaplan, Christopher Hegarty. – 2nd ed. – ARTECH HOUSE, 2006. – 703 c.

НОВЫЕ КНИГИ

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.

Витязев С.В.

Цифровые процессоры обработки сигналов / Курс лекций – М.: Изд-во «Горячая линия-Телеком», 2017 г. – 100 с.: ил.

Рассмотрены основы построения архитектур и оптимизации программного обеспечения цифровых сигнальных процессоров. Сформулированы основные задачи цифровой обработки сигналов на сигнальных процессорах. Представлено описание инструментальных и программных средств работы с цифровыми сигнальными процессорами.

Для студентов технических вузов радиотехнических и инфокоммуникационных специальностей, будет полезна преподавателям, читающим соответствующие курсы.



обработки сигналов



УДК 621.396.93

ИСПОЛЬЗОВАНИЕ ПОСЛЕДОВАТЕЛЬНОСТЕЙ ЗАДОВА-ЧУ ДЛЯ СИНХРОНИЗАЦИИ ПО КОРРЕЛЯЦИОННОЙ КРИВОЙ ЦИКЛИЧЕСКОГО ПРЕФИКСА OFDM-СИМВОЛОВ LTE ТЕХНОЛОГИИ

Киселева Т. П., аспирант кафедры радиотехнических систем Московского технического университета связи и информатики (МТУСИ), e-mail: golzev2011@yandex.ru.

USING ZADOV-CHU SEQUENCES FOR SYNCHRONIZATION ALONG THE CORRELATION CURVE OF THE CYCLIC PREFIX OFDM-SYMBOLS LTE TECHNOLOGY

Kiseleva T.P.

In this paper, the criteria for the selection of complex sequences of Zadov-Chu (ZC) for their use in the formation of cyclic prefix OFDM-symbols in LTE technology at the stage of synchronization on the correlation curve of the cyclic prefix of symbols of the frame transmitted from the base station to the mobile user are determined. The use of ZC sequences to fill the time interval of the cyclic prefix increases the speed and efficiency of synchronization, which is confirmed by the results of modeling in the MATLAB operating environment. The studied ZC sequences as part of a cyclic symbol prefix can be used to analyze subcarriers modulated by elements of these sequences, which are used by the mobile station as control ones, when detecting selective frequency distortions in the channel with fading.

Key words: LTE OFDM, Zadov-Chu sequence (ZC), correlation function, OFDM–symbol, cyclic prefix.

Ключевые слова: LTE OFDM, последовательность Задова-Чу (ZC), корреляционная функция, OFDM-символ, циклический префикс.

Введение. Краткая характеристика системы синхронизации технологии LTE OFDM

Технология LTE (Long Term Evolution – долговременное развитие) – стандарт мобильной связи, разработанный консорциумом 3GPP (3rd Generation Partnership Project) как технология дальнейшего развития системы WLAN для работы с высокоподвижными объектами (до 350...500 км/ч) и величиной занимаемой полосы в 100 МГц. Физический уровень сетей LTE реализован на базе сигнальной технологии OFDM (Orthogonal Frequency Division Multiplexing) – муль-

типлексирование с ортогональным частотным разнесением [1]. Принципиальной особенностью этой технологии является распределение информационных модулированных сигналов по поднесущим частотам, передаваемым параллельно в одном физическом канале связи. Технология OFDM позволяет высокоскоростной поток данных конвертировать в несколько параллельных битовых потоков меньшей скорости, передаваемых одновременно, причем каждый поток данных модулирует свою поднесущую. Спектры параллельных битовых потоков не перекрываются.

Принципы построения алгоритмов системы синхронизации в технологии OFDM рассмотрены в ряде таких публикаций, как [1, 2, 3, 4] и др. В данной работе вопросы построения систем синхронизации технологии LTE OFDM рассмотрены для варианта частотного разделения каналов (FDD) в нисходящем направлении от базовой станции к пользователям (DL – Down Link).

При активизации мобильного пользователя произво-

Определены критерии выбора комплексных последовательностей Задова-Чу (ZC) для их использования при формировании циклического префикса OFDM-символов в технологии LTE на этапе синхронизации по корреляционной кривой циклического префикса OFDM-символов кадра, передаваемого от базовой станции к мобильному пользователю. Применение последовательностей ZC для заполнения временного интервала циклического префикса повышает скорость и эффективность синхронизации, что подтверждается результатами моделирования в операционной среде MATLAB. Исследуемые последовательности ZC в составе циклического префикса символов можно использовать для анализа поднесущих, модулированных элементами этих последовательностей, которые используются мобильной станцией как контрольные, при выявлении селективных частотных искажений в канале с замираниями.

дится процесс синхронизации с ближайшей к нему базовой станцией (БС). В технологии LTE БС в нисходящем направлении периодически посылает два синхросигнала, построенных на последовательностях с хорошими корреляционными характеристиками – первичный (*PSS*) и вторичный (*SSS*), в которых закодирован идентификатор БС, определенный для конкретного пользователя: $N_{ID}^{cell} = 3N_{ID}^{(1)} + N_{ID}^{(2)}$,

где $N_{ID}^{(1)}$ – идентификатор группы (0...167), определяемый *SSS*; $N_{ID}^{(2)}$ – идентификатор сектора (0, 1, 2), определяемый PSS. Каждая БС формирует 504 сотовых идентификатора, организованных в 168 групп по 3 идентификатора в каждой [4]. Передача данных в нисходящем (от БС) и восходящем (к БС) направлениях организована как последовательность кадров длительностью 10 мс, включающих битовые данные пользователя, сформированные в модуляционные символы *QPSK*, 16*QAM*, 64*QAM* [4, стр.97...100], и служебную информацию, расположенную в известных на стороне приема ресурсных элементах (RE) кадра. Кадр во временной области разбивается на 20 слотов, каждый из которых состоит из 6 или 7 символов технологии OFDM (рис. 1)



Рис. 1. OFDM-символ

Для защиты от межсимвольных искажений (ISI) и интерференции поднесущих (ICI) в OFDM-символе предусмотрен защитный интервал (циклический префикс-*CP*), длительность которого, определяемая размерами соты, составляет 1/4, 1/8, 1/16 величины информационной части символа (DATA). В технологии OFDM *CP* формируется переносом в интервал *TCP* блока данных той же длительности с конца информационной части (Tdata) OFDM-символа.

Первичная синхронизация БС с пользователем включает стандартные этапы:

 тактовая синхронизация, в данном случае – определение границ OFDM-символа по корреляционной кривой CP;

– цикловая синхронизация – в данном случае это синхронизация по корреляционной кривой сигнала *PSS*, расположенного в 7-м OFDM-символе 0-го и 10-го слота передаваемого кадра [4], т.е. в начале 1 и 2 половины кадра, что позволяет получить циклическую корреляционную функцию PSS, построенного на комплексных последовательностях Задов-Чу (*ZC*(*u*,*n*)) и их циклических сдвигах. Это комплексные последовательности постоянной амплитуды с фазами, вычисляемыми по формуле, приведенной далее. Циклическая автокорреляционная функция (*ACF*) последовательности *ZC* имеет вид δ-функции с нулевым значением боковых лепестков, т.е. обладает идеальными корреляционными характеристиками [3]. Корреляционная кривая *PSS* позволяет получить грубую временную привязку к границам полукадра;

– кадровая синхронизация – в данном случае это синхронизация по корреляционной кривой *SSS*, построенных на 2-х 31-элементных *М*-последовательностях и их циклических сдвигах, размещенных в 6-м OFDM-символе 0-го и 10-го слота передаваемого кадра [4]. Поскольку М-последовательности в составе *SSS* в 0 и 10 слотах различны, корреляционные пики *SSS* повторяются через 20 слотов, т.е. с периодом длительности кадра. Это позволяет получить временную привязку к началу кадра, т.е. завершить процесс кадровой синхронизации. Компенсация частотных сдвигов поднесущих возможна поэтапно, начиная с комплексной коррелиционной характеристики *СР*, из квадратур которой можно вычислить дробную часть фазового смещения поднесущих в полосе передаваемого кадра [2].

Методы подбора последовательностей Задова-Чу для формирования циклического префикса OFDM-символов

В данной работе исследуются пути повышения эффективности синхронизации по корреляционной кривой циклического префикса. Это первый этап синхронизации, позволяющий получить временную привязку к границам OFDM-символа и началу слота. На рис. 2 показан процесс перезаписи блока информации с конца символа в область *CP* во временной области при формировании последовательности передаваемых OFDM-символов слота.



Рис. 2. Процесс перезаписи данных конца OFDM-символа во временной интервал циклического префикса последовательности символов в слоте LTE OFDM кадра

Дискретизация OFDM-символов производится с частотой выборок (Samples) f_s = 30,72 МГц, или периодом T_s – базовым временным интервалом в технологии LTE OFDM (Basic time unit) [4], равным 1/ f_s = 3,255е-8 с.

По корреляционной кривой СР в любом из вариантов проводится определение границ символа и временного положения начала слота. При прохождении канала связи с Рэлеевскими замираниями информационные данные на протяжении длительности OFDM-символа, особенно при «быстрых» замираниях, могут измениться настолько, что СР и конечный интервал символа, с которого копировалась информация при передаче кадра, не сформируют явного пика взаимнокорреляционной функции (VCF) при прохождении корреляционного или согласованного фильтра оптимального приемника. В работе [6] предложен способ улучшения корреляционных характеристик СР OFDM-символов в системе DVB-T2 путем использования М-последовательности с размещением ее в защитном интервале. Приведены результаты моделирования процесса корреляционного декодирования М-последовательности, передаваемой совместно с информацией циклического префикса. В [6] показано, что применение М-последовательностей в качестве заполнения СР повышает скорость и качество процесса синхронизации. В этой работе обращено особое внимание на короткие М-последовательности длиной 31 и 63 элемента, позволяющие уменьшить перекрестные искажения между блоком OFDM-символов и используемыми последовательностями.

В данной работе в LTE технологии предлагается использовать в качестве циклических префиксов OFDMсимволов короткие последовательности Задова-Чу (ZC), не используемые в группе стандартов, выделенных в 36-ю серию технических спецификаций 3GPP. Что касается *М*-последовательностей длиной 31 и 63 элемента, то эти *М*-последовательности и их циклические сдвиги широко используются для формирования вторичного синхросигнала (SSS) и не могут рассматриваться как формирующие СР OFDM-символов LTE технологии.

В стандарте LTE использованы следующие последовательности ZC [4]:

1. Первичный синхросигнал PSS, построенный на ZC(u,n), где корень (индекс) u = 25, либо 29 или 34, количество элементов n = 62. Математическая модель ZC(u,n)

$$d_u(n) = \begin{cases} e^{-j\frac{\pi u n(n+1)}{63}} & n = 0, 1, ..., 30\\ e^{-j\frac{\pi u (n+1)(n+2)}{63}} & n = 31, 32, ..., 61 \end{cases}$$

Λ

где $d_{u}(n) - n$ -й элемент последовательности *u*-го корня.

2. Два типа опорных сигналов восходящего (UL) направления:

– Демодулированные последовательности (Demodulation reference signal – DM-RS), передаваемые в общем канале (PUSCH) или логическом канале управления (PUCCH).

 Зондирующий сигнал (Sounding reference signal – SRS), служащий для оценки качества канала. SRS – опорный (референсный – RS) сигнал с шириной полосы, охватывающей весь диапазон поднесущих частот пользователя.

Оба рассматриваемых сигнала и их циклические сдвиги формируются из одной базовой последовательности Задова-Чу:

$$x_q(m) = \exp\left[\frac{-j\pi q m(m+1)}{N_{ZC}^{RS}}\right],$$

где корень q пропорционален значению $N_{ZC}^{RS}\left[\frac{u+1}{31}\right]$,

 $u = 0...29, N_{ZC}^{RS}$ – число элементов последовательности: $N_{ZC}^{RS} < M_{SC}^{RS}$,

где $M_{sc}^{RS} = m * N_{sc}^{RB} = m * 12$ – общее число поднесущих, доступное пользователю в текущем сеансе обмена, кратное 12, т.е. числу поднесущих N_{sc}^{RB} одного ресурсного блока (RB) [4]. Из формулы базовой последовательности для формирования SRS, DM-RS опорных сигналов следует, что число элементов этих последовательностей кратны 12 – числу поднесущих одного RB.

Преамбулы канала случайного доступа (PRACH) построены на последовательностях ZC с числом элементов N_{zc} = 139, 839.

 Преамбулы PRACH канала и их циклические сдвиги формируются из одной базовой последовательности Задова-Чу:

$$x_u(n) = \exp\left[\frac{-j\pi u n(n+1)}{N_{ZC}}\right],$$

где корень определяется из соответствующих таблиц [4], $0 \le n \le (N_{ZC} - 1)$.

При выборе последовательностей *ZC* для формирования *CP* OFDM-символов предлагается руководствоваться следующими критериями:

1. В число последовательностей *ZC* для *CP* не входят последовательности, указанные выше, и используемые в стандарте LTE [4]. 2. Если число элементов подобранной для формирования *CP* последовательности не кратно 2k (k – целое положительное число), что необходимо для выполнения быстрого Фурье-преобразования (БПФ), недостающие элементы дополняются нулевыми значениями (предполагается, что число элементов последовательности меньше ближайшего значения 2k).

З Корреляционные свойства последовательностей эффективно отражает величина отношения квадрата модуля максимума *ACF* к среднеквадратическому значению боковых лепестков [7]:

$$MF_m = \frac{R_m^2}{\left(\sum_{i=0}^N R_i^2\right) / N},$$

где R_m^2 – величина квадрата максимума (центрального пика) *ACF*, R_i^2 – величина квадрата *i*-го отсчета боковых лепестков, *N* – длина *ACF*.

Для последовательностей одинаковой длины это отношение может отличаться в зависимости от значения u – корня (индекса) ZC(u,n), что предполагает подбор последовательности с таким индексом, отношение MF_m которой будет максимальным.

Практические результаты подбора последовательностей *ZC*(*u*,*n*) для формирования циклического префикса OFDM-символов

Для проверки предположения о повышении эффективности процесса синхронизации по корреляционной кривой СР при использовании коротких последовательностей ZC вместо битовых данных CP, в операционной среде MATLAB было проведено моделирование для символов с СР из комплексных последовательностей ZC(5,9), ZC(9,19), ZC(15,31), ZC(25,37) и случайных комплексных битовых последовательностей с одинаковым количеством элементов n. Затем в среде MATLAB построены ACF всех сформированных последовательностей и вычислены значения MF_m (dB) ACF каждой последовательности, причем для случайных битовых последовательностей проведено усреднение результатов вычисленных MF_m не менее, чем по 10 вариантам случайной последовательности для каждой длины *п*. Вычислен энергетический выигрыш КЕ от применения коротких последовательностей ZC(u,n) для заполнения CP символов на первом этапе синхронизации по сравнению с классическим заполнением СР битовыми данными.

Результаты исследований приведены в табл. 1. Построен график зависимости $MF_m(dB)$ от длины последовательностей ZC(u,n) и усредненной $MF_m(dB)$ для битовых случайных последовательностей B(n).

По данным табл. 1 построен график (рис. 3).

Таблица 1

<i>п</i> -кол-во элементов последовательности	Усредненное [<i>MF_m_B</i>] (dB) для <i>СР</i> случайных битовых	<i>MF_m_ZC</i> (dB) для <i>CP ZC</i> (<i>u</i> , <i>n</i>)	$K_E = MF_m_ZC - MF_m_B (dB)$
9	9,8581	ZC(5,9) 11,5232	1,6651
31	12,7688 14,3025	ZC(9,19) 14,9381 ZC(15,31) 17,2455	2,1693
37	16,0846	ZC(19,37) 18,0392	1,9546



Рис. 3. График отношения модуля максимума ACF к среднеквадратическому значению боковых лепестков в зависимости от длины последовательности: для CP из ZC(и,n) – 1;

для СР из случайной битовой последовательности – 2

На рис. 4 представлены автокорреляционные функции случайной цифровой последовательности, формирующей *CP*:

а) из данных конца OFDM-символа с числом элементов последовательности *n* = 37;

б) последовательности *ZC*(19,37) с числом элементов последовательностей *n* = 37.

Подбор последовательности ZC(u,n) с наиболее высоким MF_m проведен на примере ZC(u,37) с различными индексами u. В табл. 2 приведены значения MF_m для последовательностей ZC(u,37).

Последовательности ZC(u,37) с разными индексами и формируют ACF с различным отношением MF_m вследствие различного уровня боковых лепестков, что отражено на рис. 5 – осциллограммы модуля ACF ZC(19,37), ZC(30,37), ZC(26,37), ZC(15,37), полученные при моделировании автокорреляционных функций данных последовательностей программой MATLAB. Из табл. 2 следует, что все величины MF_m для последовательностей ZC(u,37) больше той же величины $MF_m = 16,0846$ для ACF CP, сформированного из случайной битовой последовательности с таким же числом элементов n = 37. Путем подбора индекса и моделирования можно определить последовательность ZC(u,n) длины n с наилучшими корреляционными характеристиками. В данном случае – это ZC(19,37) – выделено в табл. 2 жирным.

Начиная с этапа грубой временной синхронизации по корреляционной кривой СР, есть возможность оценки характеристик канала связи по величине амплитуды и времени задержки корреляционных пиков СР, формирующихся с периодом, близким к длительности OFDMсимвола. Так же, как закрепляется за пользовательской станцией идентификатор N_{ID}⁽²⁾ PSS, БС может закрепить и выбранную из неиспользуемых в стандарте 3GPP TS 36.211 последовательностей ZC в качестве CP на этапе установления синхронизации. Так, из табл. 2, для числа элементов 37, можно выбрать последовательности ZC(19,37), ZC(21,37), ZC(25,37) для использования в качестве СР OFDM - символа. Число элементов n последовательности ZC(u,n) выбирается, исходя из длины СР (нормальный, расширенный, либо для слотов с тремя символами OFDM) и количества выборок на символ (samples) [3, стр. 132, рис. 5.14].

Таблица 2



Рис. 4. АСЕ СР, сформированного: а) из битовых случайных последовательностей с конца ОЕDM – символа длины n = 37; б) последовательностей ZC(19,37)

Последовательность	Величина MF_m (dB)	Последовательность	Величина MF_m (dB)
ZC(u,37)		<i>ZC</i> (u,37)	
ZC(15,37)	17,7028	ZC(26,37)	17,4088
ZC(17,37)	16,8875	ZC(27,37)	17,2435
ZC(19,37)	18,0392	ZC(29,37)	16,7371
ZC(21,37)	17,7227	ZC(30,37)	16,9138
ZC(23,37)	17,6544	ZC(31,37)	17,5867
ZC(25,37)	18,0037		

Для определения селективных частотных искажений в Рэлеевском канале в технологии LTE OFDM при завершении процесса синхронизации пользователя с БС и получения необходимой служебной информации от БС, с пользовательской станции на БС поступает зондирующий сигнал (SRS) в виде последовательности ZC, охватывающий полосу частот выделенного пользователю диапазона [4]. БС анализирует селективные искажения поднесущих SRS и пересылает информацию о наименее искажаемых каналом поднесущих пользователю. Эти функции на этапе синхронизации может выполнять и последовательность ZC в составе CP, поскольку эта последовательность известна и на БС, и пользователю и может служить контрольным «словом», поэлементно распределенным по поднесущим в полосе СР и используемым для оценки селективных частотных замираний канала связи. Для контроля качества канала можно охватить значительную часть частотной полосы пользователя, если за время передачи слота перемещать элементы последовательности в определенном порядке по всему выделенному диапазону поднесущих. Анализировать селективные частотные искажения может сам пользователь, начиная с этапа синхронизации по СР и не ожидая завершения полного цикла синхронизации и обмена с БС опорными сигналами (RS) для оценки качества канала.

Заключение

В данной работе отражены результаты моделирования в операционной среде МАТLAB корреляционных функций циклического префикса (*CP*), сформированного цифровыми битовыми данными с интервала конца символа OFDM и комплексных последовательностей ZC(u,n), заменяющих битовые последовательности на аналогичном временном интервале конца символа и *CP*. Проведено сравнение величин отношения модуля корреляционной функции к среднеквадратической величине боковых лепестков MF_m для комплексных последовательностей Задова-Чу и битовых последовательностей *CP*, сформированного стандартным методом, с тем же числом элементов n.

Последовательности *ZC*(*u*,*n*) выбирались с числом элементов, отличных от тех, которые задействованы в составе служебных сигналов технологии *LTE*. В резуль-

тате получены значения энергетического выигрыша при использовании ZC(u,n) для формирования CP символа OFDM в пределах (1,67...2,94) dB для последовательностей с числом элементов 9,19,31,37, рассмотренных в данной работе. На примере последовательности ZC с числом элементов n = 37 проведен подбор индекса последовательности u, позволяющим получить наибольшую величину MF_m . Кроме того, эти последовательности можно использовать для анализа поднесущих, модулированных элементами этой контрольной последовательности при выявлении селективных частотных искажений в канале с замираниями.

Литература

1. Казачков В.О. Исследование реализации синхронизации по сигналам Задова-Чу в стандарте Long Term Evolution для канала с замираниями. – Интернет-журнал «Науковедение» ISSN 2223-5167 http://naukovedenie.ru / Том 7, №1 (2015) УДК 621.396.94. [Электронный ресурс]. – Режим доступа: http://naukovedenie.ru/PDF/39TVN115.pdf (дата обращения 04.01.2018)

2. Гельгор А.Л., Попов Е.А. Технология LTE мобильной передачи данных: учебное пособие. СПб.: Издательство Политехнического университета, 2011. – 204 с

3. Sesia S., Toufik I., Baker M. LTE – The UMTS Long Term Evolution: From Theory to Practice. – Torquay, UK: John Wiley & Sons, 2009.

4. ETSI TS 136 211 V10.0.0 (2011-01). Technical Specification. European Telecommunications Standards Institute, 2011, 104 c. LTE; Evolved Universal Terrestrial Radio Access (E-UTRA); Physical channels and modulation (3GPP TS 36.211 version 10.0.0 Release 10).

5. Физический уровень LTE [Электронный ресурс]. – Режим доступа: http://www.russianelectronics.ru/leaderr/review/2187/doc/53411/ (дата обращения 19.12.2019).

6. Быков В.В., Аль-Мершахи С.М. Улучшение синхронизации OFDM-сигналов в системе DVB-T2 // T-Comm: Телекоммуникации и транспорт. – 2016. – Том 10. – № 6. – С. 21-26.

7 Киселева Т.П. Исследование свойств циклической автокорреляционной функции последовательности Задова-Чу в зависимости от характеристик квантования элементов последовательности. – Цифровая обработка сигналов, № 4, 2018, 40-44 с.

Уважаемые коллеги!

Для тех, кто не успел оформить подписку на первое полугодие 2020 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 621.391.13

АНАЛИЗ МЕЖДУНАРОДНОГО СТАНДАРТА DVB-S2, ОПРЕДЕЛЯЮЩЕГО ПАРАМЕТРЫ СОВРЕМЕННЫХ СИСТЕМ СПУТНИКОВОЙ СВЯЗИ

Быховский М.А., доктор технических наук, профессор МТУСИ, e-mail: bykhmark@gmail.com.

ANALYSIS OF THE INTERNATIONAL STANDARD DVB-S2, DEFINING PARAMETERS MODERN SATELLITE COMMUNICATIONS SYSTEMS

Bykhovskiy M.A.

In this paper, we analyze the DVB-S2 standard, on the basis of which satellite communication systems (SCS) are currently being created. A theoretical method for determining the reception quality of communication systems built according to the DVB-S2 standard is presented. The article provides a comparison of the noise immunity estimates of the reception of SCS systems obtained by the method of statistical modeling given in the DVB-S2 standard with those obtained on the basis of theoretical analysis. The comparison showed that the results of theoretical analysis are close to those obtained by statistical modeling.

Thus, the proposed theoretical research method makes it possible to obtain analytical formulas that determine the dependence of the reliability of receiving messages in satellite communication systems on their system parameters (on the modulation parameters of the transmitted signals, the type and parameters of the codes used in the SCS).

The next part of the article presents the results of a study of the possibilities of using optimal multidimensional surface-spherical signal ensembles. These results showed that systems in which multidimensional signals are used to transmit messages in their characteristics are significantly superior to systems created in accordance with the standard DVB-S2.

Key words: satellite communications, signal transmission methods, spectral and energy efficiency, message coding.

Ключевые слова: спутниковая связь, методы передачи сигналов, спектральная и энергетическая эффективность, кодирование сообщений.

Введение

Международные стандарты современных радиосистем связи создавались на основе результатов теории связи, в которой исследовались методы модуляции, предназначенные для передачи сигналов по каналам связи, а также помехоустойчивые коды, используемые для повышения помехоустойчивости приема сообщений. В настоящее время параметры современных систем спутникового и наземного ТВ вещания определяются во многом схожими стандартами DVB-S2 [1...3] и DVB-T2 [4]. В этих системах для повышения надежности связи применяются низкоплотностные коды (коды LDPC) [5], кодовые комбинации которых имеют большую длину (N_b = 16000 или 64000 бит). В них предусмотрено множество возможных режимов работы. В этих режимах используются двумерные ансамбли сигналов (AC), позволяющие передавать сообщения с разной удельной скоростью

 (R_f) , и коды LDPC, имеющие разную кодовую скорость $(R_{\rm LDPC})$. В системе DVB-S2 применяются сигналы QPSK $(R_f=2),\ 8\text{-PSK}\ (R_f=3),\ 16\text{-APSK}\ (R_f=4)$ и 32-APSK $(R_f=5)\ 6$ ит/сек Γ ц, а кодовая скорость составляет $R_c=1/4,\ 3/4...5/6,\ ...9/10;$ в системе DVB-T2 применяются сигналы с 4-QAM $(R_f=2),\ 16\text{-QAM}\ (R_f=4),\ 64\text{-QAM}\ (R_f=6)$ и 256-QAM $(R_f=8)\ 6$ ит/сек Γ ц, а кодовая скорость может составлять $R_c=1/2,\ 1/3...5/6.$

Отметим важный недостаток применения в системах связи двумерных сигналов. Вероятность возникновения ошибочно принятых символов на выходе спутниковой

Дан анализ стандарта DVB-S2, на основе которого в настоящее время создаются системы спутниковой связи (ССС). Представлен теоретический метод определения качества приема систем связи, построенных по стандарту DVB-S2. Приведено сравнение оценок помехоустойчивости приема систем ССС в стандарте DVB-S2, полученных методом статистического моделирования, с теми, которые получены на основе теоретического анализа. Сравнение показало, что результаты теоретического анализа близки к полученным методом статистического моделирования.

Показано, что предложенный теоретический метод исследования дает возможность получить аналитические формулы, определяющие зависимости надежности приема сообщений в спутниковых системах связи от их системных параметров (от параметров модуляции передаваемых сигналов, вида и параметров кодов, используемых в ССС).

В следующей статье представлены исследования возможностей применения оптимальных многомерных поверхностно-сферических АС, показавшие, что системы, в которых для передачи сообщений применяются многомерные сигналы, по своим характеристикам существенно превосходят системы, создаваемые в соответствии со стандартом DVB-S2.

> системы связи (ССС) должна быть весьма незначительной, так как в ней требуется обеспечить весьма высокую надежность приема передаваемых сообщений. Если при передаче сообщений помехоустойчивый код не используется, то высокая надежность приема сообщений должна обеспечиваться в демодуляторе приемника, на выходе которого должна формироваться копия двоичной последовательности символов, поступающих на передающем конце линии связи на вход модулятора. Однако, применяя двумерные сигналы, высокую надеж

ность приема можно обеспечить, только повышая уровень сигнала, поступающего на вход приемника, а, следовательно, применяя достаточно мощные передатчики, а также приемные и передающие антенны с высоким коэффициентом усиления. Повышение энергетики линии за счет применения в транспондере (ТРСД) мощных передатчиков и таких антенн приводит в ССС к существенному снижению их экономической эффективности, так как имеются большие ограничения на вес, габариты и мощность радиооборудования, которое можно разместить в спутнике, запускаемом на орбиту Земли. Энергетику ССС на трассе ТРСД—Земля можно снизить, если допустить, чтобы на выходе демодулятора надежность приема переданных символов не была бы высокой.

Чтобы обеспечить высокую надежность приема сообщений, которые будет получать пользователь системы спутниковой связи, оказывается необходимым применять в ней помехоустойчивые коды, в состав которых наряду с информационными символами входят и ряд избыточных (проверочных). Такие коды должны исправлять те ошибки, которые возникли на выходе демодулятора при обработке принимаемых сигналов. Эти коды, как правило, должны иметь весьма большую длину и их декодер является весьма сложным устройством. Существенным недостатком применения помехоустойчивых кодов является то, что наряду с информационными символами по каналу связи необходимо передавать также и избыточные. Это, во-первых, снижает скорость передачи сообщений, а во-вторых, при этом энергетика линии связи расходуется не только на передачу информационных символов, но и избыточных.

Важно отметить, что эффективные системы связи могут быть созданы без использования в них двумерных сигналов и помехоустойчивых кодов. В работах Шеннона [6, 7] показано, что при передаче сообщений возможно обеспечить, в принципе, абсолютную надежность связи (прием сообщений без ошибок), если скорость передачи меньше так называемой пропускной способность канала связи, которая определяется только отношением сигнал/шум на входе демодулятора. Шеннон установил, что для достижения абсолютной надежности приема переданных сообщений целесообразно применять [7] сигналы, имеющие весьма значительную длину. Он показал, что для передачи сообщений должны использоваться оптимальные многомерные (*N*-мерные сигналы). Значение N равно N = int[2(FT)], где (FT) – нормированная длительность сигналов, входящих в АС, F – полоса частот канала связи, T – длительность сигналов, входящих в оптимальный ансамбль сигналов (AC), a *int*(*x*) – целая часть числа *x*.

Шеннон определил минимально необходимое отношение сигнал/шум (ρ_s), которое должно быть обеспечено на входе демодулятора, для того, чтобы по каналу связи с абсолютной надежностью (безошибочно) можно было бы передавать сообщения с удельной скоростью (скоростью передачи сообщений, отнесенной к полной полосе частот канала связи, или, что тоже самое, со спектральной эффективностью), равной R_f бит/сек Гц. Величину ρ_s называют пределом Шеннона. В системе связи с рационально выбранными параметрами, в которой обеспечивается спектральная эффективность R_f , отношение сигнал/шум на входе демодулятора (ρ_{dm}) не должно существенно отличаться от ρ_s , разность $\Delta \rho = (\rho_{dm} - \rho_s)$ дБ определяет для любой системы связи ее энергетические потери относительно «идеальной» системы Шеннона.

Следует отметить, что в работах Шеннона не рассматривались вопросы, связанные с построением конкретных оптимальных АС и выбором их нормированной длительности. Эти вопросы рассмотрены в ряде статей автора данной статьи, а также в книге [8], в которой показано, что в оптимальной системе связи, энергетика которой лишь незначительно превышает предел Шеннона, нет необходимости в применении сложных помехоустойчивых кодов с большим количеством избыточных символов.

Целью данной работы является анализ стандарта DVB-S2. Кроме того, в ней представлен аналитический метод, позволяющий получить достаточно точные оценки вероятности ошибки при приеме сообщений в системах связи, в которых используются разные режимы работы и разные методы кодирования. Эффективность применения этого метода проиллюстрирована на примере стандарта DVB-S2, в котором приведены оценки вероятности ошибки, полученные эмпирическим методом математического моделирования.

В следующей статье [9] результаты данной работы будут использованы при исследовании возможности совершенствования систем спутниковой связи, в которых для передачи сообщений применяются оптимальные многомерные ансамбли сигналов. В ней также будет дано сравнение характеристик таких систем с характеристиками систем связи, построенными на основе стандарта DVB-S2.

Анализ системных параметров спутниковой связи, создаваемых по стандарту DVB-S2

В ССС, созданной по данному стандарту, предусмотрен большой набор режимов работы. Их выбор зависит вида предоставляемых в ССС услуг, а также от уровня сигнала, который может быть обеспечен на входе демодулятора в приемнике сигналов. Отметим, что ССС, построенные по стандарту DVB-S2, могут использоваться не только для передачи сигналов ТВ вещания, но и для надежной передачи в цифровом виде любой информации, в том числе IP потоков данных.

Уровень сигнала, приходящего на вход приемника, определяется условиями распространения радиоволн. В те периоды времени, когда на трассе происходит дополнительное ослабление этого уровня (например, из-за дождя), возникает необходимость изменения режима работы системы связи. Режим работы ССС определяет вид модуляции передаваемых сигналов и параметры применяемого кода LDPC, определяющие его способность исправлять ошибки. В стандарте предусмотрена возможность изменения в значительных пределах кодовой скорости этого кода (R_{LDPC}). Возможные сочетания R_{LDPC} и видов модуляции в ССС представлены в табл. 1. Отметим, что при плохих условиях распространения радиоволн, когда отношение сигнал/шум на входе

Таблица 1. Возможные сочетания R_{LDPC} и видов модуляции в ССС

R _{LDPC}	1/4	1/3	2/5	1/2	3/5	2/3	3/4	4/5	5/6	8/9	9/10
QPSK						+					
8-PSK	- +										
16-APSK	- +										
32-APSK	-							+			

демодулятора $\rho_{\rm s} < 1$ (мощность полезного сигнала меньше мощности шума), для передачи сообщений применяются сигналы QPSK и коды LDPC с небольшой кодовой скоростью – R_{LDPC} = 1/4, 1/3 и 2/5.

Если уровень сигнала, измерения которого осуществляются в приемнике, достаточно высок, то в ССС применяются сигналы с высокой спектральной эффективностью и коды LDPC с высокой кодовой скоростью. Как следует из табл. 1 в стандарте DVB-S2 предусмотрено применение 29 разных режимов работы.

На рис. 1 представлена упрощенная блок-схема системы передачи информационных сообщений в ССС, построенной в соответствии со стандартом DVB-S2.





На вход ССС, как показано на рис. 1, поступает последовательность информационных символов, которые кодируются кодом БЧХ. Кодовые комбинации с выхода кодера БЧХ поступают на вход кодера LDPC, на выходе которого, в зависимости от выбранного режима кодирования, формируются кодовые комбинации, длина которых равна N_{LDPC} = 64800 или N_{LDPC} = 16200 двоичных символов. Отметим, что коды БЧХ и LDPC образуют каскадный код, в котором внешним кодом является код БЧХ, а внутренним кодом – код LDPC.

Последовательность двоичных символов с выхода кодера LDPC поступает на вход перемежителя (ПЕРМ). В этом устройстве эта последовательность записывается в ячейки матричной памяти каскадного кода. С выхода ПЕРМ последовательность двоичных символов поступает на модулятор (МОД), а с выхода МОД на вход передатчика (ПЕР-К), в котором осуществляется усиление сигнала. Далее через передающую антенну сигнал поступает в канал связи.

После прохождения канала связи сигнал поступает на вход приемника (ПР-К), в котором он усиливается и подается на вход демодулятора (ДЕМ). В ДЕМ осуществляется формирование последовательности двоичных символов, подобных той, которая поступала на вход МОД на передающем конце линии связи. В этой последовательности ряд символов из-за действия шумов на входе приемника могут быть приняты с ошибкой, и они подлежат исправлению. Отметим, что на выходе ДЕМ возможно возникновение пачек ошибок. Ошибки могут группироваться в пакеты, т.е. быть коррелированными. К выходу ДЕМ подключен деперемежитель (ДПЕРМ), в котором над последовательностью двоичных символов осуществляются операции, обратные тем, которые были выполнены над последовательностью символов, поступивших на вход ПЕРМ с выхода кодера *LDPC*. Назначением ДПЕРМ является декорреляция ошибок, возникших при демодуляции принятого сигнала. С выхода ДПЕРМ на вход декодера *LDPC* поступает последовательность символов, подобная той, которая была сформирована в кодере *LDPC* на приемном конце линии связи.

В декодере кода *LDPC* осуществляется исправление части ошибок и формируется последовательность, подобная той, которая была сформирована выходе кодера БЧХ, при этом часть избыточных символов удаляется. В этой последовательности также могут оказаться ошибочными небольшая часть символов. Параметры кода БЧХ выбраны так, что, как это видно из табл. 2 и табл. 3, в декодере БЧХ могут быть исправлены до 12 ошибок, оставшихся после обработки принятого сигнала в декодере *LDPC*. На выходе декодера БЧХ формируется копия информационной последовательности символов, которая поступила на вход системы связи. При этом все избыточные символы удаляются.

		в станд	apme DVB-	S2 для .	$N_{LDPC} = 16$
N⁰	R _{LDPC}	k _{БЧX}	<i>N</i> _{БЧХ}	t _{БЧX}	NLDPC
1	1/4	16 008	16 200		
2	1/3	21 408	21 600		
3	2/5	25 728	25 920		
4	1/2	32 208	32 400		
5	3/5	38 688	38 880	12	
6	2/3	43 040	43 200		64800
7	3/4	48 408	48 600		
8	4/5	51 648	51 840		
9	5/6	53 840	54 000		
10	8/9	57 472	57 600	0	

58 320

Таблица 2. Параметры каскадного кода в стандарте DVB-S2 для $N_{\rm LDPC}$ = 16 200

Отметим, что вопросы, связанные с синхронизацией сигналов в данной ССС, структурой кадров и ряд других, которые описаны в [1], в данной работе не затрагиваются. Они решаются в спутниковых системах связи стандартными методами.

58 192

9/10

11

В ССС используются коды *LDPC*. Длина кодов *LDPC* равна N_{LDPC} = 64800 или N_{LDPC} = 16200, а параметры каскадного кода для этих значений N_{LDPC} приведены в табл. 2 и табл. 3.

Кодовая скорость кода для разных режимов работы указана в первом столбце табл. 2 и табл. 3. В зависимости от значений R_{LDPC} , указанных в столбце 1 этих таблиц, в кодере *БЧХ* формируется кодовая комбинация, длиной $N_{B'\!X}$, содержащая $k_{B'\!X}$ информационных символов. Из этих таблиц видно, что кодовая скорость данного кода $R_{B'\!I\!X} = (k_{B'\!I\!X}/N_{B'\!I\!X})$ весьма близка к 1 (0,988 $\leq R_{B'\!I\!X} \leq 0,998$ для $N_{LDPC} = 64800$ и 0,948 $\leq R_{B'\!I\!X} \leq 20,988$ для $N_{LDPC} = 16200$), т.е. $k_{B'\!I\!X} \approx N_{B'\!I\!X} = N_{LDPC} \cdot R_{LDPC}$. Код $B'\!I\!X$ позволяет исправить в кодовой комбинации то количество ошибок, которое указано в столбце 4 этих таблиц.

Габлица 3. Параметры каскадного ко	ода
стандарте DVB-S2 для N_{LDPC} = 64 ξ	800

№	R _{LDPC}	k _{БЧX}	<i>N</i> _{БЧХ}	<i>t</i> БЧХ	N _{LDPC}
1	1/4	3072	4050		
2	1/3	5232	5400		
3	2/5	6312	6480		
4	1/2	7032	7200		
5	3/5	9552	9720	12	16200
6	2/3	10 632	10 800	12	10200
7	3/4	11 712	11 880		
8	4/5	12 432	12 600		
9	5/6	13 152	13 320		
10	8/9	14 232	14 400		

В табл. 4 указаны параметры каскадного кода, образованного с помощью кодов БЧХ и LDPC. В зависимости от режима работы матричная память для формирования этого кода имеет параметры (количество строк и столбцов, в которых расположены ячейки памяти), указанные в табл. 4.

Сигнальные точки созвездия 8-PSK расположены на одной окружности, а у созвездия 16-APSK - на двух окружностях. Если обозначить радиус внешней окружности R₁, а внутренний R₂, то табл. 5 приведен коэффициент $\gamma = R_1/R_2$, который, согласно стандарту DVB-S2, должен быть выбран для сигналов 16-APSK в зависимости от кодовой скорости R_{LDPC}. Для созвездия 32-APSK сигнальные точки расположены на трех окружностях: на внешней, радиуса R₁, расположены 16 сигнальных точек; на средней, радиуса R₂ – 12 сигнальных точек, а на внутренней, радиуса R₃ – 4 сигнальные точки. В табл. 5 приведены значение коэффициентов $\gamma_1 = R_1/R_3$ и $\gamma_2 = R_1/R_3$, которые выбираются для разных режимов работы ССС в зависимости от кодовой скорости R_{LDPC}. В табл. 5 приведены также значения параметра $R_0 = (R_f R_{EYX} R_{LDPC})$ – удельной скорости передачи сообщений в системе связи, в которой применяется каскадное кодирование с кодами БЧХ и LDPC, имеющими разные режимы работы системы связи и разные кодовые скорости, указанные в табл. 2 и табл. 3.

Таблица 4. Параметры матрицы каскадного кода

№	Вид модуляции	Кол-во строк (N _{LDPC} =64800)	Кол-во строк (N _{LDPC} =16200)	Кол-во столбцов
1	8-PSK	21 600	5400	3
2	16-APSK	16 200	4050	4
3	32-APSK	12 960	3240	5

Запись бинарных символов каскадного кода в ячейки памяти этой матрицы осуществляется последовательно по столбцам, начиная с первого, а считывание – последовательно по строкам, начиная с первой.

В табл. 5 приведены параметры сигнальных созвездий для 16-APSK и 32-APSK.

Габлица 5. Параметр	ы сигнальных	созвездий
	для 16-APSK	u 32-APSK

Вид модуляции	16-A	PSK	32-APSK			
R_{LDPC}	R_{θ}	γ	R_{θ}	γ1	γ_2	
2/3	2,66	3,15	-	-	-	
3/4	2,99	2,85	3,74	2,84	5,27	
4/5	3,19	2,75	3,99	2,72	4,87	
5/6	3,32	2,70	4,15	2,64	4,64	
8/9	3,55	2,60	4,43	2,54	4,33	
9/10	3,59	2,57	4,49	2,53	4,30	

В соответствии со стандартом DVB-S2 на систему спутниковой связи, на ее выходе должен быть обеспечен квазибезошибочный прием сигналов. Критерием такого приема является обеспечение вероятности приема $P_{\kappa} = 10^{-7}$ пакета данных, содержащего 188 байт (1504 бит). В нем также приведены данные о значениях отношения сигнал/шум (*р*_{DVB-S2} дБ), которые должны быть обеспечены на входе демодулятора сигналов для того, чтобы выполнялись указанные выше требования к качеству приема. Как отмечено в [1], эти значения относятся к случаю, когда применяется алгоритм «мягкого» декодирования кода LDPC, состоящий из 50 итераций. Такой алгоритм, как известно [10] (стр. 390-394), дает по сравнению с «жестким» энергетический выигрыш, равный 2 дБ. Эти значения, а также значения R_f и R_0 при разных режимах работы системы указаны в табл. 6 и табл. 7.

Прочерки в строках этих таблиц указывают на то, что в соответствии со стандартом отмеченные режимы не применяются. В табл. 6 и табл. 7 указаны также значения ρ_s (в дБ) – отношение сигнал/шум для соответствующих режимов работы на входе демодулятора (предела Шеннона), рассчитанные по формуле [6-8] $\rho_s = (2^{R_0} - 1)$, которая следует из формулы Шеннона для пропускной спо-

R _{LDPC}	1/4	1/3	2/5	1/2	3/5	2/3	3/4	4/5	5/6	8/9	9/10
<i>QPSK, R_f=2</i> бит/сек. Гц											
R_{θ}	0,490	0,656	0,789	0,988	1,188	1,322	1,487	1,587	1,654	1,766	1,788
<i>р_{DVB-S2}</i> дБ	-2,35	-1,24	-0,30	1,00	2,23	3,10	4,03	4,68	5,18	6,20	6,42
<i>ρs</i> дБ	-3,93	-2,39	-1,38	0	1	1,75	2,56	3	3,3	3,8	3,9
<i>∆р</i> дБ	1,6	1,2	1	1	1,23	1,35	1,47	1,68	1,88	2,4	2,52
<i>8-РЅК, R_f=3</i> бит⁄сек•Гц											
R_{θ}	-	-	-	-	1,779	1,986	2,228	-	2,478	2,641	2,679
<i>р_{DVB-S2}</i> дБ	-	-	-	-	5,50	6,62	7,91	-	9,35	10,69	10,98
<i>ρs</i> дБ	-	-	-	-	3,86	4,7	5,66	-	6,6	7,2	7,32
<i>∆р</i> дБ	-	-	-	-	1,64	1,92	2,25	-	2,75	3,49	3,66

Таблица 6. Значения R_f и R_0 при разных режимах работы при использовании сигналов QPSK и 8-PSK

R _{LDPC}	1/4	1/3	2/5	1/2	3/5	2/3	3/4	4/5	5/6	8/9	9/10
<i>16-АРЅК, R_f=4</i> бит/сек•Гц											
R_{θ}	-	I	-	-	-	2,637	3	3,165	3,300	3,523	3,567
<i>р_{DVB-S2}</i> дБ	-	I	-	-	-	8,97	10,21	11,03	11,61	12,89	13,13
<i>ρ</i> _s дБ	-	-	-	-	-	7,17	8,33	9	9,47	10,21	10,35
<i>∆ρ</i> дБ	-	-	-	-	-	1,8	1,88	2,03	2,14	2,68	2,78
<i>32-APSK, R_f=5</i> бит⁄сек•Гц											
$R_{ heta}$	-	-	-	-	-	-	3,703	3,951	4,119	4,397	4,453
<i>р_{DVB-S2}</i> дБ	-	-	-	-	1	-	12,73	13,64	14,28	15,69	16,05
<u>ρ</u> s дБ	-	-	-	-	-	-	10,79	11,6	12,14	13	13,19
Δho дБ	-	-	-	-	-	-	1,94	2,04	2,14	2,69	2,86

Таблица 7. Значения R_f и $R_{ heta}$ при разных режимах работы при использовании сигналов 16-APSK и 32-APSK

собности канала связи. Эта формула учитывает, что система связи, в которой применяются для передачи сообщений оптимальные АС («идеальная» система Шеннона), должны иметь такую же спектральную эффективность (*R*₀), что и системы, созданные на основе стандарта DVB-S2. Отметим, что данная формула предполагает, что для передачи сигналов в системе связи применяются оптимальные АС, а использование помехоустойчивых кодов не предусматривается.

Сравнение величин ρ_{DVB-S2} и ρ_s , приведенных в табл. 6 и табл. 7, показывает, что энергетические потери $\Delta \rho = (\rho_{DVB-S2} - \rho_s)$ дБ спутниковых систем, создаваемых по стандарту DVB-S2 [4], относительно предела Шеннона зависят от кодовой скорости R_{LDPC} и при применении QPSK составляют от 1 до 1,5 дБ при 1/4 $\leq R_{LDPC} \leq 3/4$; эти потери увеличиваются до 2...2,5 дБ при $R_{LDPC} = 8/9$ и 9/10. Если для передачи сообщений применяются сигналы 8-PSK, то энергетические потери $\Delta \rho$ изменяются от 1,6 до 3,66 дБ при $3/5 \leq R_{LDPC} \leq 9/10$. Для режимов работы спутниковых систем с высокой скоростью $R_f = 4$ или 5 бит/сек-Гц, энергетические потери $\Delta \rho$ примерно составляют, как видно из табл. 7, от 2 до 2,8 дБ.

Следует отметить, что приведенные выше оценки величины потерь $\Delta \rho$ по отношению к пределу Шеннона существенно отличаются от оценок, представленных в [1], а также в [11], где указывается, что при всех режимах работы эти потери составляют всего 0,6...0,8 дБ.

Теоретические оценки надежности связи, обеспечиваемой в спутниковых системах, построенных по стандарту DVB-S2

Данные о качестве приема сообщений в системе спутниковой связи, приведенные в стандарте DVB-S2, а также в предыдущем разделе, были получены методом статистического моделирования на основе пакета программ MathLab. По сути этот метод является эмпирическим. Отметим два серьезных недостатка этого метода. Первый связан с тем, что анализ сложных систем таких, например, как ССС стандарта DVB-S2, является весьма сложным и требует расхода значительного машинного времени. Кроме того, требования к надежности приема сообщений в данной системе велики (вероятность ошибки при приеме сообщений должна иметь весьма малое значение (10⁻⁷), и поэтому для надежного определения этой вероятности для каждого режима работы ССС количество статистических испытаний (оно оценивается величиной, обратно пропорционально оцениваемой вероятности ошибки) и машинное время при моделировании должно быть весьма большим.

Вторым крупным недостатком этого метода является сложность верификации специалистами результатов выполненного с его помощью анализа сложной системы, так как она требует полного доступа к программе моделирования, которая является собственностью разработавшей ее фирмы, и независимым экспертам к этой программе практически нет доступа. Это снижает доверие к публикуемым результатам исследований, полученным методом моделирования и не подтвержденным их сопоставлением с результатами анализа, основанного на теории связи.

Аналитический метод исследования сложных коммуникационных систем, учитывающий основные положения теории связи, является в ряде случаев более простым и надежным методом получения оценок качества передачи сообщений в системах связи. Результаты такого анализа могут быть легко верифицированы специалистами, так как вытекают из основных закономерностей этой теории и представимы в аналитическом виде. Для количественного уточнения этих результатов можно использовать также и метод статистического моделирования.

В данном разделе представлен аналитический метод оценки качества приема сигналов в системах связи, построенных по стандарту DVB-S2, при «жестком» алгоритме декодирования каскадного кода, в котором код *БЧХ* является внешним, а код *LDPC* внутренним кодом. Как отмечалось выше, перед «мягким» алгоритмом этот алгоритм декодирования имеет энергетические потери порядка 2 дБ.

Изложенный ниже аналитический метод основан на развитой в [8] теории, которая позволяет получить аналитические формулы, определяющие зависимости надежности приема сообщений от системных параметров ССС (от параметров модуляции передаваемых сигналов, вида и параметров кодов). Результаты расчетов по представленным формулам оказались достаточно близки к тем данным, которые были получены методом статистического моделирования в [1], отличаясь от них на величину, при всех режимах работы не превышающую 2,8 дБ.

Как показано на блок-схеме рис. 1 рассматриваемой системы связи, между кодером *LDPC* и модулятором на

передаче расположен ПЕРМ, а на приеме между демодулятором и декодером LDPC расположен ДПЕРМ. Отметим, что как указывалось выше при обсуждении табл. 2 и табл. 3, кодовая скорость кода БЧХ (R_{БЧХ}) в системах стандарта DVB-S2 весьма близка к 1, т.е. ошибки, возникающие на выходе демодулятора, исправляются, в основном, в декодере кода LDPC. Блоки ПЕРМ и ДПЕРМ используются для повышения эффективности исправления этих ошибок в этом декодере путем разнесения во времени бит, определяющих сигнал, поступающий на вход демодулятора. По сути эти блоки осуществляют декорреляцию ошибок в битах, поступающих на вход декодера кода LDPC. В стандарте DVB-S2 эти блоки применяются только в тех режимах работы, когда для передачи сообщений используются ансамбли сигналов 16-APSK и 32-APSK. После первого декодера кода LDPC последовательность поступает на вход декодера кода БЧХ, который предназначен для устранения оставшихся в ней ошибок.

Поясним процедуру оценки вероятности ошибки декодирования каскадного кода в общем случае. Как правило, задача оценки вероятности декодирования ($P_{dec}(n,R_c,p_{dm})$) кодовой комбинации любого определенного кода не является сложной, когда код имеет определенную длину n, кодовую скорость R_c ($k=nR_c-$ количество информационных символов в декодированной последовательности), а ошибки символов, появляющихся на выходе демодулятора, происходят независимо с вероятностью, равной $p_{dm}.$

Отметим, что для кода LDPC параметры n, R_c и k в выражении для $P_{dec}(n,R_c,p_{dm})$ должны быть равны n = N_{LDPC} , R_c = R_{LDPC} и k = N_{EYX} , где значения N_{LDPC} , R_{LDPC} и N_{EYX} приведены в табл. 2 и табл. 3. Там же отмечено, что в том случае, когда N_{LDPC} = 64000 и $R_{LDPC} \leq 5 / 6$ код $\mathcal{B}\mathcal{H}X$ позволяет исправлять t_{EYX} =12 ошибок в кодовой комбинации, если R_{LDPC} = 8 / 9 или 9/10, то t_{EYX} = 8; если же N_{LDPC} = 16000, то при всех режимах работы системы код $\mathcal{B}\mathcal{H}X$ позволяет исправлять t_{EYX} =12 ошибок. Последовательность символов с выхода декодера LDPC поступает на вход декодера кода $\mathcal{B}\mathcal{H}X$. Для того, чтобы определить вероятность ошибки при декодировании этого кода, следует оценить вероятность ошибки одного двоичного знака (p_b) в этой последовательности.

Для такой оценки целесообразно использовать простой метод, предложенный Л.М. Финком [12] (стр. 125-135). Согласно этому методу, полагая, что в последовательности из $k=nR_c$ символов ошибки происходят независимо, можно приближенно считать, что вероятность их правильного приема равна вероятности правильного декодирования последовательности, поступившей на вход второго декодера, т.е. $(1-p_b)^k=[1--P_{dec}(p_{dm},R_c,n)].$ Из этого соотношения следует

$$\begin{split} p_b(p_{dm},R_c,n) &\cong 1 - [1-P_{dec}(p_{dm},R_c,n)]^{(1/k)} \cong \\ &\cong (1/nR_c)P_{dec}(p_{dm},R_c,n). \end{split}$$

Код *LDPC* относиться к эффективным кодам, у которых количество исправляемых ошибок (t_c) растет примерно линейно с увеличением их длины. Используя результаты [8] (стр. 81-86), можно оценить величину t_{LDPC}

для кодов LDPC с помощью следующих формул $t_{LDPC_n} \le t_{LDPC_v} \le t_{LDPC_v}.$ (2)

В (2) нижняя t_{LDPC_n} и верхняя оценки t_{LDPC_v} величины t_{LDPC} могут быть вычислены следующим образом $t_{LDPC_n} = N_{LDPC} \cdot \hat{t}_{LDPC_n}$, $t_{LDPC_v} = N_{LDPC} \cdot \hat{t}_{LDPC_v}$, где

$$\hat{t}_{LDPC_n} = 0,66 \left[1 - \sqrt{1 - 0,665 \left(1 - R_{LDPC}^{0,41} \right)} \right],$$

$$\hat{t}_{LDPC_v} = U(1 - U), \text{ rge } U = 2 \hat{t}_{LDPC_n}.$$
(3)

Отметим, что в качестве оценки t_{LDPC} в расчетах можно использовать среднее значение оценок \hat{t}_{LDPC_n} и \hat{t}_{LDPC_v} :

$$t_{LDPC_s} = N_{LDPC} \cdot (\hat{t}_{LDPC_n} + \hat{t}_{LDPC_v})/2.$$
(4)

Получим оценку вероятности ошибки декодирования помехоустойчивого кода [8]. Если в последовательности двоичных символов, поступающих на вход декодера с выхода демодулятора, вероятность ошибки одного символа равна $p_{\rm dm}$, то количество ошибок, которые могут появиться в этой последовательности, распределено по биномиальному закону.

В этом случае для расчета $P_{dec}(N_{LDPC},R_{LDPC},p_{dm})$ можно использовать формулу

 $\begin{array}{ll} P_{dec}(N_{LDPC},R_{LDPC},p_{dm}) &= \sum_{i=(t_{LDPC}+1)}^{N_{LDPC}} C_{N_{LDPC}}^{i} p_{dm}^{i} q_{dm}^{N_{LDPC}-i}, \mbox{(5)} \\ rge t_{LDPC} &-$ количество ошибок, которые могут быть исправлены кодом LDPC, а $q_{dm} = (1-p_{dm})$. Для того, чтобы при больших значениях N_{LDPC} получить простую формулу для расчета вероятности $P_{dec}(N_{LDPC},R_{LDPC},p_{dm})$ воспользуемся методом Чернова [10] (стр. 51). В результате получим [8] (стр. 271)

$$\begin{split} P_{dec}(N_{LDPC}, R_{LDPC}, p_{dm}) &\cong 10^{-N_{LDPC}U(p_{dm}, t_{LDPC})}, \text{ где} \\ U(p_{dm}, \hat{t}_c) &= \hat{t}_c lg\left(\frac{\hat{t}_{LDPC}}{p_{dm}}\right) + (1 - \hat{t}_c) lg\left(\frac{1 - \hat{t}_{LDPC}}{1 - p_{dm}}\right). \end{split}$$
(6)

В (6) $\hat{t}_{LDPC} = (t_{LDPC}/N)$ – относительное количество исправляемых ошибок в декодируемой последовательности символов. Отметим, что формула (6) справедлива при выполнении условия $\hat{t}_{LDPC} > p_{dm}$. Если $\hat{t}_{LDPC} \ll 1$ и $p_{dm} \ll 1$, то формула (6) упрощается и имеет вид:

$$p_{b1} \cong \left(\frac{1}{N_{LDPC}R_{LDPC}}\right) \left(\frac{p_{dm}}{\hat{t}_{LDPC}}\right)^{t_{LDPC}}.$$
(7)

Аналогичным образом найдем, что вероятность ошибочного приема одного символа в последовательности символов, сформированной на выходе декодера кода *БЧХ*, равна

$$p_{b2} \cong \left(\frac{1}{k_{EYX}}\right) \left(\frac{N_{EYX} \cdot p_{b1}}{t_{EYX}}\right)^{t_{EYX}} \cong \\ \cong \left(\frac{1}{N_{LDPC} R_{LDPC}}\right) \left(\frac{N_{LDPC} R_{LDPC} \cdot p_{b1}}{t_{EYX}}\right)^{t_{EYX}}.$$
(8)

В (8) учтено, что, в соответствии с данными о параметрах стандарта [1], приведенными в табл. 2 и табл. 3, $R_{\rm EYX}\cong 1$ и $k_{\rm EYX}\cong N_{\rm EYX}=N_{\rm LDPC}R_{\rm LDPC}.$ Учитывая (7) и (8), найдем

$$p_{b2} \cong \left(\frac{1}{N_{LDPC}R_{LDPC}}\right) \left(\frac{1}{t_{EYX}}\right)^{t_{EYX}} \left(\frac{p_{dm}}{\hat{t}_{LDPC}}\right)^{t_{EYX}\cdot\hat{t}_{LDPC}\cdot N_{LDPC}}.$$
(9)

В (9) значения \hat{t}_{LDPC} могут быть вычислены с помощью формул (3) и (4).

Как было отмечено ранее, система спутниковой связи, построенная в соответствии со стандартом DVB-S2, должна обеспечить квазибезошибочный прием сигналов, критерием которого является обеспечение вероятности приема $P_{\kappa} = 10^{-7}$ пакета данных, содержащего 188 байт ($N_k = 1504$ бит). Таким образом, исходя из (1), должно выполняться условие

$$P_k = 1 - (1 - p_{b2})^{N_k}.$$
 (10)

Используя формулы (9) и (10), рассчитаем зависимости допустимого значения вероятности ошибочного приема символов на выходе демодулятора сигналов (p_{mod}) от величины R_{LDSP} , при которой обеспечивается требуемая надежность приема сообщений в системе спутниковой связи, т.е. обеспечивается вероятность ошибки при приеме пакета данных, длиной 188 бит, равная $P_k = 10^{-7}$. При расчетах было учтено, что $t_{EYX} = 12$.

Эти зависимости приведены на рис. 2. Значения p_{dmn}, p_{dms} и p_{dmv} на этом рисунке определяются с помощью формулы (9) и зависят от значений t_{LDPC_n}, t_{LDPC_s} и t_{LDPC_v} , которые вычисляются по формулам (2) и (3). Из рис. 2 видно, что система спутниковой связи проектируется так, что, как уже отмечалось выше, вероятность ошибок, возникающих при демодуляции принимаемых сигналов, может принимать достаточно большие значения. В системах стандарта DVB-S2, значение этой вероятности лежит в интервале $10^{-2} \le p_{dm} \le 0,2$. Поэтому для исправления ошибок, возникших на выходе демодулятора, необходимо применять мощные коды, имеющие весьма большую длину.





В системах стандарта DVB-S2, как отмечалось ранее, в зависимости от выбранного оператором связи режима работы применяются следующие виды модуляции: сигналов QPSK, 8-PSK, 16-APSK и 32-APSK. Как следует из данных, приведенных в табл. 5. коэффициенты у. у1 и у2 имеют достаточно большие значения. Поэтому пикфактор этих сигналов близок к 1. Таким образом, полная энергия сигналов этих созвездий затрачивается в основном на передачу сигнальных точек, расположенных на внешней окружности их сигнальной диаграммы. По этой причине, оценивая величину p_{dm} – вероятность ошибки при демодуляции этих сигналов, можно с достаточно высокой точностью полагать, что вероятность ошибки при приеме этих сигналов равна вероятности ошибки при приеме сигналов, расположенных только на внешней окружности их сигнальной диаграммы. Эта вероятность может быть определена по формуле [10]

$$p_{\rm dm}(M,\rho_s) = 2Q(\sqrt{2\rho_s}\sin(\frac{\pi}{M})). \tag{11}$$

В (11)
$$Q(x) = Z = 2 \int_{x}^{\infty} \frac{\exp(-x^{2}/2)dx}{\sqrt{2\pi}} - функция Крампа. В$$

[8] (стр. 272) приводится формула, определяющая функцию, обратную функции Крампа, которая позволяет вычислить значение x при известном значении Z, т.е. найти зависимость $x = \Psi(Z)$. Функцию $\Psi(Z)$ с высокой точностью можно рассчитать, используя формулу [8]

$$\Psi(Z) = t(Z) - \frac{c_0 + c_1 t(Z) + c_2 t(Z)^2}{1 + d_1 t(Z) + d_2 t(Z)^2 + d_3 t(Z)^3},$$
(12)

где t(Z) = $\sqrt{2\ln[(Z)^{-1}]}$, c₀ = 2,525517, c₁ = 0,802853, c₂ = 0,010328, d₁ = 1, d₂ = 0,189269, d₃ = 0,001308. Из (11) и (12) следует, что значение отношение сигнал/шум, которое необходимо обеспечить на входе демодулятора для того, чтобы удовлетворить заданные требования к надежности приема сообщений, можно определить по формуле

$$\rho_{s} = \frac{[\Psi(p_{dm})]^{2}}{2 \sin^{2}(\frac{\pi}{M})}.$$
(13)

На рис. 3 приведены оценки сверху ($\rho_{DVB_S2_V}$) и снизу ($\rho_{DVB_S2_n}$) значения отношения сигнал⁄шум на входе демодулятора (в дБ) в зависимости от кодовой скорости (R_{LDSP}) у кода LDPC при применении для передачи сообщений сигналов QPSK, 8-PSK, 16-APSK и 32-APSK, обеспечивающих, соответственно, удельную скорость передачи сообщений в канале связи, равную R_f =2, 3, 4 и 5 бит/сек-Гц.



Рис. 3. Зависимости границ значений $\rho_{DVB_{S2}}$ от величины кодовой скорости R_{LDPC}

Оценки ρ_{DVB_S2} сверху (штриховые линии) и снизу (сплошные линии) соответствуют оценкам \hat{t}_{LDPC} снизу и сверху. Их значения определяются формулами (3). Как видно из рис. 3, разница между верхней и нижней границами величины ρ_{DVB_S2} при $R_{LDPC} \ge 0,5$ не превышает 3 дБ.

В табл. 8 приведены данные, позволяющие сравнить значения ρ_{DVB-S2} дБ, приведенные в стандарте DVB-S2 (они указаны в табл. 6 и табл. 7), с оценками, рассчитанными по формулам (8-13), определяющим оценки значения ρ_{DVB-S2} для «жесткого» алгоритма демодуляции принятых сигналов, при котором обеспечивается такая же надежность приема, что и при «мягком» алгоритме демодуляции.

Данные $\rho_{DVB-S2\,n}/\rho_{DVB-S2\,s.}/\rho_{DVB-S2\,V}$, приведенные в строках «границы ρ_{DVB-S2} дБ», определены для случаев, когда значения нормированного количества ошибок, исправляемых кодом LDPC, равны числами $\hat{t}_{LDPC_n}/\hat{t}_{LDPC_s}/\hat{t}_{LDPC_v}$, рассчитанным по формулам (3) и (4). Анализ данных табл. 8 показывает (сравниваемые данные в табл. 8 выделены жирным шрифтом), что для сигналов QPSK данные оценок ρ_{DVB-S2} , приведенных в

R _{LDPC}	1/2	3/4	5/6	9/10	
QPSK					
<i>р_{DVB-S2}</i> дБ	1	4	<u>5,18</u>	<u>6,42</u>	
границы р _{DVB-S2} дБ	-1,4⁄0,1⁄ <u>1,6</u>	2,4⁄3,3⁄ <u>4,3</u>	3,7/4,5/ <u>5,4</u>	5/5,6/ <u>6,3</u>	
8-PSK					
<i>р_{DVB-S2}</i> дБ	-	<u>7,9</u>	<u>9,35</u>	<u>10,98</u>	
границы р _{DVB-S2} дБ	-	7,6⁄8,6⁄ <u>9,7</u>	9⁄9,8⁄ <u>10,7</u>	10,4/11/ 11,7	
16-APSK					

10.2

11/12/13

12.7

13,5/14,5/15,5

_

Таблица 8. Сравнение данных табл. 6 и табл. 7 с оценками, полученными по формуле (15)

11.6

12,4/13,2/14

14,3

14,9/15,6/16,5

стандарте DVB-S2, незначительно (на 0,1...0,5 дБ) отличаются от рассчитанных оценок верхних границ этих параметров для алгоритма «жесткого» декодирования. Для других видов сигналов 8-PSK, 16-APSK и 32-APSK отличия этих оценок от данных ρ_{DVB-S2} из стандарта DVB-S2 более значительны. Они могут составить до 2,8 дБ (для режима 32-APSK и R_{LDPC} = 3/4). С учетом того, что известен выигрыш [10], который может дать применение «мягкого» декодирования (≈ 2 дБ), полученные оценки могут рассматриваться как достаточно точные оценки возможных значений ρ_{DVB-S2} , для приведенных в стандарте [1] видов модуляции и помехоустойчивых кодов.

<u>р_{DVB-S2} д</u>Б

32-APSK

*р_{DVB-S2} д*Б

границы *р*_{DVB-S2} дБ

границы *р_{DVB-S2} д*Б

Достоинством аналитических формул, связывающих параметры видов модуляции и помехоустойчивых кодов, применяемых в системах спутниковой связи, является то, что они позволяют без труда оценить надежность приема сообщений при разных параметрах сигналов и помехоустойчивых кодов.

Заключение

В работе дан анализ стандарта DVB-S2, на основе которого в настоящее время создаются системы спутникового вещания и связи. Представлен аналитический метод определения качества приема сообщений в системах, построенных по этому стандарту, но использующих «жесткий» метод демодуляции принимаемых сигналов. В статье приведено сравнение приведенных в стандарте DVB-S2 оценок помехоустойчивости приема сообщений, полученных с помощью статистического моделирования для «мягкого» метода демодуляции сигналов, с теми, которые получены на основе теоретического анализа. Сравнение показало, что результаты теоретического анализа с учетом известного энергетического выигрыша «мягкого» метода демодуляции сигналов по отношению к «жесткому», достаточно близки к полученным методом статистического моделирования.

Таким образом, предложенный теоретический метод исследования дает возможность получить аналитические формулы, определяющие зависимости надежности приема сообщений в спутниковых системах связи от их системных параметров (от параметров модуляции передаваемых сигналов, вида и параметров кодов, используемых в ССС). В статье [9] изложены результаты исследования возможностей применения оптимальных в спутниковой связи многомерных поверхностно-сферических АС, показавшие, что системы, в которых для передачи сообщений применяются многомерные сигналы по своим характеристикам существенно превосходят системы, создаваемые в соответствии со стандартом DVB-S.

13.13

13.7/14.4/15

16,05

16,2/16,8/17,5

Литература

1. European standard. ETSI EN 302 307-1 V1.4.1 (2014-11). Digital Video Broadcasting (DVB); Second generation framing structure, channel coding and modulation systems for Broadcasting, Interactive Services, News Gathering and other broadband satellite applications; Part 1: DVB-S2

2. U. Reimers, A. Morello, DVB-S2, the second generation standard for satellite broadcasting and unicasting, Int. J. Satell. Commun. Networks, 2004; vol. 22.

3. Mustafa Eroz, Feng-Wen Sun and Lin-Nan Lee. DVB-S2 low density parity check codes with near Shannon limit performance. Int. J. Satell. Commun. Network, 2004; V. 22

4. EBU Tech 3348 r4, Frequency and network planning aspects of DVB-T2, version 4.1.1. Geneva, October, 2014

5. Питерсон У., Уэлдон Э. Коды, исправляющие ошибки. М.: Мир, 1976

6. 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)

7. Shannon C. Probability of error for optimal codes in Gaussian channel. Bell System Techn. J., May, 1959. (Перевод на русский язык статьи «Вероятность ошибки для оптимальных кодов в гауссовском канале», опубликованой в книге Шеннон К. Работы по теории информации и кибернетике. М.: Из-во иностранной литературы // Под ред. Р.Л. Добрушина и О.Б. Лупанова, 1963.

 Быховский М.А. Гиперфазовая модуляция – оптимальный метод передачи сообщений в гауссовских каналах связи. М.: Техносфера, 2018.

9. Быховский М.А. Эффективные методы передачи сигналов в спутниковых системах связи (будет опубликована в журнале Цифровая Обработка Сигналов №2/2020).

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

11. Vlastimil Benovsky, Eurovision. DVB-S extension higher spectral efficiency. WBU-ISOG Forum Los Angeles, May, 2013.

12. Финк Л.М. Теория передачи дискретных сообщений. М.: Советское радио, 1970.

УДК 621.391.83

ИННОВАЦИОННЫЙ МЕТОД ИЗМЕРЕНИЯ ИМПУЛЬСНОЙ ХАРАКТЕРИСТИКИ И НЕЛИНЕЙНЫХ ИСКАЖЕНИЙ

Годына В.С., инженер-программист ИООО «Эпам Системз» (Минск, Белоруссия), e-mail: vladimir.godyna@gmail.com;

Перелыгин С.В., к.т.н., доцент кафедры аудиовизуальных систем и технологий Санкт-Петербургского государственного института кино и телевидения, e-mail: sergey.perelygin@gmail.com; Штейн Б.М., к.п.н., доцент кафедры аудиовизуальных систем и технологий Санкт-Петербургского государственного института кино и телевидения, e-mail: bomos@yandex.ru.

NOVEL APPROACH IN MEASURING IMPULSE RESPONSE AND NONLINEAR DISTORTION

Godina V.S., Perelygin S.V., Stein B.M.

This article reviews a novel method of simultaneous measurement of impulse response and nonlinear distortion, proposed by Pr. A. Farina (University of Parma). The advantages over traditional measurement methods are examined, as well as the wide possibilities for using this method in the practice of measuring the characteristics of audio devices and rooms. In addition, the accuracy of the considered method was estimated.

Key words: impulse response measurement, nonlinear distortion measurement, room acoustics, correlation method, acoustic measurement.

Ключевые слова: измерение импульсной характеристики, измерение нелинейных искажений, акустика помещений, корреляционный метод, акустические измерения.

Введение

Г

В 2000 году профессор Пармского Университета Анджело Фарина [1] предложил ориги-

нальный метод измерения импульсной характеристики с использованием тестового гармонического сигнала с экспоненциально нарастающей частотой [2] (далее ESS – Exponential Sine Sweep), имеющего вид:

$$x(t) = \sin\left[\frac{\omega_1 \cdot T}{\ln \frac{\omega_2}{\omega_1}} \cdot \left(e^{\frac{T}{t} \cdot \ln \frac{\omega_2}{\omega_1}} - 1\right)\right],$$
(1)

Г

где ω_1 , ω_2 – начальная и конечная круговая частота тестового сигнала (выбирается в соответствии с измеряемым диапазоном частот), T – длительность тестового сигнала.

Взяв производную по времени от аргумента синуса в (1), получим выражение для экспоненциально возрастающей во времени мгновенной частоты ESS-сигнала (рис. 1):

$$f(t) = f_1 \cdot e^{\frac{t}{T} \cdot \prod_{f_1}^{f_2}},$$
(2)

где f_1 , f_2 – начальная и конечная частота тестового сигнала в Герцах.

Отметим, что вследствие нелинейного закона возрастания частоты, амплитудная спектральная плотность такого сигнала неравномерна и имеет *спад* 3 дБ/окт в сторону высоких частот (рис. 2).

Приведён обзор метода одновременного измерения импульсной характеристики и нелинейных искажений, предложенного профессором А. Фарина (университет Пармы). Рассмотрены преимущества над традиционными методами измерения, а также широкие возможности по использованию данного метода в практике измерения характеристик аудиоустройств и помещений. Оценена точность рассматриваемого метода.



Рис. 1. Временна́я зависимость меновенной частоты ESS-сигнала с параметрами:

 f_1 = 11,7 Гц, f_2 =24 кГц, T = 43 с





Особенности метода

Г

Измерение импульсной характеристики и нелинейных искажений испытуемого устройства производится следующим образом [2]:

1. Вначале необходимо подать ESS-сигнал на вход испытуемого устройства и записать его выходной сигнал.

2. Далее необходимо произвести операцию свёртки выходного сигнала со специально подготовленным «обратным» ESS-сигналом.

«Обратный» ESS-сигнал образован из исходного ESS-сигнала путём его отражения с конца в начало по временной оси. Так как амплитудные спектральные плотности и исходного, и «обратного» ESS-сигнала имеют *спад* 3 дБ/окт в сторону высоких частот, то амплитудная спектральная плотность свёртки этих двух сигналов будет иметь *спад* 6 дБ/окт в сторону высоких частот. Для компенсации этого наклона отражённый по времени ESS-сигнал необходимо промодулировать так, чтобы его мгновенная амплитуда уменьшалась во времени по экспоненциальному закону, одновременно с экспоненциально убывающей мгновенной частотой. Такой сигнал имеет вид [1]:

$$x^{*}(t) = \sin \left| \frac{\omega_{1} \cdot T}{\ln \frac{\omega_{2}}{\omega_{1}}} \cdot \left(e^{\frac{T-t}{T} \cdot \ln \frac{\omega_{2}}{\omega_{1}}} - 1 \right) \right| \cdot e^{\frac{t}{T} \cdot \ln \frac{\omega_{2}}{\omega_{1}}}.$$
 (3)

Полученный после свёртки сигнал содержит линейную (основную) импульсную реакцию и импульсные реакции нелинейных искажений испытуемого устройства в виде последовательности импульсных характеристик (ИХ). Расположение импульсных реакций на временной оси иллюстрирует рис. 3 (аббревиатура IR_HD образована от первых букв словосочетаний Impulse Response и Harmonic Distortion). Последняя (правая) из импульсных реакций является линейной импульсной реакцией устройства, ей предшествует во времени импульсная реакция на квадратичную нелинейность устройства, перед ней – реакция на нелинейность 3-го порядка т.д. На практике удаётся зафиксировать нелинейности до 20 порядка.



Рис. 3. Пример сигнала, полученного в результате свёртки тестового сигнала с выхода устройства и отражённого во времени, промодулированного ESS-сигнала

Подробное объяснение такого эффекта содержится в [2], [3].

При этом именно ESS-сигнал обладает свойством разделять во времени импульсные реакции нелинейных искажений от линейной импульсной реакции.

Кроме этого, ESS-метод обладает следующими преимуществами [2], [3]:

Повышенная устойчивость к случайному шуму: гармонический сигнал «сканирует» очень узкий частотный промежуток в каждый интервал времени, что позволяет уменьшить уровень случайного шума в результирующих импульсных реакциях. Чем длиннее ESS-сигнал и, соответственно, чем медленнее возрастает частота, тем заметнее эффект снижения уровня шума.

Простота использования: декодирование сводится к процедурам вычисления быстрого преобразования Фурье, что является очень недорогой и быстрой операцией на современных компьютерах.

Возможность автоматизации процесса измерений: расположение импульсных реакций искажений относительно линейной импульсной реакции на временной оси можно найти по следующему правилу [2]: если частота ESS-сигнала растёт со скоростью 1 октава в секунду, то реакция на нелинейность 2-го порядка будет расположена за 1 секунду до линейной импульсной реакции; реакция на нелинейность 4-го порядка будет расположена за 1 секунду до реакции на нелинейность 2-го порядка и т.д.

Несмотря на очевидные преимущества, ESS-метод имеет и определённые границы применения. Сам автор описываемого метода перечисляет следующие ограничения [3]:

Чувствительность к широкополосным помехам, попадающим в диапазон частот от ω_1 до ω_2 в (1) и (2).

Собственная погрешность метода в частотной области (рис. 4), вызванная тем, что во временной области ESS-сигнал можно считать промодулированным прямоугольным окном. В результате, спектр ESS-сигнала искажается спектром прямоугольного окна (по свойству свёртки спектров). Наибольшая погрешность проявляется на краях диапазона ($\omega_1 \div \omega_2$).



Рис.4. Собственная погрешность ESS-метода при измерении АЧХ

В работе [4] рассмотрена проблема, связанная с тем, что каждая следующая октава, отсчитываемая относительно начальной частоты ω_1 ESS-сигнала (1), начинается не в нулевой фазе. Этот фактор вносит дополнительную погрешность в измерение гармоник испытуемого устройства. Для решения проблемы в [4] предложен следующий закон изменения ESS-сигнала:

$$x(n) = \sin\left[\frac{\pi \cdot L}{2^{p} \cdot P \cdot \ln 2} \cdot e^{\frac{n}{N} \cdot P \cdot \ln 2}\right], \quad \frac{\pi \cdot L}{2^{p} \cdot P \cdot \ln 2} = 2 \cdot \pi \cdot M.$$
(4)

где *P* – целое число октав, *L* – действительное (нецелое) число отсчётов ESS-сигнала, *N* – округлённое до

целого L, M – положительный целочисленный множитель.

Данная формула описывает ESS-сигнал в дискретном виде. Верхняя частота такого сигнала равна половине частоты дискретизации.

Экспериментальные исследования

Погрешность метода

Был рассчитан и сгенерирован ESS-сигнал по формуле (4) длительностью 43 секунды, диапазоном в 11 октав (11,7 ÷ 24000 Гц) с частотой дискретизации 48 кГц. Для оценки погрешности ESS-метода была произведена свёртка исходного ESS-сигнала с отражённым по времени и промодулированным ESS-сигналом по формуле (3). В результате была получена импульсная реакция идеального линейного всепропускающего безынерционного устройства. На рис. 4 представлена амплитудная спектральная плотность полученной импульсной реакции. По форме она близка к равномерной в диапазоне частот 11,7 ÷ 24000 Гц и имеет заметную неравномерность только на краях диапазона. Эту неравномерность можно считать собственной погрешностью ESS-метода при измерении АЧХ. Приемлемым диапазоном измерения частот можно считать 15 ÷ 21000 Гц при погрешности, не превышающей 0,4 дБ.

Исследование характеристик громкоговорителя

ESS-метод был также применён при измерении акустических характеристик громкоговорителя. В качестве испытуемого устройства был выбран студийный монитор (электроакустический агрегат со встроенным усилителем) промышленного производства.

Измерение производилось с использованием следующих компонентов:

– Измерительный микрофон RFT MV201 с капсюлем MK201 [7].

 Компьютер со звуковой картой – в качестве источника аналогового аудиосигнала, оцифровщика сигнала, полученного с измерительного микрофона, а также устройства для вычисления характеристик испытуемого громкоговорителя.

– Заглушённое студийное помещение объёмом 80 м³.

Процедура физической части измерения заключалась в следующем:

 – ESS-сигнал воспроизводился с выхода звуковой карты компьютера и подавался на студийный монитор.

 Акустический отклик монитора регистрировался с помощью микрофона, который также был подключён к звуковой карте компьютера.

Записанный сигнал (отклик монитора) обрабатывался программно (осуществлялась его свёртка с отражённым промодулированным ESS-сигналом). В результате был получен сигнал, состоящий из последовательности импульсных реакций (рис. 5).

Линейная импульсная реакция изображена в правой части графика (самая последняя). Импульсные реакции нелинейностей 2-го, 3-го и 5-го порядков находятся на временной отметке 5,4; 3,1 и 0,2 сек соответственно. Паузы между импульсными реакциями содержат шум, ниже уровня которого данные измерений недействительны.





Далее эти импульсные реакции были разделены и сохранены в виде отдельных сигналов, а затем преобразованы из временной области в спектральную, что дало возможность получить АЧХ и частотные зависимости нелинейных искажений студийного монитора (рис. 6).

Также, удалось оценить нижнюю границу измерения (порог шумов). Выяснилось, что помимо импульсных реакций результат вычисления свёртки (рис.5) содержит также некоторый уровень шума, маскирующий совсем малые импульсные реакции. Для оценки уровня порога шумов был выделен участок такого шума между импульсными реакциями и оценена его спектральная плотность (серый график на рис. 6).



Рис. 6. АЧХ и частотные характеристики нелинейностей 2, 3 и 5-го порядков студийного монитора. Серый график соответствует порогу шумов измерения

В данном примере приведены измерения АЧХ и частотных зависимостей нелинейных искажений студийного монитора (электроакустического агрегата) в ближнем поле при уровне звукового давления 90 дБ относительно 20 мкПа. Учитывая, что приведённый собственный шум измерительного микрофона Gefell MV201 составляет 22 дБА [7], ожидаемым авторами отношением сигнал/шум должно было стать значение 90 – 22 = 68 дБ. Но как видно из графика, в диапазоне 300 ÷ 20000 Гц отношение сигнал/шум составило более 90 дБ.

Другими примерами применения ESS-метода могут стать измерение АЧХ и нелинейных искажений аудиоустройств (в том числе музыкального оборудования), а также измерение акустических характеристик помещений.

Заключение

Метод, предложенный А. Фарина, открывает возможность производить измерения характеристик электронных и электроакустических устройств, а также акустических свойств помещений с высокой точностью. ESS-метод отличается универсальностью применения, высокой помехоустойчивостью, а также возможностью автоматизировать процесс измерений. Известно, что данный метод уже успешно применяется в программах акустический измерений, таких как CLIO [5] и REW [6].

Дальнейшей работой будет оптимизация исходного ESS-сигнала для уменьшения частотных и временных искажений самого метода.

Литература

1. Персональная страница профессора А. Фарина [Электронный ресурс]. – URL: http://pcfarina.eng.unipr.it (дата обращения 10.09.2019).

 Метод одновременного измерения импульсной характеристики и нелинейных искажений с помощью гармонического сигнала экспоненциально изменяющейся частоты. [Электронный ресурс] – URL: http://citeseerx. ist.psu.edu/viewdoc/download?doi=10.1.1.33.1614&rep=rep 1&type=pdf (дата обращения 10.09.2019).

3. Достижения в измерении импульсной характеристики с помощью гармонического сигнала экспоненциально изменяющейся частоты. [Электронный ресурс] URL: http://pcfarina.eng.unipr.it/Public/Papers/226-AES122.pdf (дата обращения 10.09.2019).

4. Реализация метода измерения импульсной характеристики с помощью сигнала экспоненциально качающейся частоты [Электронный ресурс]. URL: https://www.uniweimar.de/kunst-undgestaltung/wiki/images/ExpoChirpToolbox_a_Pd_implementation_of_ESS_impulse_response_me asurement.pdf (дата обращения 10.09.2019).

5. CLIO software. [Электронный ресурс]. – URL: http://www.audiomatica.com (дата обращения 10.09.2019).

6. REW software. [Электронный ресурс]. URL: https://www.roomeqwizard.com (дата обращения 10.09.2019).

7. Характеристики измерительного капсюля Gefell MK-201. [Электронный ресурс]. – URL:https://www.micro-techgefell.de/mikrofonkapsel?wl=472-MK201# (дата обращения 10.09.2019).

НОВЫЕ КНИГИ





ПРОГРАММНО-Конфигурируемые сети





ЮНФИГУРИРУЕМОГО РАДИО



Программно-конфигурируемые сети / Учебник для вузов - М.: Изд-во «Горячая линия-Телеком», 2020 г. – 288 с.: ил.

Рекомендован федеральным учебно-методическим объединением в системе высшего образования по укрупненной группе специальностей и направлений подготовки 09.00.00 – «Информатика и вычислительная техника» в качестве учебника для студентов, обучающихся по основным образовательным программам высшего образования по направлению подготовки бакалавров и магистрантов 09.00.00 – «Информатика и вычислительная техника».

Рассмотрены теоретические и практические основы построения, проектирования и поддержки компьютерных сетей нового поколения – программно-конфигурируемых сетей (ПКС). Особое внимание уделено различным перспективным решениям динамического реконфигурирования ПКС, вопросам разработки инструментальных средств и программных приложений, обеспечивающих высокую скорость и гибкость перепрограммирования сетевой инфраструктуры.

Фокин Г.А.

Технологии программно-конфигурируемого радио / Учебное пособие для вузов - М.: Изд-во «Горячая линия-Телеком», 2019 г. – 316 с.: ил.

Изложены теоретические и практические сведения, необходимые для получения навыков модельно-ориентированного проектирования (МОП) систем радиосвязи на основе программно-конфигурируемого радио (ПКР). Аппаратный инструментарий предполагает использование плат USRP Ettus Research и RTL-SDR. Программный инструментарий включает специальное программное обеспечение Matlab/Simulink. Представленные в пособии примеры и задачи позволяют развить компетенции программной реализации алгоритмов и методик передачи, приема и обработки сигналов в современных и перспективных системах радиосвязи, а также овладеть способностями оценки показателей функционирования реализованных алгоритмов средствами эмуляции и эксперимента. Разработанные модели Matlab/Simulink, использованные в пособии, доступны на сайте издательства www.techbook.ru.

Для студентов, обучающихся по направлениям подготовки 11.04.02 – «Инфокоммуникационные технологии и системы связи» (магистратура), 11.04.01 – «Радиотехника» (магистратура). Будет полезно аспирантам и специалистам, занятым вопросами научно-исследовательской и опытно-конструкторской деятельности в области современной радиоэлектроники. УДК 621.396

КОРРЕЛЯТОР С АДАПТИВНЫМ ПОРОГОМ

Бартенев В.Г., д.т.н., профессор Российского технологического университета (МИРЭА), e-mail: bartenev_v@mirea.ru.

THE CORRELATOR WITH ADAPTIVE THRESHOLD

Bartenev V.G.

The problem of detecting a correlated signals in noise is considered. Two correlation detectors are introduced. One detector may be realized using multiplication and coherent addition with constant threshold. Another just the same, but with adaptive threshold multiplication a binary decision device. Numerical study is presented analytically for false alarm probability but detection probability obtained using models of correlators in MATLAB.

Key words: signal correlation detection, adaptive threshold, constant false alarm rate, probability of false alarm and probability of detection.

Ключевые слова: корреляционные обнаружители сигналов, стабилизация ложных тревог, адаптивный порог, вероятности ложной тревоги и правильного обнаружения.

Введение

Задача обнаружения коррелированных сигналов на фоне некоррелированных случайных процессов по дискретным выборкам конечного объема возникает во многих технических приложениях. Известен способ корреляционного обнаружения принимаемых сигналов, когда две

выборки наблюдения принятых на двух несущих частотах перемножаются, их произведение накапливается и модуль накопленного произведения сравнивается с фиксированным порогом [1]. Полученная таким образом оценка модуля межчастотного коэффициента корреляции сравнивается с порогом, на основании чего принимается решение о наличии принятых коррелированных сигналов. Хотя данный способ позволяет осуществлять эффективное обнаружение коррелированных сигналов, тем не менее, данному способу свойственен недостаток, проявляющийся в отсутствии стабилизации ложных тревог при изменении уровня шума, на фоне которого производится обнаружение.

С целью обеспечения стабилизации ложных тревог при корреляционном обнаружении предлагается способ, который включает в себя формирование оценки модуля коэффициента корреляции на основе выборок наблюдений, принятых на двух несущих частотах, и сравнение этой оценки с порогом, который с целью стабилизации ложных тревог при изменении уровня шума делают адаптивным, формируемым как произведение коэффициента, определяющего вероятность ложной тревоги на суммарную оценку мощности шума на двух несущих частотах. Как правило, анализ их эффективности производился с помощью статистического моделирования, так как нелинейная операция умножения приводит к изменению вида распределений на выходе этих устройств и существенному усложнению их анализа с

Рассматривается задача корреляционного обнаружения флюктуирующих коррелированных сигналов на фоне некоррелированного шума. Один корреляционный обнаружитель реализуется с помощью умножения и когерентного накопления сигналов с фиксированным порогом. В другом корреляторе после умножения и когерентного накопления используется адаптивный порог. Расчет порогов для вероятностей ложных тревог на выходе этих обнаружителей для малых выборок наблюдения произведен аналитически, а вероятности правильного обнаружения рассчитаны моделированием в системе MATLAB. Данные результаты в радиолокационной практике получены впервые.

> помощью аналитических выкладок особенно для малых выборок наблюдения и низких вероятностей ложных тревог. Однако, если при нахождении характеристик обнаружения точность расчета вероятности правильного обнаружения допускает моделирование, то для малых вероятностей ложных тревог точность расчета с помощью статистического моделирования становится недопустимо низкой. По этой причине и была предпринята попытка впервые найти аналитические выражения для расчета низких вероятностей ложных тревог для нелинейных устройств с умножителем на входе при использовании малых выборок наблюдений и адаптивным порогом.

Вероятность превышения порога огибающей шума на выходе умножителя с когерентным накопителем и фиксированным порогом

Рассмотрим коррелятор с фиксированным порогом, и покажем, что при изменении уровня шума изменяется вероятность ложной тревоги на его выходе. Для расчета вероятности ложной тревоги для коррелятора с фиксированным порогом воспользуемся следующим выражением

$$\hat{R} = |(\sum_{j=1}^{N} Z\mathbf{1}_{j} * Z\mathbf{2}_{j}^{*})| = \sqrt{(\sum_{j=1}^{N} x\mathbf{1}_{j} * x\mathbf{2}_{j} + y\mathbf{1}_{j} * y\mathbf{2}_{j})^{2} + (\sum_{j=1}^{N} x\mathbf{2}_{j} * y\mathbf{1}_{j} - x\mathbf{1}_{j} * y\mathbf{2}_{j})^{2},}$$
(1)

где R – оценка модуля коэффициента корреляции,

V N – число накоплений по независимым выборкам. $Z1_j = x1_j + iy1_j, Z2_j = x2_j + iy2_j$ – комплексные выборки сигналов на входе умножителя разнесенных по частоте в виде аддитивной смеси шума и коррелированного сигнала. Квадратурные компоненты шума имеют нормальное распределение, при этом их мощность (дисперсия) равна σ^2 и среднее 0. Обнаружение сигналов в корреляторе с фиксированным порогом осуществляется путем сравнения полученной оценки модуля коэффициента корреляции с порогом $R_{ПОР}$, $\hat{R} > R_{ПОР}$. Покажем, что изменение мощности шума σ^2 приводит к изменению вероятности ложной тревоги. Для этого применяя методику из [2] нахождения вероятности ложной тревоги $F(R_{ПОР})$, получим

$$F(R_{\Pi OP}) = \frac{(R_{\Pi OP} / \sigma^2)^N K_N (R_{\Pi OP} / \sigma^2)}{2^{N-1} \Gamma(N)}.$$
 (2)

В данное выражение входит гамма функция $\Gamma(N)$,

модифицированная функция Бесселя $K_N(\hat{R}_{\Pi OP})$ порядка N и мощность шума σ^2 .



Рис. 1. Вероятность ложной тревоги для корреляционного обнаружения с фиксированным порогом в зависимости от порога L для N = 4. Ромбики (моделирование) и кружочки (аналитика) на графиках соответствуют мощности шума 0 дБ, квадратики (моделирование) и звездочки (аналитика) мощности шума 3 дБ

Расчеты по формуле (2) для *N* = 4, приведенные на рис. 1, показывают, что даже незначительные изменения мощности шума на входе от 0 до 3 дБ приводит к заметному изменению (росту) вероятности ложной тревоги. Для верификации аналитических расчетов на графике имеются результаты и моделирования коррелятора с фиксированным порогом в МАТLAB. Совпадение аналитики и моделирования подтверждают отсутствие стабильной вероятности ложной тревоги в корреляторе с фиксированным порогом.

Вероятность ложной тревоги на выходе умножителя с когерентным накопителем и адаптивным порогом

Для преодоления указанного недостатка предлагается производить дополнительно оценки мощности шума на двух несущих частотах, т.е. *z*₁ и *z*₂

$$z_{1} = \sum_{i=1}^{N-1} \operatorname{Re}(Z_{1i}) \operatorname{Re}(Z_{1i}) + \operatorname{Im}(Z_{1i}) \operatorname{Im}(Z_{1i})$$
(3)

$$z_{2} = \sum_{i=1}^{N-1} \operatorname{Re}(Z_{2i}) \operatorname{Re}(Z_{2i}) + \operatorname{Im}(Z_{2i}) \operatorname{Im}(Z_{2i}).$$
(4)

Суммирование оценок мощности $Z_s = (z_1 + z_2)$ и умножение на коэффициент определяющий вероятность ложной тревоги α , позволяет сделать порог адаптивным $\hat{R} > \alpha Z_s$.

Считая независимыми оценки модуля коэффициента корреляции и оценки мощности шума, можно получить выражение для вероятности ложной тревоги предложенного адаптивного коррелятора.

$$F(\alpha) = \int_{0}^{\infty} P(Z_{S}) dZ_{S} \int_{\alpha Z_{S}}^{\infty} P(R) dR.$$
 (5)

Считая, что оценка мощности Z_s имеет распределение χ^2 , вероятность ложной тревоги $F(\alpha)$ примет вид

$$F(\alpha) = \int_{0}^{\infty} \frac{Z_s^{N-1} e^{-z_s/2\sigma^2}}{\Gamma(N)(2\sigma^2)^N} \cdot \frac{(\alpha Z_s/2\sigma^2) K_N(\alpha Z_s/2\sigma^2)}{\Gamma(N)(2)^{N-1}} dZ_s.$$
 (6)

После взятия интеграла получаем

$$F(\alpha) = \sqrt{\pi} a^{2N} \frac{\Gamma(3N)}{2^{(4N-1)}} \Gamma\left(2N + \frac{1}{2}\right) \Gamma(N)_{2} \times \\ \times_{2} F_{1}\left(\frac{3N+1}{2}, \frac{3N}{2}; 2N + 1/2; 1 - 4\alpha^{2}\right),$$
(7)

где $_2F_1$ гипергеометрическая функция. Полученное выражение (7) говорит о главном – в нем не присутствует мощность шума σ^2 . Ниже приводится таблица порогов для корреляторов с фиксированным порогом (два левых столбца) и с адаптивным порогом (два правых столбца) при разных ложных тревогах и разных *N*.





Дальнейший анализ производился не только аналитическим расчетом по полученной формуле, но и для верификации моделированием корреляционного обнаружения с адаптивным порогом в МАТЛАБ.

Результаты аналитических расчетов и моделирования показали хорошее совпадение (см. рис. 2), что позволяет сделать вывод о корректности полученного аналитического выражения (7). И, главное, изменение уров-



Рис. 3. Вероятность правильного обнаружения PD для корреляторов с фиксированным и адаптивным порогом для N = 4 в зависимости от сигнал/шум (dB) для вероятности ложной тревоги 0,1 и коэффициента корреляции 0,9. Кружочки на графиках соответствуют коррелятору с фиксированным порогом,



квадратики коррелятору с адаптивным порогом



ня шума в корреляторе с адаптивным порогом не влияет на вероятность ложной тревоги.

Коррелятор с адаптивным порогом по эффективности сравнивался с коррелятором с фиксированным порогом расчетом характеристик обнаружения флюктуирующего коррелированного сигнала. Это было сделано с помощью моделирования в системе MATLAB. На рис. 3-6 приводятся кривые для вероятности правильного обнаружения флюктуирующего сигнала с коэффициентом корреляции 0,9 для двух рассматриваемых устройств при N = 8 и вероятности ложной тревоги 0,1 и при N = 8 и вероятности ложной тревоги 0,0001 Показано, что эффективность в пороговом сигнале для вероятности правильного обнаружения 0,5 и вероятности ложной тревоги 0,1 и 0,0001 несколько выше у коррелятора без стабилизации ложных тревог. Это своего рода плата за инвариантные свойства адаптивного коррелятора к изменениям мощнос-ти шума, обеспечивая стабилизацию вероятности ложной тревоги на выходе. Следует заметить, что эти потери снижаются при увеличении выборки наблюдений.



Рис. 4. Вероятность правильного обнаружения PD для корреляторов с фиксированным и адаптивным порогом для N = 4 в зависимости от сигнал/шум (dB) для вероятности ложной тревоги 0,0001 и коэффициента корреляции 0,9. Кружочки на графиках соответствуют коррелятору с фиксированным порогом, квадратики коррелятору с адаптивным порогом



Рис. 6. Вероятность правильного обнаружения PD для корреляторов с фиксированным и адаптивным порогом для N = 8 в зависимости от сигнал/шум (dB) для вероятности ложной тревоги 0,0001 и коэффициента корреляции 0,9. Кружочки на графиках соответствуют коррелятору с фиксированным порогом, квадратики коррелятору с адаптивным порогом

Заключение

Таким образом, проведенное исследование в системе MATLAB полностью подтверждает положительный эффект от применения предложенного коррелятора со стабилизацией ложных тревог. Важно подчеркнуть, что полученное аналитическое выражение для вероятности ложной тревоги адаптивного коррелятора позволит более обстоятельно исследовать его свойства для разных малых выборок наблюдения в широком диапазоне вероятностей ложных тревог.

Литература

1. Бартенев В.Г. Анализ эффективности обнаружителей коррелированных сигналов в шуме для малых выборок наблюдения // Цифровая обработка сигналов. 2016. № 4. – С. 35-39

2. Бартенев В.Г., Бартенев М.В. Способ нахождения вероятностных характеристик на выходе нелинейных систем // Цифровая обработка сигналов. 2013. № 4.– С. 42-44

3. Потемкин В.Г. «Справочник по MATLAB» Анализ и обработка данных. http://matlab.exponenta.ru/ml/book2/ chapter8.

УДК 519.7+004.522+004.934.1

АКУСТИЧЕСКОЕ И ЯЗЫКОВОЕ МОДЕЛИРОВАНИЕ В СКВОЗНЫХ СИСТЕМАХ РАСПОЗНАВАНИЯ РЕЧИ

Чучупал В.Я., к.ф.-м.н., ведущий научный сотрудник Федерального исследовательского центра «Информатика и Управление» РАН, e-mail: v.chuchupal@gmail.com.

ACOUSTIC AND LANGUAGE MODELING IN END-TO-END SPEECH RECOGNITION SYSTEMS

Chuchupal V.J.

End-to-end speech recognition systems have appeared recently, but they already have recognition accuracy comparable to the conventional state-of-the-art hybrid systems based on hidden Markov models and deep neural networks. The use of homogeneous network structures for acoustic, pronunciation, and language modeling in end-to-end systems, simplification of decoding algorithms, and replacement of expert knowledge with those obtained by machine learning greatly simplified the architecture of speech recognition systems. The presence of open tools and datasets greatly facilitated the entry of new teams into this scientific and technical field. As a fee for simplifying the architecture of recognition systems, one can consider the need to use a very big, accordingly to the usual concepts, datasets for training models. Collection, annotating and augmentation audio and text data has become an important task. The lack of theoretical results to justify the optimality of the choice of models and training methods significantly complicates the development of systems. Nevertheless, the existing results give reason to believe that in the nearest future this technology will become a standard for building speech recognition systems.

Key words: automatic speech recognition, deep neural networks, end to end speech recognition systems, acoustic modeling, language models.

Ключевые слова: автоматическое распознавание речи, глубокие нейронные сети, сквозные системы распознавания, акустическое моделирование, модели языка.

Гибридные системы распознавания речи

К началу прошлого десятилетия достигнутый уровень технологии распознавания устной речи позволил создавать коммерчески успешные продукты с функциями автоматического распознавания речи.

Подход к распознаванию речи соответствовал известной с 70-х годов вероятностной формулировке задачи распознавания речи [1]: если $X = \{x_u\}, t = 1, ..., T$ – наблюдаемая последовательность параметров речевого сигнала, а $W = \{w_i\} i = 1, ..., N$ – некоторая последовательность слов, то наиболее вероятная последовательность слов W^* определяется путем оптимизации выражения:

 $W^* = \arg \max_{W} P(W \mid X) =$ = $\arg \max_{W} \frac{P(X \mid W)P(W)}{P(X)} =$ = $\arg \max_{W} P(X \mid W)P(W) =$ = $\arg \max_{W} P(W) \sum_{T} P(X \mid T)P(T \mid W),$

где T – множество всех фонемных транскрипций слов из W. В критерии (1) вероятности определяются на основе разных типов моделей: P(X|T) – акустических, P(T|W) – моделей произношения и P(W) – языковых. Таким образом, задача решается с использованием

(1)

Сквозные (end-to-end) системы распознавания речи появились совсем недавно, но уже имеют показатели качества распознавания, сравнимые с лучшими продуктовыми системами, основанными на методах скрытых марковских цепей и глубоких нейросетей. Использование в сквозных системах распознавания однородных сетевых структур для акустического, произносительного и языкового моделирования, упрощение алгоритмов декодирования и замена экспертных знаний на оценки параметров, полученные методами машинного обучения, существенно редуцировало архитектуру систем распознавания речи. Наличие открытых инструментариев и корпусов данных значительно облегчило вход в эту научно-техническую область новым коллективам. Как плату за упрощение архитектуры систем распознавания можно рассматривать необходимость использования огромных, по привычным понятиям, корпусов данных для оценки параметров моделей. Сбор, аннотирование и обогащение аудио и текстовых данных стало отдельной и важной задачей. Отсутствие теоретических результатов, с помощью которых можно обосновать оптимальность выбора моделей или методов их обучения, приводит к появлению большого количества моделей, понимание причин эффективности которых не совсем очевидно. Тем не менее, уже имеющиеся результаты дают основание считать, что в ближайшее время эта технология станет общепринятой для построения систем распознавания речи.

трех уровней моделирования речевого сигнала: акустического, произносительного и языкового.

Произношение последовательности слов моделируется как последовательность произнесения контекстнозависимых вариантов фонем (аллофонов) из фонемных транскрипций. Произнесения аллофонов, в свою очередь, представляются как реализации скрытых марковских моделей (CMM), HMM (hidden Markov models), где в качестве функций плотности вероятности распределения параметров чаще всего (до начала 2000 годов) использовались модели смесей нормальных распределений, GMM (gaussian mixture models). Построенные на таком подходе системы распознавания назывались HMM-GMM системами.

Хотя в выражение (1) три основные модели: акустическая, произносительная и языковая входят равноправно, на практике для минимизации уровня ошибок наиболее критичным оказалось качество акустического моделирования.

В конце 80-х – начале 90-х годов, на волне общего подъема интереса к искусственным нейросетям, были предложены модели, ориентированные на работу с речевыми сигналами, например, модель TDNN (timedelay-ed neural network) [2]. Нейросети начали использоваться в архитектуре НММ-GMM ограниченно: вместо GMM для оценки вероятности наблюдений параметров P(X|T), то есть в акустической модели. Эта архитектура получила название гибридной HMM-MLP (MLP-multilayered perceptron, многослойный персептрон). В 2003 г. система распознавания CU-HTK, построенная на гибридной архитектуре, на совместных испытаниях в рамках европейского проекта SQUALE опередила конкурентные системы НММ-GMM архитектуры [3] по качеству распознавания, при том, что имела более простую архитектуру с меньшим числом параметров. В то время проявившиеся недостатки нейросетей: их параметры - это массив весов сети, поэтому нужно обучать все модели сразу, нужны соответствующие вычислительные мощности, а также требование большого количества обучающих данных, не позволили в полной мере воспользоваться преимуществами нейросетей.

В следующем десятилетии, с ростом вычислительных возможностей компьютеров и появлением больших корпусов данных началось успешное массовое использование нейросетей и до настоящего времени гибридные системы HMM-DNN определяют мировой уровень работ в этой области. Аббревиатура DNN означает deep neural network, т.е. глубокая нейронная сеть с числом слоев более трех. С точки зрения терминологии DNN отличались от многослойных персептронов наличием дополнительной нейросети (DBN, deep belief network) для оптимизации выбора начальных значений параметров. Поскольку используются градиентные методы оптимизации, правильный выбор начальных значений параметров играет большую роль. Фактически оказалось [4], что при наличии достаточно больших выборок данных наличие процедуры предобучения в виде DBN может не играть большой роли в отличие от других, позднее предложенных методов предобучения, например послойного (layer-wise pretraining) обучения [5].

На рис. 1. представлена упрощенная схема HMM-DNN системы распознавания речи

Из рис. 1 видно, что структура систем распознавания включает несколько уровней представления, реализованных в модульном виде на основе собственных методов и моделей. Обучение и успешная работа системы связаны с оценкой параметров моделей (в соответствии с тремя основными уровнями критерия (1)) и гиперпараметров, которые регулируют баланс между ними для выработки согласованного решения. Оптимальный выбор моделей, методов оценки их параметров и гиперпараметров являются в данном случае отдельной и нетривиальной задачей.



Рис.1. Схема гибридной DNN-HMM архитектуры системы распознавания речи

В качестве примера можно привести процедуру оценки акустических параметров в самом известном пакете с открытым кодом Kaldi [6]: для обучения GMM-HMM или GMM-DNN моделей нужно пройти 7 или 9, соответственно, стадий обработки корпуса данных, каждая из которых представляет собой достаточно сложную итерационную процедуру обучения, которая уточняет оценки, полученные на предыдущей стадии. Для оптимального выбора значений гиперпараметров, таких как отношение весов языковой и акустической моделей выполняется перебор возможных вариантов.

Основные компоненты GMM-HMM и GMM-DNN систем распознавания речи, методы и модели, алгоритмы декодирования были определены к концу прошлого века. Дальнейшая динамика улучшения показателей эффективности распознавания, например, уровня пословных ошибок распознавания WER (word error rate) была связана с ростом объема обучающих данных и применением эффективных, дополняющих скрытые марковские модели, методов и технологий (нейросети с большим числом слоев, использование дискриминантных оценок параметров акустических моделей, методов адаптации к голосу диктора и каналам связи, в том числе идентификационных векторов (i-vectors), специальных структур марковских моделей (chain models)), которые в то же время в целом еще более усложняли структуру систем распознавания.

Сложность многоуровневого и многомасштабного представления данных отчасти удалось компенсировать разработкой и использованием единого аппарата для их компактного представления в виде композиции конечных вероятностных преобразователей (weighted finite state transducer, WFST). Тем не менее одновременное использование разных моделей, методов, источников знаний и необходимость оптимальной настройки их совместной работы существенно усложняет процедуры обучения систем, их понимания, отладки или адаптации к прикладным областям.

Принципиальные недостатки использования скрытых марковских моделей как механизма представления акустики звуков речи были хорошо известны с самого начала применения этого аппарата: в первую очередь это предположение о статистической независимости параметров сигнала на соседних кадрах анализа и неадекватность моделирования длительностей состояний. Эти недостатки компенсируются использованием дополнительно производных от параметров по времени или сегментных параметров, полученных агрегированием параметров на сегменте сигнала. Успехи акустических моделей на основе HMM-DNN технологии связаны как раз с возможностями нейросетей моделировать плотности распределения вероятностей параметров сигнала сразу на достаточно длинных сегментах сигнала, т. е. фактически не используя предположений о независимости параметров и моделей длительности.

На сегодняшний день технология HMM-DNN может рассматриваться как стандарт для разработки успешных продукционных систем распознавания речи. Из пяти лучших (по показателю WER) результатов на данных открытого корпуса LibriSpeech [7] четыре (в том числе три первых) принадлежат гибридным системам [8]. Отметим, что точность распознавания MP3 кодированной речи у них заметно выше, чем у человека. Например, уровень пословной ошибки для человека составил 5,83 % на чистой речи (тестовая часть LibriSpeech «testclean») и 12,69 % на речи с помехами (тестовая часть LibriSpeech «test-other») [9], при этом на гибридной системе HMM-DNN [10] уровни ошибок на этих же тестах составили 2,3 % и 4,9 % соответственно.

Сквозные модели как дальнейшее развитие нейросетевого подхода

Успешное использование нейросетей в гибридных системах стало толчком к расширению сферы использования нейросетей в системах распознавания речи. В последние годы разработаны и исследованы архитектуры сквозных (end-to-end) систем распознавания речи, в которых отсутствуют в явном виде почти все модули архитектур HMM-GMM и HMM-DNN, разве что за исключением моделей языка.

Сквозные системы можно рассматривать как одну нейросеть, которая преобразует входной сигнал (в параметрическом виде, например, векторов мел-спектральных параметров или непосредственно как РСМ сигнал) в последовательность символов: букв, морфов (частей слов) или слов.

В архитектуре сквозных систем обычно можно выделить структурные элементы, например, слои сети, которые решают задачи кодирования, декодирования и т.п. При этом эти слои являются органической частью всей нейросети, которая обучается как единое целое, обычно с использованием градиентных методов оптимизации.

Очевидным преимуществом сквозных систем является то, что они не требуют алгоритмов или экспертных правил преобразований буквенных записей в фонемные, которые необходимы для построения произносительного лексикона в HMM-DNN/GMM системах, более того, произносительный словарь тут не используется. Аналогично не требуется алгоритмов или правил для вычисления алфавитов контекстных моделей фонем и вероятностных преобразователей для их использования. Все это сильно упрощает архитектуру систем и уменьшает объем знаний о речи, необходимых разработчикам.





Очевидным недостатком сквозных систем является необходимость использования большого объема обучающих данных. Экспертные знания отсутствуют, их нужно находить из данных. Поэтому нужны большие данные и соответствующие вычислительные ресурсы.

Существенное упрощение архитектуры, требуемых экспертных знаний, наличие готовых решений с открытым кодом и появление больших доступных корпусов данных упростило разработку систем и облегчило вхождение в эту область новым коллективам. На сегодняшний день предложен целый ряд конкретных решений на базе сквозных моделей, которые, по-видимому, почти не уступают лучшим гибридным системам по эффективности распознавания [8]. Эти решения используют комбинации нескольких базовых моделей. К таким базовым моделям относятся модель сетевой временной классификации СТС [11, 12], ее модификация с рекуррентной моделью языка T-RNN [13], модель Wave2Letter [14], модель кодера-декодера с вниманием [15, 16] и модель трансформера [17].

Модель сетевой временной классификации

Исторически первой была предложена модель сетевой временной классификации – СТС (connectionist temporary classification) [11]. Модель СТС можно рассматривать как нейросетевой аналог методов оценки параметров состояний марковских моделей с использованием процедур прямого и обратного хода.

Пусть через X обозначен входной сигнал в виде последовательности его векторизованных параметров $X = \{x_1, x_2, ..., x_T\}$. Пусть Y – его транскрипция (или разметка) – соответствующая последовательность выходных символов, например, фонем или букв. Для обучения задано множество пар (X, Y), а мерой качества распознавания является среднее значение редакторского расстояния, минимизирующего число ошибок между корректными и распознанными последовательностями символов.

Модель СТС реализована в виде глубокой двунаправленной рекуррентной нейросети, которая преобразует вектора признаков *x* непосредственно в выходные символы: фонемы, буквы, морфы в зависимости от типа использованной при обучении разметки. СТС имеет число входов, равное размерности векторов признаков *x* и число выходов, равное размерности алфавита разметки плюс один т.н. пустой символ. Значения выходов генерируются синхронно параметрам x_i , причем значение

k-го выхода в момент $t - z_k^t$ интерпретируется как вероятность *k*-го символа алфавита разметки в момент *t*. В этом смысле функционально СТС похожа на нейросети в гибридных DNN-GMM системах. Однако, поскольку СТС состоит из двунаправленных рекуррентных элементов, то значения z_k^t на любом шаге *t* зависят от всех x_k из *X*.

Для последовательности параметров $X = \{x_1, x_2, ..., x_T\}$ с разметкой Y, назовем сегментацией последовательность символов Y_1^T выходного алфавита длины T, которая может отличаться от разметки Y только повторами символов или вставками пустого символа. Как правило T >> |Y|, в крайнем случае T = |Y|. Если Y_1^T – сегментация параметров X, то ее вероятность в СТС определяется как:

$$P(Y_1^T \mid X) = \prod_{t=1}^T P(z_k^t = y_t \mid X).$$
 (2)

Таким образом, в отличие от моделей HMM-GMM и DNN-GMM в (2) полагаются независимыми не наблюдения параметров, а выходные символы.

Поскольку при оценке функции потерь важна только корректность последовательности символов (число повторений символов и вставки пустого символа при этом не учитываются), вероятности сегментаций, которые соответствуют одинаковым последовательностям символов, суммируются при вычислении полной вероятности:

$$P(Y \mid X) = \sum_{Y_1^T \in S(Y,T)} P(Y_1^T \mid X),$$
(3)

где S(Y,T) множество всех сегментаций транскрипции Y на интервале длины T

Параметры модели СТС могут оцениваться с использованием различных функций стоимости [12], чаще всего минимизацией обратного логарифма вероятности для корректной транскрипции *У* речевого высказывания, представленного параметрами *X*:

$$CTC(X) = -\sum_{Y_1^T \in \mathcal{S}(Y,T)} \log P(Y \mid X).$$
(4)

Наряду с критерием (4) также широко используется критерий минимизации вероятности ошибок в транскрипции:

$$CTC(X) = -\sum_{Y} P(Y \mid X) L(X, Y).$$
(5)

В критерии (5) функция L(X, Y) обозначает число ошибок в разметке Y и заданных параметрах X, суммирование осуществляется по всем возможным разметкам.

Предположение о независимости (2), подход к вычислению полной вероятности последовательности символов (3) и критерий оптимизации (4) являются аналогами известных методов оценки параметров HMM.

Сеть СТС на каждом кадре анализа генерирует вектор вероятностей выходных символов. В качестве результата распознавания требуется один символ, при этом подпоследовательности из одинаковых символов и пустые символы должны быть сокращены до символа. Поэтому на выходе сети используется декодер, который в простейшем случае может просто в каждый момент времени выбирать наиболее вероятный символ и фильтровать повторы и пустые символы, тем не менее, лучшие результаты получаются при использовании более сложных декодеров с памятью [11].

Модель СТС широко используется при создании сквозных систем распознавания речи. Известные решения, построенные с ее использованием включают системы DeepSpeech [9], ESPnet [18], EESen [19].

Сравнивая СТС (и другие сквозные решения) по эффективности с гибридными HMM-DNN, необходимо отметить, что, поскольку модель СТС основана на рекуррентной двунаправленной сети, она использует результаты анализа сигнала как в прямом времени, так и в обратном, то есть она «знает будущее». Замена бинаправленных элементов в сети СТС (как и в других моделях) на однонаправленные, например, для реализации обработки в реальном времени, приводит к заметному ухудшению качества распознавания.

Предположение о независимости выходных символов (2) в методе СТС фактически означает отсутствие модели языка, что должно негативно сказаться на эффективности. Это практически и происходит: подключение адекватной внешней языковой модели, пусть на уровне символов или морфов, обеспечивает снижение уровня пословной ошибки на 25-50 % относительно исходного.

Точность распознавания речи, достигаемая при использовании модели СТС существенно зависит от характеристик внешней модели языка.

В табл.1 ниже представлены значения уровня пословной ошибки распознавания WER для системы DeepSpeech-2 (основана на модели СТС) и человека на нескольких корпусах данных. В частности, значение WER на тестовых частях корпуса LibriSpeech составило для чистой речи (часть test-clean) 5,33 %, для речи с помехами (test-other) – 13,25 %, что практически соответствует точности распознавания этого же материала человеком: 5,83 % и 12,69 % соответственно. Ошибка увеличилась при распознавании акцентной речи, где DeepSpeech-2 начинает проигрывать около 50 % (относительного значения) WER человеку, а также распознавании шумной речи, где уровень ошибки DeepSpeech2 уже в 2 раза выше, чем у человека.

Тестовый корпус	DeepSpeech2	Человек	Тип речи
LibriSpeech test-clean	5,33	5,83	Читаемая
LibriSpeech test-other	13,25	12,69	читаемая + помехи
VoxForge	7,55	4,85	акцентная
CHiME evaluation	21,79	11,84	Шумная

Таблица 1. Точность распознавания речи системой DeepSpeech2 и человеком [9]

Рекуррентный нейросетевой преобразователь (RNN-T)

Модель СТС не включает языковой модели, даже информации о вероятностях следования выходных символов. Этот недостаток устранен в модели рекуррентного нейросетевого преобразователя RNN-T (recurrent neural network transducer) [13], в которой вероятности в (4), (5) вычисляются с использованием двух моделей. Первая, фактически акустическая модель, определяет вероятность появления выходных символов при заданных параметрах речевого сигнала. Оно определяется также, как и в модели СТС, то есть с помощью многослойной двунаправленной многослойной рекуррентной сети, называемой транскрипционной сетью (transcription network). Вторая модель определяет условную вероятность появления выходного символа в зависимости от предыдущего, ее можно интерпретировать как простую языковую модель. Она вычисляется сетью прогноза (prediction network), однонаправленной рекуррентной сетью с одним скрытым слоем, которая идентична по структуре обычным одношаговым рекуррентным моделям языка, с той лишь разницей, что позволяет генерировать также и пустые символы.

Выходы обоих сетей используются для определения итоговой вероятности выходных символов. В первоначальном варианте вероятности транскрипционной сети и сети прогноза суммировались, позднее [21] был предложен более удачный вариант, в котором итоговая вероятность получалась как выход еще одной, объединяющей (joint) нейросети, которая использовала выходы транскрипционной и прогнозной сетей в качестве входных признаков.

По сравнению с моделью СТС усовершенствованная модель RNN-T обладает большей точностью распознавания [21], но за счет увеличения объема вычислений и усложнения обучения, что потребовало ввода процедуры пред-обучения. Эти проблемы частично устранены в дальнейших улучшениях модели [21].

Глубокие сверточные сети. Модель Wave2Letter

Рекуррентные нейросети успешно используются в качестве основы сквозных систем распознавания. Принципиальным недостатком рекуррентных сетей является последовательный порядок вычислений и, как следствие, невозможность организации параллельных вычислений.

Вариант модели СТС с заменой рекуррентных сетейна сверточные реализован в модели Wave2Letter [14], которая реализована на основе глубоких сверточных сетей DCNN (deep convolutional neural network), аналогично одной из первых моделей нейросетей для распознавания речи TDNN [2].

Многослойная (до 12 слоев) сеть имеет простой вид однонаправленной сети без прореживающих (pooling), как обычно в сверточных сетях, слоев. Вместо них для сжатия признаков используется смещение ядра сети с шагом больше 1. Каждый слой осуществляет преобразование входных сигналов х в выходные у вида:

$$y_t^i = b_i + \sum_{j=1}^{a_x} \sum_{k=1}^{k_w} w_{i,j,k}, x_{dw^*(t-1)+k}^j \quad 1 \le i \le d_y.$$
(6)

В качестве нелинейной функции для сжатия выходных сигналов элементов используется сигмоидальная функция. В формуле (6) x^{i}, y_{i} , обозначают j и i компоненты входного и выходного вектора, d_{y}, d_{x} – их размерности, d_{w} – шаг окна (ядра) анализа, kw – его длина, т. е. размерность входного слоя равна $kw^{*}d_{x}$.

В отличие от модели СТС сеть Wave2Letter кроме вероятностей появления символов также обучается вероятностям перехода между ними, фактически биграммной модели языка символов и предусматривает возможность интеграции модели языка с большим контекстом, в том числе для слов.

Оценка параметров осуществляется оптимизацией дискриминантной функции стоимости:

$$ASG(X) = -\log_{Y \in S_{corr}(Y,T)} \times (7)$$

$$\times (\sum_{t=1}^{T} (\log P(y_t \mid X) + \log P(y_t \mid y_{t-1}, X)) + \log Y \in S_{corr}(Y,T) (\sum_{t=1}^{T} (\log P(y_t \mid X) + \log P(y_t \mid y_{t-1}, X))).$$

В выражении (7) X обозначают параметры речевого высказывания, S(Y,T) – множество сегментаций всех транскрипций Y на интервале длины T: $Y = y_{l}, y_{2}, ..., y_{T}$, При этом $S_{corr}(Y,T)$ обозначает подмножество S(Y,T) из сегментаций правильных (для X) транскрипций.

С точки зрения точности распознавания сверточные сети не проигрывают другим методам, в частности, рекуррентным сетям в модели СТС. В следующей табл. 2. приведены значения величины WER на корпусе Librispeech для модели Wave2letter и конкурентных методов, относящимися к лучшим современным. Системы CAPIO и Seq2Seq относятся к гибридным системам распознавания.

ний. На величину ошибки заметно влияет качество модеетей-[14], мы используют разные модели языка. При использоваичных нии нейросетевой модели языка модель Wave2Letter на *Таблица 2. Значения показателя пословной ошибки распознавания WER*

для нескольких систем распознавания речи, полученные на корпусе LibriSpeech

Аббревиатура «4-граммн. ЯМ» означает официальную 4-граммную языковую модель LibriSpeech

Модель/Корпус данных	LibriSpeech	LibriSpeech	LibriSpeech	LibriSpeech
	dev-clean	Dev-other	test-clean	test-other
САРІО (DNN-HMM), 4-граммн. ЯМ [34]	3,02	8,28	3,56	8,58
DeepSpeech2[9]	-	-	5,83	12,69
Seq2Seq[5], 4-граммн. ЯМ	4,79	14,31	4,82	15,30
Seq2Seq[5], рекуррентная ЯМ	3,54	11,52	3,82	12,76
Wave2Letter[14], 4-граммн. ЯМ	4,26	13,80	4,82	14,54
Wave2Letter[14], ЯМ на сверточной сети	3,13	10,61	3,45	11,92

чистых (test-clean) данных демонстрирует лучшие показатели эффективности, на более «сложных» (test-other) проигрывает лучшей модели около 30 % относительного значения WER. Отметим, что ограничения в количестве обучающих данных LibriSpeech (960 часов) в большей степени влияют на характеристики сквозных систем, чем гибридных.

С точки зрения сложности сети, в том числе вычислительной, даже в версии, когда входным сигналом является РСМ сигнал и признаки вычисляются самой сетью, она имеет 12 слоев с общим числом параметров 23 млн., что существенно меньше, чем у системы DeepSpeech2 [9] с моделью СТС, которая имеет более 100 млн. параметров.

За счет многослойности элементы верхнего слоя Wave2Letter соответствуют сегменту сигнала длительностью около 2 с., что достаточно для учета любого фонетического контекста.

Простота сверточных сетей и возможность распараллеливания вычислений дает модели Wave2Letter явные преимущества в памяти и в скорости обработки речевого сигнала (при лучшем его качестве) по сравнению с другими моделями. Она работает на два порядка быстрее [23], чем, например модель ESPNET, которая использует интерполированные решения от моделей СТС и кодера-декодера.

Сквозные системы с использованием модели кодера-декодера с вниманием

Модель кодера-декодера [24] с вниманием (encoderdecoder with attention), изначально была предложена [15] для решения задачи автоматического перевода текстов. Схематически архитектура кодера-декодера представлена на следующем рис. 3. Это глубокая рекуррентная сеть со слоями, в которых выделены три компоненты, отдельные модели: кодер, внимание и декодер.



Рис. 3. Схема модели кодера-декодера с вниманием

Входом сети, как и в предыдущих моделях, является последовательность кратковременных параметров речевого сигнала $x = \{x_1, x_2, ..., x_T\}$, выходом – наиболее вероятная последовательность символов (слов или морфов) $y^* = \{y_1, y_2, ..., y_S\}$, т.е.

$$y^* = \arg\max P(y \mid x) \tag{8}$$

Вероятность (8) вычисляется путем аппроксимации:

$$P(y \mid x) = \prod_{i} P(y_i \mid y_{i-1}, y_i, \dots, y_1, x_1, x_2, \dots, x_T)$$
(9)

т.е. вероятность текущего символа y_i вычисляется с учетом контекстной информации, включающей кодированные представления входных значений и предыдущих выходных символов.

Кодер, реализованный как рекуррентная (как в модели СТС) сеть, преобразует со сжатием последовательность входных параметров x в последовательность «высокоуровневых» признаков $h = \{h_1, h_2, ..., h_N\}$, которые определяются как величины активации (или выходы) элементов скрытых слоев сети:

$$h_t = f(x_t, h_{t-1}), (10)$$

где h_{t-1} – выход скрытого слоя в предыдущий момент времени, x_t – выход предыдущего слоя (скрытого или входного). Значение *N*<<*T*, то есть схема на рис. З имеет «пирамидальный» вид.

Рекуррентная сеть кодера состоит из двунаправленных клеток LSTM (долговременно-кратковременной памяти, long short time memory), их выходы вычисляются как в прямом, так и в обратном времени:

$$h_{t}^{+} = f(x_{t}, h_{t-1}),$$

$$h_{t}^{-} = g(x_{t}, h_{t+1}),$$

$$h_{t} = [h_{t}^{+}, h_{t}^{-}],$$
(11)

где h_t – выход в момент t, x_t – выход предыдущего слоя, а h_{t+1} и h_{t-1} – выходы в следующий и предыдущий моменты времени.

Декодер вычисляет решение (8), (9), которое для *i*-го выходного символа вычисляется однослойной рекуррентной нейросетью как:

$$P(y_i \mid x) = g(y_{i-1}, s_i, c_i),$$
(12)

где c_i – усредненное значение признаков h, формируемое моделью внимания, а s_i – так называемое состояние декодера, которое также вычисляется рекуррентной сетью:

$$s_i = f(s_{i-1}, y_{i-1}, c_i).$$
 (13)

Модель внимания оценивает c_i – среднее значение выходов кодера h для i-го выходного символа y_i , используя для этого обучаемые распределения вероятностей $a_{i,j}$.

$$c_{i} = \sum_{j=1}^{N} \alpha_{i,j} h_{j}; \quad \alpha_{i,j} = \frac{\exp(e_{i,k})}{\sum_{k=1}^{N} \exp(e_{i,k})}; \quad e_{i,k} = a(s_{i-1}, h_{k}). \quad (14)$$

В выражениях (14) функция а обозначает модель выравнивания (alignment model), которая реализуется направленной однослойной полносвязной нейросетью.

Модель кодера-декодера с вниманием активно используется в решениях компании Google и по своим характеристикам (включая точность распознавания) не уступает текущим продукционным системам на основе гибридных моделей.

Нужно отметить, что даже по сравнению с другими сквозными решениями, модель кодера-декодера с вни-

манием требует наличия больших корпусов данных, величина ошибки распознавания явно зависит от размера корпуса. При распознавании поисковых запросов и размере обучающего датасета 2 тыс. часов пословная ошибка для модели LAS (listen, attend and spell) [25] с использованием внешней языковой модели была 10,3 % [24], а при использовании корпуса в 12,5 тыс. часов ошибка понизилась до 5,6-6,9 % [25, 26].

Модель Трансформера

Указанный ранее недостаток рекуррентных сетей: сложность с распараллеливанием вычислений и вытекающий отсюда большой объем вычислений делают перспективными аналоги модели кодера-декодера с вниманием на основе сверточных или направленных полносвязных сетей. Таким аналогом является модель трансформера [17], адаптированная для задачи распознавания речи [28].

Модель трансформера, изображенная на рис. 4, в целом повторяет архитектуру кодера-декодера с вниманием, но здесь слои внимания используются повсюду между слоями кодера и декодера.



Рис. 4. Схема модели Трансформера ([16])

Кодер и декодер содержат стеки из 6 композитных слоев. Каждый слой стека кодера (изображен на рис. 5) состоит из подслоя само-внимания (self-attention) и поточечного полносвязного (positional-wise feed-forward network) подслоя.

Модель самовнимания 8-фокусная (multi-head attention), вычисляет веса как в качестве параметров *s* и *h*) выходы предыдущего слоя (т.е. элементы *h* одного и того же слоя). Мультифокусность означает, что входные вектора-признаков 8 раз линейно проектируются (элементы проекционных матриц обучаются совместно с сетью) на пространства меньшей размерности и модель внимания или само-внимания применяется раздельно к этим проекциям.

Поточечная сеть – это двуслойная (т.е. с одним скрытым слоем) сверточная сеть с длиной ядра 1, так, что каждый элемент входного слоя преобразуется независимо от других, при этом функция преобразования одна и та же. Число фильтров в модели трансформера [17] равно 4 и при 512-мерных векторах признаков поточечный слой реализует четыре различных преобразования над компонентами каждого вектора признаков. Выход верхнего слоя кодера соединен со всеми слоями декодера.



Рис. 5. Схема слоя кодера

Декодер также состоит из 6-слойного стека (изображен на рис. 6) с композитными слоями. Каждый слой декодера включает 3 подслоя: два из которых (самовнимание и полносвязный) идентичны соответствующим подслоям в кодере. Дополнительно присутствует подслой внимания для входных признаков, приходящих из кодера.



Рис. 6. Схема слоя декодера

Использование модели само-внимания в трансформере приносит заметные вычислительные преимущества по сравнению с рекуррентными вариантами слоев. Хотя количество операций, приходящихся на каждый слой внимания может быть выше, чем в рекуррентных сетях вычислительная сложность при использовании само-внимания, рекуррентного и сверточного слоев оценивается как [33]: $O(n^2 d)$, $O(nd^2)$ и $O(knd^2)$, где n – длина последовательности векторизованных признаков, *d* размерность каждого вектора, *k* – длина ядра свертки), но при этом для подслоев само-внимания и свёрточных поточечных реализуется полное распараллеливание вычислений, когда на каждый слой придется постоянное число операций, О(1), в то время как для рекуррентной сети это число растет линейно с длиной последовательности, т.е. O(n).

Модель трансформера обеспечивает высокую точ-

ность распознавания: лучший из опубликованных на конец 2019 года результат на корпусе LibriSpeech принадлежал сквозной системе распознавания, построенной на модели трансформера [28].

О модели внимания

Как и модель кодера-декодера, модель внимания сначала была предложена [15] как часть модели сквозной системы машинного перевода

По сути, «внимание» - это операция усреднения значений параметров или признаков. Обработка нейросетью длинных последовательностей данных, как в случае речевого сигнала, порождает соответствующую последовательность векторизованных признаков, выходов элементов некоторого слоя сети. Для оценки условных вероятностей типа (9), зависящих от контекстов в виде *п*-грамм признаков намного эффективнее использовать вместо длинной последовательности признаков ее компактное представление. Таким представлением может быть усреднение той части элементов последовательности, которая наиболее соответствуют текущей ситуации. Поскольку усреднение - это суммирование взвешенных значений, остается определить эффективный способ вычисления весов, которые в общем случае зависят как от значений признаков, так и от текущих значений на входе и выходе сети. Такими способами являются, в случае моделей кодера-декодера и трансформера, использование, соответственно, обучаемых совместно однослойных нейросетей и поточечных нормированных произведений для соответствующих значений.

Экспериментально показано, что существенный дополнительный выигрыш в точности распознавания может быть получен при использовании одновременно нескольких функций внимания («multi-head attention», много-фокусное внимание, описанное в предыдущем разделе).

Специфическими недостатками модели внимания и систем, построенных с ее применением, являются отсутствие монотонности и зависимость от длительности обучающих фраз. Отсутствие монотонности приводит к тому, что последовательным по времени сегментам входного сигнала могут соответствовать выходные символы, представленные в другой последовательности. Это нормально для машинного перевода, но не для распознавания речи, и является причиной ошибок. А наблюдаемая зависимость от длительности обучающих предложений приводит к тому, что система, обученная на коротких предложениях (например, поисковых запросах), может плохо распознавать длинные [16]. В качестве средства компенсации этого недостатка в модели Трансформера используется позиционное кодирование: текущий вектор параметров складывается с позиционным вектором, координаты которого кодируют позицию этого вектора параметров относительно других [17].

Сопоставление моделей сквозных сетей

Сравнение моделей сквозных систем, например, на основе показателя величины WER оказывается нетривиальным поскольку для получения достоверных результатов требуются большие датасеты размером в десятки тысяч часов. Поскольку таких открытых датасетов пока нет, сравнить в равных условиях модели затруднительно, но на небольшом для таких моделей корпусе LibriSpeech сравнения делались [5], причем с использованием одинаковой, внешней, стандартной для Libri-Speech 4-граммной языковой модели. Из представленных в табл. 3. результатов видно, что качество распознавания разных моделей для сквозных систем отличается, но не очень существенно. Модель кодера-декодера уступает на «test-other» данных, но для этой модели размеров корпуса LibriSpeech явно недостаточно.

Показатель пословной ошибки распознавания важен для оценки качества распознавания речи, также существенное значение имеют и полнота модели, все ли уровни системы распознавания в (1) учитываются моделью, возможность работы в реальном времени, объем корпусов данных, требования к вычислительным ресурсам.

С точки зрения полноты моделирования речевого сигнала очевидный недостаток модели СТС в том, что она не включает языковую модель. Кроме этого, моменты генерации выходных символов в СТС не соответствуют началам соответствующих сегментов сигнала, что важно для некоторых приложений.

Языковой модели нет и в сети Wave2Letter, но ее архитектура создана с расчетом на простую интеграцию внешней языковой модели [14].

Модель кодера-декодера лучше структурирована и в явном виде содержит рекуррентную модель языка. Обычно эта модель для языка букв или морфов, гибридного словаря из морфов и частотных слов. Наличие внешней модели языка для слов оказывается существенным (для системы [27] внешняя языковая модель дала примерно 0,8 % абсолютного уменьшения пословной ошибки).

Показанная экспериментально необходимость использования внешней по отношению к нейросети модели языка означает, что формально сквозные системы распознавания речи полностью таковыми не являются, поскольку основной признак: совместное обучение всех модулей сети как единого целого до конца не выполня-

Таблица 3. Значение показателя пословной ошибки распознавания для систем распознавания с разными типами моделей [5]. На корпусе Librispeech, с использованием внешней 4-граммной модели языка LibriSpeech

Тип Модели	Тип ЯМ	WER test-clean	WER test-other
Гибридная DNN-HMM	4-граммные, слова	5,51	13,97
CTC	4-граммные, слова	5,33	13,25
ASG	4-граммные, слова	4,80	14,50
Кодер-декодер с вниманием	4-граммные, слова	4,82	15,30

ется, модель языка обучается отдельно, также отдельно настраивается соотношение весов языковой и акустической моделей.

С точки зрения латентности распознавания речи модели СТС и кодера-декодера, которые используют двунаправленные клетки LSTM, по сути, не являются моделями реального времени: оценка активаций скрытых слоев предполагает, что известен весь сигнал, от начала до конца. Замена двунаправленных элементов на однонаправленные в этом случае ухудшает точность распознавания [27].

Тенденции развития сквозных систем

Преимущества описанных выше моделей сквозных систем по сравнению с гибридными HMM-DNN архитектурами начинают проявляться при использовании больших, по привычным представлениям, обучающих данных, которые необходимы для добывания аналогов экспертных знаний, широко используемых в гибридных системах. Поэтому для относительно небольших датасетов, таких как LibriSpeech (960 часов), гибридные системы в среднем на момент написания этого обзора демонстрируют лучшие результаты.

Увеличение размеров обучающих данных даже до экстремальных значений (известны результаты на датасетах размером 160 тысяч и даже 1 миллион часов речи [29, 30]) позволяет уменьшать уровень ошибок распознавания, причем в акустико-фоновых условиях, представляющих особый интерес для практических приложений (акцентная речь, использование разных каналов связи, наличие шума и реверберации). Пока существует возможность заметно улучшать результаты за счет увеличения размера обучающих данных, повидимому, будут использоваться существующие модели и методы.

В то же время проблемы со сбором больших корпусов данных и возможности их обрабатывать в разумные сроки вызывают интерес к более сложным и физически обоснованным нейросетевым моделям. Усложнение архитектуры моделей, увеличение времени их обучения, отсутствие гарантий результатов возвращают старые вопросы о границах применимости сквозных моделей и методов [31]. Методы оптимизации на основе градиентного спуска находят локальные экстремумы целевых функций, поэтому важны удачно выбранные начальные условия. Их современный выбор случайным образом выглядит неоптимальным. Способ выбора начальных условий с помощью глубокой доверительной сети [32] (deep belief network) для относительно простых сетей и при наличии больших данных оказался малоэффективным [33], но аналогичные методы могут оказаться необходимыми в случае использования сложных моделей. Для таких моделей возможно потребуется менять и методы обучения, переходя к более хорошо структурированным методам наподобие процедуры послойного обучения сетей в [5], которая оказалась выигрышнее даже в случае достаточно простой однородной сети.

В этом смысле показательным является продолжающееся использование внешних моделей языка: даже с учетом очень больших размеров обучающих акустических корпусов их текстовое содержание оказывается существенно меньше, чем объем специализированных текстовых корпусов данных, на которых обучается внешняя модель языка.

Заключение

Таким образом идея сквозных систем как одной большой нейросети, параметры которой оцениваются градиентным спуском все сразу и одновременно, пока остается неизменной, но сам подход к построению сети становится более физически обоснованным, что видно на примере моделей внимания, кодера-декодера и трансформера.

Развитие моделей сквозных систем также оказало влияние на совершенствование основного на сегодняшний день гибридного HMM-DNN подхода. Поскольку сквозные системы интегрируют в одной структуре уровни акустического и произносительного моделирования, перенос этого свойства в гибридные архитектуры, где нейросеть оценивает правдоподобия не состояний марковских моделей звуков речи, а непосредственно графем, также заметно упрощает архитектуру и приводит к лучшим пока результатам в тестах на корпусе LibriSpeech [35].

Сквозные системы распознавания речи имеют недолгую, но уже достаточно впечатляющую историю. На сегодняшний день эти системы не уступают лучшим гибридным системам распознавания по качеству распознавания речи, проигрывая пока по латентности распознавания. Судя по интенсивности исследований и результатам работ в этом направлении можно ожидать, что в ближайшее время эта технология станет стандартной для создания мощных продуктовых систем распознавания речи.

Литература

1. Jelinek F. Statistical Methods for Speech Recognition // Cambridge, Massachusetts The MIT Press, 1997.

2. Waibel A., Hanazawa T., Hinton G. at al. Phoneme Recognition Using Time-Delay Neural Networks. IEEE Transactions on Acoustics, Speech, and Signal Processing, vol. 37, no. 3, pp. 328-339, 1989.

3. S.J. Young, Adda-Decker M., Aubert X. at al. Multilingual large vocabulary speech recognition: the European SQUALE project, Computer Speech and Language, pp. 73-89, vol. 11, 1997.

4. Yu D. and Deng Li. Deep neural network-hidden markov model hybrid systems, in Automatic Speech Recognition. Springer, 2015, pp. 99-116.

5. Zeyer A., Irie K., Schluter R., Ney H. Improved training of end-to-end attention models for speech recognition // ArXiv:1805. 03294v1, 2018. URL: https://www.arxiv.org/pdf/ 1805.03294.

6. Povey D., Ghoshal A., Boulianne G. at al. The Kaldi Speech Recognition Toolkit. IEEE 2011 Workshop on Automatic Speech Recognition and Understanding, 2011.

7. Panayotov V., Chen G., Povey D., Khudanpur S. Librispeech: an ASR corpus based on public domain audio books. Proc. ICASSP-2015, pp. 5206-5210, 2015.

8. Were We Are. [Электронный ресурс] URL: https://github.com/syhw/were_are_we (дата обращения: 12.02.2020).

9. Amodei D., Anubhai R., Battenberg E. at al. Deep Speech 2:

End-to-End Speech Recognition in English and Mandarin // arXiv:1512.02595v1 [cs. CL], 2015. URL: http://arxiv.org/pdf/1512. 02595.pdf. (дата обращения: 12.02.2020).

10. Lüscher C., Beck E., Irie K., Kitza M. at al. RWTH ASR Systems for LibriSpeech: Hybrid vs Attention – w/o Data Augmentation // arXiv:1905.03072v3 [cs.CL]. URL: http://arxiv.org/pdf/1905.030 72.pdf (дата обращения: 12.02.2020).

11. Graves A., Fernandez S., Gomez F., Schmidhuber J. Connectionist Temporal Classification: Labelling Unsegmented Sequence Data with Recurrent Neural Networks. Proc. of the International Conference on Machine Learning, ICML 2006: pp. 369-376.

12. Graves, A., Jaitly N. Towards end-to-end speech recognition with recurrent neural networks. // Proc. International Conference on Machine Learning, ICML-2014, pp.1764-1772.

13. Graves A. Sequence Transduction with Recurrent Neural Networks // [Электронный ресурс]: arXiv: 1211. 3711v1 [cs.NE], 14 Nov 2012, URL: www.arxiv.org/pdf/ 1211.3711.pdf (дата обращения: 12.02.2020).

14. Collobert R., Puhrsch C., Synnaeve G. Wav2Letter: an End-to-End ConvNet-based Speech Recognition System // [Электронный ресурс]: arXiv:1609.03193v2 [cs. LG], 13 Sep 2016, URL: www.arxiv. org/pdf/1609.03193.pdf (дата обращения: 12.02.2020).

15. Bahdanau D., Cho K., and Bengio Y. Neural Machine Translation by Jointly Learning to Align and Translate.// International Conference on Learning Representations, 2015. [Электронный pecypc]: arXiv:1409.0473v7 [cs.CL], URL: www.arxiv.org/pdf/1409. 0473v7.pdf (дата обращения: 12.02.2020)

16. Chorowski, J.K.; Bahdanau, D.; Serdyuk, D. at al. Attention-Based Models for Speech Recognition // [Электронный ресурс]: ar-Xiv:1506.07503v1 [cs.CL], URL: www.arxiv.org/pdf/1506.075 03.pdf (дата обращения: 12.02.2020).

17. Vaswani A., Shazee N., Parmar N. at al. Attention is all you need // [Электронный ресурс]: arXiv:1706.03762v5 [cs. CL] 2017, URL: www.arxiv.org/pdf/1706.03762.pdf (дата обращения: 12.02.2020).

18. Watanabe S., Hori T., Karita S. at al. ESPnet: End-to-End Speech Processing Toolkit // [Электронный ресурс]: arXiv:1804. 00015v1 [cs.CL] 2018, URL: www.arxiv.org/pdf/1804.0001.pdf (дата обращения: 12.02.2020).

19. Miao Y., Gowayyed M., Metze F. at al. EESEN: End-to-End Speech Recognition using Deep RNN Models and WFST-based Decoding // [Электронный pecypc]: arXiv: 1507.08240v3 [cs. CL], 2015. URL: www.arxiv.org/pdf/ 1507.08240.pdf (дата обращения: 12.02.2020).

20. Graves A., Mohamed A.-R. and Hinton G. Speech Recognition with Deep Recurrent Neural Networks // [Электронный pecypc]: arXiv:1303.5778v1 [cs. NE], 2013. URL: www.arxiv.org/ pdf/1303.57 78.pdf (дата обращения: 12.02.20).

21. Dong L., Zhou S., Chen W., Xu B. Extending Recurrent Neural Aligner for Streaming End-to-End Speech. Recognition in Mandarin // [Электронный pecypc]: arXiv:1806.06342v2 [cs. SD] 2019, URL: www.arxiv.org/ pdf/1806.06342.pdf (дата обращения: 12.02.2020). 22. Zeghidour N., Xu Q., Liptchinsky V., at al. Fully Convolutional Speech Recognition // [Электронный pecypc]: arXiv: 1812. 06864v2 [cs.CL] 2019, URL: www.arxiv.org/pdf/ 1812.06864.pdf (дата обращения: 01.01.2020).

23. Pratap V., Hannun A., Xu Q. Wav2letter++: The Fastest Open-source Speech Recognition System // [Электронный ресурс]: ar-Xiv:1812.07625v1 [cs.CL], 2018. URL: www.arxiv.org/pdf/1812. 07625. pdf (дата обращения: 01.01.2020).

24. Sutskever I., Vinyals O. and Le Q. Sequence to sequence learning with neural networks // [Электронный ресурс]: arXiv:1409. 3215v3 [cs. CL], 2014. URL: www.arxiv.org/pdf/14 09.3215.pdf (дата обращения: 01.01.2020).

25. Chan W., Jaitly N., Quoc Q., Vinyals O. Listen, Attend and Spell // [Электронный pecypc]: arXiv:1508.012 11v2 [cs. CL], 2015. URL: www.arxiv.org/pdf/1508.01211.pdf (дата обращения: 01.01.2020).

26. Prabhavalkar R.P., Sainath T.N., Wu Y. at al. Minimum Word Error Rate Training for Attention-Based Sequence-to-Sequence Models // [Электронный pecypc]: arXiv: 1712.018 18v1 [cs.CL], 2017. URL: www.arxiv.org/ pdf/1712.01818.pdf (дата обращения: 12.02.2020).

27. Chiu C.C., Sainath T.N., Wu Y. at al. State-of-the-Art Speech Recognition With Sequence-to-Sequence Models // [Электронный pecypc]: arXiv:1712.01769v6 [cs.CL], 2018. URL: www.arxiv.org/pdf/ 1712.01769.pdf (дата обращения: 12.02.2020).

28. Synnaeve G., Xu Q., Kahn J. at al. End-to-End ASR: From Supervised to Semi-Supervised Learning with Modern Architectures // [Электронный ресурс]: arXiv:1911.08460v1 [cs.CL] 2019, URL: www.arxiv.org/pdf/1911.08460.pdf (дата обращения: 12.02.2020).

29. Narayanan, Misra A., Sim K.C. at al. Toward Domain-Invariant Speech Recognition Via Large Scale Training // [Электронный pecypc]: ArXiv:1808.05312v1 [cs.CL], 2018. URL: www.arxiv.org/pdf/1808.053 12.pdf (дата обра-щения: 12.02.2020).

30. Parthasarathi SHK., Strom N. Lessons from Building Acoustic Models with a Million Hours of Speech // [Электрон-ный ресурс]: ArXiv:1904.01624v1 [cs.LG] 2019. URL: www.arxiv.org/pdf/1904. 01624.pdf (дата обращения: 12.02.2020).

31. Glasmachers T. Limits of End-to-End Learning. Proceedings of Machine Learning Research, ACML, vol. 77, pp. 17-32, 2017.

32. Hinton G., Deng L., Yu D. at al. Deep Neural Networks for Acoustic Modeling in Speech Recognition: The Shared Views of Four Research Groups. IEEE Signal Processing Magazine. Vol. 29(6), pp. 82-97. – 2012.

33. Maas A.L., Peng Qi P., Xie Z, Hannun A.Y. at al. Building DNN Acoustic Models for Large Vocabulary Speech Recognition // [Электронный pecypc]: ArXiv:1406.7806v2 [cs.CL], 2015, URL: www.arxiv.org/pdf/1406. 7806.pdf (дата обращения: 12.02.2020).

34. Han K.J., Chandrashekaran A., Kim J., Lane I. The CAPIO 2017 Conversational Speech Recognition System // [Электронный pecypc]: arXiv:1801.00059v2 [cs. CL], 2018, URL: www.arxiv.org/ pdf/1801. 00059.pdf (дата обра-щения: 12.02.2020).

35. Wang Y., Mohamed A., Duc Le, Liu C., at al. Transformerbased Acoustic Modeling for Hybrid Speech Recognition // [Электронный ресурс]: ArXiv:1910.09799v1 [cs. CL], 2019. URL: www.arxiv.org/pdf/1910.09799.pdf (дата обращения: 12.02.20270). УДК 519.6

ПРИМЕНЕНИЕ ПОДХОДОВ, СВЯЗАННЫХ С МИНИМИЗАЦИЕЙ ДИСПЕРСИИ, В ЗАДАЧАХ КЛАССИФИКАЦИИ МНОГОМЕРНЫХ ДАННЫХ БИОМЕДИЦИНСКОЙ ПРИРОДЫ

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

Борзунов С.В., к.ф.-м.н., доцент кафедры цифровых технологий Воронежского государственного университета, e-mail: sborzunov@gmail.com;

Белобродский В.А., аспирант кафедры цифровых технологий Воронежского государственного университета, e-mail belobrodsky@yandex.ru.

DISPERSION MINIMIZATION APPROACHES IN BIOMEDICAL MULTIVARIABLE DATA QUALIFICATION

Turovskyi Ya.A., Borzunov S.V., Belobrodskyi V.A.

The focus of this article is an innovative method of digital data processing based on factor analysis, developed and tested to solve the task of data categorization using customized principal component analysis (PCA). The method searches the n-space for a new coordinate system, where component variance of a class is minimal. The result is that a data cluster is formed, in which probability density of a given class data is significantly higher than the one of the data belonging to other classes.

The accuracy of the method was verified by sequential enumeration of all possible angles of a data cloud rotation in a given coordinate system.

Key words: digital processing of signals, customized principal component analysis, biomedical signal classification.

Ключевые слова: цифровая обработка сигналов, адаптированный метод главных компонент, классификация биомедицинских сигналов.

Введение

Анализ значительного числа разнородных данных является одной из традиционных задач большинства научных исследований в самых разнообразных сферах деятельности [1-3]. Для её решения разработан широчайший арсенал методов, доказавших свою эффективность [4-10]. Тем не менее, совершенствование уже существующих и поиск новых методов является

одним из приоритетных направлений в области обработки результатов исследований, включая сигналы различной природы. Традиционные подходы классификации данных основаны на разделении наблюдений, относящихся к разным классам, т.е. выделении границ (возможно с использованием нечёткой логики), разделяющих в пространстве области, относящиеся к разным классам, и, одновременно, объединении наблюдений, расположенных «близко» в заданной, по сути, созданной под задачи исследования, метрике.

Целью работы является создание и апробация метода цифровой обработки данных на основе подходов минимизации дисперсии класса при использовании метода главных компонент.

Пусть исследуемая выборка представляет собой конечное множество объектов $\{x_i\}, i = 1, 2, ..., N$, каждый из которых представляется вектором вещественнозначных признаков x_i . Как хорошо известно, см., например, [11], в

Разработан и апробирован метод цифровой обработки данных на основе факторного анализа, обеспечивающий решение задачи классификации данных на основе адаптированного метода главных компонент. Суть метода заключается в поиске в п-мерном пространстве новой системы координат, дисперсия проекций данных одного из классов на ось которой будет минимальна. При этом формируется кластер, плотность вероятности нахождения наблюдений данного класса в котором существенно выше плотности вероятности нахождения в этом кластере данных из других классов. Произведена верификация метода при помощи альтернативного решения задачи методом последовательного перебора всех возможных углов поворота облака данных в существующей системе координат.

> случае успеха процедура классификации разбивает выборку на непересекающиеся группы (кластеры):

$$\{\boldsymbol{x}_j\} = \bigcup_{i=1}^{K} C_i, \forall i \neq (C_i \cap C_j = \emptyset)$$
(1)

При этом, анализируется вся исследуемая выборка, часть из которой используется для обучения классификатора, а часть – для кросс-проверки, обеспечивая избегание переобученности. Таким образом, по сути, можно сформулировать, что один из подходов к классификации сводится к тому, чтобы найти область (области) в уже существующем пространстве, или получить новые координаты в пространстве, в которых вероятность нахождения объектов одного класса будет статистически значимо больше, чем в других областях.

В отличие от широко распространённого подхода с разделением выборки, содержащей все исследуемые классы, на «обучающую» и «кросс-проверочную» под-

выборки, в рассматриваемом подходе будет использован иной алгоритм. Пусть из множества классов в массиве наблюдений имеется один, выделение которого значимо для целей и задач исследования. Или же имеется несколько групп, значимость нахождения конкретного наблюдения в каждой из которых, не равнозначна исходя из целей и задачей исследования. Рассмотрим единственный кластер С1, который в дальнейшем применим для обучения классификатора. Следовательно, задача сводится к преобразованию существующих координат наблюдений в новые, обеспечивающие выполнение условий минимизации дисперсии точек кластера С₁. Преобразуем координаты существующего пространства так, чтобы в новых координатах проекция интересующего нас класса на одну из новых осей была минимальной. Т.е. рассмотрим другой подход, задачей которого является максимизация числа наблюдений одного из классов на единицу новой оси пространства наблюдений (в общем случае гиперплоскости). Иными словами, необходимо так изменить исследуемые координаты переменных, чтобы проекции всех векторов одного класса на новую ось координат (гиперплоскость) имели бы минимальный вариационный размах:

$$\frac{1}{N} \sum_{i=1}^{N} (\mathbf{x}_{i}, \mathbf{w}_{k})^{2} \to \min \text{ для } k = 1, 2, ... n$$
(2)

где x_i – центрированные векторы данных (i = 1, ..., N), w_k – векторы, характеризующие искомые оси новой системы координат, n – размерность пространства векторов исходных данных.

Для верификации статистической значимости различия величин дисперсии будем использовать критерий Ливена, который, по сравнению с критерием Бартлетта, менее чувствителен к отличиям распределения выборки от нормального распределения. Согласно критерию Ливена для выборки $\{x_i\}, i = 1, 2, ..., N$, разделенной на классы $\{x^{(i)}_i\}, i = 1, 2, ..., k$ вычислим величину

$$W = \frac{N-k}{k-1} \frac{\sum_{i=1}^{k} N_{j} (\overline{Z}_{i\bullet} - \overline{Z}_{\bullet\bullet})^{2}}{\sum_{i=1}^{k} \sum_{j=1}^{N_{j}} (\overline{Z}_{ij} - \overline{Z}_{i\bullet})^{2}},$$
(3)

где $Z_{ij} = |(\mathbf{x}_{i}^{(j)}, \mathbf{w}_{1}) - M_{i}[(\mathbf{x}_{i}^{(j)}, \mathbf{w}_{1})]|, M_{i}[...], -$ среднее по элементам *i*-го класса, \overline{Z}_{i} – групповое среднее Z_{ij}, \overline{Z} – среднее по всей выборке. Гипотеза о равенстве дисперсий отвергается в случае $W > F_{\alpha,k-1,N-k}$, где через $F_{\alpha,k-1,N-k}$ обозначена верхняя граница *F*-распределения с *k*-1 и *N-k* степенями свободы на уровне значимости α .

При этом изменение координат векторов других классов ($C_{2..n}$) должно обеспечивать более низкую плотность вероятности их нахождения в области вариационного размаха переменной выбранной для обучения.

Таким образом, необходимо выполнение условия:

$$\frac{\sum_{i=1}^{N_j} (\boldsymbol{x}_i^{C_j}, \boldsymbol{w}_k)^2}{\left| \max(\boldsymbol{x}_i^{C_j}, \boldsymbol{w}_k) - \min(\boldsymbol{x}_i^{C_j}, \boldsymbol{w}_k) \right|} < A(C_j),$$
(4)

где $A(C_i) = \text{const.}$

Обозначим левую часть формулы (4) через $P(C_j)$ – это плотность точек класса C_j . Таким образом, вычисление этой величины методом градиентного спуска можно записать как:

$$\boldsymbol{x}_{i}^{C_{j},(k)} = \boldsymbol{x}_{i}^{C_{j},(k-1)} - \lambda^{(k-1)} \nabla P(C_{j}), \qquad (5)$$

где λ определяет скорость градиентного спуска.

Однако, возможно и применение иных методов для поиска новых координат. Как известно, примером преобразования, в результате которого могут быть изменены координаты в *N*-мерном пространстве тех или иных наблюдений, полученных экспериментально, может служить факторный анализ. Этот широко применяемый мощный метод анализа основан на поиске, т.н. латентных переменных – новых осей в уже существующем пространстве переменных, вдоль которых дисперсия облака данных максимальна [12, 13].

Следовательно, используя аппарат факторного анализа, можно получить данные о проекции облака данных на рассчитанные оси, при этом в ряде случаев дисперсия этих проекций будет последовательно увеличиваться:

$$0 < \lambda(w_1) < \lambda(w_2) < \ldots < \lambda(w_n).$$
(6)

Очевидно, что новые оси (факторы), являющиеся собственными векторами ковариационной матрицы исходных переменных формируют множество двухмерных плоскостей. При этом вектор ортогональный вектору, дисперсия проекции переменных на который максимальна, имеет, наоборот, минимальную проекцию в данной плоскости (см. (2)).

Не составляет труда найти из множества собственных векторов (факторов, латентных переменных) тот, дисперсия проекции на который является минимальной. При этом расчеты следует вести не по ковариационной матрице всех наблюдений, а только по ковариационной матрице того класса, который выбран для формирования новых координат. Пусть x_i – записанные по столбцам, образуют матрицу $X = (x_1, x_2, ..., x_N)$, и векторы, используемые для построения новых осей координат, имеют индексы j = p, p + 1, ..., N, где $1 . Ковариационная матрица <math>\tilde{\Sigma}$ системы векторов $\tilde{X} = (x_1, ..., x_{p-1})$

строится как

$$\tilde{\Sigma} = M[(x_l, x_m)]$$
 для $l, m \in [1, p-1],$ (7)

где *М* – математическое ожидание. Полученные собственные векторы этой матрицы формируют матрицу преобразования *А*, которая отражает преобразование исходного пространства в новое, так, что синхронно меняются и координаты всех переменных:

$$x' = A \cdot x$$
.

(8)

Лемма. Для точек кластера C_1 в декартовом пространстве R^2 перейдем к новой системе координат, оси которой Ox' и Oy' направлены вдоль собственных векторов w_1 и w_2 соответственно, причем этим векторам отвечают собственные значения $0 < \lambda(w_1) < \lambda(w_2)$. Тогда среднеквадратичное отклонение координат x' принимает минимальное значение среди всех возможных систем координат. Доказательство леммы следует из того факта, что в методе главных компонент при выборе собственных векторов в соответствии с убыванием отвечающих им собственных значений происходит максимизация суммы квадратов проекций на первую ось («главная компонента»). Выбор обратного порядка собственных значений приводит к минимизации дисперсии.

Обобщение леммы на случай произвольного арифметического векторного пространства R^N не составляет трудности: выбирается вектор $\lambda(w_i)$ с наименьшей (но, разумеется, большим нулю) собственным значением, причем остальные векторы из набора \widetilde{X} располагаются в произвольном порядке, например, в порядке возрастания $\lambda(w_i)$. Сформулированный алгоритм приводит к минимальному значению среднеквадратичного отклонения координат x' для требуемого класса данных.

Верификация подхода

Далее рассмотрим на практике поиск осей с минимальной дисперсией для объекта, представляющего собой три пересекающихся облака точек (см. рис. 1).



Рис. 1. Три облака точек на двумерной плоскости

Для верификации представленного алгоритма произведем два независимых друг от друга численных эксперимента (реализующие разные подходы) и сравним полученные результаты. В первом эксперименте применим подход, основанный на представленном выше алгоритме, то есть на формировании поворотной матрицы, составленной из собственных векторов ковариационной матрицы $\tilde{\Sigma}$ в порядке возрастания модуля

соответствующих собственных чисел. Во втором численном эксперименте возьмем те же самые входные



Рис. 2 а. Расположение «облаков», найденное с помощью представленного метода ($\sigma_X^2 = 92$, $\sigma_Y^2 = 648$)

данные (см. рис. 1) и реализуем следующий итерационный алгоритм:

– на каждой итерации цикла от 0 до 360⁰ с шагом в 1 градус будем производить:

 поворот 2-х мерного облака на текущий угол (от 0 до 360⁰) с помощью известной матрицы поворота (матрицы направляющих косинусов) [14]:

$$A(0) = \begin{pmatrix} \cos\theta & \mp\sin\theta \\ \pm\sin\theta & \cos\theta \end{pmatrix}$$

• вычисление дисперсии первой компоненты x_1 всего облака и запись пары (угол-дисперсия) в отдельный массив;

 после окончания итерационного процесса перебора всех возможных углов проанализируем полученные пары (угол-дисперсия) и найдем среди них угол, соответствующий минимальной дисперсии.

Исходя из взаимного расположения «облаков» (см. рис. 2 *а*, *б*) и соответствующих им числовых значений дисперсий можно сделать вывод о правильности работы алгоритма и его реализации.

При этом, как нетрудно заметить, дисперсия распределения остальных данных («облаков») в результате проекции на новую ось оказывается существенно выше, чем для выбранного класса (см. рис. 3 и рис. 4).

Заключение

Разработан новый метод классификации данных, основанный на адаптации метода главных компонент для задачи выделения ограниченного числа кластеров. Данный метод позволяет, за счёт изменения системы координат, выделить оси, проекции на которые дают минимальную дисперсию исследуемого класса. При этом, в случае успешного обучения, дисперсии остальных классов по критерию Ливена оказываются значимо выше, чем дисперсия требуемого класса. Отсюда следует, что плотность вероятности проекции данных исследуемого класса на новую ось оказывается статистически значимо выше, чем для других классов выборки. Данный метод был протестирован и с помощью альтернативного решения, основанного на итерационном переборе, и показано достижение сходных результатов.

Работа выполнена при поддержке РФФИ (грант № 19-07-01037 A).



Рис. 2 б. Расположение «облаков», найденное с помощью последовательного перебора всех возможных углов ($\sigma_X^2 = 92$, $\sigma_Y^2 = 648$, $\alpha_{\min} = 125^0$)



Литература

1. A. Ortega P. Frossard J. Kovacevi ^{*} c J. M. F. Moura and P. Vandergheynst. «Graph signal processing: Overview, challenges, and applications», Proc. IEEE, vol. 106, no. 5, pp. 808-828. May 2018.

2. Бериков В.Б. Выбор оптимальной сложности класса логических решающих функций в задачах анализа разнотипных данных: дис. д-р. тех. наук: 05.13.17. – Новосибирск, 2006. – 271 с.

3. Кручинин И.И. Моделирование и сравнительный анализ процессов распознавания и классификации многомерных объектов пересекающихся классов на основе представлений теории нечетких множеств и нейросетевых технологий: дис. канд. тех. наук: 05.13.17. – Калуга, 2003. – 209 с.

4. Sutha P., Jayanthi V. Fetal Electrocardiogram Extraction and Analysis Using Adaptive Noise Cancellation and Wavelet Transformation Techniques // J Med Syst. – 2018. № 21.

5. Белобродский В.А., Туровский Я.А., Вахтин А.А., Борзунов С.В., Кургалин С.Д. Обобщение метода цепочек локальных экстремумов для анализа сигналов различной природы // Цифровая обработка сигналов. – 2015. – № 1. – С. 35-38.

6. Пономарев А.В. Основы теории двумерной цифровой обработки сигналов в базисах Фурье с варьируе-



Рис. 3 б. Распределение проекции исходных данных на координатную ось «Y»



Рис. 4 б. Распределение проекции исходных данных на координатную ось «Y'»

мыми параметрами // Цифровая обработка сигналов. 2019. № 2. С. 12-20.

7. Утробин В.А. Методы обработки в условиях априорной неопределенности: дис. д-р. тех. наук: 05.13.17. – Нижний Новгород, 1997. – 410 с.

8. Коваленко А.П. Непараметрические методы анализа кластеров высокой плотности: дис. д-р. тех. наук: 05.13.17. – М., 1999. – 184 с.

9. Крестьянинова М.А. Применение методов контролируемой классификации для анализа биологических данных: дис. канд. физ.-мат. наук: 03.00.02: Москва, 2003. – 136 с.

10. M. Silveira, G. Figueiredo, R. Caputo New Time-Domain Approach for Digital Signal Processing: A Set of Experimental Measures for Systems with High Transmission Rates // Journal of Circuits, Systems and Computers. – 2019. Vol. 28 (05), P. 1950072.

11. Alpaydin E. Introduction to Machine Learning, MIT-Press, 2014. – 539 p.

12. Jolliffel. T. Principal Component Analysis, Series: Springer Series in Statistics, Springer, 2002. – 487 cτp.

13. Le Roux B., Rouanet H. Geometric Data Analysis: From Correspondence Analysis to Structured Data, Springer Science & Business Media, 2014, pp. 297-332.

14. Лурье А. И. Аналитическая механика. – М.: Физматлит. – 1961. – 824 с.

УДК 519.7, 520.6

ОПЕРАТИВНОЕ ОПРЕДЕЛЕНИЕ ИЗМЕНЕНИЙ НАБЛЮДАЕМОЙ СЦЕНЫ ЛОГИЧЕСКИМ ВЫЧИТАНИЕМ ЦИФРОВЫХ ИЗОБРАЖЕНИЙ

Котцов В.А., научный сотрудник Института космических исследований Российской академии наук (ИКИ РАН), e-mail: vladkott@mail.ru; Котцов П.В., инженер-программист ИКИ РАН, e-mail: kot_scorp@mail.ru.

RAPID DETECTION OF CHANGES IN THE OBSERVED SCENE BY LOGICAL SUBTRACTION OF DIGITAL IMAGES

Kottsov V.A., Kottsov P.V.

The possibility of operative determination of changes in the composition or state of objects observed on a pair of digital images by their parallel logical subtraction in the stream mode is shown.

Key words: operational comparison, differentiation, multi-channel observation, logical subtraction, stream processing.

Ключевые слова: оперативное сравнение, выделение различий, многоканальное наблюдение, логическое вычитание, потоковая обработка. Показана возможность оперативного определения изменений состава или состояния объектов, наблюдаемых на паре цифровых изображений, методом их параллельного логического вычитания в потоковом режиме.

Введение

Видеоинформационные системы сравнения изображений широко используются для наблюдения за изменением природной среды, для контроля качества продукции, обеспечения безопасности территории, выявления патологических изменений в организме, управлением движением, выявления фальсификации и многих других приложений. Многие процессы получения необходимой видеоинформации связаны со сравнением пары изображений, во многих случаях важна оперативность получаемого результата. Это могут быть снимки, полученные в разное время, при разных условиях, в разных ракурсах, и их информационное содержание определяется этими различиями. Широкое распространение видеоинформационных технологий, их использование в самых различных областях, постоянное увеличение объемов решаемых задач, делает необходимым повышение эффективности применяемых технологий. В том числе и повышения оперативности контроля изменения состояния наблюдаемых объектов или содержания сцен.

Известно множество решений подобных задач, которые используют в различных приложениях. Предлагается решение на уровне логических операциях с цифровым сигналом, которое обладает простыми технологическими особенностями, обеспечивая при этом высокое качество и оперативность получаемого результата.

Арифметическое использование логических операций для цифровых изображений

В операциях с цифровыми изображениями широко используют цифровые процессоры, обладающие большими вычислительными возможностями, но которые не всегда обеспечивают наиболее эффективный результат. Сравнение изображений может выполняться процессором на первый взгляд простой арифметической операцией вычитания. Однако, при ближайшем рассмотрении последовательности операций, ее выполнение компьютерными системами при сравнении сигналов оказывается не таким простым [1]. Известно также, что арифметика и логика являются двумя математическими областями, действия которых для бинарных сигналов дают эквивалентные решения. Это широко используется в структурах компьютерных подсистем [2].

Квантованное цифровое изображение можно представить как трехмерный многослойный массив, сформированный виде наложения набора слоев бинарных изображений, соответствующих каждому из выделенных уровней квантования. Это позволяет рассматривать составляющие цифровых изображений на каждом квантованном уровне, как логические сигналы. С этими логическими сигналами в каждом выделенном квантованном слое можно выполнять логические операции. Такое использование цифрового представления видеоинформации имеет свое преимущество.

Для прямого выполнения логических операций имеются соответствующие электронные компоненты, в то время как для выполнения арифметических операций приходится организовывать в процессорах различные, не всегда простые схемы. Кроме того, важной особенностью использования логических операций является их прямое выполнение в потоковом режиме. В то время, как при арифметическом выполнении процедур над цифровым сигналом необходимо предусматривать дополнительные операции для контроля переполнения, запоминания и переноса сигнала. При большом объеме вычислений это приводит к непредвиденным задержкам и неравномерности скорости обработки сигналов.

Как было показано нами ранее, применение логических операций позволяет создавать простые эффективные решения для реализации различных видеоинформационных систем [3]. В статье была описана возможность и эффективность выполнения операции сложения цифровых изображений логическими операциями. Поскольку суммирование логических сигналов выполняется в потоковом режиме, то такое техническое решение предложено использовать для оперативного увеличения динамического диапазона в сложных условиях наблюдения. Результат достигается путем сложения изображений, полученных на разных участках динамического диапазона. Однако, такой же подход может быть применен и при компиляции составного изображения из видеоинформации, получаемой в разных зонах спектра при дистанционном зондировании с целью повышения распознаваемости и других аналогичных задач, требующих сложения сигналов цифровых изображений.

Технология выполнения такого решения для операции сложения заключается в дополнении квантованных уровней одного изображения квантованными уровнями второго изображения, а затем выполнения логического суммирования бинарных значений между всеми парами квантованных уровней этой структуры операциями И и ИЛИ. Очевидно, что для получения окончательного результата необходимо выполнить сортировку размещения бинарных значений по уровням квантования. Схемное решение последовательности выполнения этих операций для сортировки, предложенное О.А. Ханджяном [4], обеспечивает выполнение необходимых процедур перемещения единичных значений сигналов на нижние уровни.



Рис. 1. Схема логического суммирования изображений на сдвоенных логических элементах И (черные) и ИЛИ (белые)

На рис. 1 показана описанная схема выполнения операций суммирования, организованная на логических элементах И (черные) и ИЛИ (белые), для двух цифровых изображений при четырех уровнях квантования.

В статье также приведен пример выполнения этой операции в эксперименте. В нем наглядно показано, что получаемое изображение обладает более высоким качеством, чем каждое из исходных изображений. Число одновременно складываемых изображений может быть несколько, что расширяет область применения. При этом число градаций увеличивается пропорционально числу складываемых изображений. В отличие от арифметического сложения, выполнение эквивалентной операции логическими функциями не требует при своем выполнении накоплений и переноса, что позволяет осуществлять ее выполнение оперативно, в потоковом режиме.

Сравнение цифровых изображений логическим вычитанием

Для оперативного определения изменений наблюдаемой сцены по последовательно полученным цифровым изображениям можно произвести их вычитание. Однако, выполнение арифметической операции вычитания цифровых изображений в процессоре имеет те же недостатки, что и сложение. Кроме того, для отображения полученного результата в виде изображения возникает дополнительная трудность из-за вероятности появления при вычитании отрицательных значений. Это, конечно, существенно затруднит анализ полученного результата. Чтобы выполнить сравнение двух изображений, оперируя только положительными числами, можно одно из них инвертировать, преобразовать в негативное, а затем сложить. В этом случае области с эквивалентными значениями сигналов будут нейтрализованы.

Известно, что для бинарных сигналов операция инвертирования производится выполнением логической операции НЕ. Следовательно, мы можем применить эту операцию для составляющих одного из сравниваемых изображений на каждом из его квантованных уровней. Таким образом, одно из сравниваемых изображений будет представлено в инвертированной форме. После этого выполняется описанная процедура логического суммирования двух изображений.



Рис. 2. Схема логического вычитания пары изображений на сдвоенных логических элементах И (черные) и ИЛИ (белые), с инвертированием на входе по уровням квантования одного изображения логическим элементом НЕ

На рис. 2 представлена описанная схема выполнения такой операции вычитания с инвертированием одного из сравниваемых изображений выполнением логической операции НЕ по уровням квантования и последующим сложением полученных данных известной схемой на сдвоенных логических элементах И и ИЛИ.

Очевидно, что для цифровых изображений работа со значениями на квантованных уровнях с помощью простых логических операций значительно упрощает обработку сигналов. Нет необходимости учитывать переполнение и перенос разряда, следовательно, нет промежуточного запоминания и переноса, то есть процедура обработки может выполняться в темпе поступления видеоинформации. Более того, скорость обработки данных даже увеличивается по сравнению с процессорной и остается постоянной. Сформированный в результате сравнения видеосигнал будет иметь число уровней квантования соответственно суммарному числу уровней сравниваемых изображений, что обеспечивает также потенциальное увеличение динамического диапазона, а, следовательно, и качество отображения.

Представленная технология осуществления способа сравнения цифровых изображений защищена патентом и была отмечена на Салоне изобретений Архимед-2019 в Сокольниках [5].

Экспериментальная проверка предложенного решения

Для проверки осуществимости поставленной задачи сравнения цифровых изображений и анализа особеннос-



C)

Рис. 3. Изображения наблюдаемой сцены: а) исходное состояние сцены; б) ее измененное состояние; с) результат их логического вычитания

тей получаемого результата было выполнено её моделирование. Для получения исходных данных цифровой камерой с одной точки наблюдения было получено два последовательных изображения одной и той же сцены интерьера, при этом перед получением второго кадра один из наблюдаемых предметов (ваза) был удален. После загрузки изображений в программную модель схемы, показанной на рис. 2, на её выходе было получено изображение с выделенными изменениями наблюдаемой сцены между двумя кадрами.

На рис. З а показано первое изображение, отображающее исходное состояние, при котором присутствуют все детали сцены, на рис. 3 б - второе изображение, где из состава предметов удалена только ваза с цветами. На рис. 3 с показано изображение, полученное логическим вычитанием исходных изображений, которое в эксперименте было выполнено программным путем в соответствии со схемой, представленной на рис. 2. На полученном изображении отчетливо наблюдается удаленный объект на нейтральном фоне. Можно также заметить, что наблюдается выявленное изменение фазы вращения маятника часов между кадрами. Кроме того, случайное затенение небольшого левого участка циферблата часов на первом снимке, проявилось на разностном изображении в виде негативного изображения этого фрагмента. При том, что детали выделенных на разностном изображении изменений достаточно хорошо проработаны. Стоит отметить также, что выявленные случайные изменения сцены, не предусмотренные постановкой эксперимента, при визуальном анализе сравниваемых исходных изображений были не очевидны для наблюдателя.

Как показала экспериментальная проверка предложенного решения, логическое вычитание можно эффективно использовать в задачах, требующих детального сравнения изменений в последовательности цифровых изображений. Технологически целесообразным будет использовать для выполнения этой процедуры специализированной микросхемы, например, на ПЛМ.

Заключение

Операцию сравнения цифровых изображений можно эффективно выполнять на основе использования только логических операций с применением описанного схемного решения для логического вычитания. Получаемое изображение будет содержать увеличенное число града-

ций соответственно суммарному числу градаций во взаимно дополняющих наблюдениях. Такое изображение позволяет выявить детали, которые не обнаруживаются при простом наблюдении последовательности кадров. Технологическая реализация осуществляется проще, чем при выполнении арифметических операций при цифровых вычислениях процессорами [6]. Его структура может быть легко реализована аппаратно, например, на программируемых логических элементах и может функционировать в темпе поступления видеоинформации.

Предложенная схема решения относится к классу потоковых устройств обработки информации. Их отличительной особенностью является возможность выполнения преобразования сигналов в условиях непрерывного изменения наблюдаемой сцены. Технология параллельной обработки видеосигналов обеспечивает высокую эффективность предложенного способа. Применение для решения только логических операций сближает его с процедурами, которые выполняются в живых организмах. Простота и быстрота логических методов сравнения изображений обеспечивают их высокую эффективность. Как известно, в природных биологических организмах именно этот принцип позволяет создавать рекордные по быстродействию и сложности системы [7]. Описанное техническое решение может также служить примером возможного подхода к построению арифметических узлов компьютеров следующего поколения.

Литература

1. Бруфман С.С. Цифровые элементы сравнения. М.: Энергия, 1967.

2. Поспелов Д.А. Логические методы анализа и синтеза схем. - М: Энергия, 1968.

3. Котцов В.А. Увеличение динамического диапазона видеосистемы логическим сложением цифровых изображений // Цифровая обработка сигналов № 3, 2019.

4. Ханджян О.А. Линейная фильтрация, основанная на теории симметрических функций // Радиотехника и электроника, вып. 8. 1986.

5. Котцов В.А., Котцов П.В. Способ сравнения цифровых изображений. Патент РФ 2673396 // Бюллетень изобретений № 33, 2018.

6. Карцев М.А. Арифметика цифровых машин. - М.: Наука, 1969.

7. Бергсон А. Творческая эволюция. – М.: ТЕРРА-Книжный клуб, 2001.

УДК 621.391

ОБЕСПЕЧЕНИЕ РАЗРЕШЕННОГО ОБРАЗА ПРИ ИНВЕРСНОЙ ФИЛЬТРАЦИИ СИГНАЛОВ В УСЛОВИЯХ НЕОПРЕДЕЛЕННОСТИ

Хафизов Р.Г., д.т.н., профессор, профессор кафедры радиотехнических и медико-биологических систем Поволжского государственного технологического университета, Йошкар-Ола, e-mail: HafizovRG@volgatech.net.

PROVIDING A RESOLVED IMAGE WITH INVERSE FILTERING OF SIGNALS IN CONDITIONS OF UNCERTAINTY

Khafizov R.G.

An approach is proposed to ensure a resolved image at the output of an inverse filter, specified by a finite impulse response, during cyclic processing of signals under conditions of uncertainty. It is shown that the cyclic signal processing by an inverse filter allows to obtain a zero level of side lobes at the output. The approach to eliminating the uncertainty caused by the presence of zero components in the signal spectrum is based on a change of the signal dimension. A method is proposed for minimizing the level of fluctuation noise at the output of the inverse filter by minimizing the squared norm of the filter impulse response.

Key words: inverse filter, cyclic signal procession, complex signal, signals resolving.

Ключевые слова: инверсный фильтр, циклическая обработка сигналов, комплекснозначный сигнал, разрешение сигналов.

Введение

Инверсная фильтрация сигналов является эффективным способом подавления корреляционного шума [1]. Анализ работ по инверсной фильтрации позволяет выделить два вида неопределенностей при синтезе и анализе инверсного фильтра: неопределенность, вызванная наличием нулевых компонент в спектре сиг-

нала, и неопределенность, обусловленная воздействием флуктуационного шума на входе фильтра [2-8]. Если в первом случае решаемыми проблемами являются проблемы физической реализуемости и подавления корреляционного шума на выходе фильтра, то во втором случае к указанным проблемам добавляется еще и проблема подавления флуктуационного шума на выходе фильтра. При этом задача устранения неопределенности при инверсной фильтрации решается как в частотной области путем ограничения спектра анализируемого сигнала [3, 4, 6-8], так и во временной области, например, методом деконволюции [2, 5].

Одним из методов обработки сигналов является циклическая обработка [9, 10]. Циклическая обработка сигналов дает возможность произвести расчет конечной импульсной характеристики (КИХ) фильтра и непосредственно использовать ее для реализации алгоритма фильтрации. Целью данной работы является исследование и разработка подхода к обеспечению разрешенного образа на выходе инверсного фильтра в условиях неопределенности при циклической обработке сигналов. При этом анализируемые сигналы $s(n) = \{s(n)\}, n = 0, 1, ..., N-1, и КИХ фильтра <math>\Lambda = \{\lambda(n)\}, n = 0, 1, ..., N-1, предполагаются комплекснозначными, т.е. <math>s(n = s_1(n) + is_2(n)$ и $\lambda(n) = \lambda_1(n) + i\lambda_2(n)$, где N – размерность сигнала и КИХ фильтра.

Предложен подход к обеспечению разрешенного образа на выходе инверсного фильтра, заданного конечной импульсной характеристикой, при циклической обработке сигналов в условиях неопределенности. Показано, что циклическая обработка сигналов инверсным фильтром позволяет получить на выходе нулевой уровень боковых лепестков. Подход к устранению неопределенности, вызванной наличием нулевых компонент в спектре сигнала, основан на изменении размерности сигнала. Предложен способ минимизации уровня флуктуационного шума на выходе инверсного фильтра путем минимизации квадрата нормы импульсной характеристики фильтра.

Расчет КИХ инверсного фильтра

Выражение для сигнала на выходе линейного фильтра $\mathbf{H} = \{\eta(m)\}$, заданного КИХ $\mathbf{\Lambda} = \{\lambda(n)\}$, при обработке входного сигнала $\mathbf{s} = \{s(n)\}$ [9]:

$$\eta(m) = \sum_{n=0}^{N-1} s(n)\lambda(m-n) \, .$$

Для случая инверсной фильтрации сигнала на выходе фильтра:

$$\eta(m) = \sum_{n=0}^{N-1} s(n)\lambda(m-n) = \mathbf{K}(m) ,$$

где K(m) – символ Кронекера. Данное выражение может быть записано в матричной форме:

$$S \Lambda = K$$
,

где, с учетом циклической обработки:

$$\mathbf{S} = \begin{bmatrix} s(0) & s(1) & \dots & s(N-1) \\ s(1) & s(2) & \dots & s(0) \\ \dots & \dots & \dots & \dots \\ s(N-1) & s(0) & \dots & s(N-2) \end{bmatrix},$$
$$\mathbf{\Lambda} = \begin{bmatrix} \lambda(0) \\ \lambda(1) \\ \dots \\ \lambda(N-1) \end{bmatrix}, \ \mathbf{K} = \begin{bmatrix} 1 \\ 0 \\ \dots \\ 0 \end{bmatrix}.$$

(2)

Отсчеты КИХ $\lambda(n)$ инверсного фильтра могут быть найдены в соответствии с правилом Крамера [11]:

$$\lambda(n) = \frac{\Delta_n}{\Delta},$$

где Δ – определитель матрицы S, Δ_n – определитель матрицы, полученной путем замены *n*-го столбца матрицы S матрицей-столбцом K.

Циклическая обработка инверсным фильтром с КИХ позволяет получить нулевой уровень боковых лепестков. Уровень флуктуационного шума σ_{ebix}^2 на выходе инверсного фильтра определяется [9]:

$$\sigma_{\rm sbix}^2 = R_{\rm sbix}(0) = 2\sigma_{\rm sx}^2 \left\| \mathbf{\Lambda} \right\|^2, \tag{1}$$

где $R_{eblx}(0)$ – нулевой отсчет корреляционной функции на выходе инверсного фильтра, $\|\Lambda\|$ – норма КИХ фильтра, σ_{ex}^2 – дисперсия составляющих комплекснозначного случайного процесса $\xi(n) = \xi_1(n) + i\xi_2(n)$, $n = 0, 1, \dots, N-1$, на входе фильтра. При этом $\sigma_{ex}^2 = \sigma_{\xi_1}^2 = \sigma_{\xi_2}^2$. Как показано в работе [9], комплекснозначный случайный процесс на выходе фильтра коррелирован, а мощность выходного шума пропорциональна квадрату нормы КИХ фильтра.

Рассмотрим пример расчета КИХ инверсного фильтра для сигнала $\mathbf{s} = \{s(n)\}, n = 0, 1, ..., 4$ (рис. 1,*a*): $\mathbf{s} = \{0, 3i; 1, 2 + 0, 7i; 0, 9-i; -0, 7; -0, 6 + 0, 5i\}$. Получаем:

0.9 - i0.3i1.2 + 0.7i-0,7-0,6+0,5i1,2+0,7i0,9-i-0,70,3*i* -0,6+0,5i0,9-i-0,7-0,6+0,5i1,2+0,7iS =0,3*i* -0,7-0,6+0,5i0,3*i* 1, 2 + 0, 7i0.9 - i-0,6+0,5i0,3*i* 1,2+0,7i0,9-i-0,7

 Δ = -8,03-1,984*i*; Δ_0 = -1,626 - 2,215*i*; Δ_1 = -3,207 + 2,241*i*; Δ_2 = -2,053 - 1,752*i*; Δ_3 = -2,528 + 2,86*i*; Δ_4 = 1,082 + 1,594*i*. КИХ инверсного фильтра (рис. 1, *б*):

 $\mathbf{\Lambda} = \{0, 255 + 0, 213i; 0, 311 - 0, 356i; 0, 292 +$

+0,146i; 0,214-0,409i; -0,173-0,156i.

Результат фильтрации: H = {1; 0; 0; 0; 0} (рис. 1, е).

Рассмотрим применение инверсного фильтра с КИХ для решения задачи разрешения сигналов. Пусть входной сигнал **u**_{ex} образован путем суммирования сигнала *s* и задержанной на 3 отсчета копии сигнала *s* с масштабом 4. Таким образом, поучаем сигнал:

$$\mathbf{u}_{ex} = \{0, 3i; 1, 2+0, 7i; 0, 9-i; -0, 7+$$

+1,2i; 4,2+3,3i; 3,6-4i; -2,8; -2,4+2i.

Пусть также положение сигнала не известно. Произведем ациклическую фильтрацию сигнала с КИХ (2). Результат фильтрации (рис. 2, *a*):

 $\mathbf{H}_{au} = \{0, 047 - 0, 052i; 0, 024 - 0, 244i; 0, 187 - 0, 22i; \}$

0,446-0,208i; 1,095-0,976i; 0,703-0,83i;

1,014+0,245i; 3,813+0,22i; -0,446+0,208i;

-0,095+0,976i; -0,75+0,882i; -1,038.







Рис. 2. Результат фильтрации: а – ациклическая, б – циклическая

На рис. 2, а наблюдается отклик от мощного сигнала, а отклик от слабого сигнала маскируется боковыми лепестками. В работе [9] предложен алгоритм ЧКШС, применение которого к отсчетам ациклической свертки дает результаты, аналогичные результатам циклической свертки. В соответствии с алгоритмом ЧКШС, необходимо сложить отсчеты ациклической свертки с интервалом в *N* отсчетов, т.е.:

$$\begin{split} \eta_{u}(n) &= \eta_{au}(n) + \eta_{au}(n+) + \eta_{au}(n+2N) + \dots, n = 0, 1, \dots, N-1. \\ \text{Для рассматриваемого примера получаем (рис. 2, б):} \\ \eta_{u}(0) &= \eta_{au}(0) + \eta_{au}(5) + \eta_{au}(10) = \\ &= 0,047 - 0,052i + 0,703 - 0,83i - 0,75 + 0,882i = 0; \\ \eta_{u}(1) &= \eta_{au}(1) + \eta_{au}(6) + \eta_{au}(11) = \\ &= 0,024 - 0,244i + 1,014 + 0,245i - 1,038 = 0; \\ \eta_{u}(2) &= \eta_{au}(2) + \eta_{au}(7) = 0,187 - 0,22i + 3,813 + 0,22i = 4; \\ \eta_{u}(3) &= \eta_{au}(3) + \eta_{au}(8) = 0,446 - 0,208i - 0,446 + 0,208i = 0; \\ \eta_{u}(4) &= \eta_{au}(4) + \eta_{au}(9) = 1,095 - 0,976i - 0,095 + 0,976i = 1. \\ \text{ На рис. 2, 6 можно наблюдать два отклика с соот-} \end{split}$$

на рис. 2, о можно наолюдать два отклика с соответствующими амплитудами. Положение на дистанции первого сигнала формируется в момент времени N-1 (в нашем случае это в момент n = 4). Положение на дистанции последующих сигналов относительно первого определяется:

 $t_s = \operatorname{mod}(m_s + 1, N), \tag{3}$

где t_s – положение *s*-го сигнала на дистанции, m_s – номер *s*-го отклика. При m_1 = 4 получаем t_1 = 0, а при m_2 = 2 – t_2 = 3, что соответствует условиям примера. Таким образом, циклическая обработка сигналов инверсным фильтром с КИХ позволяет разрешать сигналы и обеспечивает нулевой уровень боковых лепестков.

Если спектр $\mathbf{P} = \{\rho(m)\}, m = 0, 1, ..., N-1$, сигнала s содержит равные нулю компоненты, то определитель Δ матрицы S становится равным нулю и вычисление ИХ инверсного фильтра становится невозможным. Например, пусть задан сигнал:

 $\mathbf{S} = \{0,075 + 0,534i; 1 + 0,844i; 0,702 - 1,146i; -0,623 - 0,234i; -0,354 + 0,501i\},$ (4) CHERTP KOTOPOFO: $\mathbf{S} = \{0,075 + 0,0234i; -0,054i; -1,002i - 1,146i; -0,012i - 1,002i - 1,002i$

 $\mathbf{P} = \{0,8+0,499i; 0; 0,646-0,516i; -1,492-1,445i; 0,421+4,133i\},\$

содержит равную нулю компоненту $\rho(2) = 0$. Определитель Δ матрицы S равен нулю и расчет КИХ инверсного фильтра становится проблематичным.



Устранение неопределенности методом интерполяции спектра

В работе [12] рассмотрен метод интерполяции спектра сигнала добавлением нулей. При этом, как указано в [12], дополнение нулями не улучшает разрешающую способность этого преобразования, а позволяет получить интерполированное преобразование более сглаженной формы. В спектре дополненного нулями сигнала образуются компоненты, находящиеся между компонентами спектра исходного, т.е. не дополненного нулями, сигнала (рис. 3).

Пусть модификация сигнала s производится дополнением нулей так, что размерность сигнала s_м становится равной *M*:

$$\mathbf{s}_{MOO} = \left\{ s(0), s(1), \dots, s(N-1), \underbrace{0, 0, \dots, 0}_{M-N} \right\}.$$
 (5)

Тогда спектр $\mathbf{P}_{Mod} = \{\rho_{Mod}(m)\}, m = 0, 1, ..., M-1, моди$ $фицированного сигнала <math>\mathbf{s}_{M}$ равен:

$$\rho_{MO\partial}(m) = \sum_{n=0}^{M-1} s_{MO\partial}(n) e^{\frac{-i2\pi mn}{M}} = \sum_{n=0}^{N-1} s(n) e^{\frac{-i2\pi mn}{M}}$$

Если величина M кратна N, то наблюдается совпадение гармоник спектров исходного и модифицированного сигналов (рис. 3, *a*). Если же величина M не кратна N, то при интерполяции спектра сигнала в его составе не будут равные нулю компоненты (рис. 3, *б*).

Таким образом, если спектр сигнала содержит равную нулю компоненту и определитель Δ матрицы S также равен нулю, то требуется изменить интервал определителя сигнала, например, изменив размерность сигнала путем добавления нулей.

Запишем сигнал (4) в виде:

 $\mathbf{P}_{M} = \{0, 8 + 0, 499i; 0, 329 - 0, 268i; 0, 935 + 0, 249i; 0, 935 + 0, 249i; 0, 935 +$

0,046 - 0,721i; -3,379 + 0,152i; 1,719 + 3,293i

не содержит равных нулю компонент. КИХ инверсного фильтра для модифицированного сигнала:

 $\Lambda = \{0,607 + 0,299i; 0,453 - 0,629i; 0,229 - 0,003i; -0,073 - 0,579i; -0,343 + 0,115i; 0,026 + 0,236i\},$ (7) а результат фильтрации: **H** = {1; 0; 0; 0; 0; 0}. Квадрат нор-



Рис. 3. Интерполяция спектра сигнала с размерностью N = 5 при дополнении нулями до размерности: a) M = 10, 6) M = 7

мы КИХ такого фильтра равен 1,638. Тогда в соответствии с выражением (1) уровень шума на выходе фильтра будет определяться соотношением: $\sigma_{\rm sblx}^2 = 3,276 \, \sigma_{\rm scx}^2$.

Импульсная характеристика инверсного фильтра изменяется в зависимости от количества n_0 добавляемых нулей в сигнал. При этом суммы отсчетов сигнала \mathbf{s}_{M} и КИХ Λ фильтра независимо от количества n_0 добавляемых нулей в сигнал остаются постоянными. Для рассматриваемого примера:

$$\sum_{n=0}^{M-1} \lambda(n) = 0.9 - 0.561i, \quad \sum_{n=0}^{M-1} s_M(n) = 0.8 + 0.499i.$$

Кроме того, сумма отсчетов КИХ Λ фильтра равна величине обратной сумме отсчетов сигнала s_{M} :

$$\sum_{n=0}^{M-1} \lambda(n) = \frac{1}{\sum_{n=0}^{M-1} s_M(n)} = \frac{1}{0,8+0,499i} = 0,9-0,561i.$$

Квадрат нормы ИХ также зависит от количества добавляемых нулей в сигнал. На рис. 4 представлена полученная экспериментальным путем зависимость квадрат нормы КИХ инверсного фильтра от количества добавляемых в сигнал (4) нулей n_0 .



Рис. 4. Зависимость квадрат нормы КИХ инверсного фильтра от количества добавляемых нулей n₀

При $n_0 = 5, 10, 15$ и т.д. значение квадрат нормы КИХ инверсного фильтра стремится к бесконечности, т.к. размерность M модифицированного сигнала s_M (5) становится кратной размерности N сигнала s (4), и опреде-



литель Δ матрицы $S_{\scriptscriptstyle M}$ становится равным нулю. Получено, что для сигнала (4) минимальный уровень флуктуационного шума на выходе инверсного фильтра достигается при добавлении в сигнал двух нулей. Уровень шума на выходе фильтра при этом будет определяться соотношением: $\sigma_{\rm gbx}^2 = 2,014 \, \sigma_{\rm gx}^2$.

Рассмотрим пример разрешения сигналов инверсным фильтром с использованием метода интерполяции спектра. Пусть теперь входной сигнал \mathbf{u}_{ex} образован путем суммирования сигнала s (4) и задержанной на 1 отсчет копии сигнала 4s:

 $\mathbf{u}_{ex} = \{0,075 + 0,534i; 1,3 + 2,98i; 4,702 + 2,23i; 2,185 - 4,818i; -2,846 - 0,435i; -1,416 + 2,004i\}.$

Сигнал s (4) содержит в своем спектре равную нулю гармонику, поэтому КИХ фильтра будем формировать по модифицированному сигналу (6). Результат ациклической фильтрации сигнала с КИХ (7) (рис. 5, *a*): $\mathbf{H}_{ay} = \{-0,757 + 0,211i; -0,887 + 0,215i; 0,979 - 0,682i;$

1,459 - 0,794*i*; 1; 4,124 - 0,032*i*; 0,757 - 0,211*i*; 0,887 - 0,215*i*; -0,979 + 0,682*i*; -1,459 + 0,794*i*}.

На рис. 5, а наблюдается отклик от мощного сигнала, а отклик от слабого сигнала маскируется боковыми лепестками. Вычислим с помощью алгоритма ЧКШС результат циклической фильтрации. Размерность сигнала s_{M} увеличилась по сравнению с сигналом s, поэтому складываем отсчеты ациклической свертки с интервалом в M отсчетов. Получаем (рис. 5, $\boldsymbol{6}$):

 $\mathbf{H}_{u} = \{0; 0; 0; 0; 1; 4, 124 - 0, 032i\}.$

На рис. 5, *б* уже наблюдаются два отклика в моменты времени, определяемые выражением (3), *t*₁ = 0 и *t*₂ = 1, что соответствует условиям примера. Таким образом, циклическая обработка сигналов инверсным фильтром с КИХ, сформированной по модифицированному сигналу, позволяет разрешать сигналы.

Заключение

Рассмотрен подход к обеспечению разрешенного образа на выходе инверсного фильтра, заданного конечной импульсной характеристикой, при циклической обработке сигналов в условиях неопределенности, вызванной наличием нулевых компонент в спектре сигнала и воздействием флуктуационного шума на входе



Рис. 5. Результат фильтрации: а – ациклическая, б – циклическая

фильтра. Показано, что циклическая обработка сигналов инверсным фильтром с КИХ позволяет получить на выходе нулевой уровень боковых лепестков. При этом уровень флуктуационного шума на выходе инверсного фильтра зависит от нормы импульсной характеристики. Для формирования отчетов циклической свертки входного сигнала и импульсной характеристики фильтра по результатам ациклической свертки предложено использовать алгоритм ЧКШС.

Предложен подход к устранению неопределенности, вызванной наличием нулевых компонент в спектре сигнала, основанный на изменении размерности сигнала путем добавления нулей. В спектре дополненного нулями сигнала образуются компоненты, находящиеся между компонентами спектра исходного, т.е. не дополненного нулями, сигнала. При этом квадрат нормы ИХ зависит от количества добавляемых нулей в сигнал. Предложен подход к минимизации уровня флуктуационного шума на выходе инверсного фильтра путем минимизации квадрата нормы ИХ фильтра подбором количества добавляемых нулей в сигнал

Рассмотрен пример разрешения сигналов инверсным фильтром с использованием метода интерполяции спектра. Показано, что циклическая обработка сигналов инверсным фильтром с КИХ, сформированной по модифицированному сигналу, позволяет разрешать сигналы.

Литература

1. Василенко Г.И. Голографическое опознавание образов. – М.: Сов. радио, 1977.

2. Schneider M., Habets E.A.P. Iterative DFT-Domain Inverse Filter Optimization Using a Weighted Least-Squares Criterion // IEEE/ACM Transactions on Audio, Speech, and Language Processing. 2019. Vol. 27, № 12. Pp. 1957-1969.

3. Zhang Yo. et al. Super-resolution surface mapping

for scanning radar: inverse filtering based on the fast iterative adaptive approach // IEEE transactions on geoscience and remote sensing. 2018. Vol. 56. №. 1. Pp. 127-144. DOI: 10.1109/TGRS.2017.2743263.

4. Mudukutore A.S., Chandrasekar V., Keeler R.J. Pul-se compression for weather radars // IEEE Transactions on Geoscience and Remote Sensing. 1998. Vol. 36, № 1. Pp. 125-142.

5. Nelander A. Processing for continuous radar waveforms. 2004 International Waveform Diversity & Design Conference. Edinburgh, 2004. Pp. 1-5. DOI: 10.1109/IWDDC.20 04.8317557.

6. Абраменков В.В., Васильченко О.В., Семченков С.М., Печенев Е.А. Инверсная фильтрация импульсных сигналов // Электромагнитные волны и электронные системы. 2017. № 4. С. 42-53.

7. Семченков С.М., Печенев Е.А. Способ повышения разрешающей способности за счет инверсной фильтрации импульсных сигналов // Радиопромышленность. 2017. № 3. С. 103-109.

8. Гонсалес Р., Вудс Р. Цифровая обработка изображений. – М.: Техносфера. 2005.

9. Введение в контурный анализ и его приложения к обработке изображений и сигналов / Под ред. Я.А. Фурмана. – М.: Физматлит, 2002.

10. Хафизов Р.Г., Охотников С.А. Линейная фильтрация непрерывных контуров изображений, заданных в комплекснозначном виде // Компьютерная оптика. – 2010. – Т. 34, № 3, – С. 408-416.

11. Корн Г. А., Корн Т. М. Справочник по математике для научных работников и инженеров. М.: Наука. Главная редакция Физико-математической литературы; Издание 4-е, 1977.

12. Марпл С.Л. Цифровой спектральный анализ и его приложения. М.: Мир, 1990.

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 – 250 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. Журнал «Цифровая обработка сигналов» включен в базу данных Web of Science - Russian Science Citation Index (287-я позиция).

Планируемые сроки издания отдельных номеров журнала:

№ 2 август 2020 г. Тематический выпуск: «Цифровая обработка изображений».

№ 3 октябрь 2020 г. Тематический выпуск по материалам 22-й Международной научно-технической конференции «Цифровая обработка сигналов и ее применение-DSPA».

№ 4 декабрь 2020 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

ПАМЯТИ ВИКТОРА ПАВЛОВИЧА ДВОРКОВИЧА (18.08.1938 – 17.04.2020)



С глубоким прискорбием сообщаем, что скоропостижно ушел из жизни Виктор Павлович Дворкович – профессор, доктор технических наук, крупный специалист и ученый в области телевидения, цифровой обработки и передачи сигналов.

В 1955 г. В.П. Дворкович с серебряной медалью окончил среднюю школу в Таганроге, а в 1960 году с отличием – Таганрогский радиотехнический институт и был направлен на работу на одно из предприятий ВПК в Куйбышеве, где занимался разработкой транзисторных усилителей.

В 1962 г. видным отечественным ученым М.И. Кривошеевым он был приглашен на работу и в аспирантуру в телевизионный отдел Научно-исследовательского института радио (НИИР, Москва), в котором проработал 45 лет. В.П. Дворкович многие годы активно участвовал в создании выпускаемой в промышленных масштабах измерительной аппаратуры для отечественных систем радиорелейной и спутниковой связи, по которым

передавались телевизионные (TB) сигналы. Он был одним из ведущих сотрудников этого отдела и по выдвинутым им идеям под его руководством выполнялись важные разработки приборов для измерения параметров TB каналов. Эти приборы были внедрены в производство и выпущены в последние десятилетия XX века отечественной промышленностью в количестве более 500 тысяч экз.

Будучи сотрудником НИИР, В.П. Дворкович успешно защитил в 1967 г. кандидатскую диссертацию, а в 1990 г. – докторскую диссертацию по теме «Анализ и разработка новых методов оценки качественных показателей телевизионного канала».

В 1992 г. в НИИР им был создан новый Научно-технический отдел цифровой обработки информации и метрологии, в котором выполнялись научные исследования и велась разработка аппаратуры, предназначенной для повышения эффективности обработки и компрессии статистических и динамических цветных видеоизображений. Результатом этих исследований, к которым он привлек своего сына А.В. Дворковича, ставшего позже также видным ученым и чл.-корреспондентом РАН, стала оригинальная система, предназначенная для кодирования изображений и их передачи по каналам связи с ограниченной пропускной способностью.

Виктор Павлович и его сын Александр активно участвовали в создании в НИИР важной отечественной системы «Приморка». Эта система работает и сегодня, являясь составной частью первичной сети связи и передачи данных наземного автоматизированного комплекса управления космическими аппаратами Федерального космического агентства. За создание этой системы В.П. Дворкович, в составе творческого коллектива, был удостоен в 2004 г. премии Правительства РФ в области науки и техники.

С 2008 г. В.П. Дворкович – заместитель директора по науке Главного радиочастотного центра РФ, где было создано Управление научных исследований и разработки цифровых систем передачи информации. Основной разработкой этого этапа в научной деятельности В.П. Дворковича явилась Российская АудиоВизуальная Информационная Система вещания (РАВИС). Она позволяет в одном канале шириной 250 кГц передавать до 20 программ стереофонического звукового вещания высокого качества или видеопрограмму, которую можно смотреть на небольшом экране в движущемся автомобиле в условиях плотной городской застройки. Эта система вошла в число важнейших современных систем цифрового звукового и мультимедийного вещания, рекомендованных Международным Союзов Электросвязи для развертывания в наземных сетях вещания.

Технология РАВИС основана на использовании 8-ми патентов РФ, материалов Международной и отечественной стандартизации (стандарт МСЭ, 5 стандартов РФ). В процессе разработки недавно были проведены успешные испытания РАВИС и натурное эфирное вещание в Москве, Сочи, Казани и Ижевске.

В.П. Дворкович внес большой вклад в создание в 1999 г. одного из ведущих отечественных журналов в области современных информационных технологий – «Цифровая обработка сигналов». Большое внимание он уделял распространению новых научных знаний среди специалистов и в течение 20 лет являлся одним из организаторов Международной научно-технической конференции «Цифровая обработка сигналов и ее приложения», на которых рассматривался широкий круг научных вопросов, связанных с развитием современных технологий телекоммуникаций.

За период свой научной деятельности В.П. Дворкович опубликовал более 250 научных работ, среди них 15 книг. Им получено более 70 авторских свидетельств и патентов РФ. В начале 2020 года вышла в свет книга «Теория, практика и метрология аудиовизуальных систем», посвященная 75 годовщине Великой Победы, памяти участников и жертв войны. Данная монография является объединением и дополнением к ранее изданным ими книгам «Цифровые видеоинформационные системы (теория и практика)» (2012), «Измерения в видеоинформационных системах (теория и практика)» (2015), «Оконные функции для гармонического анализа сигналов» (2016) и др. Эти книги являются настольным учебником многих разработчиков и специалистов, аспирантов и студентов технических ВУЗов.

Большое внимание В.П. Дворкович уделял педагогической деятельности в ВУЗах, где читал лекции студентам и аспирантам по избранной специальности. На протяжении многих лет он работал профессором кафедры радиоэлектронных систем и устройств МГТУ им. Н.Э. Баумана. В 2010 г. ему было присвоено ученое звание профессора. Он внес значительную лепту в подготовку высококвалифицированных специалистов по радиотехнике и телевидению. В последние годы он плодотворно трудился заведующим кафедрой мультимедийных технологий и телекоммуникаций МФТИ, где под его руководством подготавливались магистры и высококвалифицированные специалисты.

В.П. Дворкович активно участвовал в научной общественной жизни как член Президиума РНТОРЭС им. А.С. Попова и руководитель секции «Обработка и передача изображений». С 1995 г. он был членом экспертного совета ВАК по электронике, измерительной технике, радиотехники и связи. Как видный ученый, В.П. Дворкович являлся действительным членом Международной академии информатизации и Российской медико-технической академии, был членом диссертационного совета ЗАО «МНИТИ», членом экспертных советов Российского фонда фундаментальных исследований и Российского научного фонда.

Виктор Павлович Дворкович был добрым и скромным человеком, хорошим семьянином и талантливым ученым. Он был созидателем – им были созданы измерительные приборы для контроля качества передачи ТВ сигналов по отечественным радиорелейным и спутниковым линиям связи, оригинальная система цифрового вещания РАВИС, а также выполнен ряд других разработок в области телевидения и метрологии. Светлая память о Викторе Павловиче навсегда останется в памяти и сердцах знавших его людей, специалистов, ученых, его учеников и сотрудников.

Редакция журнала