4/2011

ЦИФРОВАЯ ОБРАБОТКА СИГНАЛОВ

НАУЧНО - ТЕХНИЧЕСКИЙ ЖУРНАЛ

КОМПЛЕКСНОЕ РЕШЕНИЕ ПРОБЛЕМ ИНФОРМАЦИОННОЙ БЕЗОПАСНОСТИ:
разработка, изготовление и поставка средств защиты информации, защищенных информационных телекоммуникационных систем и комплексов

60 лет

Федеральному Государственному Унитарному Предприятию «Научно-исследовательский институт автоматики»

ISSN 1684-2634

ЦИФРОВАЯ ОБРАБОТКА СИГНАЛОВ

Научно-технический журнал
№ 1/2012

Издается с 1999 года
Выходит четыре раза в год

ГЛАВНЫЙ РЕДАКТОР
Ю.Б. ЗУБАРЕВ

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

РЕДАКЦИОННАЯ КОЛЛЕГИЯ:
Б.А. Алпатов, Б.А. Бабаян, Ю.И. Борисов,
Ю.А. Брюханов, С.А. Грибачев, Г.В. Зайцев,
Р.В. Зубарев, А.П. Кирпичников, М.И. Кривошеев,
Н.А. Кузнецов, В.Г. Мистюков, С.Л. Мишенков,
А.А. Петровский, Ю.Н. Прохоров,
Ю.Г. Сосулин, В.В. Шахгильдян, Ю.С. Шинаков

Адрес редакции:
107031 Москва, Рождественка, 6/9/20, стр. 1
Тел.: (495) 621-06-10, (495) 621-16-39
Факс: (495) 621-16-39
E-mail: nto.popov@mtu-net.ru
vityazev.v.v.@rsreu.ru
http://www.dspa.ru

Для писем:
129090 Москва, а/я 48.

Издатель:
Российское научно-техническое общество
радиотехники, электроники и связи им. А.С. Попова
Компьютерная верстка: И.А. Благодарова
Дизайн: Н.С. Осипова

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

Подписано в печать 30.01.12 г.
Формат 60x90/8.
Гарнитура «Arial». Печать офсетная.
Бумага офсетная. Печ.л. 6,5. Тираж 700 экз.

Заказ № 0394. 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

УЧРЕДИТЕЛИ:

ЗАО «Инструментальные системы»
ФГУП «НИИ радио»

В НОМЕРЕ:

Букашкин С.А.
История развития шифрующей аппаратуры в России……..2

Глинченко А.С., Комаров В.А., Тронин О.А.
Исследование случайных погрешностей спектрально-весового измерения параметров сигналов……7

Кравченко А.Н.
Методы и аппаратура кодирования и декодирования низкоплотностных квазициклических кодов……12

Овинников А.А., Бакке А.В.
Сравнительный анализ алгоритмов частотного детектирования в системах с OFDM……16

Соловьева Е.Б.
Синтез компенсаторов на основе нейронной сети Гаммерштейна……19

Новоселов С.А., Топников А.И., Савватин А.И., Приоров А.Л.
Подавление шума в речевых сигналах на основе метода нелокального усреднения……23

Мокеев В.В.
О повышение эффективности вычислений главных компонент в задачах анализа изображений……29

Бартенев В.Г., Битюков В.К., Кузеленкова Е.Г.
Инновационный курс «Программируемая радиоэлектроника» для инженерных вузов……37

Броневич А.Г., Каркищенко А.Н., Уманский В.И.
Статистический метод восстановления профиля по данным лазерного сканирования……42

Линович А.Ю.
Методы многоскоростной обработки сигналов в многоканальных системах с самоорганизующейся структурой: эквалайзеры с динамическим выбором порядка адаптивных фильтров……50

Гудкова Н.В., Гудков В.И.
Адаптивная цифровая линеаризация передаточных характеристик электронных усилительных устройств……59

Бегишев С.В.
Мониторинг спектральных изменений во время испытаний на виброустойчивость……62

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

ИСТОРИЯ РАЗВИТИЯ ШИФРУЮЩЕЙ АППАРАТУРЫ В РОССИИ

Появление средств электросвязи во второй половине XIX века и их бурное развитие в XX веке открыло перед человечеством новые возможности для повышения оперативности связи и поставило ряд трудных и принципиально новых задач в разработке аппаратуры скрытия передаваемой информации. Использование ручных шифров и машин предварительного шифрования, особенно при ведении переговоров по каналу связи, становится явно нецелесообразным. Нужно было привести время шифрования на передаче и приеме в соответствие с практически мгновенной скоростью передачи информации по каналам связи.

5 мая 1921 года декретом Совнаркома создается специальная служба для решения общих вопросов организации в стране секретной связи, разработки государственных шифров, рекомендаций по их использованию и осуществления контроля над правильным их применением.

Телефонная связь является наиболее оперативной и удобной для абонентов и дает возможность с малой затратой времени обсуждать и решить необходимые для них вопросы.

Проблема линейного засекречивания телеграфа представляла скорее техническую, чем теоретическую проблему, так как относительно небольшие скорости передачи легко могли быть реализованы на стандартном канале связи, а теория ручных шифров и машин предварительного шифрования создала хорошую базу для разработки линейного шифратора.

Засекречивание телефонной информации с её «аналоговой» формой требовало решения более сложных проблем, чем в случае с телеграфной информацией.

За начальную точку отсчета в развитии секретной телефонии следует принять 1930 г., когда под руководством молодого инженера Владимира Александровича Котельникова небольшая группа разработчиков - связистов создала новую аппаратуру, обеспечивающую засекречивание телеграфных и телефонных передач на коротковолновой линии связи.

Молодой В.А. Котельников, будущий вице-президент Академии наук, лауреат Ленинской и Государственных премий, председатель Верховного Совета РСФСР, дважды Герой Социалистического Труда, стал отцом секретной телефонии в нашей стране.

Чуть позднее, в 1931 г., в области секретной телефонии уже работало 7 организаций. Среди них НИИ Народного Комиссариата Почт и Телеграфа (НКПиТ), НИИИС Красной Армии НКО, завод им. Коминтерна, завод «Красная Заря», НИИ связи и телемеханики ВМФ, НИИ № 20 Народного Комиссариата Электротехнической Промышленности, подразделения НКВД. Еще позднее, в 1938 и 1939 гг. в ЦНИИ связи НКПиТ были организованы две лаборатории под руководством В.А. Котельникова по засекречиванию телеграфной и телефонной информации.

С началом Великой Отечественной Войны сотрудники завода «Красная Заря» были эвакуированы в г. Уфа и входили в состав Государственного союзного производственно-экспериментального института (ГСПЭИ 56), где продолжала результативно работать. Лаборатория В.А. Котельникова была разделена на две части: основная часть с ним самим была эвакуирована в ГСПЭИ 56, а другая часть была передана в НКВД СССР.

В первые послевоенные годы коллектив специалистов, работавший в ГСПИ 56, был практически расформирован. Большинство сотрудников завода «Красная Заря» возвратились в г. Ленинград, сотрудники лаборатории В.А. Котельникова были переведены в отдел правительственной связи (ОПС) МГБ СССР, сам он стал работать в МЭИ.

Одновременно с созданием аппаратуры засекречивания были начаты работы по ее дешифрованию. В 6 управлении МГБ была создана группа под руководством А.П. Петерсона, которая одновременно с работами по созданию аппаратуры шифрования занялась вопросами дешифрования. С середины 1944 г. группа А.П. Петерсона начала «сокрушать» одну аппаратуру за другой. В итоге в 28.12.1945 г. был составлен отчет, в котором постулировалось, что аналоговая аппаратура шифрования мозаичного типа теоретически дешифруема. В 1947 г. в лаборатории секретной телефонии ОПС МГБ группой под руководством К.Ф. Калачева было создано автоматическое устройство дешифрования вслед разговора. В разработке аппаратуры засекречивания телефонии наступил кризис. Для того, чтобы получить недешифруемую аппаратуру засекречивания телефонных переговоров, речь необходимо сначала перевести в цифровую форму, а затем шифровать. Было предложено использовать для этих целей вокодер. Руководство МГБ СССР, очевидно, понимало не только важность, но и сложность проблемы, возникшей в секретной телефонии. Необходимо было собрать всех имеющихся специалистов, обеспечить их всем необходимым и поручить создание аппаратуры с гарантией недешифруемости. Руководство МГБ решило поручить эту работу отделу оперативной техники (ОТТ) МГБ СССР.

Рис.1. Главный корпус Марфинской лабаратории

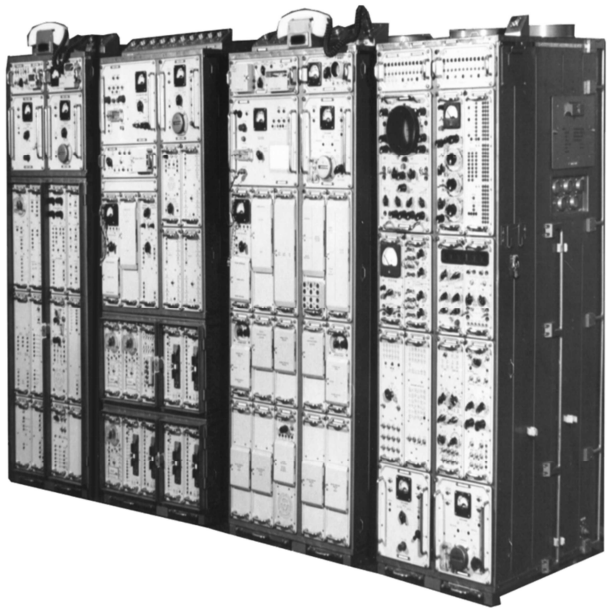


Рис.2. Аппаратура шифрования с гарантированной стойкостью телефонной информации и передачи её по каналам тональной частоты и коротковолновым каналам: год выпуска 1970; масса 1080 кг; потребляемая мощность 2,1 кВт

Было подготовлено постановление Совета Министров СССР, которое вышло 21 января 1948 г., в соответствии с которым в поселке Марфино под Москвой была организована спецлаборатория в составе ОТТ по разработке засекречивающей аппаратуры для ВЧ связи. В связи с тем, что ОТТ в своей работе использовал труд заключенных, это наложило свой отпечаток на работу спецлаборатории. Работа лаборатории, правда в достаточно негативном аспекте, отражена в книге А.И. Солженицына «В круге первом», Л.З. Копелева «Утоли мои печали» и наиболее объективно в книге К.Ф.Калачева «В круге третьем».

Использование труда заключенных было, конечно же, ошибкой, которая была исправлена через четыре года – постановлением Совмина СССР от 12.01.52 г. в Главном управлении Специальной Службы при ЦК ВКПб создается НИИ-2, ныне ФГУП «НИИ автоматики».

Работы по созданию аппаратуры шифрования начали проводиться все более широким фронтом в интересах правительственной связи, и, в том числе для Вооруженных Сил.

Первая серийная аппаратура засекреченной связи на новом предприятии была создана уже в 1954 году и установлена на самых протяженных в то время в мире линиях связи: Москва – Берлин, Москва – Пекин.

За прошедшие 60 лет коллективом «НИИ автоматики» создано несколько поколений технических средств и систем шифрованной связи. Это обеспечило решение многообразных задач засекреченной связи, боевого и государственного управления в стране.

За работы по созданию вышеуказанных средств и систем лауреатами Ленинских и Государственных премий стали 38 сотрудников института, более 400 человек награждены орденами и медалями.

В 1978 году за достижения и области науки и техники «НИИ автоматики» был награжден Орденом Ленина.

В 50–60 годы в НИИА было создано первое поколение отечественной шифраппаратуры на полупроводниковых и ферритовых элементах: М-803-5, «Лиана»,

«Алмаз», «Коралл», «Эльбрус», «Яхта», «Булава» и другие.

В 79-80 годах создаются специальные комплексы технических средств засекреченной связи и управления на базе микросхем малой и средней степени интеграции: «Кавказ», «Роса», «Интерьер».

Наряду с традиционным направлением засекречивания телефонии создается новый класс аппаратов по защите цифровой и телекодовой информации: «Стрекоза», «Лань», «Газель», «Грим», «Алибек», «Чегет», «Ветер», «Штиль», «Шторм», «Шквал», «Дуб», «Багет» и другие.

Разработка аппаратуры комплексов и систем засекреченной связи всегда опиралась на результаты фундаментальных теоретических исследований в области криптографии, проводимых институтом совместно с Институтом криптографии, связи и информатики, Академии криптографии РФ и ведущими академическими вузами страны.

С развитием информационных технологий появилась потребность в засекречивании передачи данных, факсимильной информации, видеоизображений, телеметрии и т.д. Со временем коллектив НИИ-2 стал «захлебываться» от возникающих новых задач. Для его разгрузки приказом Госкомитета по радиоэлектронике при СМ СССР от 18.01.1958 г. был создан НИИ-3. В настоящее время это ФГУП «Пензенский научно-исследовательский электротехнический институт», которому были поручены вопросы создания шифротехники для телекодовой и телеграфной информации в интересах различных заказчиков.

Новый всплеск активности по созданию образцов шифротехники пришелся на вторую половину 80-х годов прошлого столетия. Это связано с началом широкого применения персональных компьютеров и развитием компьютерных сетей. В настоящее время более 40 предприятий имеют лицензии на право разработки и изготовления шифротехники различного назначения.

Коллектив ФГУП «НИИ автоматики» по инициативе первого заместителя директора (с февраля 1999 г. - директора) С.А. Букашкина развертывает работы по освоению новой технологии проектирования аппаратуры на базе импортных процессоров цифровой обработки сигналов (DSP) фирмы Texas Instruments (США) и программируемых логических матриц фирмы Xilinx (США). Это происходит в условиях возросшей конкуренции и возрастаниям требований заказчиков к тактико-техническим характеристикам аппаратуры. Работа проводилась в тесном сотрудничестве с Научно-Исследовательским Центром (НИЦ) Федерального агентства правительственной связи и информации.

Результат не замедлил сказаться. В рекордные по старым меркам сроки удалось создать ряд принципиально новых изделий, которые по своим тактико-экономическим показателям намного превосходили предыдущее поколение шифротехники. К таким изделиям относится аппаратура Е-16, Е-9У, Е-11, «Медео» и особенно «Гамма». Демонстрация их на выставке «Связь-95» показала весьма большой интерес к ним со стороны ведущих зарубежных фирм, таких как Motorola, Siemens, NEC и многих других. Это свидетельствует о том, что удалось в короткие сроки создать продукцию, во многом соответствующую мировому уровню.

В 1995 г. первый заместитель Генерального директора РГНПО «Автоматика» д.т.н., профессор, академик АК РФ, академик МАИ Букашкин С.А. делает доклад на секции №6 «Радиоэлектроника» НТС Госкомоборонпрома по вопросу: «О состоянии и перспективах развития техники защиты информации». НТС принимает решение: «Считать правильными и безальтернативными направления работ в РГНПО «Автоматика» по созданию перспективной техники защиты информации. Это реализовано на аппаратно-программных принципах с использованием современной высокоинтегрированной элементной базы отечественного и импортного производства, а также сетей специальной связи с использованием импортных компонентов (коммутаторы, терминальное оборудование и т.д.), так как только таким способом возможно быстрое создание специальных сетей связи современного по мировым оценкам уровня обслуживания».

Дальнейшее развитие событий показало правильность выбранного направления работ.

В 1998 году проходят государственные испытания аппаратов Е-16 (рис.3) и «Гамма» (рис 4).

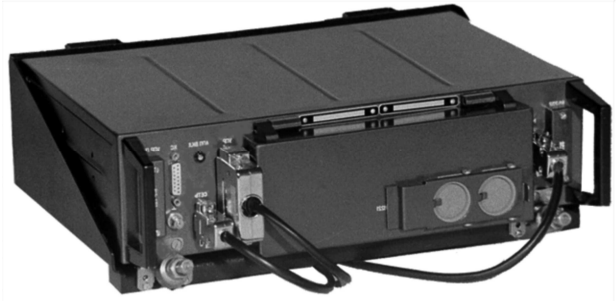


Рис.3. Аппарат Е-16

Малогабаритный аппарат для автоматического шифрования речевой, факсимильной и документальной информации в сети подвижной радиосвязи в дуплексном режиме, передаваемой со скоростью 2,4/4, 8/9, 6 Кбит/с по УКВ и ТЧ каналам

Аппарат «Гамма-М» является стандартным телефонным аппаратом с интегрированным в него цифровым шифратором. Аппарат предназначен для шифрования телефонной, факсимильной и документальной информации в дуплексном режиме с гарантированной стойкостью.

Рис.4. Аппарат «Гамма-М»

Аппарат имеет три дуплексных режима связи: режим открытой связи (работает как стандартный телефонный аппарат); режим шифрования речи; режим

шифрования данных (факсимильные аппараты и ПЭВМ подключаются по стандартному интерфейсу V.24 / V.28 (RS-232-C)). Аппарат обеспечивает работу через телефонную коммутированную сеть общего пользования по 2-х проводной линии в соответствии с рекомендациями ITU-T V32bis.

В 2002 г. успешно завершаются государственные испытания системы радиоподвижной связи «Кавказ-9», в 2004 г. комплекса аппаратуры «Исполком» (рис. 5)

Рис.5. «Аппарат - 1И»

Основной многоканальный аппарат шифрования комплекса «Исполком» для сетей засекреченной связи силовых ведомств.

Обеспечивает шифрование информации на скоростях от 1,2 до 64 Кбит/с и образование от 1 до 42 основных шифрованных каналов на скорости 1,2 Кбит/с. Позволяет образовывать каналы шифрованной связи на скоростях от 2,4 до 32 Кбит/с в различных (программируемых) комбинациях

В 2006 году завершается модернизация аппарата Е-16. Результаты работы по реализации решения НТС Госкомоборонпрома 1995 года наглядно иллюстрируются рис. 6 «Эволюция». На рисунке приведены фотографии аппаратуры аналогичного назначения разработки 1979, 1998 и 2005 годов.

Изменения в подходах и принципах создания новой техники, а также требования рыночной экономики потребовали быстрой реализации программы технического перевооружения лабораторий и основных производственных участков, решения проблем кадрового обновления и структурных преобразования в институте.

Все это позволили только за последние годы разработать и передать в производство нескольких десятков образцов качественно новой аппаратуры. На базе новых технических средств, и полной мере соответствующих международным стандартам, созданы и введены в эксплуатацию защищенные сети и системы связи. Качественно новую технику получили Вооруженные силы России, представители других силовых структур.

Научно-техническая общественность и руководство ФГУП «НИИ автоматики» понимают, что в современных условиях, при усложняющихся информационно-телекоммуникационных технологиях, проблемы безопасности управления государством и Вооруженными Силами только путем создания аппаратуры шифрования решить невозможно. Проблема обеспечения безопасности – проблема комплексная и может решаться только комплексом мер при создании автоматизированных систем управления и систем связи. Именно поэтому, по инициативе директора ФГУП «НИИ автоматики» в институте развернуты работы, напрямую не связанные с шифртехникой, но непосредственно касающиеся вопросов обеспечения безопасности в современных сетях связи и системах управления.

Современный этап развития общества характеризуется возрастающей ролью информационной сферы, которая, являясь системообразующим фактором жизни общества, активно влияет на состояние политической, экономической, оборонной и других составляющих безопасности Российской Федерации. Решения Президента Российской Федерации, Правительства Российской Федерации, направленные на ускоренное внедрение инновационных технологий в информационно-телекоммуникационное обеспечение деятельности органов государственной власти, обороны, безопасности и правопорядка, определяют значительный рост объемов обрабатываемой в технических средствах и передаваемой по каналам связи информации ограниченного доступа, в том числе, содержащей сведения, составляющие государственную тайну. Вместе с тем продолжает эксплуатироваться значительное количество ранее созданных ведомственных информационно-телекоммуникационных систем и автоматизированных систем управления в интересах обеспечения обороны, безопасности и правопорядка.

В соответствии с Концепцией национальной безопасности Российской Федерации, утвержденной Указом Президента Российской Федерации от 17 декабря 1997 г. № 1300, одной из важнейших задач являются защита государственного информационного ресурса, и, прежде всего, в федеральных органах государственной власти и на предприятиях оборонного комплекса. Основой для обеспечения гарантированной защиты информации является применение средств шифровальной техники, средств криптографической защиты информации ограниченного доступа, не содержащей сведения, составляющие государственную тайну.

На сегодня ФГУП «Научно-исследовательский институт автоматики» – головной исполнитель работ по одной из наиболее значимых Федеральных целевых программ Российской Федерации, направленной на создание и развитие информационно-телекоммуникационной системы специального назначения в интересах органов государственной власти.

Аппараты, комплексы средств шифрованной связи, защищенные сети и информационно-телекоммуникационной системы, разработанные и изготовленные в ФГУП «НИИА», установлены в Кремле, Доме Правительства, Государственной Думе, Счетной палате, Таможенном комитете, Центральном банке РФ, министерствах, пунктах управления ФСО, МО РФ. Они надежно работают на космических аппаратах, самолетах, кораблях, подводных лодках, автомобилях, бронетранспортерах и т.д.

Услугами правительственной связи пользуются должностные лица всех ветвей и уровней государственной власти. Рядом с Президентом Российской Федерации постоянно находится одно из изделий нашего института, известное всему миру как «ядерный чемоданчик».

Рис.6 «Эволюция»

В начале 90-х годов начали бурно развиваться телекоммуникационные технологии. Главным направлением развития стал переход на асинхронные сети связи с коммутацией пакетов. Стремительно развивается Интернет и связанные с ним технологии. Появляется новый класс сетей связи – мобильные сети связи. Стало ясно, что эра разработки отдельных шифровальных аппаратов закончилась – началась эра разработки криптозащищенных сетей связи, опирающихся на современные телекоммуникационные технологии.

Изменение технической политики потребовало привлечения молодых специалистов, владеющих современными технологиями систем связи и разработки РЭА. Имея в виду трудности привлечения талантливых молодых специалистов на работу в режимное предприятие Москвы, директор принял решение об организации филиалов ФГУП «НИИА» на базе Рязанского государственного радиотехнического университета и Санкт-Петербургского государственного университета. Решение оправдало себя: сотрудники филиалов принимают активное и плодотворное участие в новейших разработках.

В ФГУП «НИИА» работает базовая кафедра Московского института радиоэлектроники и автоматики (технический университет МИРЭА), ведется подготовка аспирантов и докторантов.

За последнее пятилетие проведена коренная модернизация опытного производства. Установленные импортные станки и автоматические линии повышенной точности и производительности позволяют выпускать изделия, соответствующие современному мировому уровню.

Свое шестидесятилетие ФГУП «НИИА» встречает рядом успешных разработок. К ним относятся:

– высокоскоростной криптомаршрутизатор «Никель-4», позволяющий создавать защищенные сети для передачи всех видов информации;

– аппаратные машины современных полевых сетей связи «Переселенец-5» (аппаратные П-244И и П-242И);

– станции обеспечения безопасности различных криптозащищенных сетей связи;

– подсистема мобильной (сотовой) защищенной сети связи «Сапфир»;

– телефон шифрованной связи по сети Internet «Круиз» и многое другое.

Разработан мобильный комплекс шифрованной связи для выездных мероприятий высшего руководства страны, который вобрал в себя лучшие достижения шифраторостроения. Комплекс получил высокую оценку Президента РФ и сопровождает его во всех поездках по стране и за рубеж. На базе основного аппарата комплекса созданы десятки модификаций, предназначенных для организации прямых линий для связи с руководителями иностранных государств (аппарат Е-45). Аппарат протестирован специалистами США и рекомендован ими для оснащения прямой линии Москва – Вашингтон.

Не забывает ФГУП «НИИА» и о коммерческих потребителях. Большая часть коммерческих проектов сориентирована на массового потребителя, на их применяемость в крупных корпоративных структурах. Более подробную информацию о коммерческой продукции можно найти на сайте по адресу http://niia.ru/

В целях сохранения и развития научно–производственного и технологического потенциала предприятий по производству систем и средств шифровальной техники, а также информационно–телекоммуникационных систем специального назначения в интересах обороноспособности, безопасности государства, Указом Президента РФ от 16.10.2010 г. № 1261 ФГУП «Научно-исследовательский институт автоматики» преобразуется в открытое акционерное общество «Концерн «Автоматика», 100% акций которого находится в федеральной собственности. В эту интегрированную структуру войдут, после преобразования в открытые акционерные общества, еще 7 предприятий: Башкирское ПО «Прогресс», завод «Калугаприбор», Калужский электромеханический завод, НПО «Сигнал» (г. С-Петербург), Пензенский научно-исследовательс-кий Электротехнический институт, ПО «Электроприбор» (г. Пенза) и Уфимский завод микроэлектроники «Магнетрон».

Концентрация основных разработчиков и изготовителей шифротехники и оборудования телекоммуникационных систем специального назначения в единый комплекс даст, несомненно, новые возможности по ускорению разработки и освоению в производстве современных средств комплексной защиты информации в Вооруженных силах РФ, структурах государственной власти и ключевых отраслях экономики Российской Федерации.

Директор ФГУП «НИИ автоматики»,
академик Академии криптографии РФ,
д.т.н., профессор С.А. Букашкин

УДК 621.317.332

ИССЛЕДОВАНИЕ СЛУЧАЙНЫХ ПОГРЕШНОСТЕЙ СПЕКТРАЛЬНО-ВЕСОВОГО ИЗМЕРЕНИЯ ПАРАМЕТРОВ СИГНАЛОВ

Глинченко А.С., *к.т.н., профессор кафедры «Радиоэлектронные системы» Сибирского федерального университета, e-mail: AGlinchenko@sfu-kras.ru;*
Комаров В.А., *к.т.н., доцент кафедры «Приборостроение и наноэлектроника», Институт инженерной физики и радиоэлектроники Сибирского федерального университета, e-mail: VKomarov@sfu-kras.ru;*
Тронин О.А., *ст. преподаватель кафедры «Радиоэлектронные системы» Сибирского федерального университета, e-mail: toa12@yandex.ru*

Ключевые слова: спектральная оценка, гармонические сигналы, частота, фаза, методические погрешности, методика исследований, коррекция, усреднение.

Определены и сопоставлены случайные погрешности, соответствующие разным способам спектрально-весового измерения параметров гармонических сигналов, а также возможные способы усреднения, обеспечивающие их уменьшение, которые реализованы в измерителе параметров сигналов на основе ПЭВМ.

Введение

В работах [1, 2], посвященных спектральным оценкам параметров вещественного гармонического сигнала (частоты, амплитуды, начальной фазы), основное внимание уделено анализу и способам уменьшения их методических погрешностей, связанных с наложением (спектральной утечкой) и интерполяцией. В данной работе рассмотрены случайные погрешности этих оценок, вызываемые шумом и априорной неизвестностью (случайностью) начальной фазы сигнала, без использования и с использованием процедуры усреднения для их уменьшения.

Методика исследования случайной погрешности, вызываемой шумом

Наложенный на сигнал аддитивный шум $e(n)$ полагаем гауссовым с нулевым математическим ожиданием, дисперсией $\sigma_ш^2$ и равномерной спектральной плотностью мощности $P_ш(f) = \sigma_ш^2 / f_д$, где $f_д$ – частота дискретизации сигнала.

Дисперсии оценок параметров сигнала при отношениях сигнал-шум, превышающих единицу, находим как дисперсии линеаризованной функции y случайных аргументов $x_1, x_2, \ldots, x_l$ с математическими ожиданиями $m_{x1}, m_{x2}, \ldots, m_{xl}$ с учетом их возможной корреляционной связи [3]:

$$\sigma_y^2 = \sum_{i=1}^{l} \left(\frac{\partial y}{\partial x_i} \right)_m^2 \cdot \sigma_{x_i}^2 + 2 \sum_{i<j} \left(\frac{\partial y}{\partial x_i} \right)_m \left(\frac{\partial y}{\partial x_j} \right)_m r_{ij} \cdot \sigma_{x_i} \cdot \sigma_{x_i} \qquad (1)$$

Приложение (1) к оценкам конкретных параметров сигнала включает вычисление производных функции dy/dx_i при заданных значениях (математических ожиданиях) аргументов, дисперсий случайных переменных $\sigma_{x_i}^2$ и коэффициентов корреляции между ними r_{ij}.

Случайные шумовые погрешности оценок частоты

Принятая в [1] за базовую оценка $\hat{\alpha}$ дробной части нормированной частоты сигнала $\theta = k + \alpha$ ($\alpha = 0 \pm 0{,}5$), выраженной в бинах ДПФ, находится по большему из отношений амплитуд частотных выборок (ЧВ) ДПФ $R_{k\pm1} = X_{k\pm1} / X_k$ (X_{k+1}/X_k при $\alpha > 0$ и X_{k-1}/X_k при $\alpha < 0$), с которыми она связана функцией измерения: $\hat{\alpha} = f(R_{k\pm1}) = \psi^{-1}(R_{k\pm1})$. Ей отвечает дисперсия $\sigma_{\hat{\alpha}}^2 = (\partial\alpha / \partial R_{k\pm1})^2 \cdot \sigma_{R_{k\pm1}}^2$. Полученная в соответствии с (1) дисперсия отношения амплитуд ЧВ $\sigma_{R_{k\pm1}}^2$ с учетом их корреляционной связи $r_{x_{k\pm1}, x_k}$ определяется выражением:

$$\sigma_{R_{k\pm1}}^2 = \left[1 + R_{k\pm1}^2 - 2R_{k\pm1} \cdot r_{x_{k\pm1}, x_k} \right] \frac{\sigma_{\theta\hat{e}}}{X_k^2} \qquad (2)$$

Она зависит от отношения $\delta_{шк}^2(\alpha) = \sigma_{шк}^2 / X_k^2$ дисперсии шума канала ДПФ $\sigma_{шк}^2 = \sigma_ш^2 \sum_{n=0}^{N-1} w^2(n)$ к квадрату амплитуды сигнальной ЧВ $X_k^2 = \left[(X_m/2) \cdot N \cdot W(\alpha) \right]^2$, которое выражается через амплитуду сигнала X_m (отношение сигнал-шум $SNR = X_m / \sqrt{2}\sigma_ш$) и параметры ВФ $w(n)$ длины N – ее нормированную ЧХ (преобразование Фурье) $W(\alpha)$, когерентное усиление $W(0) = W(\alpha)|_{\alpha=0}$ и эквивалентную шумовую полосу $\Delta F_{шэ}$ [4]:

$$\delta_{шк}^2(\alpha) = \sigma_{шк}^2 / X_k^2 = \frac{W^2(0) \cdot 2 \cdot \Delta F_{шэ}}{W^2(\alpha) \cdot SNR^2 \cdot N}. \qquad (3)$$

Полагая, что в математическом ожидании отношение амплитуд ЧВ $R_{k\pm1}$ равно отношению значений ЧХ ВФ $\psi(\alpha) = W(1-\alpha)/W(\alpha)$ [1], и переходя к стандартным отклонениям (СКО) $\sigma_{R_{k\pm1}}$, $\sigma_{\hat{\alpha}}$, получим:

$$\sigma_{R_{k\pm1}} = Q_R \left[SNR \cdot \sqrt{N} \right]^{-1}, \quad \sigma_{\hat{\alpha}} = Q_\alpha \left[SNR \cdot \sqrt{N} \right]^{-1}, \qquad (4)$$

$$Q_R = \sqrt{1 + \psi^2(\alpha) - 2\psi(\alpha) \cdot r_{x_{k\pm1}, x_k}} \cdot \sqrt{2\Delta F_{шэ}} \left[W(0)/W(\alpha) \right], \quad Q_\alpha = \left| \partial\alpha / \partial\psi(\alpha) \right| \cdot Q_R \qquad (5)$$

где Q_R, Q_α – коэффициенты, которые для конкретной ВФ зависят только от значения α.

Необходимые для расчетов значения $\psi(\alpha)$, $W(\alpha)$, $|\partial\alpha/\partial\psi(\alpha)|$ могут быть найдены по их аналитическим выражениям, приведены в [1, 2].

Наибольшие трудности расчета случайной погрешности оценок частоты связаны с нахождением коэффициента корреляции $r_{X_{k\pm1},X_k}$ ЧВ $X_{k\pm1}$, X_k. Он зависит от степени перекрытия ЧХ соседних каналов ДПФ, возрастающей с увеличением ширины главного лепестка ЧХ ВФ. Его значения получены с помощью статистического моделирования и составляют $r_{X_{k\pm1},X_k}$ ≈ 0,67 для ВФ Ханна и $r_{X_{k\pm1},X_k}$ ≈ 0,75 для ВФ Блэкмана.

В таблице приведены значения коэффициентов Q_R, Q_α (5) и СКО $\sigma_{R_{k\pm1}}$, $\sigma_{\hat{\alpha}}$ (4), рассчитанные для ВФ Ханна при N=128, α = 0; 0,25; 0,5, SNR = 3,16 (10дБ).

Таблица 1

α	0	0,25	0,5
Q_R	1,32	1,342	1,658
Q_α	1,76	1,37	1,243
$\sigma_{R_{k\pm1}}$	0,037	0,038	0,046
$\sigma_{\hat{\alpha}}$	0,05	0,0383	0,035

При α = 0 отношение $\sigma_{R_{k\pm1}} / R_{k\pm1}$ = 0,074 < 0,1, что подтверждает обоснованность линейного приближения функции $\alpha = f(R_{k\pm1}) = \psi^{-1}(R_{k\pm1})$.

Существует также пороговое отношение сигнал-шум $SNR_{\text{пор}}$, ниже которого возникают грубые ошибки за счет измерений по шумовым ЧВ, не соответствующим измеряемой частоте сигнала. Это возможно уже при $\delta_{\text{шк(пор)}}(\alpha) \le (0{,}5-1)$, что при подстановке в (3) дает $SNR_{\text{пор}} = (1-2)\cdot\sqrt{2\Delta F_{\text{ш}}/N}$. Тем же исходным данным, что и выше, отвечают значения $SNR_{\text{пор}} \le$ (0,3 – 0,6).

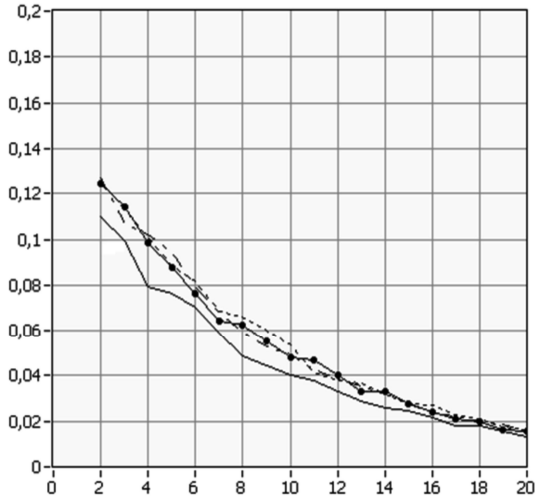
С помощью статистического моделирования получены зависимости случайной погрешности базовой оценки частоты $\sigma_{\hat{\alpha}}$ для ВФ Ханна от отношения сигнал-шум *SNR* в пределах изменения его значений от 0 до 20 дБ при α = 0 (рис. 1 *а*) и α = 0,5 (рис. 1 *б*). Они соответствуют N =128, k = 32.

Число циклов усреднения оценок частоты составляло 256 при одинаковой начальной фазе реализаций сигнала φ = 0 (когерентное усреднение).

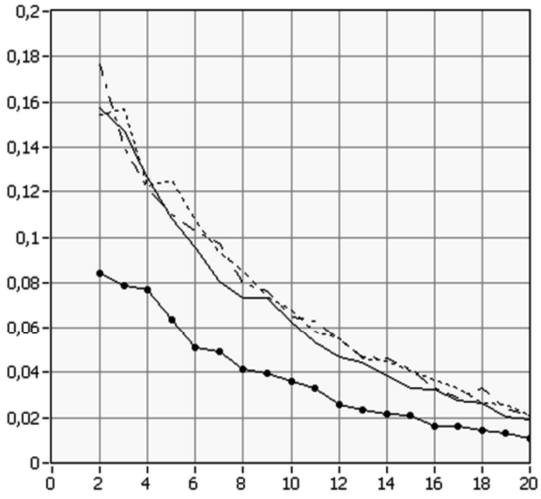
На этих же рисунках показаны графики случайных погрешностей оценок частоты по расширенным функциям измерения $\sigma_{\hat{\alpha}+}$, $\sigma_{\hat{\alpha}-}$, $\sigma_{\hat{\alpha}\text{ср}}$ [1]. При α = 0 они достаточно близки к СКО базовой оценки частоты, а при α = 0,5 уступают ей.

Сравнение расчетных случайных погрешностей и погрешностей, полученных путем моделирования, показывает их достаточное соответствие.

Интерполяционные оценки частоты [2] являются функциями амплитуд трех ЧВ $\hat{\alpha}_{\text{ин}} = f(X_{k-1}, X_k, X_{k+1})$, и их дисперсии могут быть найдены аналитически по той же методике, что и для базовой оценки. Проведенные исследования случайной погрешности интерполяционных оценок показали близость их СКО для разных видов интерполяции и соизмеримость со случайной погрешностью базовой оценки частоты. Они несколько уступают ей по пороговому отношению сигнал-шум ввиду того, что при интерполяционных измерениях в области значений α, близких к 0,5, шумовая составляющая спектра достигает значения амплитуды минимальной ЧВ сигнала $X_{(k-1)\min}$ при более высоких отношениях сигнал-шум.



а)



б)

Рис. 1. Зависимости случайных погрешностей оценок частоты от отношения сигнал-шум SNR (дБ) – $\sigma_{\hat{\alpha}}$ (); $\sigma_{\hat{\alpha}+}$ (); $\sigma_{\hat{\alpha}-}$ (); $\sigma_{\hat{\alpha}\text{ср}}$ () при α = 0 (а) и α = 0,5 (б)

Случайные шумовые погрешности оценок амплитуды и фазы сигнала

Базовыми для амплитуды и начальной фазы вещественного гармонического сигнала являются оценки, определяемые по одной ЧВ X_k наибольшей амплитуды [2]:

$$\hat{X}_{m(\text{и})(k)} = 2\cdot X_k / \left[N\cdot W(\hat{\alpha})\right];$$

$$\hat{\varphi}_{\text{и}(k)} = \arg\left\{X(j\lambda_k)\right\} + \pi/2 - \pi\cdot\hat{\alpha}.$$

Их случайная погрешность вызывается как непосредственно случайными изменениями амплитуды ЧВ, так и косвенно случайной погрешностью оценки частоты $\hat{\alpha}$, если она априорно неизвестна.

Относительная дисперсия оценки амплитуды сигнала находится в соответствии с (1) как

$$\frac{\sigma^2_{\hat{X}m}}{X_m^2} = \frac{\sigma^2_{\text{шк}}}{X_k^2} + \left(\frac{\partial W(\alpha)}{\partial\alpha\cdot W(\alpha)}\right)^2 \sigma^2_{\hat{\alpha}}. \qquad (6)$$

Она определяется относительной дисперсией шума канала (3) $\delta^2_{\text{шк}}(\alpha) = \sigma^2_{\text{шк}} / X_k^2$ и дисперсией оценки часто-

ты $\sigma^2_{\hat{\alpha}}$, влияние которой зависит от коэффициента $\frac{\partial W(\alpha)}{\partial\alpha \cdot W(\alpha)}$. Вводя коэффициент $Q_{\hat{X}m} = [W(0)/W(\alpha)]\sqrt{2\Delta F_{\text{шэ}}}$ и представляя $\sigma^2_{\hat{\alpha}}$ как $\sigma^2_{\hat{\alpha}} = Q^2_{\alpha}/(SNR^2 \cdot N)$, выражение (6) можно привести к виду, удобному для расчета случайной погрешности оценки амплитуды при произвольных значениях SNR и N:

$$\frac{\sigma_{\hat{X}m}}{X_m} = \left[\left(\frac{Q_{\hat{X}m}}{SNR\sqrt{N}}\right)^2 + \left(\frac{\partial W(\alpha)}{\partial\alpha \cdot W(\alpha)}\right)^2\left(\frac{Q_{\alpha}}{SNR\sqrt{N}}\right)^2\right]^{1/2}. \quad (7)$$

При α = 0, а также в случае, когда частота сигнала точно известна, второе слагаемое в (7) равно нулю и $\sigma_{\hat{X}m}/X_m\big|_{\alpha=0} = Q_{\hat{X}m}/(SNR\sqrt{N})$.

Значение $Q_{\hat{X}m}\big|_{\alpha=0}$ для ВФ Ханна равно $\sqrt{3}$ и $\sigma_{\hat{X}m}/X_m\big|_{\alpha=0} = 1{,}73/(SNR\sqrt{N})$.

При SNR = 3,16 (10 дБ) и N = 128 $\sigma_{\hat{X}m}/X_m\big|_{\alpha=0} = 0{,}0484$.

При α = 0,5, без учета второго слагаемого в (7), $Q_{\hat{X}m}\big|_{\alpha=0,5}$ = 2,04 и $\sigma_{\hat{X}m}/X_m\big|_{\alpha=0,5} = 0{,}057$.

Степень влияния случайной погрешности оценки частоты при α = 0,5 определим для ВФ Ханна по значению $\partial W(\alpha)/\partial\alpha \cdot W(\alpha)$ = –0,667 и значению коэффициента Q_{α} = 1,243 для базовой оценки частоты (см. табл. 1). При тех же значениях SNR и N получим $\sigma_{\hat{X}m}/X_m\big|_{\alpha=0,5}$ = 0,0615.

Как видим, доминирующее влияние на случайную погрешность оценки амплитуды оказывает случайная погрешность амплитуды ЧВ X_k.

Определим случайную погрешность оценки *начальной фазы*. Переходя от приращений начальной фазы $d\varphi_{(\text{и})} = arctg\frac{dX_k}{X_k} - \pi \cdot d\alpha_{\text{и}} \approx \frac{dX_k}{X_k} - \pi \cdot d\alpha_{\text{и}}$, вызываемых изменениями амплитуды ЧВ X_k и дробной части частоты α, к их дисперсиям, получим

$$\sigma^2_{\hat{\varphi}} = (\sigma^2_{\text{шк}}/X^2_k) + \pi^2 \cdot \sigma^2_{\hat{\alpha}}. \quad (8)$$

Отсюда находится СКО оценки начальной фазы:

$$\sigma_{\hat{\varphi}} = \left[\left(\frac{Q_{\hat{X}m}}{SNR\sqrt{N}}\right)^2 + \pi^2\left(\frac{Q_{\alpha}}{SNR\sqrt{N}}\right)^2\right]^{1/2}. \quad (9)$$

При точно известном значении частоты сигнала или при измерении разности фаз второе слагаемое в этих выражениях отсутствует, и СКО оценки начальной фазы и относительное значение СКО оценки амплитуды сигнала совпадают. Дисперсия оценки *разности фаз* при этом будет в два раза больше.

При неизвестной частоте сигнала случайная погрешность оценки начальной фазы очень чувствительна к случайной погрешности оценки частоты $\sigma_{\hat{\alpha}}$ (коэффициент π^2 во втором слагаемом (8), (9)). При исходных данных вышеприведенного примера при α = 0,5 получим $\sigma_{\hat{\varphi}}$ = 0,123.

Для уменьшения случайной погрешности оценки начальной фазы нужно существенно уменьшить случайную погрешность оценки частоты.

На рис. 2 приведены полученные с помощью моделирования графики зависимости от отношения сигнал-шум случайных погрешностей оценок амплитуды (*а*) и начальной фазы (*б*) при α = 0 для ВФ Ханна при измеренном и заданном значении частоты α.

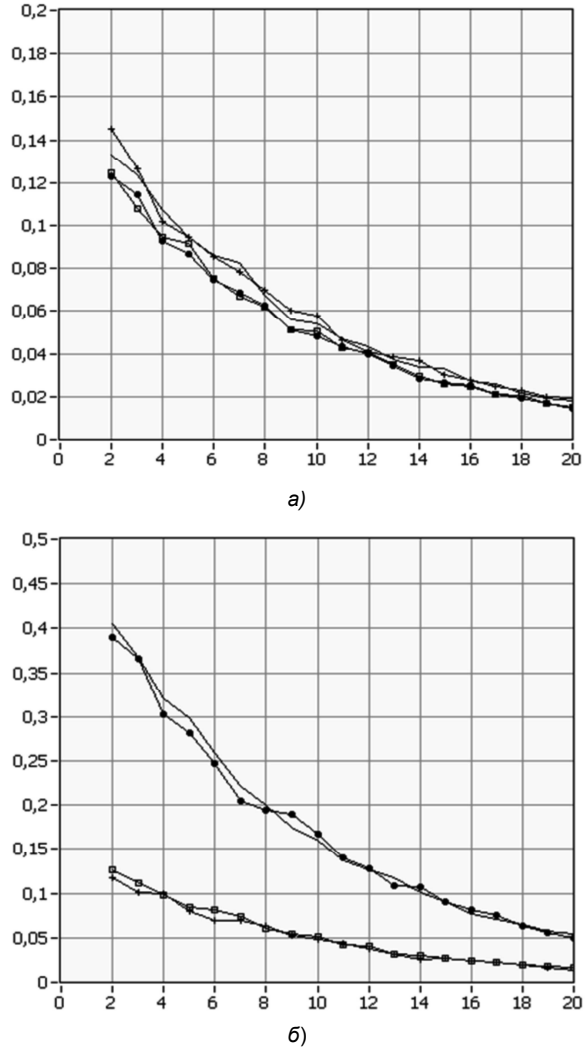


Рис. 2. Зависимости от отношения сигнал-шум случайных погрешностей оценок амплитуды (а) и начальной фазы (б) при α = 0 по измеренному ( ) и заданному ( ) значению частоты α и средневзвешенной оценки по измеренному ( ) и заданному ( ) значению частоты α

Графики показывают достаточно слабую зависимость оценок амплитуды от точности оценки частоты, которая сильно проявляет себя на оценке начальной фазы. Они хорошо согласуются с аналитическими результатами, приведенными выше.

На рис. 2 приведены также полученные путем моделирования для ВФ Ханна графики зависимости случайной погрешности от частоты для средневзвешенных оценок амплитуды и начальной фазы [2] при измеренном и заданном значениях частоты. Как видим, они практически не отличаются от случайных погрешностей этих оценок по ЧВ наибольшей амплитуды X_k.

Проведенные исследования случайной погрешности *интерполяционных оценок* амплитуды [2] показали близость их СКО для разных видов интерполяции и соизмеримость со случайной погрешностью оценки амплитуды по одной ЧВ (базовым способом).

Можно полагать, что небольшие различия значений случайных погрешностей для разных спектральных оце-

нок параметров сигнала являются следствием достаточно сильной корреляционной связи амплитуд соседних ЧВ ДПФ, по которым они находятся.

Применение усреднения при измерении параметров сигнала в шумах

Повышение точности оценок параметров сигнала в шумах может быть достигнуто как путем вычисления ДПФ (БПФ) одной реализации сигнала большой длины N, так и с помощью усреднения по множеству реализаций небольшой длины (числу циклов усреднения K_y). Применение процедуры усреднения обосновано ее более высокой вычислительной эффективностью и гибкостью. Усреднение в частотной области может выполняться [5]:

– по комплексному спектру сигнала (векторное (когерентное) усреднение);

– по амплитудному спектру сигнала;

– по спектру мощности сигнала;

– по амплитудному спектру или спектру мощности сигнала с перекрытием;

– путем усреднения оценок параметров сигнала (усреднение по параметру).

С каждым способом усреднения связаны особенности его аппаратной поддержки (синхронизации выборки и ввода сигнала), которые нами не рассматриваются, возможности измерения начальной фазы и др. При усреднении в частотной области вариации ЧВ ДПФ, по которым вычисляются параметры сигнала, уменьшаются в K_y раз по мощности или в $\sqrt{K_y}$ раз по СКО. Различие способов когерентного и некогерентного усреднения проявляется на шумовых ЧВ, не содержащих сигнальных составляющих. Усреднение по параметру практически ограничивается высоким пороговым значением отношения сигнал-шум при однократном измерении, приводящим к ошибкам обнаружения сигнала.

При измерении частоты и амплитуды сигналов по усредненным спектрам амплитуд и мощности, наряду с уменьшением случайной погрешности, возникает вопрос о *смещении* их оценок, вызываемом шумом.

Смещение амплитуд ЧВ подтверждается аналитически. Средняя амплитуда ЧВ сигнала и шума определяется как $\widetilde{X}_k = \sqrt{X_k^2 + \overline{E}_k^2} = X_k\sqrt{1+\overline{E}_k^2/X_k^2}$, где $\overline{E}_к^2 = 4\sigma_ш^2\,\Delta F_{шэ}/N$ – средний квадрат модуля амплитуд ЧВ шума. При $\overline{E}_k^2 \le (0{,}1-0{,}2)X_k^2$ смещение усредненной амплитуды ЧВ сигнала и шума составляет $\Delta_{см\,k} = \widetilde{X}_k - X_k = \overline{E}_k^2/(2X_k)$, а ее относительное смещение $\Delta_{см\,k}/X_k = \left[2F_{шэ}/(SNR^2\cdot N)\right]\cdot\left[W(0)/W(\alpha)\right]^2$. Это совпадает с относительной дисперсией шума канала (3), которой определяются случайные вариации амплитуды сигнальной составляющей спектра X_k. Однако в отличие от относительной дисперсии шума канала, уменьшающейся при усреднении спектров, относительное смещение сигнальной составляющей спектра остается неизменным.

Оценке амплитуды сигнала по ЧВ X_k соответствует относительное смещение $\Delta_{см\,Xm}/X_m = (\Delta_{см\,k}/X_k)$. Для ВФ Ханна при N=128, SNR=1 и α=0 $\Delta_{см\,Xm}/X_m$ = 0,0235, при α = 0,5 смещение составляет 0,033. Эти данные согласуются и с результатами моделирования.

Смещение оценки частоты зависит от смещения отношения амплитуд ЧВ: $\Delta\alpha_{см} = \left(\partial\alpha/\partial R_{k\pm1}\right)\Delta\widetilde{R}_{k\pm1}$. Аналитическое подтверждение смещения для этой оценки не является достаточно надежным, как и те числовые значения, полученные с помощью моделирования, которые соизмеримы с их СКО.

Кроме того, эффективность усреднения (степень уменьшения случайной погрешности) для базовой и модифицированных оценок частоты зависит от измеряемого значения α и, в частности, может существенно различаться при α = 0 и α = 0,5 или –0,5. Так, для базовой оценки эффективность ниже расчетной при α = 0, а для модифицированной базовой – при α = 0,5 или –0,5. Это различие связано с особенностью используемых функций измерения в области указанных значений α, там, где они претерпевают разрыв [1]. Для базовой оценки частоты это имеет место при α = 0, а для оценок частоты по расширенным функциям измерения – при α = ±0,5.

Поэтому при измерении в шумах можно сочетать эти оценки частоты в зависимости от измеряемого значения α.

Особенности коррекции методических погрешностей при многократных измерениях в шумах

При синхронизированном вводе реализаций сигнала, ДПФ которых усредняются, применимы все способы коррекции погрешности наложения и погрешности интерполяции (для интерполяционных оценок), рассмотренные в [1, 2]. В качестве первичных оценок параметров сигналов для коррекции используются их усредненные оценки. Далее алгоритмы коррекции применяются к каждой из считанных реализаций сигнала.

При несинхронизированном вводе возникают проблемы коррекции погрешности наложения с помощью способов, использующих первичную оценку начальной фазы (второй и третий способы). В данном случае приемлемы первый и четвертый способы, использующие для коррекции первичную усредненную оценку частоты. Алгоритмы коррекции при этом применяются к каждой строке матрицы отсчетов сигнала размерностью K_y строк и N_I = 1,5N или N столбцов.

Следует отметить, что коррекция методических погрешностей при измерении в шумах целесообразна до определенного соотношения сигнал-шум, пока случайная погрешность не является доминирующей. При случайных погрешностях, соизмеримых и больших погрешности наложения и/или интерполяции, их коррекция не дает общего существенного улучшения оценок параметров сигнала.

Влияние начальной фазы сигнала

Для изучения влияния случайной начальной фазы сигнала на рис. 3 *а, в* приведены графики зависимости от нее погрешности наложения оценок частоты (*а*) и амплитуды (*в*) для ВФ Ханна при k =2 и α, равных –0,5, 0,5, –0,25, 0,25.

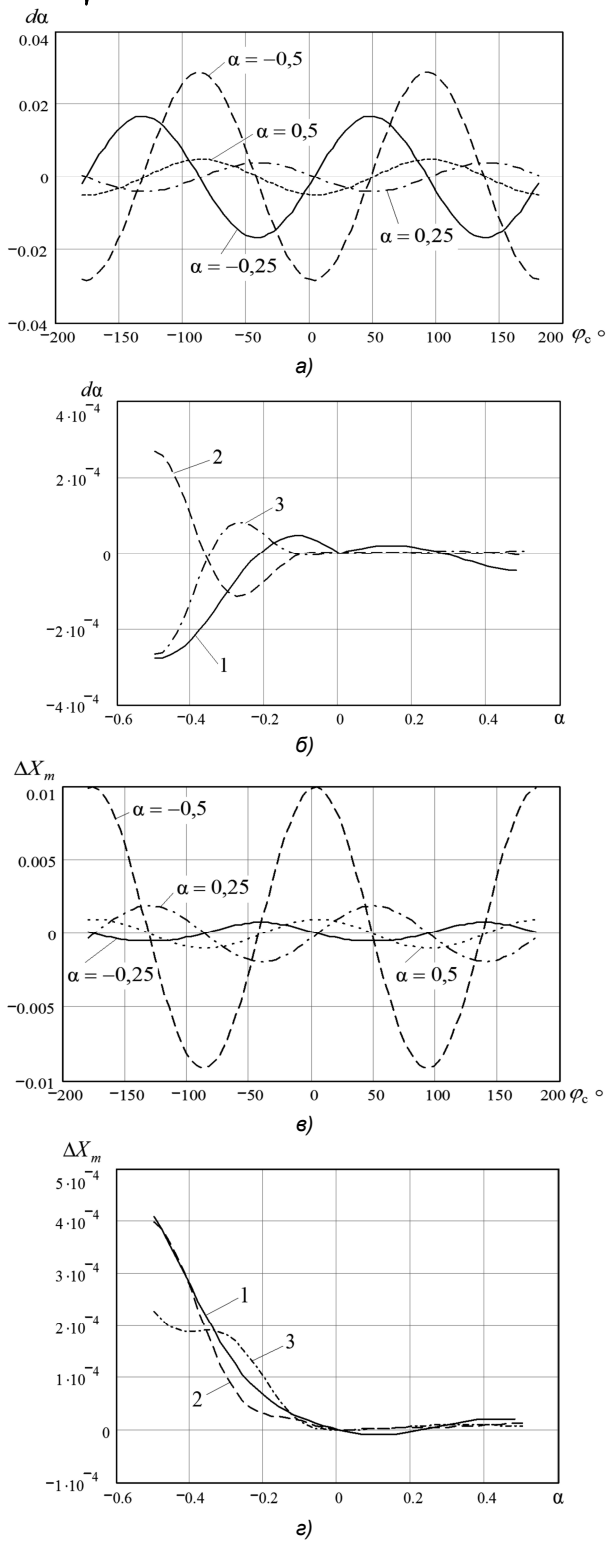


Рис. 3. Графики влияния начальной фазы сигнала на погрешность оценок частоты (а, б) и амплитуды (в, г) для ВФ Ханна, k = 2

Вызванному наложением изменению амплитуд ЧВ $\Delta X_{k(\varphi 1)} = \widetilde{X}_k - X_k$ при произвольной начальной фазе φ1 отвечает, согласно [1], примерно равное по модулю и противоположное по знаку изменение амплитуд ЧВ $\Delta X_{k(\varphi 2)} \approx -\Delta X_{k(\varphi 2)}$ при начальной фазе $\varphi_2 = \varphi_1 \pm (\pi/2)$. Так же соотносятся и погрешности оценок частоты и амплитуды при начальных фазах φ_1, φ_2. В связи с этим представляют интерес погрешности наложения, усредненные по множеству значений начальной фазы сигнала в пределах –π, π, графики которых приведены на рис. 3 *б*, *г* (графики 1). На графиках 2, 3 показаны также зависимости погрешности наложения, полученные усреднением двух оценок, соответствующих начальным фазам 0; π/2 (график 2), π/4; 3π/4 (график 3), π/2; π (график 2). Они показывают существенное уменьшение погрешности наложения оценок частоты и амплитуды при усреднении как по множеству значений начальной фазы в пределах 0,…± π, так и при попарном усреднении оценок частоты при начальных фазах, смещенных на π/2. Это же относится и к оценкам начальной фазы сигнала и может быть использовано для коррекции погрешности наложения, в том числе и при измерениях в шумах.

При начальной фазе сигнала, равновероятной в пределах ± π, закон распределения зависящей от нее погрешности наложения оценок параметров сигнала с учетом графиков рис. 3 *а*, *в* близок к арксинусу [3]. При этом СКО случайной погрешности наложения в $\sqrt{2}$ раз меньше ее максимального значения (более вероятными являются большие значения погрешности).

Заключение

Результаты работы дают необходимые для практического применения представления об эффективности исследуемых спектральных оценок параметров сигнала, об общих свойствах и различиях разных процедур усреднения и о реализации измерений в шумах с усреднением и с коррекцией методических погрешностей. Все они апробированы в разработанном исследовательском измерительном комплексе на базе ПЭВМ [1, 2].

Литература

1. Глинченко, А. С. Исследование спектрально-весового измерения частоты сигналов/ А. С. Глинченко, О. А. Тронин // Цифровая обработка сигналов. –2010. – №2. – С.22-28.

2. Глинченко, А. С. Повышение точности интерполяционных оценок параметров сигналов при спектрально-весовых измерениях/ А. С. Глинченко, О. А. Тронин // Цифровая обработка сигналов. –2011. – №1. – С.7-12.

3. В. С. Пугачев. Введение в теорию вероятностей: М., «Наука», 1968.

4. Хэррис Ф. Дж. Использование окон при гармоническом анализе методом дискретного преобразования Фурье / Ф. Дж. Хэррис // ТИИЭР, т. 66. 1978. №1.

5. Р. Лайонс. Цифровая обработка сигналов: пер с англ./ 2-е изд. М.: ООО «Бином-Пресс», 2007.

RESEARCH OF SPECTRAL-WEIGHT MEASUREMENT RANDOM ERRORS OF SIGNALS PARAMETERS

Glinchenko A.S., Komarov V.A., Tronin O.A.

In work the random errors corresponding to different ways of spectral-weight measurement of harmonic signals parameters, and also the possible ways of averaging providing their reduction which are realized in a measuring instrument of signals parameters on the base of personal computer are defined and compared.

УДК 681.391

МЕТОДЫ И АППАРАТУРА КОДИРОВАНИЯ И ДЕКОДИРОВАНИЯ НИЗКОПЛОТНОСТНЫХ КВАЗИЦИКЛИЧЕСКИХ КОДОВ

Кравченко А.Н., *к.т.н., проектировщик систем Ubiso GmbH, Germany, e-mail: alexander.kravtchenko@ubiso.com*

Ключевые слова: кодирование, декодирование, низкоплотностное, LDPC коды, проверочная матрица, кодовая скорость.

Произведен анализ проверочных матриц нерегулярных квазициклических низко-плотностных кодов WPAN (IEEE 802.15.3c) стандарта (QC-LDPC). Установлено, что проверочные матрицы являются квадратичными и обратимыми. В связи с этим к расчету порождающих матриц применен традиционный метод с приложением к квазициклическим низкоплотностным кодам. Приведены методы кодирования и декодирания QC-LDPC кодов на основе проверочных матриц. Предложена аппаратура кодирования этих кодов. Представлены результаты моделирования эффективности декодирования LDPC кодов WPAN-стандарта в канале с АБГШ для случая двоичной фазовой модуляции.

Введение

Системы связи и вещания, в частности беспроводные системы, часто подвержены помехам, в связи с чем использование помехоустойчивых кодов имеет важное значение. Среди помехоустойчивых кодов LDPC коды, предложенные Галлагером [1] и позже вновь открытые Маккеем и Нилом [2], появляются как класс кодов, которые обладают эффективной способностью исправлять ошибки в каналах связи. Превосходная эффективность декодирования этих кодов получается при использовании алгоритма декодирования, вычисляющего итеративно распределение вероятностей в граф-орентированной модели, который известен в литературе как «belief propagation algorithm» (BPA). В отечественной литературе этот метод известен как алгоритм декодирования с итеративным распространением доверия (АРД) [3]. Следует отметить, что в литературе существует большое количество модификаций этого алгоритма (алгоритмы, основывающиеся на логарифмическом отношении функций правдоподобия, упрощенные методы). Некоторые сведения о модификациях можно найти в [11]. В настоящее время LDPC коды со средней пропускной способностью используются в стандартах как DVB-S2, WiMax(IEEE 802.16e) и WLAN(IEEE 802.11n). Более того, LDPC коды с высокой пропускной способностью используются в стандарте WPAN (IEEE 802.15.3c). Следует отметить, что во всех перечисленных стандартах используются квази-циклические низкоплотностные коды (QC-LDPC).

QC-LDPC коды, основные определения

LDPC коды универсально специфицируются их проверочными матрицами. Проверочная матрица QC-LDPC кода задается как массив циркулянтов (перестановочных матриц) одного и того же размера. Циркулянт [4] представляет собой квадратную матрицу $(b\times b)$, в которой каждая последующая строка является циклическим сдвигом вправо на одно место предыдущей строки. Индекс сдвига определяет позицию «1» в первой строке матрицы. Весы строк и столбцов в циркулянте одни и те же, скажем к примеру $w=1$. Для простоты можно сказать, что циркулянт имеет вес Хэмминга w. Если $w=1$, то циркулянт есть циклическая перестановочная матрица. Циркулянт полностью характеризуется его первой строкой (или первым столбцом), которую называют генератором циркулянта. LDPC код определяется матрицей размера $n\times m$, где n длина кода и m число проверочных бит в коде. Число имформационых бит определяется как $k=n-m$. Проверочная матрица $\mathbf{H}_{qc}$ QC-LDPC кода с размерностью $t\times c$, где $n=b\times t$ и $m=b\times c$, может быть построена путем соединения $t\times c$ циклических перестановочных матриц размерности $(b\times b)$.

Матрица $\mathbf{H}_{qc}$ QC-LDPC кода для двух натуральных чисел t и c с $c\le t$ может быть определена как массив циркулянтов $\mathbf{A}_{i,j}$ в поле $GF(2)$:

$$\mathbf{H}_{qc}=\begin{bmatrix}\mathbf{A}_{1,1} & \mathbf{A}_{1,2} & \cdots & \mathbf{A}_{1,t}\\ \mathbf{A}_{2,1} & \mathbf{A}_{2,2} & \cdots & \mathbf{A}_{2,t}\\ \vdots & \vdots & \ddots & \vdots\\ \mathbf{A}_{c,1} & \mathbf{A}_{c,2} & \cdots & \mathbf{A}_{c,t}\end{bmatrix} \qquad (1)$$

Структуру QC-LDPC кода можно определить на основе его проверочной матрицы в циркулянтной форме, определенной формулой (1). На основе этой формы, каждая кодовое слово $\mathbf{v}$ может быть разделено на секции $\mathbf{v}=(v_1,v_2,\cdots,v_t)$, и каждая секция v_j включает b компонент (бит). Для $1\le j\le t$, b компонент j-ой секции соответствуют b столбам j-го столбца циркулянта в $\mathbf{H}_{qc}$.

Анализ проверочных матриц и метод кодирования

LDPC коды, используемые в системе WPAN [5], являются систематическими нерегулярными QC-LDPC кодами. Табл. 1 для примера иллюстрирует проверочную матрицу LDPC кода с кодовой скоростью R = 7/8.

Индекс -1 определяет нулевую перестановочную матрицу (матрица состоит из одних нулей). Проверочная матрица состоит из двух частей $\mathbf{H}_{qc}=[\mathbf{H}_1\ \mathbf{H}_2]$, где $\mathbf{H}_1$ субматрица соответствует информационным битам, и $\mathbf{H}_2$ представляет проверочные биты кодового слова. Кодовое слово должно удовлетворять уравнению $\mathbf{H}_{qc}\cdot\mathbf{v}^{\mathrm{T}}=0$.

Таблица 1. Проверочная матрица LDPC кода с кодовой скоростью R = 7/8

1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32
0	18	6	5	7	18	16	0	10	2	3	6	10	16	9	0	20	7	9	5	4	12	4	4	4	10	19	5	10	-1	-1	-1
5	0	18	6	0	7	18	16	6	10	2	3	0	10	16	9	5	20	7	9	4	4	12	4	5	4	10	19	19	10	-1	-1
6	5	0	18	16	0	7	18	3	6	10	2	9	0	10	16	9	5	20	7	4	4	4	12	19	5	4	10	17	19	10	-1
18	6	5	0	18	16	0	7	2	3	6	10	16	9	0	10	7	9	5	20	12	4	4	4	10	19	5	4	7	17	19	10

Анализ проверочных матриц всех кодов показывает, что $\mathbf{H}_2$ - части матриц $\mathbf{H}_{qc}$ квадратичны и обратимы, в том числе и для LDPC кода с R=7/8. В этом случае для расчета порождающей матрицы $\mathbf{G}_{qc}$ может быть использован традиционный метод [6] с применением к QC-LDPC кодам. Исходя из этого, субматрицу $\mathbf{H}_2$ можно представить в виде

$$\mathbf{H}_2 = \begin{bmatrix} \mathbf{A}_{1,t-c+1} & \mathbf{A}_{1,t-c+2} & \cdots & \mathbf{A}_{1,t} \\ \mathbf{A}_{2,t-c+1} & \mathbf{A}_{2,t-c+2} & \cdots & \mathbf{A}_{2,t} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{A}_{c,t-c+1} & \mathbf{A}_{c,t-c+2} & \cdots & \mathbf{A}_{c,t} \end{bmatrix} \quad (2)$$

Порождающая матрица может быть представлена как

$$\mathbf{G}_{qc} = \begin{bmatrix} \mathbf{G}_1 \\ \mathbf{G}_2 \\ \vdots \\ \mathbf{G}_{t-c} \end{bmatrix} = \begin{bmatrix} \mathbf{I} & \mathbf{O} & \cdots & \mathbf{O} & \mathbf{G}_{1,1} & \mathbf{G}_{1,2} & \cdots & \mathbf{G}_{1,c} \\ \vdots & \mathbf{I} & \cdots & \mathbf{O} & \mathbf{G}_{2,1} & \mathbf{G}_{2,2} & \cdots & \mathbf{G}_{2,c} \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots \\ \mathbf{O} & \mathbf{O} & \cdots & \mathbf{I} & \mathbf{G}_{t-c,1} & \mathbf{G}_{t-c,2} & \cdots & \mathbf{G}_{t-c,c} \end{bmatrix} =$$

$$= \left[\mathbf{I}_{(t-c)b} \quad \mathbf{P}\right], \quad (3)$$

где $\mathbf{I}$ - единичная матрица размерности $b \times b$, $\mathbf{O}$ - нулевая матрица размерности $b \times b$ и $\mathbf{G}_{i,j}$ - циркулянты с $1 \le i \le t-c$ и $1 \le j \le c$ и размерностью $b \times b$.

Порождающая матрица (3) состоит из двух частей: левой части $\mathbf{I}_{(t-c)b}$ и правой части $\mathbf{P}$. Левая часть матрицы $\mathbf{I}_{(t-c)b}$ с $(t-c)b \times (t-c)b$ представляет собой единичные матрицы, расположенные на главной диагонали, а правая часть - массив циркулянтов. Представление кода в форме (3) является систематическим представлением [6]. Левая часть отображает информационные биты кодового слова, правая часть - его проверочные биты. Необходимым и достаточным условием для существования порождающей матрицы $\mathbf{G}_{qc}$ является условие $\mathbf{H}_{qc}\mathbf{G}_{qc}^{\mathrm{T}} = \mathbf{O}$. Допустим $\mathbf{g}_{i,j}$ представляет собой генератор циркулянта $\mathbf{G}_{i,j}$. Когда $\mathbf{g}_{i,j}$ известны, можно сформировать все циркуланты порождающей матрицы $\mathbf{G}_{qc}$.

Таким образом $\mathbf{G}_{qc}$ характерезируется полностью группой $c(t-c)$ генераторов. Допустим $\mathbf{u} = (1, 0, \cdots, 0)$ есть группа из b бит с '1' в первой позициии и $\mathbf{0} = (0, 0, \cdots, 0)$ - группа, которая содержит только нули. Для $1 \le i \le t-c$ первый ряд субматрицы $\mathbf{G}_i$ определяется следующим образом

$$\mathbf{g}_i = (\mathbf{0} \cdots \mathbf{0}\ \mathbf{u}\ \mathbf{0} \cdots \mathbf{0}\ \mathbf{g}_{i,1}\ \mathbf{g}_{i,2} \cdots \mathbf{g}_{i,c}), \quad (4)$$

где группа $\mathbf{u}$ находится в i - ой позиции $\mathbf{g}_i$.

Обозначим $z_i = (\mathbf{g}_{i,1}\ \mathbf{g}_{i,2} \cdots \mathbf{g}_{i,c})$ и $\mathbf{H}_1^{(i)} = \left[\mathbf{A}_{1,i}^{\mathrm{T}} \cdots \mathbf{A}_{c,i}^{\mathrm{T}}\right]^{\mathrm{T}}$. Тогда выражение $\mathbf{H}_{qc}\mathbf{g}_i^{\mathrm{T}} = \mathbf{0}$ дает следующее равенство:

$$\mathbf{H}_1^{(i)}\mathbf{u}^{\mathrm{T}} + \mathbf{H}_2\mathbf{z}_i^{\mathrm{T}} = \mathbf{0}. \quad (5)$$

Поскольку $\mathbf{H}_2$ есть квадратная матрица и имеет полный ранг, то она обратима. Тогда из (5) следует

$$\mathbf{z}_i^{\mathrm{T}} = \mathbf{H}_2^{-1}\mathbf{H}_1^{(i)}\mathbf{u}^{\mathrm{T}}. \quad (6)$$

Решение (6) для $1 \le i \le t-c$ дает все генераторы $\mathbf{g}_{i,j}$ порождающей матрицы $\mathbf{G}_{qc}$.

Обозначим $\mathbf{a} = (a_1, a_2, \cdots, a_{(t-c)b})$ как информационную последовательность бит $(t-c)b$, подлежащих кодированию. После деления информационых бит на секции $(t-c)$ равной длины $\mathbf{a} = (\mathbf{a}_1, \mathbf{a}_2, \cdots, \mathbf{a}_{t-c})$, кодовое слово можно представить в форме $\mathbf{v} = \mathbf{a}\mathbf{G}_{qc}$, которое имеет систематическое представление: $\mathbf{v} = (\mathbf{a}, \mathbf{p}_1, \mathbf{p}_2, \cdots, \mathbf{p}_c)$. Здесь $\mathbf{p}_1, \mathbf{p}_2, \cdots, \mathbf{p}_c$ - секции равной длины b, где каждая секция представляет собой проверочные биты $\mathbf{p}_j = (p_{j,1}, p_{j,2}, \cdots, p_{j,b})$. Проверочные биты в каждой секции вычисляются в соответствии с выражнием

$$\mathbf{p}_j = \mathbf{a}_1\mathbf{G}_{1,j} + \mathbf{a}_2\mathbf{G}_{2,j} + \cdots \mathbf{a}_{t-c}\mathbf{G}_{t-c,j}. \quad (7)$$

На основании рассмотренного метода были вычислены генераторы для всех порождающих матриц стандарта. Например, рис. 1 иллюстрирует генераторы, вычисленные для LDPC кода с кодовой скоростью R=7/8.

```
 0  000200  000011  01100c  048088
 1  001000  000208  008031  01904c
 2  000008  009000  020248  048071
 3  000010  010008  049080  0a02c8
 4  000004  004200  014021  020004
 5  001000  00000c  00c220  01c061
 6  004000  001020  02008c  02c320
 7  000200  004001  000024  021084
 8  100000  000808  00a044  0640c4
 9  000080  180000  000c09  00a446
10  000040  040080  080200  000e08
11  000008  008040  0600c0  0c0240
12  100000  000a00  002004  000000
13  004000  100020  020a80  022104
14  000001  005000  104028  028a88
15  000200  000000  004004  105020
16  000400  000012  012009  043094
17  000004  004400  010032  032029
18  000001  001004  000408  01803a
19  000010  010001  041084  080488
20  000020  020020  0a0120  1e0000
21  040000  000220  020821  0a1121
22  000020  060000  080320  120921
23  000020  020020  0e0100  180220
24  000020  020010  090900  040184
25  100000  000820  022014  094904
26  000800  100004  004830  026034
27  000010  010800  140084  0848b0
```

Рис. 1. Генераторы порождающей матрицы для LDPC кода с R=7/8

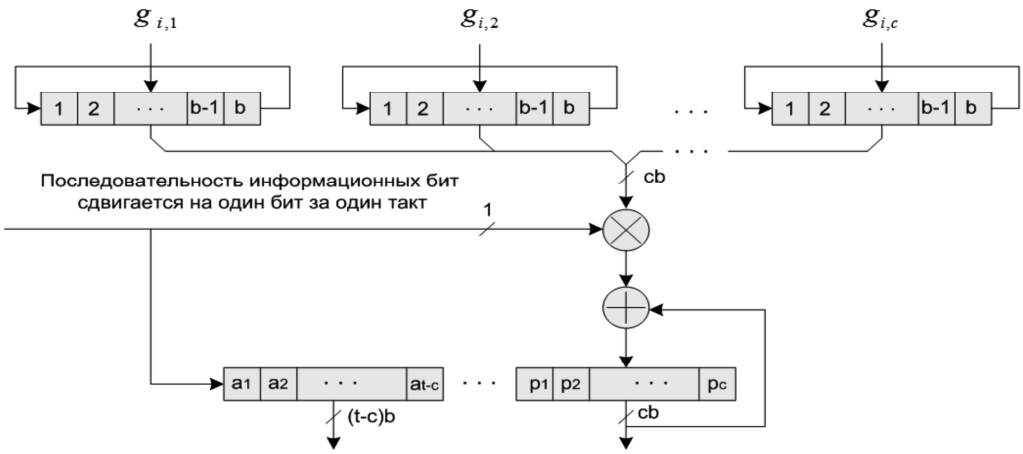


Рис. 2. Структура кодера

Представление $\mathbf{G}_{qc}$ в виде генераторов очень практично при разработке аппаратуры. В основе кодера лежит циклический сдвигающий регистр, который первоначально загружается генератором и в процессе последовательных сдвигов реализует циркулянт. Рис. 2 представляет структуру кодера. Кодер передает информационные биты на выход, как биты кодового слова. Одновременно кодер вычисляет проверочные биты $\mathbf{p}_j = (p_{j,1}, p_{j,2}, \cdots, p_{j,b})$ путем последовательного выполнения умножения циркулянтов на группы информационных бит. Непосредственное осуществление матричного умножения показано на рис. 2, как это предлагает Лин [7].

Циклические сдвигающие регистры, каждый длины b, загружаются генераторами (первая строка рис. 1). Для каждого информационного бита a_m $(m = 1, \cdots, (t-c)b)$, в свою очередь, эти регистры циклически сдвигаются один раз. Если a_m = 1, то выполняется операция XOR и результат загружается в выходной регистр (проверочные биты). Когда операция перемножения циркулянтов на информационные биты группы завершена, следующие генераторы $g_{1,1}, g_{1,2}, \cdots, g_{1,c}$ загружаются в регистры сдвига, и операция кодирования повторяется. Процесс кодирования завершается, когда последняя группа циркулянтов (27, рис. 1) перемножается на последнюю группу информационых бит (27). После кодирования все проверочные биты находятся в выходном регистре.

Декодирование QC-LDPC кодов

Для декодирования QC-LDPC кодов широко используется алгоритм послойного декодирования (layered decoding), развитый в работах [8-9]. Применение этого алгоритма для декодирования кодов в стандартах DVB-S2, WiMax(IEEE 802.16e) и WLAN(IEEE 802.11n) изложено в работе [10]. В настоящей работе этот алгоритм применен для декодирования кодов стандарта WPAN (IEEE 802.15.3c). Коротко остановимся на особенностях алгоритма:

1. Декодер, спроектированный при использовании этого алгоритма, требует меньшего объема памяти чем стандартый декодер, реализирующий стандартный двухфазный метод декодирования – «сумма произведение» (two-phase message-passing, TPMP algorithm) [10].

2. Скорость сходимости алгоритма существенно выше чем у «TPMP» (алгоритм выполняет меньше итераций для достижения равной эффективности декодирования).

При описании алгоритма воспользуемся метрикой, называемой логарифмическим отношением функций правдоподобия (log-likelihood ratio - LLR), которая определяется выражением

$$LLR(x \mid y) = \ln\left[\frac{p(y \mid x = 0)}{p(y \mid x = 1)}\right].$$

Допустим, $x = [x_1, x_2, \ldots, x_n]$ обозначает кодовое слово, которое модулируется при использовании двоичной фазовой модуляции, и модулированные значения x передаются по каналу с белым аддитивным гауссовым шумом (АБГШ). Допустим $y = [y_1, y_2, \ldots, y_n]$ обозначает входную последовательность принятых сигналов (символов). Демодулятор принимает входную последовательность сигналов и вычисляет соответствующие LLR значения для: $j = 1, 2, \ldots, n$

$$\lambda_j = LLR(y_j \mid x_j) = \ln\left[\frac{p(y_j \mid x_j = 0)}{p(y_j \mid x_j = 1)}\right]. \quad (8)$$

Обычно при двоичной биполярной передаче по каналу LLR значения вычисляются как

$$\lambda_j = \frac{2}{\sigma^2} * y_j \quad (9)$$

где σ^2 - дисперсия АБГШ.

При описании алгоритма воспользуемся следующими обозначениями:

λ_j - «LLR» значение принятого *j*-го символа, $R_{ij}[k]$ – внешние (extrinsic) «LLR» значения, $Q_{ij}[k]$ – внутренние (intrinsic) «LLR» значения, Λ_j - апостериорные «LLR» значения.

Выполнение алгоритма осуществляется в следующем порядке:

1. Инициализация:

$$R_{ij}^0 = 0, \Lambda_j^0 = \lambda_j. \quad (10)$$

2. Горизонтальное сканирование (check node update rule):

$$Q_{ji}^{(k)} = \Lambda_j^{(k-1)} - R_{ij}^{(k-1)}; \quad (11)$$

$$R_{ij}^{(k)} = \prod_{j \in R\{i\} \setminus j} sign(Q_{ji}^{(k)}) * \{MIN_{j \in R\{i\} \setminus j} \mid Q_{ji}^{(k)} \mid\}; \quad (12)$$

$$\Lambda_j^{(k)} = Q_{ji}^{(k)} + R_{ij}^{(k)}, \quad (13)$$

где k – текущая итерация.

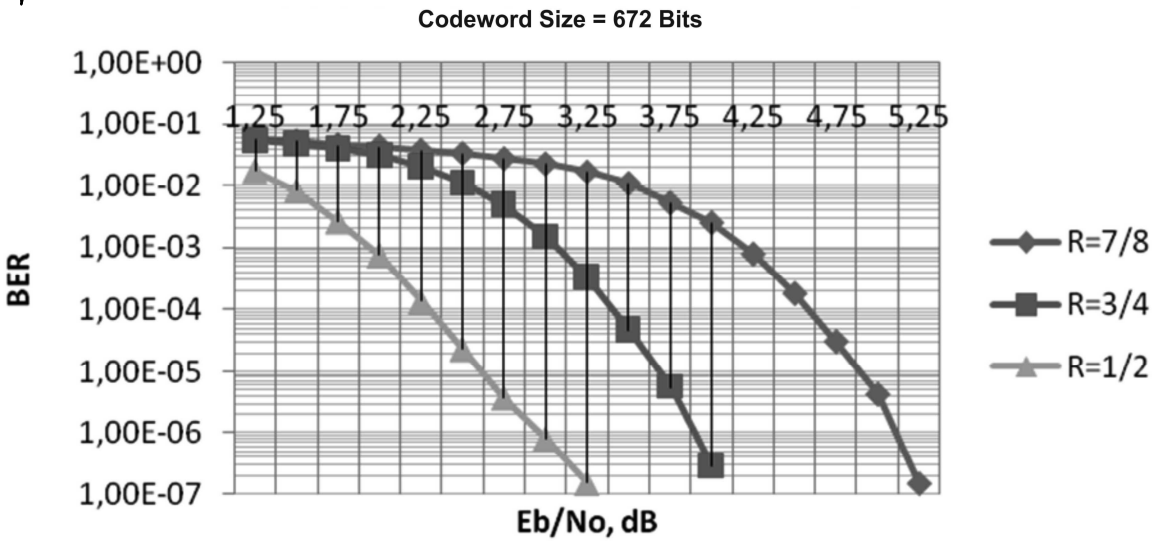


Рис. 3. Ефективность декодера для кодовых скоростей – 1/2, 3/4, 7/8

Твердое решение декодера (hard decision) после каждой итерации осуществляется на основе проверки условия - $\nu = 1$, если $\Lambda_j^{(k)} \le 0$, иначе $\nu = 0$.

Шаг 2 выполняется до тех пор, пока не выполнится условие - $\mathbf{H}_{qc} \cdot \mathbf{v}^{\mathrm{T}} = 0$ или будет достигнуто максимальное значение итераций. При вычислении $R_{ij}[k]$ используется метод, основанный на минималиной сумме ($Min-sum$ - аппроксимация) [11]. Данный алгоритм хорошо реализируется в аппаратуре декодера [12]. На рис. 3 представлены результаты моделирования эффективности декодирования LDPC кодов WPAN-стандарта в канале с АБГШ для случая двоичной фазовой модуляции.

При моделировании декодера использовалась арифметика с фиксированной запятой. Для представления мягких решений (LLR значений) демодулятора и сообщений внутри декодера использовался прямой код с числом разрядов b = 8.

Литература

1. R. G. Gallager, Low-Density Parity-Check Codes, MIT Press, Cambridge, MA, 1963. - 90 p.

2. D.J.C. MacKay and R.M. Neal. Near Shannon limit performance of low density parity check codes. Electron. Lett., Aug. 1996, vol.32, no.18. - pp.1645-1646.

3. Р. Морелос-Сарагоса. Искусство помехоустойчивого кодирования. Методы, алгоритмы, применение. Перевод с английского В.Б. Афанасьева. М.: Техносфера, 2005. - 319 с.

4. Z. Li, L. Chen, L. Zeng, S. Lin, and W.-H. Fong. Efficient Encoding of Quasi-Cyclic Low-Density Parity-Check Codes. IEEE Trans. Commun., Jan. 2005, vol. 54, no. 1. - pp. 71–81.

5. IEEE P802.15.3c/July 2007, Wireless Personal Area Network (WPAN) Standard Physical Layer (PHY) specifications (Draft).

6. S. Lin and D.J. Costello Jr., Error Control Coding: Fundamentals and Applications, 2nd ed. Upper Saddle River, NJ: Prentice-Hall, 2004. - 1260 p.

7. S. Lin, Qusi-Cyclic LDPC Code. CCSDS working group white paper, Oct. 2003.

8. D. Hocevar. A Reduced Complexity Decoder Architecture via Layered Decoding of LDPC Codes. In EEE Workshop on Signal Processing System, SISP 2004, pp. 107-112.

9. M. Mansour and N.R. Shanbhad. High-throughput LDPC decoders. IEEE Trans. VLSI Syst., Dec. 2003 , vol. 11, no. 6. - pp. 976-996.

10. T. Brack, M. Alles, T. Lehning-Enden, F. Kienle, N. Wehn, N.E. L'Insalata, F. Rossi, M. Rovini, L. Fanucci. Low Complexity LDPC Code Decoders for Next Generation Standards. Design, Automation & Test in Europe Conference & Exhibition, 2007. DATE '07.

11. Кравченко А.Н. Снижение сложности декодирования низкоплотностного кода. «Цифровая обработка сигналов». 2010, № 2, С.35-41.

12. Кравченко А.Н. Методы и аппаратура кодирования и декодирования систематического нерегулярного кода повторения - накопления (IRA) для DVB-S2 и DVB-T2 демодуляторов. «Цифровая обработка сигналов». 2009, № 4, С.41-47.

METHODS AND APPARATUS FOR ENCODING AND DECODING QUASI-CYCLIC LOW DENSITY PARITY-CHECK CODES

Kravchenko A.N.

The analysis of the parity check matrices of irregular quasi-cyclic low-density codes (QC-LDPC) of WPAN (IEEE 802.15.3c) standard is performed. It is established that the parity check matrices are quadratic and invertible. In this connection, for the calculation of generator matrices can be used the traditional method with application to quasi-cyclic low-density codes.

The encoding and decoding methods are proposed based on the parity check matrices. The apparatus of a hardware encoding of these codes are suggested. Result of performance simulation of the decoder for QC-LDPC codes of WPAN standard over AWGN channel is presented.

УДК 621.3.018.51

СРАВНИТЕЛЬНЫЙ АНАЛИЗ АЛГОРИТМОВ ЧАСТОТНОГО ДЕТЕКТИРОВАНИЯ В СИСТЕМАХ С OFDM

Овинников А.А., *аспирант Рязанского государственного радиотехнического университета, e-mail: ovinnikovalexey@gmail.com;*
Бакке А.В., *к.т.н., доцент Рязанского государственного радиотехнического университета, e-mail: usr37ru@yandex.ru*

Рассмотрены алгоритмы оценки частотного смещения в системе синхронизации OFDM сигналов во временной и частотной областях. В результате сравнительного анализа представленных алгоритмов в многолучевых каналах были сделаны выводы о точности получаемых оценок смещения несущей частоты и о возможностях практического применения подобных методов в системах с OFDM.

Ключевые слова: мультимедийное вещание, многолучевой канал, частотное детектирование, синхронизация, мультиплексирование.

Введение

Большинство современных телекоммуникационных систем мобильной связи и цифрового мультимедийного вещания используют в основе радиоинтерфейса физического уровня технологию ортогонального частотного разделения с мультиплексированием потока данных (OFDM). Обеспечивая высокую спектральную эффективность, устойчивость к многолучевому распространению и межсимвольной интерференции, технология OFDM при этом демонстрирует весьма существенную чувствительность к относительно небольшим частотным сдвигам между принимаемым сигналом и гетеродином приемного устройства.

Этот сдвиг проявляется в смещении всех поднесущих OFDM сигнала на величину рассогласования (carrier frequency offset - CFO), что может быть вызвано как расстройкой задающих генераторов приемного и передающего устройств, так и эффектом Доплера. Возникающее частотное рассогласование является причиной снижения эффективности функционирования приемника OFDM, состоящее в росте числа срывов синхронизации и в существенном увеличении вероятности битовой ошибки.

Для устранения этой проблемы в системах связи с OFDM применяются различные алгоритмы [2-11], как правило, основанные на содержащейся в OFDM сигнале избыточной информации: специальных вспомогательных символах [6] или циклическом префиксе [3,5,11]. Эффективно решая рассматриваемую проблему, подобные алгоритмы, вместе с тем, не могут применяться в системах с высокой спектральной эффективностью, в составе которых может отсутствовать подобного рода избыточность [1].

Основная цель настоящей работы состоит в сравнительном анализе эффективности функционирования наиболее распространенного способа оценки частотного сдвига, использующего циклический префикс, с методом оценки, основанным на анализе статистических свойств принимаемого сигнала OFDM. Последний метод не требует наличия в структуре сигнала OFDM циклического префикса или каких-либо вспомогательных символов и, следовательно, может быть применен в OFDM системах с высокой спектральной эффективностью.

Определение CFO на основе статистических характеристик OFDM сигнала

В простейшем случае одиночный OFDM символ во временной области можно представить в виде суммы из N гармонических сигналов с частотами ω_k, $k=\overline{1,N}$, каждый из которых взвешен комплексным квадратурно-модулированным отсчётом $x(k)$ [10]:

$$s(t)=\left[\sum_{k=1}^{N}x(k)e^{j\omega_k t}\right]\times g(t)= \\ =\left[\sum_{k=1}^{N}x(k)e^{j2\pi k\Delta f\, t}\right]\times rect\left(\frac{t}{NT_M}\right), \quad (1)$$

где T_M – период следования комплексных отсчетов $x(k)$, $\omega_k=2\pi k\Delta f=\frac{2\pi k}{NT_M}$ – частота k-й поднесущей, $\Delta f=1/NT_M$ – расстояние между поднесущими, $g(t)=rect(t/NT_M)$ – функция, формирующая огибающую OFDM символа. Нетрудно заметить, что (1) с точностью до постоянного множителя совпадает с выражением обратного преобразования Фурье, что означает непосредственное использование этого преобразования или его быстрых форм для практического формирования OFDM сигналов. Фурье преобразование от $g(t)$ представляет собой свертку функции sinc с дельта-функцией $\delta(f-f_k)$ и для k-й поднесущей может быть представлено в виде:

$$S_k(f)=\frac{\sin(\pi NT_M f/2)}{\pi NT_M f/2}*\delta(f-f_k)=\frac{\sin\left(\pi\frac{f-f_k}{\Delta f}\right)}{\pi\frac{f-f_k}{\Delta f}}. \quad (2)$$

При демодуляции OFDM символа над принятым сигналом выполняется прямое преобразование Фурье. При этом, случаю совпадения опорных частот в приемном устройстве с реальными значениями поднесущих ω_k ($k=\overline{1,N}$) в OFDM сигнале будут соответствовать максимальная амплитуда выделяемых сигналов всех поднесущих и отсутствие их взаимного влияния (рис.1).

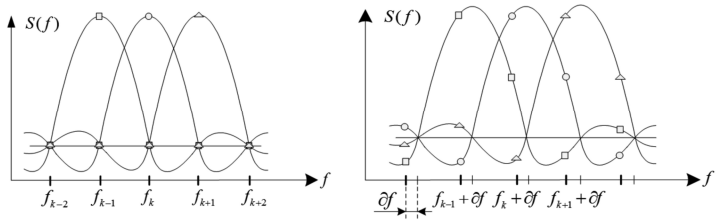


Рис.1

Если между опорными частотами приемного устройства и значениями поднесущих в принимаемом OFDM сигнале имеется рассогласование $f_k + \partial f$, то получаемые в результате демодуляции символы $\tilde{x}(k)$ будут искажены интерференцией (рис.1, справа).

Амплитуда интерференционной составляющей на *n*-й поднесущей при наличии рассогласования $f_k + \partial f$ будет определяться как

$$I_n = \left| \sum_{k \neq n} \tilde{x}(k) \cdot S_k(f_n + \partial f) \right| .$$

Поскольку комплексные отсчеты принятого квадратурно-модулированного сигнала $\tilde{x}(k)$ являются случайными величинами, то и амплитуда интерференционной составляющей на *n*-й поднесущей I_n также будет случайной величиной. Дисперсия интерферирующих поднесущих может быть представлена выражением [11]:

$$\sigma^2_{I_n} = \sum_{k \neq n} \sigma^2_{x(k)} \cdot \left| S_k(f_n + \partial f) \right|^2 ,$$

где $\sigma^2_{x(k)}$ - дисперсия отсчетов $\tilde{x}(k)$. Предполагая, что $\tilde{x}(k)$ поступают из канала связи с гауссовским белым шумом с дисперсией σ^2_n, нетрудно заметить, что наличие интерференционной составляющей в спектре OFDM сигнала приводит к отклонению формы закона распределения спектральных отсчетов от нормального с тенденцией уплощения вершины распределения [10] пропорционально росту ∂f (рис.2). Это обстоятельство позволяет использовать для оценки величины частотного смещения ∂f коэффициент эксцесса, отражающий степень островершинности эмпирического распределения по отношению к нормальному.

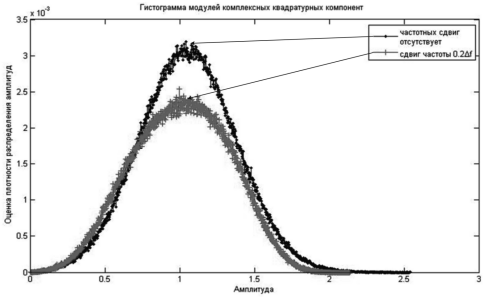


Рис. 2.

В простейшем случае, алгоритм оценки частотного смещения, основанный на использовании коэффициента эксцесса ($J(s)$) (далее алгоритм ST), можно представить в следующем виде [10]:

$$\hat{\varepsilon} = \min_{s \in (-N,N)} \{ J(s) \}, \tag{3}$$

$$\tilde{x}_{i,s}(k) = \sum\nolimits_{n=1}^{N} r'_i(n) e^{-j2\pi(f_n+s)k},$$

$$J(s) = \frac{\sum_{i=1}^{L} \sum_{k=1}^{N} \left| \tilde{x}_{i,s}(k) \right|^4 / L}{\left(\sum_{i=1}^{L} \sum_{k=1}^{N} \left| \tilde{x}_{i,s}(k) \right|^2 / L \right)^2}. \tag{4}$$

где $\hat{\varepsilon}$ - нормированное к Δf значение частотного сдвига; N – общее количество поднесущих в OFDM символе, L – количество символов OFDM, используемых для оценки, $r'_i(k)$ - комплексные отсчёты i – го символа OFDM после процедуры удаления циклического префикса, s – величина варьируемого частотного сдвига.

С целью упрощения расчетов для нахождения (3) можно воспользоваться методом аппроксимации (4) кривыми [10]. Для этого необходимо вычислить значения (4) для трех разных последовательностей $\tilde{x}_{i,s}(k)$: непосредственно принятой - $J(0)$ и смещенных на величину $\pm \Delta f / 4$ относительно $f_k = k\Delta f$ ($k = \overline{1, N}$) - $J(-1/4)$, $J(1/4)$. В этом случае решение (3) можно определить в виде [10]:

$$\hat{\varepsilon} = \begin{cases} \dfrac{1}{2\pi} \sin^{-1}(b), \; a \geq 0, \\ \dfrac{1}{2} - \dfrac{1}{2\pi} \sin^{-1}(b), \; a < 0, \; b \geq 0, \\ -\dfrac{1}{2} - \dfrac{1}{2\pi} \sin^{-1}(b), \; a < 0, \; b \leq 0. \end{cases} , \tag{5}$$

где $b = \left(J(1/4) + J(-1/4) \right) / 2$, $a = b - J(0)$.

Определение CFO на основе использования циклического префикса

Согласно [5], с целью нахождения $\hat{\varepsilon}$ для каждого принятого OFDM символа необходимо определить максимум следующей функции (далее алгоритм ML):

$$\Lambda(\hat{\varepsilon}, i) = \left| \gamma(i) \right| \cos(2\pi\hat{\varepsilon} + \angle\gamma(i)) - \rho\Phi(i) ,$$

$$i = \overline{1, (2 \cdot N_g - 1)} \tag{6}$$

где оператор $\angle$ обозначает аргумент комплексного числа,

$$\gamma(i) = \sum_{k=0}^{N_g - 1} r(k+i) r^*(k+i+N) , \tag{7}$$

$$\Phi(i) = \frac{1}{2} \sum_{k=0}^{N_g - 1} \left| r(k+i) \right|^2 + \left| r(k+i+N) \right|^2 , \tag{8}$$

$$\rho = \frac{\sigma^2_{x(k)}}{\sigma^2_{x(k)} + \sigma^2_n} , \tag{9}$$

где $r(k)$ - отсчеты принятого OFDM сигнала во временной области. Функция $\gamma(i)$ определяет взаимную корреляцию между N_g отсчётами циклического префикса и соответствующей части OFDM символа, а $\Phi(i)$ позволяет компенсировать влияние на решение средней мощности сигнала. Количество отсчётов на один OFDM символ определяется параметром N. Из (6) следует, что $\Lambda(\hat{\varepsilon}, i)$ достигает максимума при выполнении равенства

$$\hat{\varepsilon} = -\frac{1}{2\pi} \angle\gamma_{\max} . \tag{10}$$

Связь нормированной оценки $\hat{\varepsilon}$ с искомым значением частотного сдвига $\partial\hat{f}_{ML}$ определяется следующим выра-

жением [11]:

$$\partial\hat{f}_{ML} = \frac{(N+N_g)}{N}\hat{\varepsilon}\Delta f \,. \qquad (11)$$

Эксперимент

Моделирование алгоритмов оценивания частотной нестабильности было выполнено в среде Matlab. В качестве исследуемых OFDM сигналов использовались символы длительностью T_s=2,25 мс, включая циклический префикс T_g = Ts/8, состоящие из 256 поднесущих с Δf =440 Гц и модуляцией QPSK на каждой поднесущей.

В качестве каналов связи использовались модели многолучевого стационарного канала связи с аддитивным белым гауссовским шумом (АБГШ), нормированные профили которых представлены на рис. 3. Вычисление значения частотного сдвига выполнялось для двух рассмотренных выше методов. С целью повышения достоверности результатов использовалось усреднение оценки сдвига частоты по L = 10000 OFDM символам.

Точность определения оценок смещения частоты рассчитывалась в соответствии со следующим выражением [9]:

$$e = \frac{\sum_{i=1}^{N_{экс}}\left(\partial f_i - \partial\hat{f}_i\right)^2}{N_{экс}\Delta f^2}, \qquad (12)$$

где ∂f_i и $\partial\hat{f}_i$ - истинное и вычисленное в ходе эксперимента значения частотного сдвига.

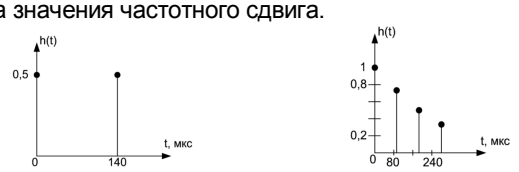


Рис. 3. а) двухлучевая модель; б) экспоненциальная модель

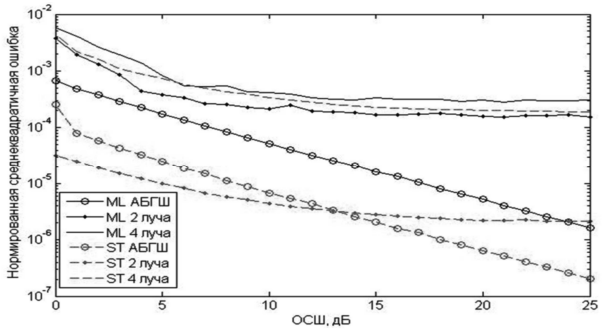


Рис. 4.

Графики зависимости нормированной среднеквадратичной ошибки e оценки смещения частоты от отношения сигнал-шум в канале связи для рассматриваемых методов представлены на рис. 4. Из анализа полученных зависимостей можно сделать вывод о том, что применение алгоритма ST в каналах с АБГШ по сравнению с ML даёт выигрыш в ОСШ порядка 10 дБ при нормированной СКО определения частотного сдвига $e = 10^{-5}$. В 2-х лучевом канале алгоритм ST оказывается значительно эффективнее и даёт в 100 раз более точную оценку смещения частоты, однако с увеличением количества лучей оба исследуемых алгоритма дают сопоставимые результаты.

Заключение

В настоящей статье проведено исследование точности определения частотного смещения OFDM сигнала по двум методам оценки. Был приведён сравнительный анализ представленных алгоритмов для различных моделей многолучевых каналов с АБГШ. Показано, что в условиях АБГШ и двухлучевой модели канала связи метод оценки частотного смещения, основанный на расчете коэффициента эксцесса, предоставляет несомненное преимущество по сравнению методом оценки, основанным на циклическом префиксе. Это наблюдение позволяет сделать вывод о том, что алгоритм ST может с успехом быть применен для решения задачи оценки частотного смещения в системах с OFDM, не использующих циклический префикс.

Литература

1. R.W. Chang, Synthesis of band-limited orthogonal signals for multichannel data transmission, Bell Systems Technical Journal, vol. 46, P. 1775–1796, 1966.

2. Schmidl, T.M. and Cox, D.C. Robust frequency and timing synchronization for OFDM. IEEE Trans. Commun., vol. 45(12), P. 1613–1621, 1997.

3. F. Daffara and A. Chouly, Maximum likelihood frequency detectors for orthogonal multicarrier systems, in Proceedings of the IEEE International Conference on Communications (ICC'93), P. 766-771 1993.

4. Скляр Б. Цифровая связь. Теоретические основы и практическое применение, Изд. 2-е, испр.: Пер. с англ. – М.: Издательский дом «Вильямс», 2004. – 1104 с.

5. J. J. van de Beek, M. Sandell, and P.O. Bojesson, ML estimation of time and frequency offset in OFDM systems, IEEE Transactions on Signal Processing, vol. 45, no. 7, P. 1800-1805, 1997.

6. Moose, P.H. A technique for orthogonal frequency division multiplexing frequency offset correction. IEEE Trans. Commun., vol 42, P. 2908–2914, 1994.

7. Classen, F. and Myer, H. Frequency synchronization algorithm for OFDM systems suitable for communication over frequency selective fading channels. IEEE VTC'94, P. 1655–1659, 1994.

8. H. Liu and U. Tureli. A high-efficiency carrier estimator for OFDM communications. IEEE Commun. Lett., 2(4):104-106, 1998.

9. F. Yang, K. H. Li, and K. C. Teh, A carrier frequency offset estimator with minimum output variance for OFDM systems, IEEE Communication Letters, vol. 8, no. 11, P. 677-679, 2004.

10. Yingwei Yao, Georgios B. Giannakis; Blind carrier frequency offset estimation in SISO, MIMO, and multiuser OFDM systems, IEEE Trans. Communications, vol. COM-53, pp. 173 - 183, 2005.

11. L. Hanzo, T. Keller. OFDM and MC-CDMA. A Primer. – IEEE Communicational Society. Sponsor John Wiley & Sons, Ltd.

RELATIVE ANALYSIS OF FREQUENCY DETECTION ALGORITHMS IN OFDM SYSTEMS

Ovinnikov A., Bakke A.

Two algorithms of carrier frequency offset estimation in time, based on the redundant information contained within the cyclic prefix, and "blind" assay in frequency domain for OFDM signal are considered. Analysis of there efficiency in different multipath channels allows to suppose that this synchronization techniques may be used in practical systems with a small modifications.

УДК 621.396.4

СИНТЕЗ КОМПЕНСАТОРОВ НА ОСНОВЕ НЕЙРОННОЙ СЕТИ ГАММЕРШТЕЙНА

Соловьева Е.Б., *д.т.н., завкафедрой теоретических основ электротехники Санкт-Петербургского государственного электротехнического университета «ЛЭТИ» им. В.И. Ульянова (Ленина), e-mail: selenab@hotbox.ru.*

Ключевые слова: нелинейный компенсатор, нейронная сеть, персептрон, модель Гаммерштейна, канал связи.

Рассмотрены модели нелинейных компенсаторов: многомерный полином, двухслойный персептрон, сеть Элмана, нейронная модель Гаммерштейна. Показано, что нейронный компенсатор с рекуррентной моделью Гаммерштейна, применяемый для подавления нелинейных искажений сигналов в Винеровском канале связи, превосходит модели-аналоги по точности обработки сигналов и простоте реализации. Представлены оценки точности компенсации при действии в канале связи гауссовского шума.

Введение

Нелинейные искажения порождают многие нежелательные эффекты и являются причиной ухудшения ряда показателей качества радиоэлектронных систем. Вызывая появление высших и интермодуляционных составляющих в реакции устройств, они ухудшают качество звучания электроакустической аппаратуры, создают препятствия на пути решения проблемы электромагнитной совместимости в приборостроении, снижают точность воспроизведения сигналов, помехозащищенность, разрешающую и пропускную способности в системах связи, технике СВЧ и других областях радиотехники.

Традиционные методы борьбы с нелинейными искажениями, такие, как линейная инверсия, линеаризация характеристик отдельных нелинейных элементов и т. п., не всегда эффективны, что обусловлено высокой степенью интеграции технических устройств, расширением области задач прецизионной обработки сигналов.

На практике предпочтение отдается методам нелинейной компенсации, применение которых позволяет повысить качество технических устройств при заданном уровне развития технологии их производства. Методы синтеза компенсаторов можно разделить на два класса:

– методы слепой линеаризации, выполняемой без «обучения» компенсатора. К данным методам относятся: инверсия высокого порядка, метод фиксированной точки, корней уравнения Вольтерры, итерационно-операторным методом [1].

– методы линеаризации устройств с «обучением» компенсатора, основанные на применении функциональных рядов и полиномов [1], [2]; нейронных сетей (персептронных [3], [4], рекуррентных [5], [6], радиальных [7] и т. д.).

Нелинейный компенсатор (НК) может быть установлен либо на выходе искажающей системы (посткомпенсатор), либо на ее входе (предкомпенсатор). Сформулируем задачу синтеза НК.

Пусть нелинейная система (НС) описана операторным уравнением

$$y(n) = H[x(n)],$$

где $x(n)$, $y(n)$ – входной и выходной сигналы НС соответственно, H – нелинейный оператор системы. Необходимо построить нелинейный оператор F компенсатора, действующий на операторное уравнение НС так, чтобы выполнялись соотношения:

$$\tilde{x}(n) = F[y(n)] = F[\,H[x(n)]\,] = R[x(n)] \approx x(n) \qquad (1)$$

в случае посткомпенсатора;

$$\tilde{x}(n) = H[\,F[x(n)]\,] = R[x(n)] \approx x(n)$$

в случае предкомпенсатора. Здесь R – линейный оператор результирующей системы (каскадного соединения НС и НК либо НК и НС). Для простоты преобразований полагаем: $R \cong 1$.

Таким образом, цель компенсации – ввести нелинейный оператор F так, чтобы результирующая система описывалась линейным оператором R.

В работе предлагается использовать нейронную модель Гаммерштейна для компенсации нелинейных искажений сигналов в цифровом канале связи (КС), описываемом низкочастотной моделью Винера; выполняется сравнение моделей НК по точности восстановления сигналов и вычислительным затратам на компенсацию.

Полиномиальная и нейронные модели компенсаторов

Операторные методы синтеза НК инвариантны к схеме подключения компенсатора, поэтому рассмотрим нелинейные модели НК на примере посткомпенсатора с уравнением (1).

Многомерный полином (МП) [1], [2]:

$$\tilde{x}(n) = P^{I}[\vec{y}(n)] =$$

$$= \sum_{i_1=0}^{I_1} \sum_{i_2=0}^{I_2} \cdots \sum_{i_m=0}^{I_m} C_{i_1 i_2 \ldots i_m} [y_1(n)]^{i_1} [y_2(n)]^{i_2} \ldots [y_m(n)]^{i_m},$$

где I – степень полинома, $I = I_1 + I_2 + \ldots + I_m$.

Двухслойный персептрон (ДП) [3], [4]:

$$\tilde{x}(n) = G(u^{2}(n)),$$

где $u^{(2)}(n) = \sum_{k=0}^{J} c_k \mathrm{net}_k^{(1)}(n)$; $\mathrm{net}_0^{(1)}(n) = 1$;

$$\mathrm{net}_k^{(1)}(n) = G(u_k^{(1)}(n)); \; u_k^{(1)}(n) = \sum_{l=0}^{m} w_{kl} y_l(n), \; k = 1, 2, \ldots, J;$$

J – число нейронов в скрытом слое; G – функция активации (гиперболический тангенс).

Блок-схема ДП изображена на рис. 1.

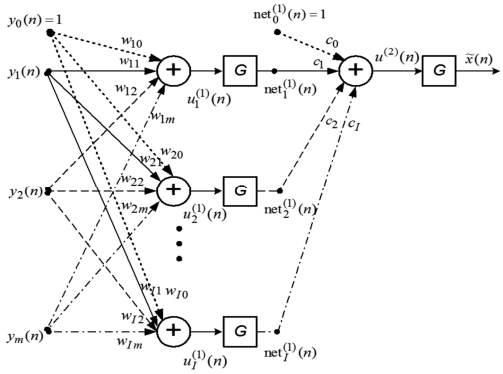


Рис. 1.

Нейронная сеть Элмана (НСЭ) [8], описываемая уравнением

$$\tilde{x}(n) = G\big(g(n)\big),$$

где $g(n) = \sum_{k=0}^{I} c_k v_k(n)$; I – число нейронов в скрытом слое; $v_0(n) = 1$; $v_k(n) = G\big(s_k(n)\big)$;

$s_k(n) = \sum_{l=1}^{I} w_{kl} v_l(n-1) + \sum_{l=I+1}^{I+m} w_{kl} y_{l-I}(n)$, $k = 1, 2, ..., I$, и имеющая структуру, изображенную на рис. 2.

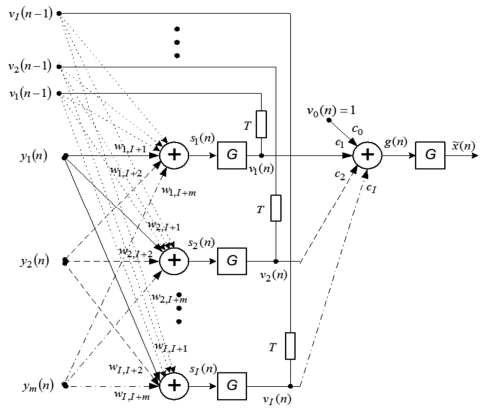


Рис. 2.

Нейронная модель Гаммерштейна (НМГ) [9]:

$$\tilde{x}(n) = \sum_{r_b=0}^{R_b} (\mathrm{vb})_{r_b} \mathrm{net}^{(2)}(n - r_b) + \sum_{r_a=1}^{R_a} (\mathrm{va})_{r_a} \tilde{x}(n - r_a),$$

где $\mathrm{net}^{(2)}(n) = \sum_{k=0}^{I} c_k \mathrm{net}_k^{(1)}(n)$; I – число нейронов в 1-м слое; $\mathrm{net}_0^{(1)}(n) = 1$; $\mathrm{net}_k^{(1)}(n) = G\big(u_k^{(1)}(n)\big)$;

$u_k^{(1)}(n) = \sum_{l=0}^{m} w_{kl} y_l(n)$, $k = 1, 2, ..., I$; $\max\{R_a, R_b\}$ – длина памяти модели.

Подобно НМГ строится нейронная модель Винера (НМВ) [9], в которой каскадно соединены линейная рекурсивная цепь и безынерционная нелинейность. Отметим, что в НМГ очередность следования указанных частей обратная.

Вектор воздействий для ДП, МП, НСЭ и НМГ часто формируется с помощью линии задержки, на выходе которой вектор $\vec{y}(n)$ имеет вид

$$\vec{y}(n) = \big[y_0(n), y_1(n), y_2(n), ..., y_m(n)\big] =$$
$$= \big[1, y(n), y(n-1), ..., y(n-(m-1))\big], \qquad (2)$$

где $(m-1)$ – длина памяти.

НСЭ, НМГ, НМВ являются рекуррентными нейронными моделями компенсаторов.

Блок-схема НМГ показана на рис. 3.

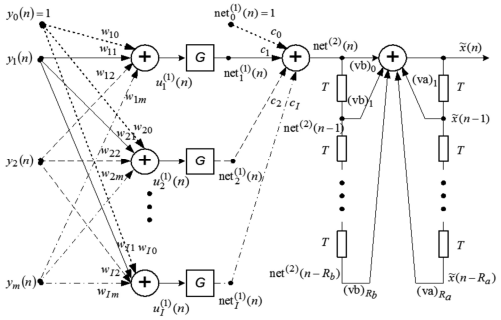


Рис. 3.

Компенсация нелинейных искажений сигналов в модели Винера канала связи

Модель Винера КС описывает каскадное соединение двух блоков [6]:

– линейной динамической цепи с передаточной функцией

$$H(z) = (1{,}0119 - 0{,}7589j) + (-0{,}3796 + 0{,}5059j) \cdot z^{-1};$$

– безынерционной нелинейности

$$y(n) = d_1 \vartheta(n) + d_2 \vartheta^2(n) + d_3 \vartheta^3(n),$$

где $\vartheta(n)$, $y(n)$ – выходные сигналы линейного блока и модели Винера КС соответственно; $d_1 = 1$; $d_2 = 0{,}2$; $d_3 = 0{,}1$.

Для компенсации нелинейных искажений сигналов в модели Винера КС использованы: ДП ($m = 5$), МП ($m = 10$), НСЭ ($m = 7$), НМГ ($m = 1$, $R_a = R_b = 1$).

Исследования выполнены на классах низкочастотных сигналов с восьмипозиционной фазовой модуляцией (8PSK) и четырехпозиционной квадратурной амплитудной модуляцией (4QAM).

При компенсации нелинейных искажений сигналов оценивались:

– равномерная погрешность

$$\delta(n) = \tilde{x}(n) - x(n), \; n \in [Q_n, Q];$$

– максимальная абсолютная погрешность

$$\delta_m = \max_{n \in [Q_n, Q]} |\delta(n)|;$$

– среднеквадратичная погрешность

$$\varepsilon = \frac{1}{Q - Q_n} \sqrt{\sum_{n=Q_n}^{Q} |\delta(n)|^2},$$

где $Q_n = 3$, $Q = 1000$, $x(n)$ – низкочастотная огибающая модулированного входного сигнала КС.

При 8PSK-воздействии КС зависимости $\delta_m(I)$ и $\varepsilon(I)$, полученные разными моделями компенсаторов: ДП (кривая *1*), МП (кривая *2*), НСЭ (кривая *3*), НМГ (кривая *4*), показаны на рис. 4, *а*, *б* соответственно. При 4QAM-воздействии КС аналогичные кривые представлены на рис. 5, *а*, *б*.

На рис. 6, *а–г* при 8PSK-воздействии КС изображены созвездия выходных сигналов следующих моделей компенсаторов: ДП с $I = 5$, $m = 5$; МП с $I = 3$, $m = 10$; НСЭ с $I = 5$, $m = 7$; НМГ с $I = 3$, $R_a = R_b = 1$ при $\vec{y}(n) = \big[1, y(n)\big]$ ($m = 1$ в (2)). В центре комплексных плоскостей на рис. 6 изображены равномерные погрешности $\delta(n)$, $n \in [Q_n, Q]$.

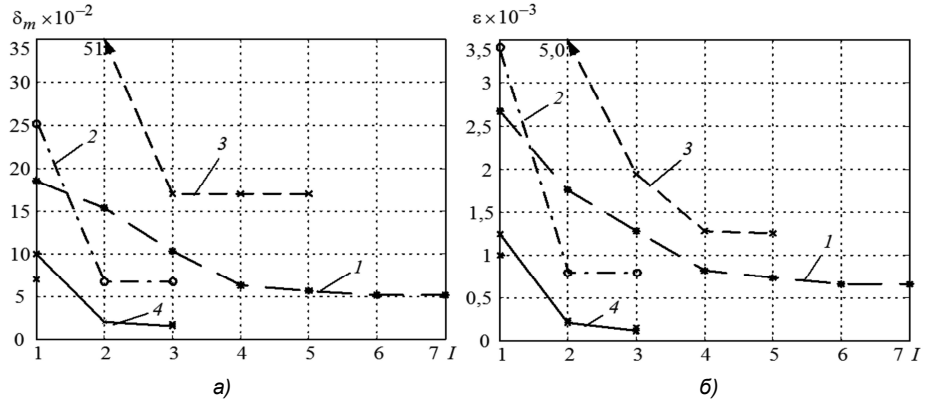


Рис. 4.

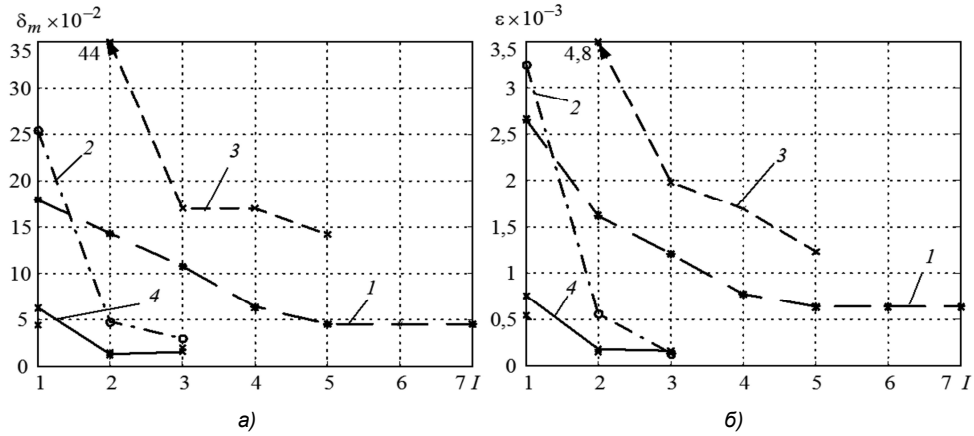


Рис. 5.

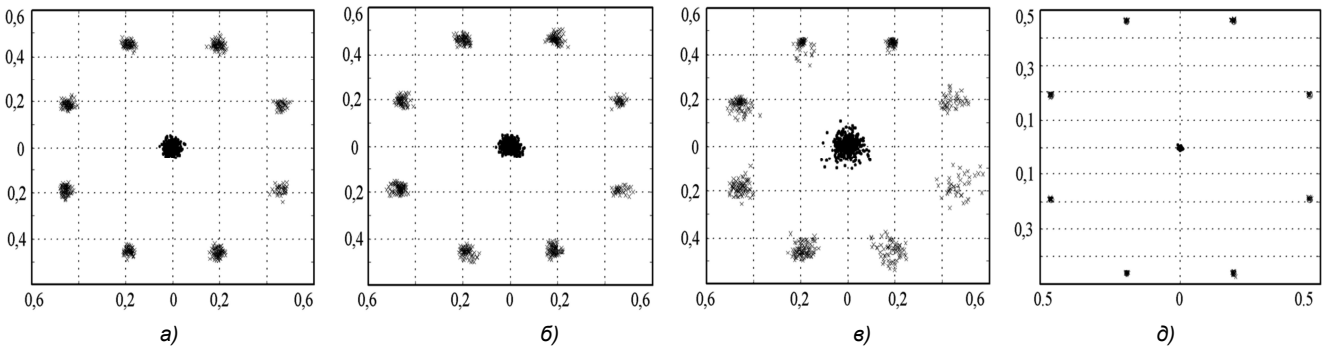


Рис. 6.

Из анализа рис. 4–6 следует, что НМГ с $I=3$, $R_a=R_b=1$, $m=1$ дает наименьшие погрешности компенсации в равномерной и среднеквадратичной метриках.

В табл. 1 указано число параметров в используемых моделях компенсаторов.

Таблица 1.Число параметров в моделях компенсаторов

ДП с $I=5$, $m=5$	МП с $I=3$, $m=10$	НСЭ с $I=5$, $m=7$	НМГ с $I=3$, $R_a=R_b=1$, $m=1$
36	286	66	16

Из табл. 1 видно, что НМГ является более простой моделью по сравнению с ее аналогами.

Компенсация нелинейных искажений сигналов в низкочастотной модели Винера выполнена при действии в КС аддитивной комплексной гауссовской помехи $\xi(n)$. Компенсаторы синтезированы в виде ДП, МП и НМГ (НСЭ исключена из рассмотрения, поскольку не обеспечивает высокую точность компенсации, см. рис. 4–6).

В табл. 2, 3 представлены среднеквадратичная погрешность ε и частота появления ошибочных битов BER (Bit Error Rate), вычисляемая по формуле:

$$\mathrm{BER}=\frac{q}{(Q-Q_n)\log_2 M},$$

где q – количество ошибочных битов в последовательности, полученной при детектировании выходного сигнала $\tilde{x}(n)$ компенсатора; $(Q-Q_n)$ – длина последовательности $\tilde{x}(n)$; M – позиционность модуляции, $M=8$ при 8PSK и $M=4$ при 4QAM. Значения ε и BER указаны при разных SNR, где SNR – отношение сигнал/шум (Signal to Noise Ratio), дБ:

$$\mathrm{SNR}=20\lg\left(\sqrt{\sum_{n=Q_n}^{Q}|\tilde{x}(n)|^2}\Big/\sqrt{\sum_{n=Q_n}^{Q}|\xi(n)|^2}\right).$$

Как следует из анализа табл. 2, 3:

При 8PSK-воздействии канала связи

– НМГ дает более высокую точность компенсации по

сравнению с другими рассмотренными моделями;

– указанный результат сохраняется при действии в КС гауссовской помехи с $\mathrm{SNR} > 20$ дБ; при $\mathrm{SNR} \le 20$ дБ погрешности, полученные НМГ, ДП и МП приблизительно одинаковые;

– на практике целесообразно использовать НМГ, так как она более проста по сравнению с моделями-аналогами.

При 4QAM-воздействии канала связи

– НМГ и МП дают одинаковую точность компенсации, более высокую по сравнению с ДП и НСЭ;

– указанный результат сохраняется при действии в КС гауссовской помехи с $\mathrm{SNR} > 25$ дБ. В этом случае целесообразно применять НМГ –более простую модель компенсатора, чем МП;

– при $\mathrm{SNR} \le 25$ дБ меньшую погрешность компенсации обеспечивает ДП.

Таблица 2. ε и BER при 8PSK-воздействии КС

SNR, дБ	ДП с $I=5$, $m=5$; МП с $I=3$, $m=10$		НГМ с $I=3$, $R_a=R_b=1$, $m=1$	
	ε, $\times 10^{-3}$	BER, $\times 10^{-2}$	ε, $\times 10^{-3}$	BER, $\times 10^{-2}$
10	6,09	6,22	6,02	6,02
15	3,81	0,98	3,74	0,94
20	2,31	0	2,20	0
25	1,43	0	1,26	0
30	1,03	0	0,76	0
35	0,82	0	0,48	0

Таблица 3. ε и BER при 4QAM-воздействии КС

SNR, дБ	ДП с $I=5$, $m=5$		МП с $I=3$, $m=10$; НГМ с $I=3$, $R_a=R_b=1$, $m=1$	
	ε, $\times 10^{-3}$	BER	ε, $\times 10^{-3}$	BER, $\times 10^{-2}$
10	2,15	0	4,39	0,91
15	0,47	0	2,49	0
20	0,90	0	1,37	0
25	0,72	0	0,82	0
30	0,66	0	0,53	0
35	0,66	0	0,36	0

Заключение

Наряду с хорошо известными полиномиальными моделями Винера и Гаммерштейна, в которых безынерционная нелинейность описывается полиномом, существуют нейронные типы указанных моделей, содержащие нейронную сеть в качестве безынерционной нелинейности. Нейронные модели Винера и Гаммерштейна включают также линейную динамическую цепь, в общем случае являющуюся рекурсивной системой. В результате модели Винера и Гаммерштейна могут быть построены в виде рекуррентных нейронных сетей.

В процессе исследований установлено, что компенсатор в виде нейронной модели Гаммерштейна, применяемый для подавления нелинейных искажений сигналов в Винеровском канале связи, превосходит традиционные полиномиальные и нейронные модели по точности обработки сигналов и простоте реализации.

Литература

1. Соловьева Е. Б. Макромоделирование нелинейных цепей и синтез операторов.– СПб.: Изд-во СПбГЭТУ «ЛЭТИ», 2010.

2. Соловьева Е. Б. Компенсация нелинейных искажений сигналов в каналах связи методом расщепления // Цифровая обработка сигналов.– 2005.– № 4.– С. 2–8.

3. Martens J.-P., Weymaere N. An equalized error back-propagation algorithm for the on-line training of multilayer perceptrons // IEEE Trans. Neural Networks. – 2002.– Vol. 13, № 3.– P. 532–541.

4. Ланнэ А. А., Соловьева Е. Б. Моделирование нелинейных дискретных систем на основе персептрона с расщепителем // Цифровая обработка сигналов.– 2006.– № 3.– С. 2–7.

5. Choi J., Bouchard M., Yeap T. H. Decision feedback recurrent neural equalization with fast convergence rate // IEEE Trans. Neural Networks.– 2005.– Vol. 16, № 3.– P. 699–708.

6. Нам Л. Х. Рекуррентный нейросетевой эквалайзер с алгоритмом расширенного фильтра Калмана // Нейрокомпьютеры: разработка и применение.– 2006.–№ 2.– С. 71–79.

7. Jianping D., Sundararajan N., Saratchandran P. Communication channel equalization using complex-valued minimal radial basis function neural networks // IEEE Trans. Neural Networks.– 2002.– Vol. 13, № 3.– P. 687–696.

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

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

SYNTHESIS OF COMPENSATORS ON THE BASIS OF NEURAL HAMMERSTEIN NETWORK

Solovyeva E. B.

Models of nonlinear compensators are considered as a multidimensional polynomial, two-layer perceptron, Ellman network, neural Hammerstein model. The neural Hammerstein model, used for suppression of nonlinear distortions in Wiener communication channel, is shown to surpass other considered models in accuracy of signal processing and simplicity of realization. The estimation of compensation accuracy are represented at Gaussian noise in communication channel.

УДК 621.395

ПОДАВЛЕНИЕ ШУМА В РЕЧЕВЫХ СИГНАЛАХ НА ОСНОВЕ МЕТОДА НЕЛОКАЛЬНОГО УСРЕДНЕНИЯ

Новоселов С.А., *аспирант кафедры динамики электронных систем Ярославского государственного университета им. П.Г. Демидова, e-mail: sergnovoselov@mail.ru;*

Топников А.И., *аспирант, ассистент кафедры динамики электронных систем Ярославского государственного университета им. П.Г. Демидова, e-mail: topnikov@gmail.com;*

Савватин А.И., *аспирант кафедры динамики электронных систем Ярославского государственного университета им. П.Г. Демидова, e-mail: dcslab@uniyar.ac.ru;*

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

Описывается метод нелокального усреднения, анализируется возможность его применения в задаче подавления шума в речевых сигналах. Для демонстрации эффективности подхода представлены результаты моделирования, полученные с использованием алгоритма, основанного на нелокальном усреднении. Проведено сравнение полученных результатов с результатами, полученными для алгоритма, реализующего метод спектрального вычитания.

Ключевые слова: подавления шума, речевые сигналы, гауссовский шум, нелокальное усреднение.

Введение

Речевые сигналы, с которыми приходится иметь дело на практике, всегда в той или иной степени зашумлены. При наличии внешних шумов значительной интенсивности результаты анализа и распознавания речи существенно ухудшаются, поэтому разработка методов шумоподавления для речевых сигналов является актуальным направлением исследований.

На сегодняшний день в области цифровой обработки речевых сигналов существует множество различных алгоритмов шумоподавления. Наиболее распространенной моделью шума для рассматриваемых алгоритмов является аддитивный белый гауссовский шум (АБГШ). Классификацию методов повышения качества и разборчивости речи удобнее всего проводить согласно основной концепции и идее, лежащей в основе метода. На современном этапе наибольшее распространение получили следующие группы методов [1-4]:

– методы, основанные на использовании авторегрессионной модели речевого сигнала;

– методы, основанные на обработке речевого сигнала с использованием аппарата скрытых марковских моделей;

– методы, основанные на искусственных нейронных сетях;

– методы, основанные на оценке параметров шума, минимизации среднеквадратичной ошибки и пороговой обработке в области трансформант.

Существуют и другие подходы. Например, сравнительно недавно в области цифровой обработки речевых сигналов появилось ещё одно направление – слепое разделение речевых сигналов (Blind Speech Separation), объединяющее методы, позволяющие получить оценки сигналов источников, основываясь лишь на знании сигналов смесей и некоторых общих предположениях о свойствах сигналов источников и процессе смешивания [5-8]. В ряде случаев эти методы позволяют достичь значительного улучшения качества и разборчивости речи. В области цифровой обработки изображений (ЦОИ) также существует множество методов шумоподавления [9-14], часть из которых можно распространить и на обработку речевых сигналов. Предложенный в статье алгоритм базируется на хорошо зарекомендовавшем себя в области ЦОИ [11-14] методе нелокального усреднения (Non-Local Means, NLM).

Метод нелокального усреднения

Кратко опишем метод нелокального усреднения, используемый в задаче подавления шума в изображениях [14]. Пусть цифровое изображение $\mathbf{x}$ искажено аддитивным белым гауссовским шумом $\mathbf{n}$ с нулевым математическим ожиданием и дисперсией σ^2.

1. Для обрабатываемого пикселя i зашумленного изображения $\mathbf{y} = \mathbf{x} + \mathbf{n}$ описываем квадратную окрестность фиксированного размера, центрированную на этот пиксель.

2. Определяем подобность обрабатываемого пикселя i зашумленного изображения $\mathbf{y}$ с пикселем j того же изображения, используя взвешенное евклидово расстояние $\left\| y(N_i) - y(N_j) \right\|_{2,a}^2$, где N_i и N_j – квадратные окрестности, центрированные на пиксели i и j соответственно, a – положительное число, определяющее среднеквадратическое отклонение гауссова ядра, используемого для вычисления взвешенного евклидова расстояния.

3. Определяем вес подобного к i пикселя j в итоговой оценке пикселя i:

$$w(i,j) = \frac{1}{Z(i)} \cdot e^{-\frac{\left\| y(N_i) - y(N_j) \right\|_{2,a}^2}{h^2}}, \quad Z(i) = \sum_j e^{-\frac{\left\| y(N_i) - y(N_j) \right\|_{2,a}^2}{h^2}},$$

где h – параметр, влияющий на степень фильтрации цифрового изображения.

4. Формируем итоговую оценку пикселя i на основе следующего выражения:

$$\hat{x}(i) = \sum_j w(i,j) y(j).$$

Так как спектрограмму звукового сигнала можно рассматривать как изображение, то это позволяет адаптировать алгоритмы шумоподавления, предназначенные для изображений, для обработки звуковых сигналов. Подобный подход позволяет, в частности, значительно снизить уровень «музыкального шума», который является основным артефактом работы алгоритмов, реализующих метод спектрального вычитания [15]. Также возможно применение метода нелокального усреднения как к реальной, так и к мнимой частям спектра звукового сигнала для шумоподавления с целью повышения эффективности системы распознавания [16].

Однако структура речевого сигнала во временной области позволяет сделать предположение о возможности применения метода нелокального усреднения непосредственно к одномерному сигналу. Действительно, в пределах интервала стационарности речевой сигнал состоит из последовательности похожих фрагментов, аналогично тому, как похожие фрагменты могут располагаться в различных частях изображения. Обоснование возможности применения метода нелокального усреднения к речевым сигналам, а также достижимость положительного эффекта в задаче шумоподавления были продемонстрированы в работе [17].

Настоящая работа направлена на дальнейшее развитие и исследование возможности подавления шума в речевых сигналах на основе применения метода нелокального усреднения во временной области. Особое внимание уделяется методикам поиска похожих фрагментов речевого сигнала, так как качество выполнения этого этапа в значительной степени определяет конечную эффективность алгоритма.

Описание предлагаемого алгоритма

Структура отрезков речевого сигнала, приведенная на рис.1, позволяет сделать вывод о том, что речь по своей природе, в некотором смысле, является примером «помехоустойчивого кода». Как видно из осциллограмм, помехоустойчивость речи достигается за счет избыточного повторения «информационных блоков» (это свойство особенно характерно для гласных звуков). Возможно, звуковой анализатор человека использует подобную избыточность и благодаря этому способен выделять информацию на фоне достаточно интенсивных помех (-10 дБ). Эти положения позволяют рассматривать и применять следующую модель речевого сигнала: речь – периодический на интервалах стационарности процесс, состоящий из повторяющихся «информационных блоков» или фрагментов. В рамках этой модели для решения задачи шумоподавления в речевых сигналах эффективно использовать алгоритм нелокального усреднения. При адаптации алгоритма нелокального усреднения для задачи подавления АБГШ в речевом сигнале учитываются особенности речи, благодаря чему получившийся алгоритм значительно отличается от своего двумерного аналога. Например, усреднение производится непосредственно для фрагмента длиной несколько десятков отсчетов, а не для одного отсчета. Рассмотрим особенности работы предлагаемого алгоритма подавления АБГШ в речевых сигналах.

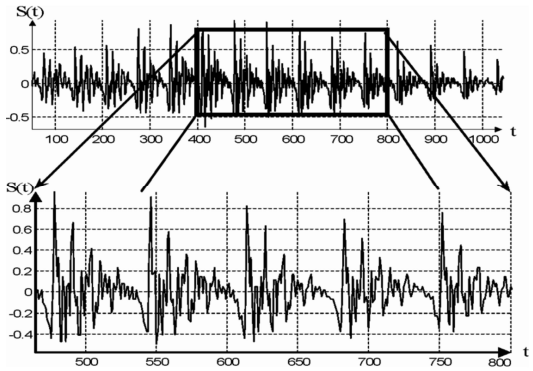
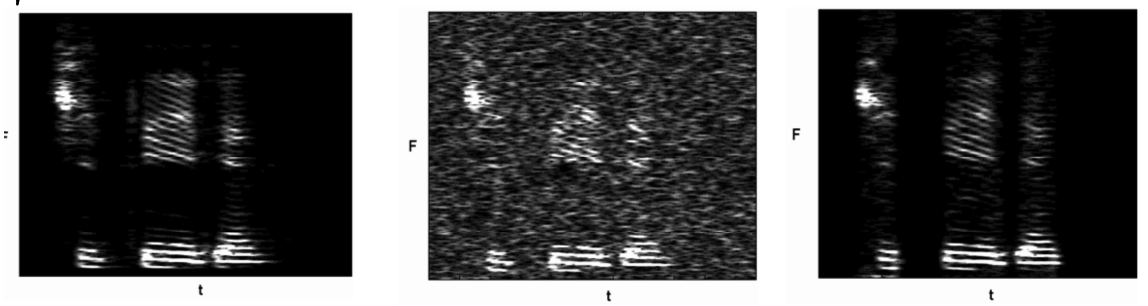


Рис. 1. Структура речевого сигнала

Обработка речевого сигнала осуществляется во временной области. Можно выделить два основных этапа: поиск похожих фрагментов и усреднение этих фрагментов, позволяющие значительно снизить дисперсию аддитивного шума. Длина фрагмента выбирается в пределах интервала 40-80 отсчетов (установлено экспериментально). Поиск похожих фрагментов ведется в пределах окна длиной 400 отсчетов, так как при частоте дискретизации 8000 отсчетов в секунду это соответствует 50 мс, то есть сравнимо с интервалом стационарности речевого сигнала, который примерно равен 15-30 мс. Для определения степени близости между фрагментами используется евклидово расстояние. Как будет показано ниже, локальные минимумы функции расстояния между выбранным блоком и остальными соответствуют положениям наиболее похожих на данный блок фрагментов. Теоретически чем больше число идентичных фрагментов усредняется, тем сильнее подавляется аддитивный шум, однако для речевых сигналов их число ограничено, к тому же их нельзя считать строго идентичными. На практике число таких фрагментов для речевого участка длиной 50 мс равно 4-12 (рис. 1), так как усреднение меньшего числа фрагментов малоэффективно, а поиск большего числа похожих фрагментов ограничен структурой речевого сигнала. После того как интересующие фрагменты найдены, происходит их усреднение и замена отсчетов сигнала, соответствующих этим отрезкам, на отсчеты, полученные в результате процесса усреднения. При низких значениях отношения сигнал / шум (ОСШ) процесс поиска похожих фрагментов значительно осложнен. Для повышения эффективности алгоритма имеет смысл формировать опорный сигнал путем подавления шума в исходном сигнале при помощи метода спектрального вычитания. Этот сигнал необходим лишь для поиска похожих фрагментов.

Выбор опорного сигнала. В предлагаемом алгоритме шумоподавления определяющим фактором является выбор опорного сигнала, по которому проводится поиск похожих фрагментов на интервалах речи длительностью 50 мс. Исследования показывают, что если в качестве опорного сигнала взять сам исходный малозашумленный сигнал (ОСШ > 30 дБ) и выбирать похожие фрагменты на основе простого порогового сравнения евклидового расстояния, то можно добиться значительного снижения уровня шума при сохранении почти всех особенностей речевого сигнала (рис. 2). Этот факт подтверждает адекватность выбранной модели сигнала.



a) исходный сигнал *b) зашумленный сигнал (ОСШ = 0 дБ)* *c) нелокальное усреднение зашумленного сигнала на основе исходного*

Рис. 2. Спектрограммы сигналов

На практике использовать исходный незашумленный сигнал не представляется возможным. В работе предлагается использовать в качестве опорного сигнал, полученный из зашумленного сигнала путем шумоподавления с помощью метода спектрального вычитания (Spectral Subtraction, SS), описанного в статье [2]. При этом коэффициент сглаживания β в этом алгоритме следует выбирать так, чтобы максимально сохранить особенности речевого сигнала (шипящие звуки, глухие согласные и др.) при вынужденном сохранении остаточного шума. Экспериментально установлено, что целесообразно выбирать значения β ≈ 0.7-0.8. При таких значениях, ослабление исходного шума составляет 7-8 дБ. Остаточный шум почти равномерно распределен по сигналу и по свойствам схож с АБГШ (рис. 3).

Поиск похожих фрагментов. Фрагментом или блоком стационарного отрезка речевого сигнала будем называть участок речи, протяженностью в 5 мс, находящийся внутри этого отрезка. При частоте дискретизации 8000 отсчетов в секунду 5 мс соответствует 40 отсчетам. Фрагменты выбираются с максимальным перекрытием в 39 отсчетов. Как уже отмечалось, длина самого отрезка равна 50 мс, что соответствует 400 отсчетам. Таким образом, в одном отрезке содержится L = 361 фрагментов.

Итак, опорный и зашумленный сигнал одинаковым образом разбиваются на перекрывающиеся отрезки длиной 50 мс. Величина перекрытия составляет 90% от размера окна.

Для общности обозначим фрагменты текущего отрезка речевого сигнала векторами $\vec{S}_i$, соответствующие им шумовые фрагменты текущего отрезка опорного сигнала – векторами $\vec{n}_i$, а соответствующие им шумовые фрагменты текущего отрезка исходного обрабатываемого сигнала – векторами $\vec{N}_i$.

Похожесть речевых фрагментов-векторов определяется евклидовым расстоянием между ними. Речевые фрагменты $\vec{S}_i$ и $\vec{S}_j$ будем считать похожими, если евклидово расстояние между ними не превышает некоторого порогового значения:

$$\left\|\vec{S}_j - \vec{S}_i\right\| \le t_{исх}.$$

Этот порог соответствует допустимым незначительным отклонениям похожих речевых блоков друг от друга.

Требуется для некоторого выбранного вектора $\vec{S}_j + \vec{n}_j$, $j \in [1;L]$, из всего набора фрагментов $\{M_{\vec{n}} : \vec{S}_i + \vec{n}_i\}$ опорного сигнала найти такое множество векторов $\{O_{\vec{n}} : \vec{S}_o + \vec{n}_o\}$, что $\forall \vec{S}_o : \left\|\vec{S}_j - \vec{S}_o\right\| \le t_{исх}$, т.е. любой $\vec{S}_o$ является похожим на $\vec{S}_j$. Рассмотрим евклидово расстояние между зашумленным фрагментом $\vec{S}_j$ и произвольным вектором $\vec{S}_i$. Используя неравенство треугольника, получим:

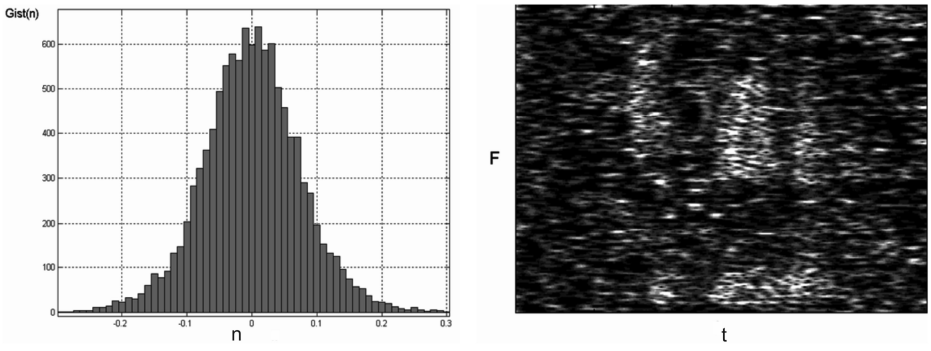
$$\left\|(\vec{S}_j + \vec{n}_j) - (\vec{S}_i + \vec{n}_i)\right\| \le \left\|\vec{S}_j - \vec{S}_i\right\| + \left\|\vec{n}_j - \vec{n}_i\right\|.$$

Если фрагменты $\vec{S}_j$ и $\vec{S}_i$ похожи, то положим, что $\left\|\vec{S}_j - \vec{S}_i\right\| \le t_{исх} << \left\|\vec{n}_j - \vec{n}_i\right\|$. Дополнительно, учитывая стационарность шума (рис. 3b) и используя решение детектора речевой активности, можно оценить максимальную амплитуду шумового вектора $\vec{n}_i$ на интервалах, где отсутствует речевая активность $\left\|\vec{n}_i\right\| \le \left\|\vec{n}\right\|_{\max}$. Тогда имеем:

$$\left\|(\vec{S}_j + \vec{n}_j) - (\vec{S}_i + \vec{n}_i)\right\| \le \left\|\vec{n}_j - \vec{n}_i\right\| \le 2\left\|\vec{n}\right\|_{\max}.$$

Последнее неравенство определяет условия похожести двух зашумленных речевых фрагментов и задает порог, который необходимо применять для нахождения множества $\{O_{\vec{n}} : \vec{S}_o + \vec{n}_o\}$:

$$\vec{S}_i + \vec{n}_i \in \{O_{\vec{n}}\} : \left\|(\vec{S}_j + \vec{n}_j) - (\vec{S}_i + \vec{n}_i)\right\| \le 2\left\|\vec{n}\right\|_{\max}.$$



а) *б)*

Рис. 3. Гистограмма а) и спектрограмма, b) остаточного шума

Такой метод поиска приводит к некоторому приближению множества похожих фрагментов. Однако следует иметь в виду, что смежные блоки выбранного фрагмента речевого сигнала являются «ложно» похожими на сам фрагмент из-за высокой корреляции соседних отчетов сигнала. Если их не исключить из рассмотрения и произвести оценку незашумленного блока путем нелокального усреднения по найденному множеству зашумленных векторов, то получим сглаженную копию искомого, потеряв при этом часть информации. Для исключения «ложных» фрагментов из множества $\{O_{\vec{n}} : \vec{S}_o + \vec{n}_o\}$ предлагается находить позиции локальных минимумов функции расстояния между векторами. Эти позиции соответствуют положениям пригодных для усреднения фрагментов (рис. 4). Таким образом, из рассмотрения исключаются окрестности минимумов, соответствующие «ложным» фрагментам.

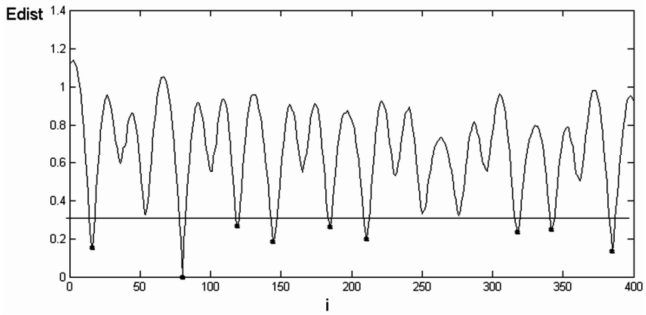


Рис. 4. Евклидово расстояние между текущим и другими фрагментами зашумленного отрезка речевого сигнала

Нелокальное усреднение. После того как множество $O_{\vec{n}}$ для заданного фрагмента $\vec{S}_j + \vec{n}_j$ найдено, с помощью гауссова ядра w_o реализуется процесс взвешенного усреднения множества векторов $\{O_{\vec{N}} : \vec{S}_o + \vec{N}_o\}$ отрезка обрабатываемого зашумленного сигнала, соответствующего векторам множества $\{O_{\vec{n}} : \vec{S}_o + \vec{n}_o\}$ опорного сигнала, с целью нахождения оценки $\vec{S}_j$:

$$\vec{S}_j{}' = \sum_{\forall o} w_o (\vec{S}_o + \vec{N}_o),$$

где $w_o = \dfrac{e^{-\frac{\left\|\vec{S}_j + \vec{n}_j - (\vec{S}_o + \vec{n}_o)\right\|^2}{h^2}}}{\sum_o e^{-\frac{\left\|\vec{S}_j + \vec{n}_j - (\vec{S}_o + \vec{n}_o)\right\|^2}{h^2}}}$,

$$\vec{S}_j{}' = \sum_{\forall o} w_o (\vec{S}_o + \vec{N}_o) \approx \sum_{o=1}^{N} \frac{1}{N} (\vec{S}_j + \vec{N}_o) \approx \vec{S}_j + \sum_{o=1}^{N} \frac{1}{N} \vec{N}_o .$$

Здесь N – количество векторов множества $O_{\vec{n}}$, h – параметр, влияющий на степень фильтрации (в работе $h = 1$). При этом $h^2 >> 2\left\|\vec{n}_{\max}\right\|^2$. Полагая, что $\{\vec{N}_o\}$ – набор случайных независимых векторов с постоянным вектором математического ожидания, равным $\vec{\mu}$, получим $\vec{S}_j \approx \vec{S}_j + \vec{\mu}$. Подобный вывод подтверждается тем, что отсчеты АБГШ независимы друг от друга, и математическое ожидание шума не зависит от времени. Таким образом, можно получить оценку $\vec{S}_j$ незашумленного фрагмента $\vec{S}_j$ с помощью выражения:

$$\hat{\vec{S}}_j = \vec{S}_j{}' - \vec{\mu} \approx \vec{S}_j .$$

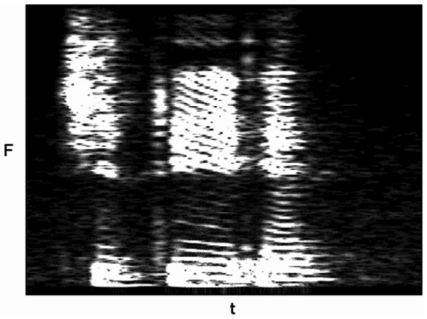
После проведения оценки всех фрагментов незашумленного отрезка сигнала, можно получить оценку всего отрезка. Затем по оценкам всех отрезков сигнала можно получить оценку и самого незашумленного речевого сигнала. Все это осуществляется с помощью процедуры восстановления сигнала с применением усреднения перекрывающихся данных.

Результаты моделирования

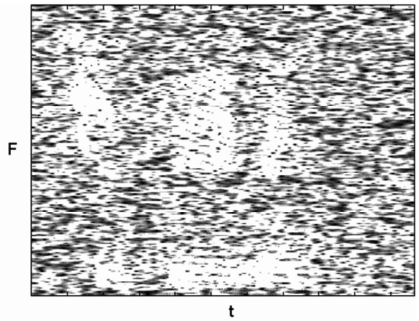
Для проверки эффективности предложенного алгоритма проведены практические исследования. Алгоритм реализован в среде математического моделирования Matlab. Речевые сигналы, оцифрованные с частотой 8000 отсчетов в секунду и разрешением 8 бит на отсчет, зашумляются АБГШ. После этого сигнал поступает на вход алгоритма подавления шума.

Субъективная оценка позволяет сделать вывод, что предложенный алгоритм способен значительно снизить уровень шума, при этом улучшается разборчивость речи и не возникает артефактов, получивших название «музыкальный шум». Сравнение спектрограмм речевого сигнала (рис. 5) до добавления шума, после, а также при обработке с помощью предложенного алгоритма и алгоритма, реализующего метод спектральных вычитаний, подтверждают сделанные ранее выводы. В частности видно, что при использовании предложенного алгоритма (рис. 5d) хорошо сохраняются все значимые детали спектра и отсутствуют нежелательные всплески, приводящие к возникновению «музыкального шума», что характерно для методов фильтрации в спектральной области (рис. 5c).

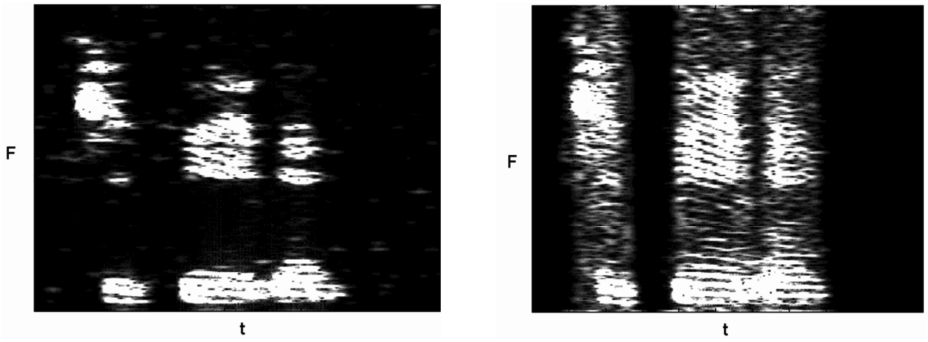
Для объективной оценки речевого сигнала на выходе алгоритма шумоподавления используется сегментное отношение сигнал / шум (СегОСШ, segmental SNR, SSNR). Результаты моделирования показывают, что при высоких уровнях шума предложенный алгоритм обладает большей эффективностью по сравнению с алгоритмом, реализующим метод спектральных вычитаний (рис. 6).



а) исходный сигнал



б) зашумленный сигнал (ОСШ = 0 дБ)



в) сигнал, обработанный *г) сигнал, обработанный предложенным алгоритмом*

Рис. 5. Спектрограммы речевого сигнала

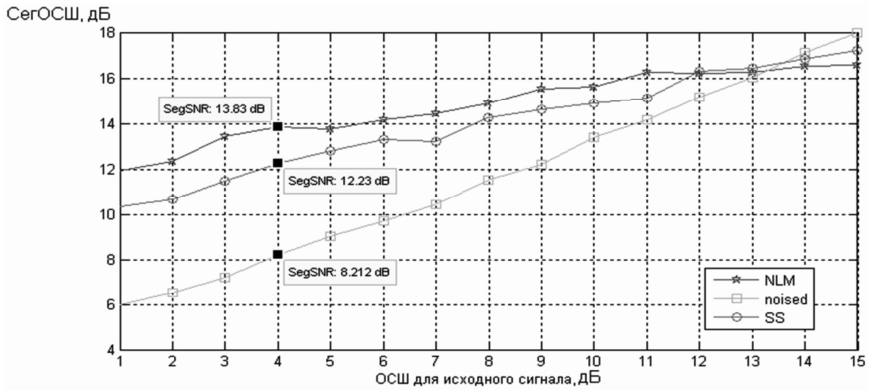


Рис. 6. Зависимость сегментного ОСШ от ОСШ исходного сигнала

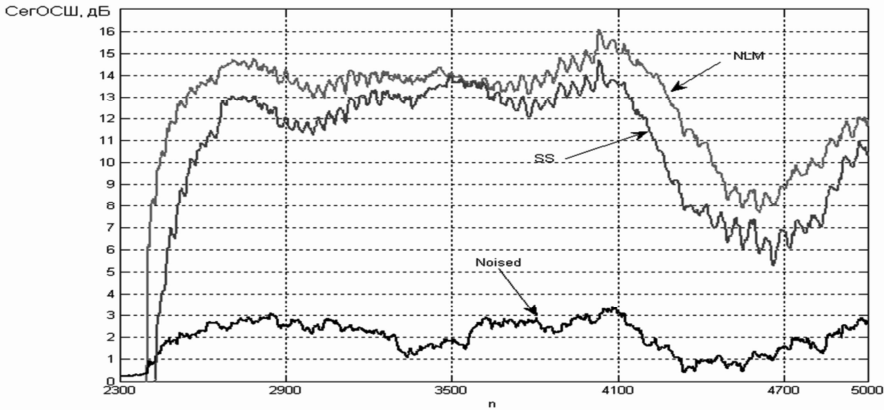


Рис. 7. Сегментное ОСШ для участка речевого сигнала до и после обработки алгоритмами шумоподавления

Наибольший выигрыш достигается при обработке участков речевого сигнала, отвечающих гласным звукам (рис. 7). Это объясняется особенностями алгоритма и структуры речевого сигнала: гласные звуки содержат больше похожих фрагментов (имеют структуру, близкую к периодической), на поиске которых и основан алгоритм нелокального усреднения. Предложенный алгоритм увеличивает сегментное ОСШ по сравнению с зашумленным сигналом на 10-12 дБ (рис. 7). Для сравнения, метод спектрального вычитания уступает предложенному методу на 1-1.5 дБ.

Ограничениями предложенного алгоритма можно считать то, что участки речевого сигнала, соответствующие шипящим звукам и глухим согласным, искажаются, что негативно влияет на качество звучания речи. Исходя из этого, актуальной является задача разработки специального алгоритма для фильтрации фрикативных звуков речи. Например, можно использовать комбинацию двух алгоритмов: гласные обрабатываются алгоритмом нелокального усреднения, а участки, отвечающие согласным, – методом спектрального вычитания.

Заключение

В работе предложен метод, позволяющий решать задачу подавления аддитивного белого гауссовского шума в речевых сигналах. К достоинствам алгоритма можно отнести высокую эффективность при работе с сильнозашумленными сигналами (ОСШ ~ 0-7 дБ), а также отсутствие «музыкального шума» в обработанном сигнале. Основным недостаткам алгоритма можно считать относительно высокую вычислительную сложность. Для повышения быстродействия алгоритма рекомендуется использовать параллельные вычисления.

Литература

1. Чучупал В.Я., Чичагов А.С., Маковкин К.А. Цифровая фильтрация зашумленных речевых сигналов: Сообщения по программному обеспечению ЭВМ. М.:

Вычислительный центр РАН, 1998. 52 с.

2. Scalart P., Filho J.V. Speech enhancement based on a priori signal to noise estimation // IEEE international conference on acoustics, speech and signal processing (ICASSP-96). 1996. P. 629–632.

3. Ephraim Y., Malah D. Speech enhancement using a minimum mean square error short time spectral amplitude estimator // IEEE transactions acoustics, speech and signal processing. 1984. P. 1109–1121.

4. Ephraim Y., Malah D. Speech enhancement using a minimum mean square error log spectral amplitude estimator // IEEE transactions acoustics, speech and signal processing. 1985. P. 443–445.

5. Hyvärinen A., Karhunen J., Oja E. Independent component analysis. – Wiley-Interscience, 2001. 504 p.

6. Yilmaz O., Rickard S. Blind separation of speech mixtures via time-frequency masking // IEEE transactions on signal processing. 2004. V. 52, No. 7. P. 1830–1846.

7. Jourjine A., Rickard S., Yilmaz O. Blind separation of disjoint orthogonal signals: demixing N sources from 2 mixtures // IEEE international conference on acoustics, speech and signal processing (ICASSP2000). Istanbul, Turkey, 2000. V. 5. P. 2985–2988.

8. Araki S., Makino S., Blin A., Mukai R., Sawada H. Blind separation of more speech than sensors with less distortion by combining sparseness and ICA // International workshop on acoustic echo and noise control (IWAENC2003). 2003. P. 271–274.

9. Моисеев А.А., Кобелев В.Ю., Волохов В.А. Курвлет-преобразование в задаче подавления шума в изображениях // Цифровая обработка сигналов. 2008. № 1. С. 43–50.

10. Приоров А.Л., Хрящев В.В., Голубев М.Н. Удаление импульсного шума со случайными значениями импульсов из изображений // Радиотехника. 2010. № 5. С. 72-79.

11. Buades A., Coll B., Morel J.M. A review of image denoising algorithms, with a new one // Multiscale modeling and simulation. 2005. V. 4. P. 490–530.

12. Buades A., Coll B., Morel J.M. Nonlocal image and movie denoising // International journal of computer vision. 2008. V. 76, No. 2. P. 123–139.

13. Katkovnik V., Foi A., Egiazarian K., Dabov K. From local kernel to nonlocal multiple-model image denoising // International journal of computer vision. 2010. V. 86, No. 8. P. 1-32.

14. Волохов В.А., Сергеев Е.В., Мочалов И.С., Приоров А.Л. Модификация алгоритма нелокального усреднения для удаления шума из цифровых изображений // Доклады 13-й международной конференции и выставки «Цифровая обработка сигналов и ее применение» (DSPA-2011). Москва, 2011. Т. 1. С. 267–270.

15. Ткаченко М.С., Лукин А.С. Многомасштабный метод спектрального вычитания для подавления шумов в аудиосигналах // Доклады 12-й международной конференции и выставки «Цифровая обработка сигналов и ее применение» (DSPA-2010). Москва, 2010. Т. 1. С. 223–226.

16. Haitian X., Zheng-Hua T., Dalsgaard P., Lindberg B. Robust speech recognition by nonlocal means denoising processing // IEEE signal processing letters. 2008. V. 15. P. 701-704.

17. Szlam A. Non-local means for audio denoising // Recent UCLA computational and applied mathematics reports. 2008. 5 p.

SPEECH DENOISING BY NON-LOCAL MEANS

Novoselov S., Topnikov A., Savvatin A., Priorov A.

A method of noisy speech enhancement based on non-local means algorithm is proposed. Non-local means algorithm is well-proven in the field of digital image processing. A description of the original two-dimensional method is performed and the possibility of its adaptation in one-dimensional case for noise suppression task is analyzed. Speech signals features takes into account in the method adapting so that the resulting algorithm is much differ from its two-dimensional analog. For example, the averaging is performed directly for a portion of a several tens of samples and not for a single one.

In this paper the suppression of additive white Gaussian noise in the speech signal is considered. There are two main steps in the proposed method of speech enhancement: the search for similar fragments and the average of these fragments to significantly reduce the variance of additive noise. The length of the fragment is chosen within the interval 40-80 samples. The similar fragments search is made in the window with 400 samples length. To determine the degree of similarity between segments Euclidean distance is used. In practice, the number of such fragments for 50 ms speech section is equal to 4-12, as averaging a smaller number of fragments is inefficient, and the search for a larger number of similar fragments is limited by the speech signal structure. After fragments of interest are found the procedure of its averaging is performed. The signal samples corresponding to these segments are changed to the samples obtained from the averaging process.

The article presents the simulation results of speech enhancement obtained using a non-local means algorithm. The effectiveness of the approach is shown. The results are compared with results obtained for the algorithm that implements the method of spectral subtraction. Segmental signal to noise ratio is used for an objective evaluation of enhanced speech quality. The proposed algorithm is more efficient than one that implements spectral subtraction method at high noise levels. The greatest benefit is achieved in vowel sounds processing that can be easily explained based on the characteristics of the algorithm and the speech signal structure.

The advantages of the proposed algorithm are high efficiency when working with low signal to noise rations and the lack of "musical noise" in the processed signal. The main drawback of the algorithm is relatively high compu-tational complexity. To improve the performance of the method parallel computing utilization is recommended.

УДК 004.93

О ПОВЫШЕНИЕ ЭФФЕКТИВНОСТИ ВЫЧИСЛЕНИЙ ГЛАВНЫХ КОМПОНЕНТ В ЗАДАЧАХ АНАЛИЗА ИЗОБРАЖЕНИЙ

***Мокеев В.В.,** д.т.н., Южно-Уральский государственный университет, e-mail: mokeyev@mail.ru*

Ключевые слова: анализ изображений, собственный вектор, матрица ковариаций, главные компоненты, линейная конденсация.

Рассматривается задача вычисления главных компонент для наборов содержащих большое число изображений. Для вычисления собственных векторов ковариационной матрицы предлагается использовать метод линейной конденсации. Описывается алгоритм многоуровневой линейной конденсации. Исследуется точность и быстродействие разработанного алгоритма.

Введение

В настоящее время активно развиваются биометрические технологии, которые лежат в основе систем идентификации человека на основе таких признаков как: отпечатки пальцев, рисунок радужной оболочки глаза, изображения лица. Использование подобных систем уменьшает количество преступлений, связанных с несанкционированным доступом. Поиск в базах данных по фотопортретам человека, автоматизированный контроль удостоверений личности особенно актуальны для правоохранительных органов большинства стран – все это позволяет снизить опасность террористических угроз.

В последнее время широкое распространение получили подходы, работающие с изображениями как с массивом признаков, каждый из которых отражает значение одного пикселя изображения (например, яркость отдельного пикселя). Так как вектор каждого изображения имеет достаточно высокую размерность, задача обработки большого числа изображений является не тривиальной. Большинство систем анализа изображений основано на методах, понижающих размерность изображений. Проблема сокращения размерности важна еще и потому, что сложность большинства алгоритмов увеличивается экспоненциально с ростом размерности задачи. Одним из наиболее распространенных методов сокращения размерности изображений является метод главных компонент [1].

Идея использования главных компонент для решения задачи распознавания лиц была изучена в работе [2]. В работе было показано, что изображение можно представить в виде линейной комбинации собственных векторов ковариационной матрицы. Собственные векторы имеют ту же размерность, что и само изображение, но являются линейно независимыми, что позволяет достаточно точно реконструировать изображение. Собственные значения определяют степень вклада каждого собственного вектора в реконструкцию изображения. Главными компонентами называются собственные векторы с наибольшими собственными значениями. Каждое изображение может быть описано взвешенной комбинацией главных компонент. Таким образом, для каждого изображения достаточно хранить только набор весовых коэффициентов (главных факторов), но каждый фактор уже отражает не отдельный пиксель, а группу пикселей, которую можно представить в виде картинки.

Обычно, решение задачи распознавания изображений, основанное на методе главных компонент, включает следующие действия [3-5]: а) нормализацию исходного изображения, б) вычисление главных компонент; в) классификацию изображений в редуцированном пространстве (кластеризация, метрика, нейронная сеть и т.п.). При нормализации изображений последние представляются как отцентрированные и приведённые к единому масштабу изображения лиц. Процесс распознавания заключается в сравнении главных факторов искомого изображения с главными факторами всех остальных изображений. При этом предполагается, что изображения лиц, соответствующих одному человеку, сгруппированы в кластеры в пространстве изображений. Результатом поиска являются те изображения, которые имеют наименьшее расстояние до искомого изображения.

Эксперименты по распознаванию лиц были проведены различными исследователями [4-7]. В работе [6] эффективность распознавания лиц исследовалась на базе данных, состоящей из 7563 изображений приблизительно 3000 людей. Главные компоненты для реконструкции изображений этой базы данных были получены на обучающем наборе из 128 изображений лиц. Распознавание было выполнено с использованием первых 20 собственных векторов. Во время эксперимента с 200 отобранными изображениями был получен достаточно высокий коэффициент распознавания, равный 95%. Коэффициент распознавания численно равен проценту правильно распознанных изображений из общего числа предъявленных и является характеристикой точности распознавания. Высокое качество распознавания связано с тем, что отобранные изображения представляли собой изображения типа фотографий для удостоверений, т.е. фронтальные изображения одного и того же размера, полученные при одинаковых условиях освещенности. В этой же работе были проведены эксперименты по исследованию качества распознавания наборов изображений, полученных при различных условиях съемки. Коэффициенты распознавания при этом менялись от 50% до 90%. В работе [7] описываются эксперименты с изображениями базы данных Feret. База данных Feret создана в 96-97 годах и включает более 3000 изображений лиц, полученных при различных ракурсах,

размерах и условиях освещения. Различные эксперименты показывают, что коэффициент распознавания изображений, сделанных разными аппаратами и при разном освещении и ракурсах, как правило, лежит в диапазоне от 40% до 95% [8].

Таким образом, набор главных компонент, полученный один раз на основе обучающей выборки, используется для описания всех остальных изображений, Считается, что с помощью ограниченного количества главных компонент можно получить хорошую реконструкцию исходного изображения. Если в обучающей выборке нет изображения лица при аналогичных условиях (ракурса, поворота, освещенности и т.п.), то система распознавания изображений может ошибаться. Для обеспечения требуемого качества процедуры распознавания лиц вычисление главных компонент необходимо проводить на представительных обучающих выборках. В случае, если поиск требуется выполнять на увеличивающемся числе изображений, то будет расти как размер обучающей выборки, так и порядок ковариационной матрицы. При этом растет и время вычисления главных компонент. Ситуация усложняется, когда система распознавания изображений работает в режиме реального времени. Для вычисления главных компонент используются методы решения задачи собственных значений.

Среди методов решения проблемы собственных значений можно выделить следующие группы методов: итерационные и преобразования подобия [9]. Распространенным методом определения небольшого числа наибольших или наименьших собственных значений является метод их вычисления прямыми или обратными итерациями. Метод прямых итераций или степенной метод позволяет вычислять наибольшее по модулю собственное значение матрицы и соответствующий собственный вектор только с помощью умножения матрицы на вектор; никаких преобразований самой матрицы при этом не требуется. Главный недостаток этого метода заключается в том, что на скорость сходимости решений влияет отношение величины искомого собственного значения к ближайшему собственному значению. Также скорость сходимости итерационного процесса зависит от того, насколько удачно выбран начальный вектор. Методы преобразований подобия применяются с целью получить из исходной матрицы новую матрицу с теми же собственными значениями, но более простого вида. Наиболее известными являются методы Якоби, Гивенса и Хаусхолдера [9]. Метод Хаусхолдера позволяет получить требуемый результат быстрее, чем метод Гивенса, так как связан с выполнением меньшего числа, хотя и более сложных преобразований.

Методы понижения порядка матриц (конденсации) можно условно отнести к методам преобразования подобия, так как целью конденсации является получение матрицы меньшего порядка, которая была бы подобна исходной матрице в том смысле, что собственные значения этих матриц в заданном диапазоне совпадали бы с заданной точностью. Методы конденсации базируются на предположении о зависимости одних (вспомогательных) переменных от других (основных), что позволяет исключить вспомогательные переменные. Одним из первых представителей методов конденсации является метод статической конденсации или метод Гайана [10], предложенный в 1965 году для решения обобщенной задачи собственных значений, возникающей при расчете колебаний конечно-элементных моделей. Метод определяет наименьшие собственные значения и требует предварительного выбора удерживаемых переменных, который базируется на опыте и интуиции расчетчика. Метод статической конденсации показал невысокую точность при решении динамических задач. Поэтому различные схемы конденсации (динамическая конденсация [11], частотно-динамическая конденсация [12], частотная конденсация [13]) были предложены для улучшения точности. В работах [14-15] используются итерационные схемы, в которых понижение порядка матриц с помощью динамической конденсации повторяется до тех пор, пока желаемая точность не будет получена.

Таким образом, в настоящее время существует достаточное количество методов, которые хорошо зарекомендовали себя для вычисления собственных векторов матриц небольшого размера. Но для больших матриц ситуация иная. Разработан ряд методов, но они, как правило, применимы для задач конечно-элементного анализа, когда матрицы являются разреженными.

Основная цель статьи описать технологию понижения порядка матриц применительно к задачам расчета главных компонент больших наборов изображений. Предлагаемая технология реализуется в виде метода линейной конденсации и используется для вычисления собственных векторов стандартной задачи собственных значений (одна из матриц является единичной) для случая плотных, а не разреженных матриц.

Анализ изображений методом главных компонент

Пусть имеется набор изображений, каждое из которых описывается вектором x_i. где i – номер изображения (i = 1,2,3,..., m). Размерность вектора x_i равняется числу пикселей изображения n. Таким образом, все изображения можно представить в виде матрицы $\mathbf{X}$, строками которой являются векторы x_i. Размерность пространства изображений определяется произведением $n \times m$. Основная трудность анализа такого пространства заключается в его высокой размерности для реальных задач.

Для построения пространства меньшей размерности используется метод главных компонент. Обработка изображений начинается с операции центрирования пространства изображений относительно среднеарифметического вектора, который вычисляется по формуле:

$$\bar{x} = \frac{1}{m}\sum_{i}^{m} x_i \qquad (1)$$

При этом осуществляется переход к новому пространству изображений $\mathbf{X}^0$, строками которого являются векторы $x_i^0 = x_i - \bar{x}$.

Основная формулировка метода главных компонент строится на расчете ковариационной матрицы $\mathbf{A}^* = 1/_m(\mathbf{X}^0)^T\mathbf{X}^0$, которая имеет порядок $n \times n$, и вычислении собственных векторов уравнения

$$(\mathbf{A}^* - \lambda\mathbf{I})\mathbf{v}_0 = 0 \qquad (2)$$

где $\mathbf{I}$ – единичная матрица, $\mathbf{v}_{0i}$ – собственный вектор,

λ_i – собственное значение, индекс «Т» означает транспонирование матрицы. Матрица главных компонент формируется из r собственных векторов, которым соответствуют наибольшие собственные значения

$$\mathbf{V} = \left[\mathbf{v}_{01}\ \mathbf{v}_{02}...\mathbf{v}_{0r}\right] \quad (3)$$

В связи с тем, что матрица $\mathbf{A}$ имеет высокий порядок, решение уравнения (2) представляет существенные трудности. Поэтому при обработке изображений используется дуальная формулировка метода главных компонент, в соответствии с которой вычисляется матрица Грамма $\mathbf{A} = \frac{1}{m}\mathbf{X}^0(\mathbf{X}^0)^T$. Матрица $\mathbf{A}$ имеет порядок m, который существенно меньше n. Для вычисления главных компонент необходимо найти собственные векторы матричного уравнения

$$(\mathbf{A} - \lambda\mathbf{I})\mathbf{u}_0 = 0 \quad (4)$$

где u_{0i} – собственный вектор, λ_i – собственное значение. Следует отметить, что собственные значения уравнений (2) и (4) совпадают. После сортировки собственных значений формируется матрица главных компонент из r первых собственных векторов.

$$\mathbf{U} = \left[\mathbf{u}_{01}\ \mathbf{u}_{02}...\mathbf{u}_{0r}\right] \quad (5)$$

Известно, что матрицу $\mathbf{X}^0$ можно представить в виде сингулярного разложения

$$\mathbf{X}^0 = \mathbf{U}\,\mathbf{T}\,\mathbf{V}^T \quad (6)$$

где $\mathbf{V}$ – матрица собственных векторов уравнения (2), столбцы которой ортонормированны $\mathbf{V}^T\mathbf{V} = \mathbf{I}$, $\mathbf{U}$ – матрица собственных векторов уравнения (4) с ортонормированными столбцами $\mathbf{U}^T\mathbf{U} = \mathbf{I}$, $\mathbf{T}$ – диагональная матрица порядка $r \times r$, диагональные элементы которой $\eta_1, \eta_2, ..., \eta_r$ называются сингулярными числами матрицы $\mathbf{X}^0$. Чаще всего $r = m$, но приводимый ниже результат охватывает общий случай, он справедлив и при условии $r < m$. Собственные значения матрицы $\mathbf{X}^0$ связаны с собственными значениями матрицы $\mathbf{A}$ соотношением $\eta_i = (m\lambda_i)^{1/2}$. В соответствии с терминологией, принятой в линейной алгебре, матрицу $\mathbf{V}$ можно назвать матрицей правых собственных векторов, а матрицу $\mathbf{U}$ – матрицей левых собственных векторов.

Матрицы левых и правых собственных векторов связаны друг с другом соотношением (6). Для получения формулы вычисления матрицы $\mathbf{V}$ умножим выражение (6) справа на транспонированную матрицу $\mathbf{U}$ и матрицу T^{-1}. Далее, используя свойство ортогональности матрицы $\mathbf{U}$, получим

$$\mathbf{V}^T = \mathbf{T}^{-1}\mathbf{U}^T\mathbf{X}^0 \quad (7)$$

Каждый вектор пространства $\mathbf{X}^0$ теперь может быть представлен в редуцированном пространстве небольшим числом главных факторов, образующих вектор $\mathbf{z}_i$, т.е.

$$x_i^0 = \mathbf{V}\mathbf{z}_i \quad (8)$$

Векторы $\mathbf{z}_i$ образуют матрицу $\mathbf{Z}$, которая может быть вычислена по следующей формуле

$$\mathbf{Z} = \mathbf{X}^0\mathbf{V} \quad (9)$$

где $\mathbf{Z}$ – матрица главных факторов пространства $\mathbf{X}^0$.

Матрица главных факторов $\mathbf{Z}$ представляет новое редуцированное пространство изображений. Но при этом каждому главному фактору соответствует группа пикселов, которые можно представить в виде картинки. Поэтому описанный подход часто называют методом собственных лиц. Размерность редуцированного пространства равна $p \times m$, что меньше исходной размерности в n/p раз.

После выделения главных факторов изображения требуется произвести распознавание. Набор главных факторов представляет собой точку в пространстве признаков. Процедура классификации основывается на разделении пространства признаков на области, относящиеся к одинаковым классам, и процедуре поиска на основе вычисления расстояния от неизвестного изображения до всех остальных при помощи какой-либо метрики (Махаланобиса, евклидово расстояние).

Метод линейной конденсации

Переход к редуцированному пространству изображений с помощью метода главных компонент требует вычисления наибольших собственных значений уравнения (4). Эта задача может быть сведена к поиску наименьших собственных значений уравнения вида

$$(\mathbf{I} - \mu\,\mathbf{A})\mathbf{v}_0 = 0, \quad (10)$$

где $\mu = 1/\lambda$.

Если число изображений велико, то решение уравнения (10) представляет определенные трудности. Для преодоления этих трудностей предлагается использовать понижение порядка матриц, которое базируется на предположении о зависимости одних (вспомогательных) признаков от других (основных), что позволяет исключить вспомогательные признаки. Разделим вектор $\mathbf{v}_0$ на две составляющие: вектор исключаемых признаков и вектор удерживаемых признаков. Уравнение (10) в соответствии с этим делением перепишется в виде

$$\begin{Bmatrix}\mathbf{v}_{0r}\\ \mathbf{v}_{0s}\end{Bmatrix} - \mu\begin{bmatrix}\mathbf{A}_{rr} & \mathbf{A}_{rs}\\ \mathbf{A}_{sr} & \mathbf{A}_{ss}\end{bmatrix}\begin{Bmatrix}\mathbf{v}_{0r}\\ \mathbf{v}_{0s}\end{Bmatrix} = 0 \quad (11)$$

где индекс r относится к основным признакам, а индекс s – к вспомогательным признакам. Используя нижнее уравнение системы (11), можно определить связь вспомогательных и основных признаков

$$\mathbf{v}_{0s} = \mu(\mathbf{I}_{ss} - \mu\mathbf{A}_{ss})\mathbf{A}_{sr}\mathbf{v}_{0r}. \quad (12)$$

После исключения вспомогательных признаков с помощью соотношения (12) уравнение (11) имеет вид

$$(\mathbf{I}_{rr} - \mu\,\mathbf{A}_{rr} + \mathbf{D}_{rr}(\mu))\mathbf{v}_{0r} = 0 \quad (13)$$

где $\mathbf{D}_{rr}(\mu) = \mu^2\,\mathbf{A}_{rs}\left(\mu\mathbf{A}_{ss} - \mathbf{I}_{ss}\right)^{-1}\mathbf{A}_{rs}^{\ T}$.

Понижение порядка матриц осуществляется путем аппроксимации матрицы $\mathbf{D}_{rr}(\mu)$ выражением вида

$$\mathbf{D}_{rr}(\mu) \approx \mathbf{I}_{rr}^* - \mu\mathbf{A}_{rr}^*. \quad (14)$$

Коэффициенты матриц $\mathbf{A}_{rr}^*$ и $\mathbf{I}_{rr}^*$ соотношения (14) определяются из условия наилучшего приближения матриц $\mathbf{D}_{rr}(\mu)$ и $\mathbf{I}_{rr}^* - \mu\mathbf{A}_{rr}^*$ в диапазоне $\mu_1 \ldots \mu_2$. Например, из условия совпадения матриц $\mathbf{D}_{rr}(\mu)$ и $\mathbf{I}_{rr}^* - \mu\mathbf{A}_{rr}^*$ в граничных точках искомого диапазона μ_1 и μ_2. Это соответствует линейной аппроксимации матрицы $\mathbf{D}_{rr}(\mu)$. В этом случае матрицы $\mathbf{A}_{rr}^*$ и $\mathbf{I}_{rr}^*$ определяются по формулам

$$\mathbf{I}_{rr}^* = \mathbf{D}_{rr}(\mu_1) - \mu_1\frac{\mathbf{D}_{rr}(\mu_1) - \mathbf{D}_{rr}(\mu_2)}{\mu_1 - \mu_2}$$

$$\mathbf{A}^{*}_{rr} = -\frac{\mathbf{D}_{rr}(\mu_1) - \mathbf{D}_{rr}(\mu_2)}{\mu_1 - \mu_2} \qquad (15)$$

Если $\mu_1 = 0$, то выражения (15) можно переписать в виде

$$\mathbf{I}^{*}_{rr} = 0$$

$$\mathbf{A}^{*}_{rr} = -\frac{\mathbf{D}_{rr}(\mu_2)}{\mu_2} \cdot$$

На рис. 1 показано изменение коэффициентов матрицы $\mathbf{D}_{rr}(\mu)$ и линейная аппроксимация этих коэффициентов выражением $i_{rr} - \mu a_{rr}$. Параметр μ_s обозначает наименьшее собственное значение матрицы $\mathbf{A}_{SS}$. Если μ_s попадает в диапазон $\mu_1 \ldots \mu_2$, то погрешность аппроксимации матрицы $\mathbf{D}_{rr}(\mu)$ для $\mu = \mu_S$ становится бесконечной.

После понижения порядка матриц уравнение (10) представляется в виде редуцированного уравнения

$$(\mathbf{I}_r - \mu^{*}\mathbf{A}_r)\mathbf{v}^{*}_{0r} = 0\,, \qquad (16)$$

где $\mathbf{A}_r = \mathbf{A}_{rr} + \mathbf{A}^{*}_{rr}$, $\mathbf{I}_r = \mathbf{I}_{rr} + \mathbf{I}^{*}_{rr}$. Если μ_1=0, то решение уравнения (16) представляет стандартную задачу собственных значений, а в случае $\mu_1 > 0$ – обобщенную задачу собственных значений.

Решая уравнение (16), мы получаем приближенное решение уравнение (10). Причиной появления ошибок в решении является линейная аппроксимация матрицы $\mathbf{D}_{rr}(\mu)$. В общем случае погрешность аппроксимации матрицы $\mathbf{D}_{rr}(\mu)$ определяется как

$$\mathbf{E}_{rr}(\mu) = \mathbf{D}_{rr}(\mu) - (\mathbf{I}^{*}_{rr} - \mu\mathbf{A}^{*}_{rr})\,. \qquad (17)$$

Можно связать погрешность собственных значений с матрицей $\mathbf{E}_{rr}(\mu)$. Для этого вычтем из уравнения (13) уравнение (16) и при допущении $\mathbf{v}_{0r} = \mathbf{v}^{*}_{0r}$ получим

$$(\Delta\mu\mathbf{A}_r + \mathbf{E}_{rr}(\mu))\mathbf{v}_{0r} = 0\,, \qquad (18)$$

где $\Delta\mu = \mu - \mu^{*}$ – погрешность собственных значений.

Значение погрешности, вносимое в собственные значения при понижении порядка матриц, можно оценить по формуле

$$\Delta\mu \leqslant \left|\mathbf{A}_r^{-1}\,\mathbf{E}_{rr}(\mu)\right|. \qquad (19)$$

Если собственные значения матрицы $\mathbf{A}_{SS}$ попадают внутрь диапазона $\mu_1 \ldots \mu_2$, то коэффициенты матриц $\mathbf{D}_{rr}(\mu)$ и $\mathbf{E}_{rr}(\mu)$ принимают бесконечно большие значения. Поэтому одним из основных условий выбора исключаемых переменных является

$$\mu_s^{\min} > \mu_2\,, \qquad (20)$$

где $\mu_s^{\min}$ – минимальное собственное значение матрицы исключаемых переменных $\mathbf{A}_{SS}$. Если условие (20) выполняется, то коэффициенты матрицы погрешности равны нулю в граничных значениях диапазона $\mu_1 \ldots \mu_2$ и достигают максимального значения внутри диапазона при $\mu = \mu_м$. Приближенно в качестве $\mu_м$ можно использовать середину диапазона $\mu_м = \mu_{ср}$ (см. рис. 1).

Алгоритм многоуровневой линейной конденсации

Метод линейной конденсации реализуется в виде алгоритмов одноуровневой и многоуровневой линейной конденсации. Суть алгоритма многоуровневой линейной конденсации состоит в том, что вспомогательные признаки исключаются небольшими группами в рамках многоуровневой схемы. При этом матрицы уравнения (10) на каждом уровне уменьшаются на размер группы исключаемых признаков. В частном случае мы получаем алгоритм одноуровневой конденсации, который группирует все исключаемые признаки в одном блоке. Преимущество многоуровневой схемы заключается в том, что выбор исключаемых признаков можно осуществлять в процессе понижения порядка матриц, в то время как одноуровневая конденсация требует делать такой выбор до начала процедуры понижения порядка матриц.

Алгоритм многоуровневой линейной конденсации включает пять шагов. Первый шаг представляет собой многоуровневую процедуру, на каждом уровне которой решается вопрос об исключении группы вспомогательных признаков. Известно, что сумма диагональных коэффициентов матрицы равна сумме ее собственных значений [9]. Поэтому в качестве верхней границы максимального собственного значения $\lambda_{\max}$ матрицы $\mathbf{A}_{SS}$ можно использовать сумму диагональных коэффициентов матрицы исключаемых признаков, а в качестве нижней границы минимального собственного значения $\mu_{\min}$ матрицы $\mathbf{A}_{SS}$ обратную сумму диагональных коэффициентов

$$\mu^{*}_{\min} = 1/\sum_{i=1}^{s} a_{ii}\,. \qquad (21)$$

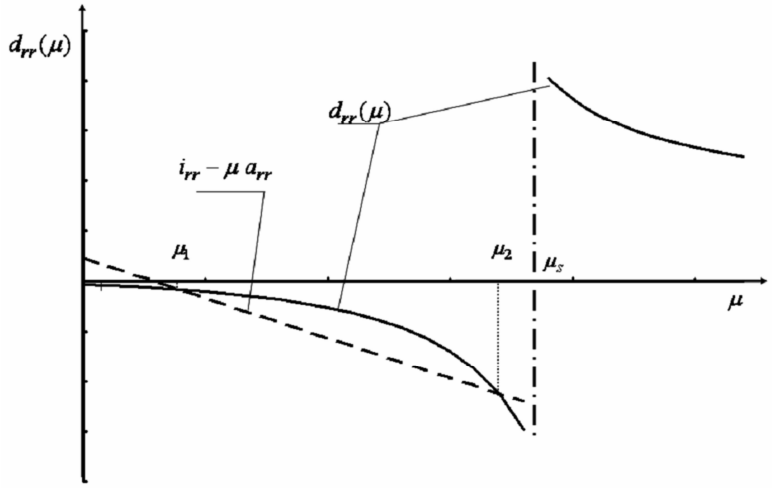


Рис.1. Линейная аппроксимация коэффициентов матрицы $\mathbf{D}_{rr}(\mu)$

Процедура понижения порядка матрицы начинается с того, что все признаки сортируются в порядке убывания диагональных коэффициентов ковариационной матрицы $\mathbf{A}$. На первом уровне выбирается группа признаков с минимальными диагональными коэффициентами a_{ii}. В качестве критерия, по которому принимается решение об исключении группы признаков, используется либо наименьшее собственное значение ($\mu_c = \mu_{\min}$) блока исключаемых переменных, либо параметр обратной суммы $\mu_c = \mu^*_{\min}$, Если выполняется условие

$$\mu_c > k_c \mu_2, \tag{22}$$

то принимается решение об исключении переменных. Параметр k_c называется параметром отсечения. Если в качестве критерия отбора используется наименьшее собственное значение ($\mu_{\min}$), то k_c лежит в диапазоне от 1 до 4; в случае если используется обратная сумма диагональных коэффициентов ($\mu^*_{\min}$), то k_c меняется от 0.5 до 2.

Далее вычисляются коэффициенты аппроксимирующих матриц $\mathbf{A}^*_{rr}$ и $\mathbf{I}^*_{rr}$, после чего осуществляется переход ко второму уровню понижения порядка матриц. На каждом новом уровне формируется группа вспомогательных признаков, кандидатов на исключение. Критерием отбора является выполнение условия (22). Если оно выполняется, то осуществляется понижение порядка матриц и переход на следующий уровень. Понижение порядка матриц прекращается в случае нарушения условия (22). Для повышения эффективности алгоритма полезно перед понижением порядка матриц делать их предварительную обработку, направленную на уменьшение недиагональных коэффициентов матрицы. Для этого можно использовать обычные вращения Гивенса [9].

На втором шаге вычисляются собственные значения и собственные векторы матричного уравнения (16). В связи с тем, что матрица не является единичной, решение уравнения (16) относится к классу обобщенных задач на собственные значения. В этом случае можно использовать, например, метод обратных итераций или метод Якоби [9].

На третьем шаге матрица редуцированных собственных векторов восстанавливается до полной матрицы, т.е. для каждого редуцированного собственного вектора вычисляется вектор вспомогательных признаков по формуле

$$x^*_{0si} = \mu^*_i (\mathbf{I}_{ss} - \mu^*_i \mathbf{A}_{ss}) \mathbf{A}_{sr} x^*_{0ri}. \tag{23}$$

В результате получаем матрицу восстановленных собственных векторов

$$\Delta^* = \begin{bmatrix} x^*_{0r1} & x^*_{0r2} \cdots x^*_{0rp} \\ x^*_{0s1} & x^*_{0s2} \cdots x^*_{0sp} \end{bmatrix}. \tag{24}$$

Восстановленные собственные векторы, полученные на третьем шаге, не являются ортогональными. Поэтому алгоритм включает четвертый и пятый шаги, в результате выполнения которых получается матрица ортогональных собственных векторов.

На четвертом шаге вычисляются приведенные матрицы

$$\mathbf{A}_p = \Delta^{*T} \mathbf{A} \Delta^*$$

$$\mathbf{I}_p = \Delta^{*T} \Delta^*. \tag{25}$$

На пятом шаге решается задача собственных значений матричного уравнения

$$(\mathbf{I}_p - \mu \mathbf{A}_p)\mathbf{q} = 0 \tag{26}$$

и уточняется матрица собственных векторов

$$x_{0i} = \Delta^* \mathbf{q}_i. \tag{27}$$

Решение уравнения (26) представляет собой задачу собственных значений, и здесь опять можно использовать такие известные методы, как метод Якоби или Хаусхолдера.

Исследование эффективности метода линейной конденсации

Любой численный метод должен в той или иной мере удовлетворять следующим требованиям: надёжность, точность, быстродействие. В связи с тем, что численный метод можно рассматривать как функцию, отображающую вход в выход, надежность метода можно связать с областью применения метода, т.е. с множеством входов (матриц), для которых метод работает. Итерационные методы оказываются чувствительными к особенностям матрицы, с которой они работают. Например, при плохой обусловленности матриц методы могут давать осечку. Метод линейной конденсации обладает высокой надежностью, так как не накладывает никаких ограничений на матрицы, с которыми он работает. Другими важными критериями численного метода являются быстродействие и точность получаемых с его помощью результатов. Для исследования быстродействия и точности метода линейной конденсации используется программа PEGAS, предназначенная для вычисления главных компонент различными методами. Исследования проводятся на примерах задач расчета главных компонент наборов, состоящих из различного числа изображений.

Точность метода линейной конденсации исследуется для задачи нахождения первых восьми главных компонент набора из 518 изображений, каждое из которых имеет размерность 112x196. Исследуется два варианта. В первом варианте в качестве критерия отбора используется обратная сумма диагональных коэффициентов. Рассматриваются три варианта понижения порядка матрицы, отличающихся друг от друга значением параметра отсечения: 0.5, 1, 1.5. В качестве эталонного решения используются главные компоненты, полученные методом Хаусхолдера. Погрешность главных компонент определяется по формуле

$$e_i = \frac{\sum_{k=1}^{m} (\mathbf{v}_{0ik} - \mathbf{v}^*_{0ik})^2}{\sum_{k=1}^{m} \mathbf{v}_{0ik}^{\ 2}}$$

где $\mathbf{v}^*_{0ik}$ – главные компоненты, вычисленные методом линейной конденсации, $\mathbf{v}_{0ik}$ – главные компоненты, полученные методом Хаусхолдера. Погрешность главных компонент представляет собой отношение нормы вектора погрешности к норме вектора $\mathbf{v}_{0i}$.

В табл. 1 показаны погрешности первых восьми главных компонент. Из таблицы видно, что наибольшие погрешности получены для параметра отсечения 0.5.

Таблица 1. Погрешность главных компонент: критерий обратной суммы

k_c		0,5	1	1,5
Число основных переменных		100	110	125
Номер собственного вектора	1	0.00026	0.00004	0.00001
	2	0.00101	0.00020	0.00007
	3	0.0016	0.0003	0.0001
	4	0.00439	0.00055	0.00026
	5	0.0080	0.0010	0.0005
	6	0.0166	0.0022	0.0014
	7	0.0152	0.0019	0.0008
	8	0.0114	0.0023	0.0009

Для всех вариантов изменения параметра отсечения погрешность вычисления собственных векторов незначительна, однако, чем больше параметр отсечения, тем выше точность. Увеличивая параметр отсечения, можно повысить точность вычисления главных компонент, но при этом степень сжатия матриц будет ниже, порядок матриц редуцированного уравнения (16) выше и, следовательно, времени для решения этого уравнения потребуется больше.

Во втором варианте в качестве критерия отбора исключаемых признаков используется наименьшее собственное значение блока вспомогательных переменных. Рассматриваются четыре варианта понижения порядка матрицы, отличающихся друг от друга значением параметра отсечения: 2.0, 2.5, 3.0. В качестве эталонного решения используются главные компоненты, полученные методом Хаусхолдера. В табл. 2 показаны погрешности первых восьми главных компонент.

Таблица 2. Погрешность главных компонент: критерий наименьшее собственное значение

k_c		2	2,5	3
Число основных переменных		65	70	75
Номер собственного вектора	1	0.000054	0.000042	0.000042
	2	0.00050	0.00017	0.00017
	3	0.00040	0.00034	0.00034
	4	0.00064	0.00038	0.00038
	5	0.00503	0.00421	0.00421
	6	0.00278	0.00160	0.00160
	7	0.00854	0.00660	0.00660
	8	0.01216	0.00437	0.00437

На рис. 2 показано типичное изменение погрешности главных компонент от величины собственного значения. Как видно из рисунка максимальное значение погрешности достигает в районе μ_{cp}.

Сравнение различных критериев отбора удаляемых переменных показывает, что они дают примерно одинаковую точность, но выбор исключаемых признаков по наименьшему собственному значению является более эффективным, так как степень сокращения матрицы наиболее высокая.

Таким образом, метод линейной конденсации может получать решения с различной степенью точности, однако повышение точности связано с дополнительными временными затратами. Точность вычисления главных компонент степенным методом обычно задается в пределах 0.001. Для предотвращения потери ортогональности на каждой итерации степенного метода выполняется процедура ортогонализации. Чем выше точность, тем больше итераций требуется выполнить для ее достижения. Исследование быстродействия метода осуществляется путем сравнения времени вычисления левых главных компонент методом линейной конденсации, степенным методом и методом Хаусхолдера. Под временем вычисления понимается время расчета левых главных компонент u_{0i} путем решения уравнения (4) на персональном компьютере (AMD Athlon, Dual Core Processor, 2010 MHz).

В табл. 3 представлено время (в секундах) расчета различного количества левых главных компонент набора, состоящего из 528 изображений, с помощью метода линейной конденсации и степенного метода. При вычислении главных компонент методом линейной конденсации параметр отсечения устанавливается равным 1.5.

Как видно из табл. 3, время решения степенного метода увеличивается с ростом числа вычисляемых главных компонент, в то время как время расчета методом линейной конденсации увеличивается в 1,5 раза при увеличении числа вычисляемых векторов в 6 раз, в то время как время расчета степенным методом увеличивается в 18 раз.

В табл. 4 представлено время (в секундах) вычисления 67 главных компонент, полученных при обработке наборов из 512, 1188, 1977, 3000 и 4832 изображений.

Как видно из табл. 4, время расчета главных компонент степенным методом существенно увеличивается с ростом порядка матриц, в то время как метод линейной конденсации требует гораздо меньше времени для вычислений. Для матрицы размерностью 3000 метод линейной конденсации вычисляет главные компоненты в 20 раз, а для матриц размерностью 4832 – в 25 раза быстрее, чем степенной метод. Метод Хаусхолдера, несмотря на то, что вычисляет все собственные векторы уравнения (4), показывает более высокую скорость, чем степенной метод. Это связано с тем, что скорость сходимости степенного метода существенно снижается при высокой плотности собственных значений. Для порядка матрицы 4832 метод линейной конденсации показывает быстродействие на порядок выше быстродействия метода Хаусхолдера и степенного метода.

Главные компоненты, вычисленные различными методами, используются для распознавания изображений базы данных Feret, которая сформирована из изображений лиц, полученных при различных условиях освещения, различных ракурсах и имеющих различные размеры. В результате экспериментов, проведенных рядом авторов для изображений базы данных Feret, установлено, что для фронтальных изображений, сделанных в один и тот же день, точность распознавания, как правило, составляет 95%, для изображений, сделанных разными аппаратами и при разном освещении и ракурсах, точность падает до 80%, для изображений, сделанных с разницей в год, точность распознавания составляет примерно 50% [8].

Таблица 3. Время решения (с) задачи собственных значений

Число собственных значений/векторов	Линейная Конденсация	Степенной Метод
12	15	39
21	15	83
38	17	300
67	22	648

Таблица 4. Время решения (с) задачи собственных значений

Порядок матрицы	Линейная конденсация	Степенной метод	Метод Хаусхолдера
512	3	102	6
1188	16	648	95
1977	67	1427	445
3000	213	4893	1740
4832	1092	24096	11778

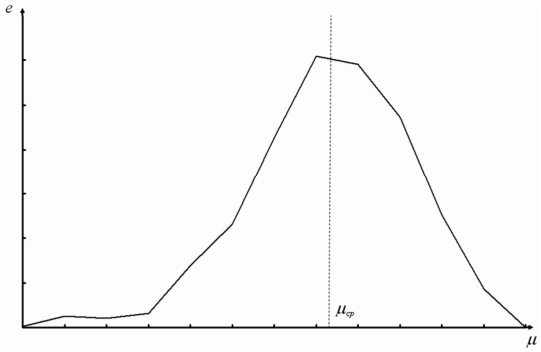


Рис.2. Изменение погрешности главных компонент от величины собственного значения

При проведении исследований решается вопрос, как влияет точность вычисления главных компонент методом линейной конденсации на качество распознавания изображений лиц. Эксперименты проводятся для набора из 3000 изображений, полученных при различных условиях освещений, ракурсах, сделанных с разницей в один год. Все изображения получены из базы данных Feret. Изображения учебной выборки выбираются из базового набора, а ее размер принимает значения 1000 и 2000. Тестовый набор формируется из оставшихся изображений. В ходе экспериментов для каждой учебной выборки различными методами вычисляются до 400 главных компонент. Для распознавания лиц используется процедура PCA+LDA, которая включает два шага [16]. Первым шагом является использование метода главных компонент (PCA) для сокращения размерности изображений. Вторым шагом является применение линейного дискриминантного анализа (LDA) для преобразованных данных. Основная идея состоит в том, что после применения метода главных компонент размерность ковариационных матриц внутри-классовых и межклассовых различий для преобразованных данных не является высокой. С помощью линейного дискриминантного анализа удается получить подпространство небольшой размерности, в котором кластеры изображений пересекаются минимально. Число классов соответствует количеству персон (людей). Один класс включает все изображения одного человека. Для распознавания изображений используется классификатор минимального расстояния неизвестного изображения до цента класса. В качестве центра класса используется среднеарифметический вектор изображений, входящих в данный класс.

В табл. 5 представлены коэффициенты распознавания, полученные с использованием главных компонент, вычисленных для учебных выборок разного размера алгоритмом многоуровневой линейной конденсации. Точность вычисления главных компонент зависит от параметра отсечения, который принимает значение: а) 1.5, б) 2.5 и с) 3. При этом погрешность вычисления главных компонент меняется от 0.08 до 0.003. Самая высокая точность вычисления главных компонент соответствует варианту (с). В этой же таблице для сравнения приведены коэффициенты распознавания, полученные при использовании метода Хаусхолдера для расчета главных компонент. Коэффициенты распознавания определяются для тестового (K_{test}) и учебного (K_{learn}) наборов изображений.

Таблица 6. Коэффициенты распознавания

Размер учебной выборки	Линейная конденсация						метод Хаусхолдера	
	а		б		с			
	K_{learn}	K_{test}	K_{learn}	K_{test}	K_{learn}	K_{test}	K_{learn}	K_{test}
1000	96.12	71.3	96.21	70.75	96.21	70.8	96.21	70.8
2000	98.12	81.3	98.22	82.21	98.12	81.7	98.12	81.7

Как видно из таблицы, коэффициенты распознавания, полученные методом линейной конденсации (вариант (с)) и методом Хаусхолдера, совпадают. Однако главные компоненты вариантов (а) и (б), хотя и получены с более низкой точностью, тем не менее, их использование не приводит к существенному снижению точности распознавания, а в ряде случаев даже может повышать ее.

Выводы

Рассмотрена задача вычисления главных компонент, возникающая при распознавании изображений. Для расчета главных компонент обучающих наборов изображений предлагается использовать метод линейной конденсации. Метод использует понижение порядка матриц, позволяющее достаточно точно вычислять собственные векторы, собственные значения которых находятся в заданном интервале. Описан алгоритм многоуровневой линейной конденсации. Проведено исследование точности и быстродействия разработанного алгоритма. Показано, что метод линейной конденсации позволяет вычислять главные компоненты с хорошей точностью на порядок быстрее, чем такие известные методы, как метод Хаусхолдера и степенной метод.

Литература

1. Jolliffe I.T. Principal Component Analysis. Springer-Verlag, New York, 1986.

2. Kirby M. and Sirovich L. Application of the KL procedure for the characterization of human faces, IEEE Trans. Pattern Anal. Mach. Intell. 12 (1) (1990) 103-108.

3. Swets D.L and Weng J., Using discriminant eigenfeatures for image retrieval, IEEE Trans. Pattern Anal. Mach. Intell. 18 (8) (1996) 831-836.

4. Кухарев Г.А. Биометрические системы: Методы и средства идентификации личности человека. СПб.: Политехника, 2001. 240 с.

5. Sung K.K., Poggio T., Example-based learning for view-based human face detection, IEEE Trans. Pattern Anal. Mach. Intell., 20 (1) 1998 39-51.

6. Pentland A. and Moghaddam B. and Starner T. View-based and modular eigenspaces for face recognition, Proceedings of the IEEE Conference on Computer Vision and Pattern recognition, vol. 1, Seattle, WA, 1994, pp. 84-91.

7. Moon H., Phillips P.J. Computational and Performance aspects of PCA-based Face Recognition Algorithms, Perception, Vol. 30, 2001, pp. 303-321

8. Pentland A., Choudhury C. Face recognition for smart environments. IEEE Computer, 33(2) 2000, pp. 50-55

9. Парлетт В. Симметричная проблема собственных значений. М.: Мир, 1983. 384с

10. Guyan R.J., “Reduction of stiffness and mass matrices,” AIAA J. 3 (1965) 380.

11. Miller C.A., Dynamic Reduction of Structural Models, Journal of Structural Division. 106 (10) (1980) 2097-2108

12. Гриненко Н.И., Мокеев В.В. О задачах исследований колебаний конструкций методом конечных элементов.// Прикладная механика. 21 (3) (1985) 25-30.

13. Мокеев В.В. О задаче нахождения собственных значений и векторов больших матричных систем.//Журнал вычислительной математики и математической физики. 32 (10) (1992) 1652-1657

14. Z. Qu and F. Fu, New structural dynamic condensation method for finite element models, AIAA J. 36 (7) (1998) 1320-1324

15. Y. Jung, Z. Qu, and D. Jung, Structural Dynamic Condensation Method with an Iterative Scheme, Journal of Civil Engineering, 8. (2) (2004) 205-211

16. P.N. Belhumeur, J.P. Hespanha, D.J. Kriegman, Eigenfaces vs.Fisherfaces: recognition using class specific linear projection, IEEE Trans. Pattern Anal. Mach. Intell. 19 (1997) 711–720.

ON EFFECTIVENESS INCREASE OF PRINCIPAL COMPONENTS COMPUTATION IN IMAGE ANALYSIS PROBLEM

Mokeyev V.V.

A problem of principal component analysis for large image set is considered. To compute eigenvectors of covariance matrix, the linear condensation method is used. The algorithm for multy-level linear condensation is described. The accuracy and high performance of the developed algorithm is also evaluated.

Уважаемые коллеги!

Для тех, кто не успел оформить подписку на первое полугодие 2012 года через ОАО «Роспечать», сохраняется возможность приобретения журналов непосредственно в редакции по адресу:
107031, г. Москва, Рождественка, 6\9\20, стр. 1,
Российское научно-техническое общество радиотехники, электроники и связи им. А.С. Попова, или оформить заказ в соответствии с требованиями, выставленными на сайте журнала: www.dspa.ru.

Справки по телефонам: (495) 621-71-08, 621-06-10.
Факс: (495) 621-16-39. E-mail: nto.popov@mtu-net.ru

УДК 004.67+004.9

ИННОВАЦИОННЫЙ КУРС «ПРОГРАММИРУЕМАЯ РАДИОЭЛЕКТРОНИКА» ДЛЯ ИНЖЕНЕРНЫХ ВУЗОВ

***Бартенев В.Г.**, д.т.н., профессор кафедры теоретической радиотехники и радиофизики Московского государственного технического университета радиотехники, электроники и автоматики (МИРЭА), e-mail: syntaltechno@mail.ru*

***Битюков В.К.**, д.т.н., профессор, заведующий кафедрой теоретической радиотехники и радиофизики МИРЭА, e-mail: bitukov@mirea.ru*

***Кузеленкова Е.Г.**, инженер кафедры теоретической радиотехники и радиофизики МИРЭА, e-mail: Kuzelenkova@mirea.ru*

Ключевые слова: программируемая радиоэлектроника, микроконтроллер, сигнальный процессор, программируемая логиическая интегральная схема, программируемые цифровые фильтры, программируемые устройства оценки параметров помех, программируемые обнаружители сигналов.

Выполнен анализ основных этапов развития радиоэлектроники: от ее возникновения и до настоящего времени. Констатировано, что за последние двадцать лет произошел революционный переворот в этой области техники, связанный с переходом к программируемым радиоэлектронным встраиваемым системам и устройствам с использованием микроконтроллеров, сигнальных процессоров и программируемых логических интегральных схем. Это предопределило необходимость в постановке и преподавании в инженерных высших ученых заведениях инновационного учебного курса «Программируемая радиоэлектроника». Этот курс целесообразно читать бакалаврам, студентам и магистрам, включив его в базовую часть профессионального цикла федеральных государственных образовательных стандартов. На факультете радиотехнических и телекоммуникационных систем МИРЭА был успешно апробирован новый учебный курс «Программируемая радиоэлектроника».

Введение

Анализ современного состояния отечественной радиоэлектроники показал, что по ряду причин Российской Федерации не удалось в полном объеме осознать произошедший в мире за последние двадцать лет революционный переворот в этой области техники. Во многом это произошло из-за развала СССР, своеобразия российских экономических реформ, в основу которых положена сырьевая составляющая, и недальновидности государственных чиновников, считающих вполне нормальным повсеместное внедрение импортной элементной базы радиоэлектроники (даже в системы и комплексы ОПК).

Отсутствие современной отечественной элементной базы грозит многим отраслям промышленности нашей страны потерей конкурентоспособности, как на внутреннем, так и внешнем рынке. Более того, это сказывается на развитии инновационных технологий, модернизации экономики и, самое главное, угрожает обороноспособности России. Чтобы лучше понять, какого этапа развития достигла современная радиоэлектроника, следует вспомнить, как она возникла и развивалась.

Этапы развития радиоэлектроники в мире

Радиоэлектроника есть продолжение и модернизация в расширенном толковании радиотехники - науки о методах и средствах передачи и приема сообщений на расстояние посредством электромагнитных волн, распространяющихся в свободном пространстве. В рамках радиоэлектроники имеют дело с устройством и функционированием радиотехнических устройств и систем (РТУС) с применением электронных приборов – полупроводниковых, электровакуумных, акустоэлектронных, оптоэлектронных и других [1]. Три человека стоят у истоков радиотехники: русский ученый Александр Степанович Попов (1859 - 1906 г.г.), итальянский изобретатель Гульельмо Маркони (1874 - 1937 г.г.) и американец сербского происхождения Никола Тесла (1856 - 1943 г.г.). Важно отметить и предшественников этих великих изобретателей.

В 1873 г. английский ученый Джеймс Клерк Максвелл опубликовал работу «Трактат по электричеству и магнетизму». Как следствие из составленных им уравнений, следовал вывод о возможности распространения электромагнитных волн в свободном пространстве со скоростью света. Но полученному теоретическим путем открытию мало кто поверил. Однако спустя 15 лет немецкий ученый Генрих Рудольф Герц экспериментальным путем доказал справедливость теории Максвелла. От опытов Герца, опубликовавшего результаты своих экспериментальных исследований, отталкивались и Попов, и Маркони, и Тесла.

7 мая 1895 г. А.С. Попов впервые продемонстрировал на заседании физического отделения Русского физико-химического общества свой радиоприемник, принимавший колебания, излучаемые видоизмененным осциллятором Герца. Этот день в нашей стране и отмечается, как День радио. 24 марта 1896 г. на заседании того же общества А.С. Попов помимо радиоприемника демонстрирует и созданный им искровой радиопередатчик, передав из одного здания в другое азбукой Морзе

первую в мире радиотелеграмму. Текст ее был краток: «ГЕНРИХ ГЕРЦ» [2]. Этой телеграммой А.С. Попов продемонстрировал дань уважения своему предшественнику.

Другой изобретатель радио - Г. Маркони - в 1896 г. получил первый английский патент на систему беспроволочной телеграфии [3]. В июле 1897 г. он основал компанию «The Wireless Telegraph & Signal Company Limited» (в 1900 г. она была переименована в «Marconi`s Wireless Telegraph Company Limited»). Маркони в Англии зарекомендовал себя не только великим изобретателем, но крупным предпринимателем, сумевшим быстро и эффективно внедрить в промышленное производство изобретенные им радиотелеграфные аппараты, приносившие созданной им компании большую прибыль. А что же А.С. Попов? В 1898 г. в научных исследованиях профессора А.С. Попова наступил важный этап. Построив и установив приемо-передающие устройства на учебном судне «Европа» и крейсере «Африка», он довел дальность беспроволочной связи до 5 км. Эти опыты подтверждают его приоритет, как изобретателя радио. Это же подтверждает и официальный отказ Морского ведомства в выдаче патента Г. Маркони на его систему связи в России. В том же 1898 г. в Петербурге выпускается первая партия приемников конструкции А.С. Попова для военно-морского флота. Но самое главное состоит в том, что от телеграфной беспроводной связи А.С. Попов переходит к радиотелефонной. На построенный им «приемник депеш» с телефонными трубками А.С. Попов получает привилегию (Российский патент) №6066 с приоритетом от 14(26) июля 1899 г. На свое изобретение А.С. Попов получает также патенты в Англии и Франции. Более того, в Париже на Всемирной выставке в 1900 г. приемник А.С. Попова был удостоен золотой медали [2].

Маркони, получив отказ в выдаче патента во Франции, устремляется в Америку. В США он сближается с изобретателем лампы накаливания Томасом Эдисоном. Из тысячи изобретений Эдисона только одно может быть отнесено к беспроволочному телеграфу. 14 мая 1885 г. за десять лет до знаменитой демонстрации первого в мире радиоприемника А.С. Поповым Эдисон подал заявку для получения патента на «передачу без проводов сигналов азбуки Морзе». Однако это была низкочастотная индуктивная связь, действующая на малом расстоянии. В 1903 г. Эдисон, стремясь поддержать предприимчивого Маркони в его стремлении проникнуть со своим изобретением на американский рынок, передал ему этот свой патент. Однако этот благородный поступок Эдисона не помог Маркони заявить о себе в США, как об изобретателе радио. Свой приоритет в изобретении радио в США сумел отстоять выдающийся изобретатель Никола Тесла. После длительной тяжбы в Верховном суде США уже после смерти Теслы в 1943 г. он был официально признан изобретателем радио. Непризнанный в США Маркони, тем не менее, добивается больших успехов в своей деятельности и в 1901 г. устанавливает радиосвязь через Атлантический океан с Америкой, а в 1918 г. - с Австралией.

История радиоэлектроники, насчитывающая немногим более 100 лет, весьма показательна и является одним из блестящих примеров бурного развития научно-технического прогресса. Эту историю можно разбить на ряд характерных этапов. В каждом из них радиоэлектроника достигала своего известного развития, но в то же время появлялось нечто новое, что и являлось основой следующего этапа.

Первый этап развития начался тогда, когда такого понятия как радиоэлектроника еще и не существовало. Это было время, как уже было сказано, первых изобретений беспроводного телеграфа А.С. Поповым, Г. Маркони, Н. Теслой и др. В этот период зарождалась радиотехника, и в научных исследованиях происходил переход от сильно затухающих колебаний к незатухающим колебаниям. Искровые передатчики и кристаллические детекторы на этом этапе постепенно достигли технического совершенства. В завершении первого периода, который приходился на конец первой Мировой войны, появились дуговые передатчики и машины высокой частоты.

Казалось бы, все возможности радиотехники на этом этапе были исчерпаны и ее развитие должно бы остановиться. Однако появляются новые приборы – электронные лампы. Именно радиолампы внесли принципиально новое качество в развитие радиоэлектроники. Поэтому второй этап в развитии радиоэлектроники характеризовался широким использованием электронных ламп, как в качестве детекторов и усилителей в приемниках, так и генераторов в передатчиках. Радиолампы усовершенствовали радиотелефонию. Все это предопределило формирование новых отраслей промышленности - радиовещание, телевидение, радиолокация, автоматика, телемеханика и вычислительная техника.

Первые отечественные радиолампы были созданы в Нижегородской радиолаборатории в трудные первые годы советской власти. В разруху, голод и гражданскую войну коллективу Нижегородской радиолаборатории во главе с М.А. Бонч-Бруевичем удалось создать мощные радиолампы, не имевшие в то время аналогов и за рубежом [4]. За годы первых советских пятилеток наметился бурный скачок в развитии отечественной радиоэлектроники. К началу Великой Отечественной войны советская промышленность уже серийно выпускала средства связи, радиостанции, пеленгационные и радиолокационные системы, ни в чем не уступавшие зарубежным аналогичным образцам [5]. В это время электронные приборы получили массовое распространение, их номенклатура расширилась от усилительных и генераторных радиоламп разных типов и характеристик до электроннолучевых трубок разных размеров и очень чувствительных фотоэлектронных приборов. Опять наступило насыщение в развитии радиоэлектроники. Нужен был новый качественный скачок для ее новой модернизации.

Появление к концу второй Мировой войны полупроводниковых усилительных приборов – транзисторов указало на начало третьего этапа в развитии радиоэлектроники. Полупроводники были известны давно и широко использовались в качестве детекторов и выпрямителей. Исторически первым устройством, в котором полупроводник заменил электронную лампу, был изобретенный в 1922 г. сотрудником Нижегородской радиолабора-

тории О.В. Лосевым приемник, известный под названием «кристадин» [6]. Однако недостаточное в то время теоретическое и экспериментальное знание свойств полупроводников не позволило этому изобретению оказать заметное влияние на развитие радиоэлектроники. Только изобретение транзистора в США в 1948 г. [7], способного заменить электронную лампу, привело к широкому проникновению полупроводников в радиоэлектронику. Полупроводниковые приборы становятся непременной и существенной частью радиоэлектронных устройств и систем. Их применение приводит к коренным изменениям в радиоэлектронной аппаратуре как схемотехническим, так и конструктивным. Появляются полупроводниковые источники электрической энергии - это и фотоэлектрические преобразователи, и термоэлектрические генераторы. Широко используются полупроводниковые источники света: светодиоды и полупроводниковые лазеры. Совершенствование полупроводниковой технологии позволило существенно уменьшить габариты разрабатываемых радиоэлектронных систем сначала за счет микромодульных конструкций, а затем и микросборок, используя монтаж с бескорпусными полупроводниковыми элементами.

Четвертый этап в развитии радиоэлектроники наступил одновременно в СССР (НИИ «Пульсар») [8] и США (Texas Instruments) [9], где были созданы первые твердотельные интегральные схемы. Интегральная радиоэлектроника хотя и медленно, но стала набирать обороты. Интегральные микросхемы постепенно вытесняют транзисторы, как в аналоговых, так и цифровых устройствах. Появляются настольные вычислительные машины на цифровых интегральных микросхемах. Новое направление в радиоэлектронике - цифровая обработка сигналов, благодаря цифровым интегральным схемам и первым аналого-цифровым преобразователям в интегральном исполнении, получило практическую реализацию.

За десятилетие своего развития интегральная радиоэлектроника достигает таких высот, что в 1971 г. в компании Intel (США) создается первый в мире программируемый однокристальный микропроцессор I4004 с 2300 транзисторами на одном кристалле. А еще через 8 лет эта же фирма создает первый программируемый сигнальный процессор I2920 на одном кристалле, который мог подаваемый на его вход аналоговый сигнал преобразовывать в цифровой код, подвергать код цифровой обработке по запрограммированному алгоритму и преобразовывать результат в аналоговую форму, выдавая его на выход. Несколько позже в СССР появляется аналог американского микропроцессора I8080 с серийным названием 580ИК80, а в 80-х годах двадцатого столетия - усовершенствованный аналог сигнального процессора I2920 с серийным названием КМ1813ВЕ1 [10].

Последнее десятилетие прошедшего столетия характеризовалось наступлением пятого этапа развития радиоэлектроники - эры программируемой радиоэлектроники. Полезно отметить характерные особенности этого нового современного этапа развития радиоэлектроники. Программируемая радиоэлектронная система – это не только цифровая, но и реконфигурируемая система, способная к постоянному усовершенствованию и модернизации только за счет смены программного обеспечения. Концепция программируемой радиоэлектроники отражает главное изменение в современной конструкторской парадигме, для которой соотношение аппаратно-программных средств выбирается с явным преобладанием программных средств, что и обеспечивает возможность быстрого изменения тактико-технических характеристик РТУС в соответствии с изменяющимися требованиями и возможностями. Эта концепция распространяется практически на все разрабатываемые современные радиоэлектронные устройства и системы, начиная с сотовых телефонов и заканчивая радиолокационными станциями.

Разработчиками современных РТУС становятся, главным образом, программисты. Создание программируемых устройств существенно сокращает период разработки новых моделей радиоэлектронных систем. Их аппаратная реализация требовала макетирования, отладки и трудоемкой регулировки изделия в целом. В то же время для одного и того же воздействия сигнальный процессор с одной и той же программой даст на выходе один и тот же отклик. Это существенно снижает затраты на изготовление и проверку программируемого устройства.

Быстрый рост и сменяемость требований, предъявляемых к современным радиоэлектронным устройствам и системам, использование различных сигналов требуют реализации нескольких режимов работы проектируемых устройств. Такая многорежимность РТУС, в свою очередь, требовала набора конкретных аппаратных средств для каждого режима. Это увеличивало объем аппаратуры и стоимость изделия в целом. А программируемая система производит смену режима простой перезагрузкой программного обеспечения. Более того, программируемая система может быть модернизирована уже на этапе испытаний для корректировки непредвиденных проблем или для улучшения тактико-технических характеристик. Другим преимуществом программируемой системы является ее способность к глубокой самодиагностике, как всей системы, так и отдельных устройств, что повышает надежность изделия и требует минимального его обслуживания.

В то время как в России в 90-е годы была практически полностью разрушена электронная промышленность, на западе, а затем в Юго-Восточной Азии и Китае была создана мощная инфраструктура производств комплектующих для программируемой электроники. К началу нового XXI века там уже были разработаны несколько поколений микропроцессоров, сигнальных процессоров и программируемых логических интегральных схем (ПЛИС).

Появление целых семейств высокопроизводительных сигнальных процессоров, созданных фирмами-лидерами в этой области Texas Instruments и Analog Devices, а также супер ПЛИС из Xilinx и Altera, открыло новые горизонты в развитии программируемых радиоэлектронных систем.

Рис. 1. Первый отечественный программируемый сигнальный процессор КМ1813ВЕ1 на фоне справочного листка с его описанием

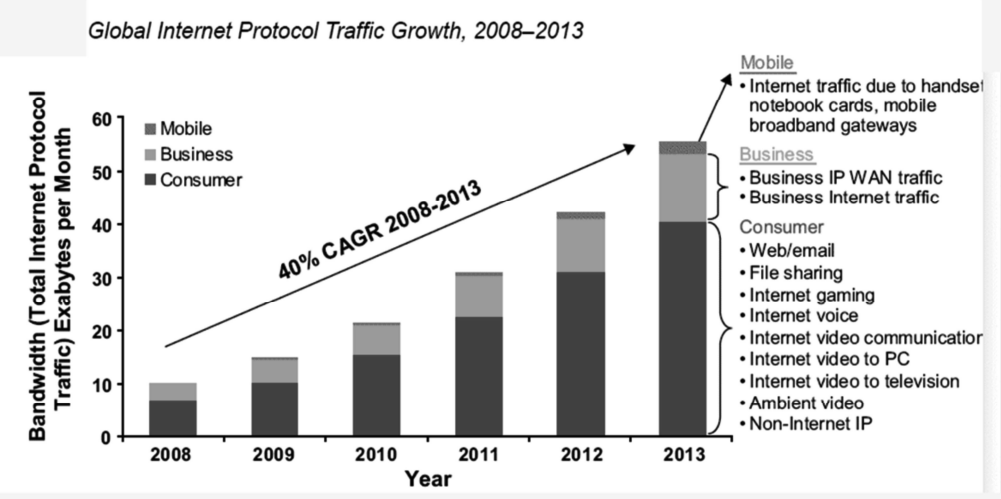


Рис. 2. Рост объема информации по годам в глобальном Интернете

Проекты, реализованные на сигнальных процессорах и ПЛИС, уже применяются в таких областях как адаптивная цифровая фильтрация, адаптивное обнаружение и накопление сигналов, управление фазированными антенными решетками, программируемое формирование различных генерируемых сигналов сложной формы с большой частотно-временной базой и их программируемое сжатие [11]. Сформировалось целое направление в создании программируемых радиоприемных устройств (Software Defined Radio). Эти идеи программируемой обработки потоковых данных на сигнальных процессорах и ПЛИС широко используются в системах цифрового радиовещания и телевидения, сотовой связи и Интернете. О том, какой поток информации (глобальный Интернет трафик в экзабайт в месяц) к 2013 г. будут обрабатывать программируемые радиоэлектронные системы, показан на рис. 2, заимствованном на сайте фирмы Altera [12]. Полезно отметить, что приставка экза (exa) означает 10^{18}.

Инновационный курс
«Программируемая радиоэлектроника»

Чтобы в должной мере ответить на этот новый современный технологический вызов в развитии радиоэлектроники необходимо иметь высококвалифицированные инженерные кадры. Для этого в инженерных высших учебных заведениях, наряду с традиционными дисциплинами (например, «Промышленная электроника», «Схемотехника аналоговых электронных устройств», «Цифровые устройства и микропроцессоры») стали читать студентам новый учебный курс «Программируемая радиоэлектроника». Основные цели и задачи курса состоят в подготовке обучающихся построению программируемых устройств на сигнальных процессорах и ПЛИС, а также показе на конкретных примерах всех преимуществ программируемой радиоэлектроники.

На факультете радиотехнических и телекоммуникационных систем МИРЭА был поставлен и прочитан студентам третьего и четвертого курсов, обучающимся по специальностям 210302 Радиотехника, 210304 Радиоэлектронные системы и 210201 Проектирование и технология радиоэлектронных средств, инновационный учебный курс «Программируемая радиоэлектроника».

Инновационный курс «Программируемая радиоэлектроника» включал в себя следующие разделы: программируемые цифровые фильтры; программируемые устройства оценки параметров помех; программируемые обнаружители сигналов.

Большое внимание в новом курсе уделено приобретению на лабораторных занятиях студентами практических навыков, умений и профессиональных компетенций по разработке программируемых устройств. Алгоритмы программируемых устройств вначале моделируются в системе МАТЛАБ, а затем программа работы этих устройств после соответствующей компиляции загружается в ПЗУ сигнальных процессоров или конфигурирует ПЛИС. При этом для синтеза цифровых фильтров, в программу курса включено изучение таких специализированных пакетов проектирования как Digital Filter Design Pack (DFDP) и dsPIC Filter Design. Верификация правильной работы программируемых устройств осуществляется сравнением предварительных результатов, полученных на модели в системе МАТЛАБ, с результатами работы программируемого устройства.

В отличие от сигнальных процессоров, для которых основной этап проектирования устройств ЦОС сводится к разработке программы, для ПЛИС начальный этап - это разработка алгоритма, т.е. выбор вычислительных процедур, обеспечивающих реализацию оператора обработки с учетом особенностей FLEX микросхем. При этом можно воспользоваться DSP конструктором (например, DSP TOOL KIT фирмы Altera), с готовыми заготовками АЛУ и умножителя как с фиксированной, так и плавающей запятой или мегафункциями включающими в себя наиболее распространенные алгоритмы ЦОС в том числе БПФ, КИХ и БИХ фильтры и т.д. В частности, этот пакет можно дополнить и своими мегафункции такими как: адаптивные фильтры, устройства выделения квадратур (X,Y), адаптивное пороговое устройство и т.д. Хотелось бы подчеркнуть, что курс «Программируемая радиоэлектроника» не дублирует уже читаемые курсы по цифровой обработке сигналов, хотя в нем и дается краткий обзор основных принципов ЦОС. В данном курсе ЦОС - это всего лишь инструмент, и поэтому цель курса развить интуитивное понимание принципов ЦОС на программном уровне. Поэтому чрезмерная детализация предмета в части ЦОС не целесообразна. Для этого существуют другие специальные курсы, углубляющие знания в отдельных направлениях ЦОС.

В данном курсе минимизировано и рассмотрение аппаратного обеспечения, хотя во всех примерах обсуждаются схемы включения DSP и ПЛИС на базовых платформах. Основное внимание слушателей сфокусировано в направлении разработки отдельных программируемых устройств на этих базовых платформах с максимальным использованием программного обеспечения. Студентам демонстрируются все достоинства инновационного подхода к построению программируемых встраиваемых радиотехнических и телекоммуникационных устройств и систем. В том числе и такое важное преимущество, как сокращение периода разработки новых моделей программируемых устройств.

К сожалению, сейчас в России производство электронных компонент для программируемой радиоэлектроники не отвечает современным требованиям. Правительство РФ принимает экстренные меры по преодолению отставания в радиоэлектронике. В частности, Правительством РФ для возрождения отечественной электроники в 2006 г. была принята Федеральная целевая программа «Национальная технологическая база» с подпрограммой «Развитие электронной компонентной базы» на 2007-2011 годы. Однако эта программа не выполнена. Если зарубежные фирмы уже предлагают программируемые микросхемы, изготовленные по технологии 28 нм, то наши же организации еще не смогли освоить производство программируемых микросхем даже с технологией 90 нм. Тем не менее, Правительство РФ принимает новую Федеральную целевую программу «Развитие электронной компонентной базы и радиоэлектроники» до 2015 года [13].

В настоящее время, изучая курс «Программируемая радиоэлектроника», студенты, к сожалению, разрабатывают программы пока для импортных сигнальных процессоров и ПЛИС. Однако введение в системную подготовку инженеров инновационного учебного курса «Программируемая радиоэлектроника» обеспечит готовность дипломированных специалистов к широкому и эффективному использованию отечественных программируемых микросхем при разработке систем различного назначения.

Литература

1. Каганов В.И., Битюков В.К. Основы радиоэлектроники и связи. – М.: Горячая линия-Телеком, 2006. – 542 с.

2. А.С. Попов в характеристиках и воспоминаниях современников. - М.-Л.: АН СССР, 1958.

3. Marconi G. Brit. Patent No. 12039. June 2, 1896.

4. Арнаутов Л.И., Карпов Я.К. Прорыв в грядущее. Страницы жизни М.А. Бонч-Бруевича. - М.: Московский рабочий, 1986.

5. Бартенев В.Г. К 70 летию создания первых РЛС дальнего обнаружения. - Современная электроника. 2010, №3.

6. Лосев О.В. Детектор-усилитель; детектор-генератор. - ТиТбп. НРЛ, 1922, №14.

7. Brattain W.H., Bardeen J. Three Electrode Circuit Element Utilizing Semiconductive Materials. - U.S. Patent 2,524,035. Oct.3, 1950. Filed June 17, 1948.

8. Малин Б.В. Создание первой отечественной микросхемы. - НИИ-35 (теперь НИИ «Пульсар»). www.chip-news.ru.

9. Kilby J.S. Miniaturized Electronic Sircuits. - U.S. Patent 3,138,743. June 23, 1964. Filed Feb. 6, 1959.

10. Однокристальная перепрограммируемая микроЭВМ с аналоговыми устройствами ввода и вывода КМ1813ВЕ1А,Б. Спровочный листок ЯТ09360, ИЭВТ, 1980.

11. Bartenev V. Software Radar: New Reality. - Report on the International conference RADAR 2006. China, 2006.

12. Stratix V. FPGAs: Built for Bandwidth. Altera's 28-nm FPGAs deliver the industry's highest bandwidth, highest level of system integration». www.altera.com.

13. Концепция федеральной целевой программы «Развитие электронной компонентной базы и радиоэлектроники на 2008-2015 годы». Утверждена распоряжением Правительства РФ №972-р от 23 июля 2007. www.minprom.gov.ru.

УДК 519.254

СТАТИСТИЧЕСКИЙ МЕТОД ВОССТАНОВЛЕНИЯ ПРОФИЛЯ ПО ДАННЫМ ЛАЗЕРНОГО СКАНИРОВАНИЯ

***Броневич А.Г.**, д.ф-м.н., главный научный сотрудник Центра перспективных фундаментальных и прикладных исследований Научно-исследовательского и проектно-конструкторского института информатизации, автоматизации и связи на железнодорожном транспорте (ОАО «НИИАС»), e-mail: brone@mail.ru .*
***Каркищенко А.Н.**, профессор, руководитель Центра перспективных фундаментальных и прикладных исследований Научно-исследовательского и проектно-конструкторского института информатизации, автоматизации и связи на железнодорожном транспорте (ОАО «НИИАС»), e-mail: karkishalex@gmail.com*
***Уманский В.И.**, кандидат технических наук, Генеральный директор ЗАО «ИнтехГеоТранс», e-mail: umanvi@yandex.ru*

Ключевые слова: профиль земляного полотна, лазерное сканирование, случайная последовательность, функция правдоподобия, дискриминантная функция, разладки случайного процесса.

Статья посвящена восстановлению поперечного профиля земляного полотна по зашумленным данным лазерного сканирования. Приведена формальная постановка задачи. Показано, что задача может быть сведена к анализу случайной последовательности измерений. Дано решение задачи с помощью дискриминантной функции Фишера. Показано, как с помощью построенной дискриминантной функции локализуются точки переключения в кусочно-параметризованной случайной последовательности, и предложен вычислительно эффективный способ приближенного отыскания таких точек. Приведен пример реализации метода на моделируемых данных.

Введение

В настоящее время развиваются методы автоматического мониторинга профиля земляного полотна и верхнего строения железнодорожного пути на протяженных участках дороги. Под воздействием динамических поездных нагрузок и влиянием природно-климатических факторов происходит деформация полотна, так что со временем оно может не удовлетворять требованиям нормативной документации и условиям дальнейшей эксплуатации. В этом случае предпринимаются меры по поддержке земляного полотна. Постоянная диагностика земляного полотна, учитывая протяженность существующих железнодорожных линий, является трудоемкой, требующей постоянного внимания процедурой [1].

Эффективность диагностики во многом определяется такими свойствами как: объективность, достоверность, простота проведения, оперативность, дешевизна. При этом требуется мониторинг целых участков и направлений, применяющийся для деформаций массового характера на большом протяжении. Особую ценность приобретают методы мониторинга, главными свойствами которых являются быстрота проведения и дешевизна. Наиболее простыми, но важными являются наблюдения за соответствием сложившихся профилей земляного полотна с типовыми нормативными профилями для заданных грунтов, слагающих насыпи или выемки. На основании таких наблюдений и их сравнительного анализа требуется фиксировать такие параметры как:

– ненормативная ширина плеча балластной призмы;

– ненормативная ширина обочины земляного полотна;

– места с завышенной крутизной откосов;

– нарушение структурных линий полотна, представляющих собой продольные линии излома поверхности земли в полосе отвода и др.

В связи с этим возникает необходимость разработки методов автоматического сравнения поперечных профилей на основе использования специальных технических средств получения информации о поверхностном рельефе земли. Для решения этой задачи перспективным является использование современных лазерных сканеров, которые, по сути, представляют собой лазерные дальномеры. Технически это сводится к тому, что посредством геометрических преобразований оказывается возможным восстановить в некоторой координатной системе положение большого количества точек (облака точек) поверхности. Анализ полученных данных позволяет сделать выводы о поверхностном строении просматриваемого участка. Современные лазерные сканеры импульсного типа способны сканировать поверхность объекта со скоростью до 10000 точек в секунду, а сканеры фазового типа - до 500000 точек.

Стандартный способ получения информации о профиле состоит в создании технических комплексов наземного лазерного сканирования и обработки информации непосредственно на перемещающемся поезде [2]. Технологически данная задача решается путем установки двух лазеров на задней стенке последнего вагона. В процессе движения вагона каждый лазер сканирует свою часть полупространства – соответственно, слева и справа относительно движения, формируя совокупность замеров (рис. 1).

В данной работе приведена формальная постановка задачи восстановления поперечного профиля земляного полотна по зашумленным данным лазерного сканирования, описан подход к решению задачи путем сведения ее к анализу случайной последовательности измерений. Дано решение данной задачи с помощью дискрими-

нантной функции Фишера в полярной системе координат путем локализации точек переключения в параметризованной случайной последовательности.

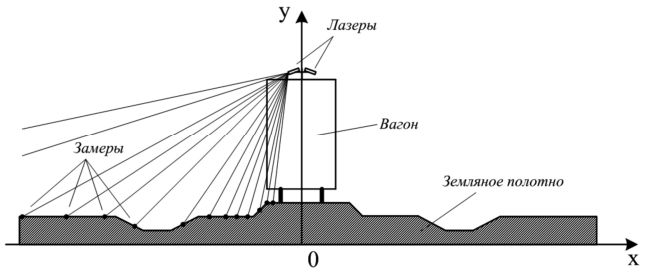


Рис. 1. Схема измерений земляного полотна

Описание и общая постановка задачи

При организации лазерного сканирования следует принимать во внимание ряд особенностей:

– совокупность измерений (замеров) представляет собой дискретное множество точек, количество которых определяется частотой излучения лазера и может варьироваться в зависимости от его технических возможностей;

– за один проход лазеров получается совокупность замеров по общему углу охвата в 360 градусов (по 180 градусов на каждый лазер);

– распределение замеров зависит от высоты расположения лазера над поверхностью земли;

– так как сканирующий луч перемещается дискретно во времени, а значит, и в пространстве, через равный угловой шаг, то замеры расположены относительно равномерно по угловой мере;

– замеры располагаются неравномерно в линейном измерении, поскольку происходит отражение от поверхностей, расположенных под разным углом по отношению к лучу лазера и находящихся на разном расстоянии от источника излучения (при этом замеры сгущаются при приближении луча к балластной призме);

– данные реальных измерений оказываются сильно зашумленными; ошибки измерений возникают из-за погрешностей собственно лазерного дальномера, а также в большей степени ввиду естественной неровности отражающей поверхности земляного полотна, наличия на нем посторонних предметов, растительности и пр.

Ниже рассматривается задача восстановления по лазерным замерам формы земляного полотна и балластной призмы на достаточно протяженном участке дороги. Теоретически решение данной задачи позволит провести автоматическое сравнение реального рельефа полотна с рельефом, соответствующим требованиям нормативных документов.

Рассмотрим один цикл съема информации о земляном полотне с помощью лазерного датчика. Под циклом будем понимать период лазерного сканирования, соответствующий обзору лазера сверху-вниз, т.е. на 180 градусов. Далее будем рассматривать информацию, получаемую только при сканировании земляного полотна, и не будем принимать во внимание точки, возникающие при отражении луча сканера от объектов, находящихся выше полотна. В результате сканирования получается совокупность замеров в поперечном сечении, т.е. в плоскости, перпендикулярной направлению пути[1]. Если точки соединить последовательно отрезками прямых линий, то можно получить представление о форме реального профиля. Полученные замеры будут лишь приближенно описывать земляное полотно. Даже в том случае, если оно соответствует требованиям действующих нормативов, измерения будут содержать случайную составляющую, появление которой вызвано различными причинами: неровностями поверхности земли; различными предметами, отражающими лазерный луч; растительным покровом; снегом; собственно погрешностями лазерного дальномера; продольными и поперечными колебаниями вагона и др.

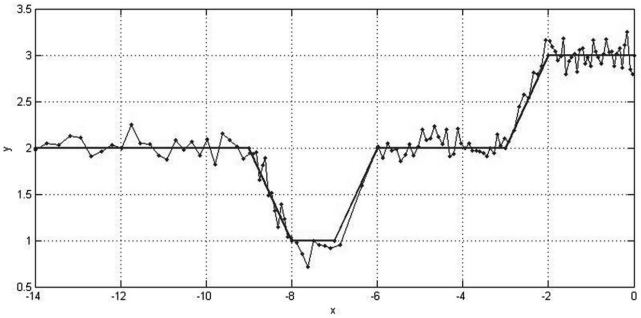


Рис. 2. Пример последовательности замеров

Будем считать, что выполняются следующие предположения:

1. Ошибка каждого измерения в одном цикле является суммой большого количества независимых или слабо зависимых факторов, поэтому она может быть достаточно точно описана случайной величиной, подчиненной нормальному закону распределения;

2. Случайные величины имеют короткий «промежуток зависимости». Иными словами, возмущения слабо связаны с их предысторией, и поэтому статистическая зависимость двух измерений быстро уменьшается с ростом расстояния между ними. Это условие гарантирует, что получаемые на основе таких измерений статистические оценки являются асимптотически гауссовскими в смысле центральной предельной теоремы [3].

В результате одного цикла измерений получается случайная последовательность замеров, которая примерно отражает рельеф земляного полотна слева и справа от железнодорожного пути. Пример такой последовательности, полученной в результате моделирования с гауссовским зашумлением слева от дальномера, приведен на рис. 2. Отметим, что точки показаны в декартовой системе координат, в которой ось абсцисс расположена перпендикулярно направлению перемещения дальномеров, как показано на рис. 1. На рисунке видно, что точки сгущаются по мере приближения к началу координат.

[1] На самом деле из-за движения поезда точки замеров будут лежать на пересечении винтовой линии, описываемой лучом лазера, и поверхности земли, пересекающейся с этой винтовой линией. Будем считать в дальнейшем, что замеры на участке одного цикла путем геометрических преобразований приведены к рассматриваемой перпендикулярной плоскости.

Задача обработки полученных данных состоит в том, чтобы по последовательности измерений восстановить исходную форму профиля полотна (показанную на рис. 2 более толстой линией – графиком кусочно-линейной функции).

Формальное описание статистического анализа профиля полотна

Последовательность замеров можно рассматривать как реализацию некоторого случайного процесса, который является нестационарным, т.е. меняет статистические характеристики на различных участках сканирования. При этом могут быть участки, на которых эти характеристики примерно постоянны, но могут быть и такие, на которых они меняются по неизвестному неслучайному закону. Выявление этого закона или функции, описывающей эти изменения, а также определение точек, в которых вид этой функции меняется, позволяют провести анализ реального профиля. Задачи в подобной или похожей формулировке рассматриваются во многих областях, связанных с обработкой статистических данных (случайные процессы и последовательности, временные ряды, конкретные приложения). Многообразие методов решения указанной задачи определяется тем, какие априорные предположения допускаются в каждом конкретном случае.

Рассмотрим вначале широко используемую и удобную модель кусочно-стационарного случайного процесса, в котором изменение свойств происходит скачком. Такой процесс можно представить в виде последовательности стационарных участков, на которых его свойства одинаковы. Обнаружение стационарных участков можно рассматривать как классическую задачу математической статистики – оценивания точек, в которых происходит изменение свойств кусочно-стационарного случайного процесса. Подобные случайные процессы возникают в разных прикладных задачах – в системах управления в пространстве состояний, при мониторинге окружающей среды, анализе сейсмической активности и др. С этой точки зрения стационарные участки представляют собой устойчивые и возможно периодически меняющиеся в процессе наблюдения состояния порождающего их объекта. Мы будем рассматривать случайные процессы, стационарные в широком смысле [4, 5].

В практических задачах информация об исследуемом случайном процессе обычно представляет собой последовательность измерений $(y_1, y_2,..., y_N)$ в некоторых точках $(x_1, x_2,..., x_N)$, а вероятностной моделью описания данных измерений является последовательность $(\xi_1, \xi_2,..., \xi_N)$ случайных величин. Таким образом, считается, что измерения $(y_1, y_2,..., y_N)$ являются реализациями данных случайных величин. При описании последовательностей $(\xi_1, \xi_2,..., \xi_N)$ можно использовать цепи Маркова, различного рода авторегрессионные модели, стохастические разностные уравнения [6].

В силу сделанного выше предположения о коротком «промежутке зависимости» ограничимся в дальнейшем наиболее простой моделью такого типа, а именно, будем считать, что последовательность состоит из независимых случайных величин, т.е. принимаем предположение, что случайные величины ξ_i попарно независимы.

Рассматриваемая ниже модель последовательности независимых случайных величин ориентирована на использование

1) при небольших по объему обучающих выборках $(y_1, y_2,..., y_N)$, которые не могут гарантировать получение оптимальных статистических выводов при использовании стандартных вероятностных методов;

2) когда вычислительные ресурсы не позволяют реализовать применение более трудоемких и сложных вероятностных моделей.

Известно, что даже при использовании модели последовательности независимых случайных величин в задаче обнаружения многократных изменений свойств (или разладок) случайного процесса возникают труднопреодолимые проблемы вычислительного характера при решении ее классическими методами [7].

Предположим, что последовательность измерений $(y_1, y_2,..., y_N)$ в точках $(x_1, x_2,..., x_N)$ можно описать с помощью последовательности одномерных независимых случайных величин $(\xi_1, \xi_2,..., \xi_N)$. Будем вначале считать для простоты, что в этой последовательности можно выделить некоторое число n стационарных участков $[\chi_1, \chi_2 - 1]$, $[\chi_2, \chi_3 - 1]$, ... , $[\chi_n, \chi_{n+1} - 1]$, $\chi_1 = x_1$, $\chi_{n+1} - 1 = x_N$. Внутри каждого стационарного участка $[\chi_m, \chi_{m+1} - 1]$ случайные величины ς_i и ς_j, $x_i, x_j \in [\chi_m, \chi_{m+1} - 1]$, должны обладать постоянными математическими ожиданиями и дисперсиями, т.е. $E(\xi_i) = E(\xi_j)$, $D(\xi_i) = D(\xi_j)$.

Рассмотрим оценивание точек нарушения стационарности χ_i, $i = 1, 2,..., n$, по методу максимального правдоподобия. Будем считать, что на каждом стационарном участке $[\chi_m, \chi_{m+1} - 1]$ функцию $f_m(y)$ плотности распределения случайных величин можно задать параметрически вектором параметров a_m, т.е. $f_m(y) = f(y, a_m)$. В этом случае логарифм функции правдоподобия Фишера имеет вид:

$$\Phi(a_1,...,a_n,\ \chi_1,...,\chi_n) = \sum_{m=1}^{n}\sum_{i=\chi_m}^{\chi_{m+1}-1} \ln[f(y_i, a_m)].$$

Таким образом, для оценивания точек $\chi_1,...,\chi_N$ необходимо найти максимум логарифмической функции правдоподобия относительно аргументов $a_1,...,a_n$, $\chi_1,..., \chi_N$.

В более общем случае в последовательности одномерных независимых случайных величин на каждом из n участков может наблюдаться тренд математического ожидания, описываемый неслучайной функцией $g(x)$, при постоянной дисперсии, т.е. для ξ_i и ξ_j, таких, что $x_i, x_j \in [\chi_m, \chi_{m+1} - 1]$ имеет место $E(\xi_i) = g_m(x_i)$, $D(\xi_i) = D(\xi_j)$. Тогда для оценки точек переключения χ_i, $i = 1, 2,..., n$, по методу максимального правдоподобия следует считать, что на каждом участке $[\chi_m, \chi_{m+1} - 1]$ функция $f_m(y)$ плотности распределения случайной величины ξ_i имеет вид $f_m(y) = f(y, g_m, (x_i), a_m)$. В частности, если функции, описывающие тренд, принадлежат одному классу (например, классу линейных или квадратичных функций), то $f_m(y) = f(y, g(x_i, b_m), a_m)$,

где b_m - вектор параметров, описывающих конкретную функцию из данного класса. В дальнейшем такой случайный процесс будем называть кусочно-параметризованным процессом. Заметим, что при $g_m(x) = c_m = const$ данный случай сводится к предыдущему случаю стационарных в широком смысле участков. Логарифм функции правдоподобия Фишера имеет вид:

$$\Phi(a_1,...,a_n,\ b_1,...,b_n,\ \chi_1,...,\chi_n) = \\ = \sum_{m=1}^{n}\sum_{i=\chi_m}^{\chi_{m+1}-1}\ln[f(y_i, g(x_i, b_m), a_m)]$$

Следовательно, для оценивания точек $\chi_1,...,\chi_N$ необходимо найти максимум логарифмической функции правдоподобия относительно большего, чем в предыдущем случае, числа аргументов, а именно, $a_1,...,a_N$, $b_1,...,b_N$, $\chi_1,...,\chi_N$.

Как правило, построить рекуррентный алгоритм максимизации функционала Φ не удается, поэтому остается единственная возможность – проводить полный перебор всех возможных положений точек переключения χ_m. Эта задача для случая многократных разладок является недопустимо трудоемкой – требуется оценивать функцию правдоподобия Фишера порядка N^n раз. Эти выводы согласуются с результатами, полученными в [8], где показано, что применение методов максимального правдоподобия позволяет построить эффективные решающие правила для оценки моментов разладки, но приводят к алгоритмам очень высокой сложности, трудно реализуемым на практике. Для решения этой проблемы в [9] предлагается непараметрический метод обнаружения многократных нарушений стационарности, основанный на идее многократного разрезания и проверке однородности отдельных частей общей выборки. Этот метод использует достаточно сложную двумерную статистику, что затрудняет его применение. В работе [10] предложен алгоритм апостериорного обнаружения многократных разладок случайной последовательности по изменению моментной функции первого порядка. В нем используется одномерная статистика, имеющая смысл разности оцениваемых математических ожиданий случайного процесса на отрезках различной длины.

Идея описываемого ниже метода состоит в предварительной локализации участков, содержащих «нестационарные» точки, т.е. точки, в которых происходит изменение характеристик случайной последовательности. Кратко суть метода состоит в следующем. На множестве независимых наблюдений вводится дискриминантная функция, позволяющая судить о стационарности случайной последовательности в произвольной точке x. По априорным сведениям о возможной длине и отличительных свойствах параметризованного участка делаются правдоподобные предположения о поведении этой функции и вводится понятие нестационарных точек. После этого определение таких участков сводится к задаче кластеризации нестационарных точек и определению точек разладки.

Дискриминантная функция для случая кусочно-параметризованной последовательности

Как отмечалось выше, распределение замеров в условиях лазерного сканирования характеризуется неравномерностью (см. рис. 2). Это усложняет использование статистических методов оценивания параметров случайной последовательности. Вместе с тем, в данном случае можно избежать указанной неравномерности, если воспользоваться равномерностью распределения точек по угловой мере. Для этого целесообразно перейти к полярной системе координат, связанной с точкой подвеса лазерного дальномера. На рис. 3 показано соотношение декартовой и полярной системы координат.

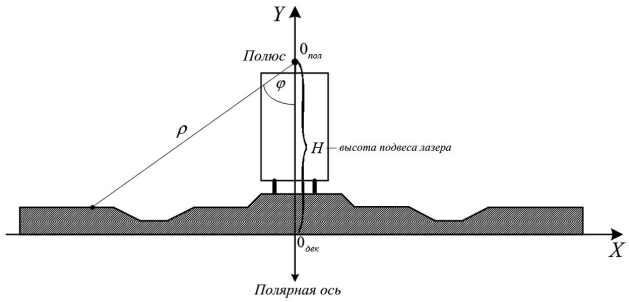


Рис. 3. Полярная система координат, связанная с лазерным дальномером

Из рисунка видно, что формулы, связывающие координаты (x, y) некоторой точки в декартовой и (p, φ) в полярной системах координат, имеют вид: $x = p\sin\varphi$, $y = H - p\cos\varphi$, где H – высота подвеса лазера, т.е. расстояние между полюсом полярной и началом декартовой системы координат.

На рис. 4 видно, как кусочно-линейная функция, задающая некоторый идеальный профиль и представленная в декартовой системе координат (слева), выглядит в полярной системе (справа). На правом графике по горизонтальной оси отложены значения полярного угла, а по вертикальной – расстояние от точки до полюса. Далее будем предполагать, что все статистические данные представлены именно в полярной системе координат.

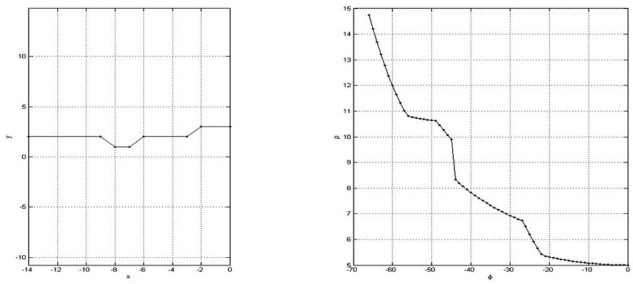


Рис. 4. Кусочно-линейная функция, описывающая «идеальный» профиль, в декартовой (слева) и полярной (справа) системах координат

Поскольку в соответствии с существующими нормативами все регламентированные профили земляного полотна являются непрерывными кусочно-линейными функциями, то при переходе к полярной системе координат они будут представлять собой непрерывную кусочно-параметризованную кривую, состоящую из кусков прямых линий, заданных в полярной системе: $\rho = \dfrac{r_m}{\cos(\varphi - \psi_m)}$, $m = 1,2,...,n$, где r_m – расстояние от прямой, описывающей m-й участок, до полюса, а ψ_m – угол, который образует нормаль к прямой с полярной осью. Поэтому наблюдаемый профиль в поляр-

ной системе координат будет реализацией $(p_1, p_2,..., p_N)$ кусочно-параметризованной последовательности случайных величин $(\eta_1, \eta_2,..., \eta_N)$, которая на m-м участке описывается трендом вида $E(\eta_1) = \frac{r_m}{\cos(\varphi_i - \psi_m)}$. Построим дискриминантную функцию для анализа таких случайных последовательностей в полярной системе координат.

Основное требование, которое предъявляется к дискриминантной функции, состоит в том, что она должна принимать максимальные значения в областях изменения свойств последовательности независимых случайных величин η_t и минимальные – там, где эти свойства относительно неизменны или подчинены регулярной зависимости, т.е. внутри параметризованных участков. Для построения дискриминантной функции потребуется условие минимально возможной «длины» произвольного параметризованного участка. Будем предполагать, что количество точек-замеров на каждом параметризованном участке не меньше некоторого значения $2S$, где S – натуральное число, т.е. для любого $i \in \{1, 2,..., n\}$ выполняется неравенство $|\{\varphi_t \mid \chi_i \le \varphi_t < \chi_{i+1}\}| \ge 2S$. Это условие представляется менее жестким, чем априорно известное число стационарных участков, правдоподобные предположения о котором в реальных задачах сделать достаточно трудно.

Рассмотрим последовательность независимых случайных величин $(\eta_1, \eta_2,..., \eta_N)$ с функциями плотности $f_i(\rho)$ и математическими ожиданиями, подчиняющимися закону $E(\eta_1) = g(\varphi_t, r_m, \psi_m) = \frac{r_m}{\cos(\varphi_i - \psi_m)}$ при $\chi_i \le \varphi_t \le \chi_{i+1}$ с неизвестными и подлежащими определению параметрами ρ_m и ψ_m. Пусть в точке φ_t происходит изменение статистических свойств этой последовательности, т.е.

$$f_i(\rho) = \begin{cases} f(\rho, g(\varphi, r_1, \psi_1), a_1), & i > t, \\ f(\rho, g(\varphi, r_2, \psi_2), a_2), & i > t. \end{cases}$$

Локализуем момент разладки по реализации случайной последовательности $(\rho_1, \rho_2,..., \rho_N)$ с помощью дискриминантной функции, строящейся по логарифмической функции правдоподобия:

$$L(k) = \max_{r_1, \psi, a_1} \sum_{i=k-s+1}^{k} \ln f(\rho_i, g(\varphi_i, r_1, \psi_1), a_1) + \max_{r_2, \psi_2, a_2} \sum_{i=k+1}^{k+s} \ln f(\rho_i, g(\varphi_i, r_2, \psi_2), a_2) + \max_{r_3, \psi_3, a_3} \sum_{i=k-s+1}^{k+s} \ln f(\rho_i, g(\varphi_i, r_3, \psi_3), a_3).$$

Данную функцию можно интерпретировать как логарифм отношения правдоподобия двух конкурирующих гипотез: H_1 – в точке φ_k произошла разладка случайного процесса (первое и второе слагаемое), против гипотезы H_2 – отрезок $[\varphi_{k-s+1}, \varphi_{k+s}]$ принадлежит кусочно-параметризованному участку (третье слагаемое). При этом параметры r_1, ψ_1, a_1, r_2, ψ_2, a_2, r_3, ψ_3, a_3 случайной последовательности на отрезках $[k-s+1, k]$, $[k+1, k+s]$, $[k-s, k+s]$ оцениваются по методу максимального правдоподобия. Можно предположить, что построенная таким образом дискриминантная функция будет обладать достаточно хорошими статистическими свойствами, например, можно ожидать, что момент разладки вероятнее всего произошел в тех точках случайной последовательности, в которых функция $L(k)$ принимает относительно большие значения, и наоборот – значения этой функции будут сравнительно невысоки внутри стационарных участков.

Для нормально распределенных случайных величин $\eta_1, \eta_2,..., \eta_N$ дискриминантная функция принимает вид:

$$L(k) = \max_{r_1, \psi, \sigma_1} \sum_{i=k-s+1}^{k} \ln\left\{\frac{1}{\sqrt{2\pi}\sigma_1} \exp\left[-\frac{1}{2\sigma_1^2}\left(\rho_i - \frac{r_1}{\cos(\varphi_i - \psi_1)}\right)^2\right]\right\} + \max_{r_2, \psi_2, \sigma_2} \sum_{i=k+1}^{k+s} \ln\left\{\frac{1}{\sqrt{2\pi}\sigma_2} \exp\left[-\frac{1}{2\sigma_2^2}\left(\rho_i - \frac{r_2}{\cos(\varphi_i - \psi_2)}\right)^2\right]\right\} + \max_{r_3, \psi_3, \sigma_3} \sum_{i=k-s+1}^{k+s} \ln\left\{\frac{1}{\sqrt{2\pi}\sigma_3} \exp\left[-\frac{1}{2\sigma_3^2}\left(\rho_i - \frac{r_3}{\cos(\varphi_i - \psi_3)}\right)^2\right]\right\}.$$

Для построения дискриминантной функции необходимо, таким образом, решить три экстремальные задачи относительно неизвестных параметров r_j, ψ_j, σ_j, $j = 1, 2, 3$, соответственно. Рассмотрим решение первой из них. Приравнивая частные производные по переменным r_1, ψ_1, σ_1 к нулю, после эквивалентных преобразований получаем

$$\begin{cases} \sum_{i=k-s+1}^{k} \left(\rho_i - \frac{r_1}{\cos(\varphi_i - \psi_1)}\right) \frac{1}{\cos(\varphi_i - \psi_1)} = 0, \\ \sum_{i=k-s+1}^{k} \left(\rho_i - \frac{r_1}{\cos(\varphi_i - \psi_1)}\right) \frac{\sin(\varphi_i - \psi_1)}{\cos^2(\varphi_i - \psi_1)} = 0, \\ \sum_{i=k-s+1}^{k} \left(\rho_i - \frac{r_1}{\cos(\varphi_i - \psi_1)}\right)^2 = s\sigma_1^2, \end{cases}$$

откуда следуют выражения:

$$\sigma_1^2 = \frac{1}{s} \sum_{i=k-s+1}^{k+s} \left(\rho_i - \frac{r_1}{\cos(\varphi_i - \psi_1)}\right)^2,$$

$$r_1 = \frac{\sum_{i=k-s+1}^{k} \frac{\rho_i}{\cos(\varphi_i - \psi_1)}}{\sum_{i=k-s+1}^{k} \frac{1}{\cos^2(\varphi_i - \psi_1)}},$$

$$\sum_{i=k-s+1}^{k} \sum_{j=k-s+1}^{k} \frac{\rho_i \sin(\varphi_i - \varphi_j)}{\cos^2(\varphi_i - \psi_1)\cos^3(\varphi_j - \psi_1)} = 0.$$

Из данных равенств находим r_1^*, ψ_1^*, σ_1^*. Последнее уравнение можно решить относительно ψ_1 каким-либо итерационным методом[2]. Можно показать, что на отрез-

[2] Например, данное уравнение можно решить в MatLab с помощью процедуры fzero. Для значения $s = 10$ время отыскания корня ψ_1^* с точностью 10^{-4} равно примерно 0,009 с.

ке $\left[-\frac{\pi}{2},\frac{\pi}{2}\right]$ хорошим начальным приближением будет $\psi_1^{(0)}=\arcsin\frac{\operatorname{sgn} b}{\sqrt{1+W^2}}$, где

$$W=\frac{\sum_{i=k-s+1}^{k}\rho_i\cos\varphi_i\left(s\rho_i\sin\varphi_i-\sum_{i=k-s+1}^{k}\rho_i\sin\varphi_i\right)}{\sum_{i=k-s+1}^{k}\rho_i\cos\varphi_i\left(s\rho_i\cos\varphi_i-\sum_{i=k-s+1}^{k}\rho_i\cos\varphi_i\right)},$$

$$b=\sum_{i=k-s+1}^{k}\rho_i(\sin\varphi_i-W\cos\varphi_i).$$

Аналогично находятся решения r_2, ψ_2, σ_2, r_3, ψ_3, σ_3 второй и третьей задач:

$$\sigma_2^2=\frac{1}{s}\sum_{i=k+1}^{k+s}\left(\rho_i-\frac{r_2}{\cos(\varphi_i-\psi_2)}\right)^2,$$

$$r_2=\frac{\sum_{i=k+1}^{k+s}\frac{\rho_i}{\cos(\varphi_i-\psi_2)}}{\sum_{i=k+1}^{k+s}\frac{1}{\cos^2(\varphi_i-\psi_2)}},$$

$$\sum_{i,\,j=k+1}^{k+s}\frac{\rho_i\sin(\varphi_i-\varphi_j)}{\cos^2(\varphi_i-\psi_2)\cos^3(\varphi_j-\psi_2)}=0,$$

$$\sigma_3^2=\frac{1}{2s}\sum_{i=k-s+1}^{k+s}\left(\rho_i-\frac{r_3}{\cos(\varphi_i-\psi_3)}\right)^2$$

$$r_3=\frac{\sum_{i=k-s+1}^{k+s}\frac{\rho_i}{\cos(\varphi_i-\psi_3)}}{\sum_{i=k-s+1}^{k+s}\frac{1}{\cos^2(\varphi_i-\psi_3)}}$$

$$\sum_{i,\,j=k-s+1}^{k+s}\frac{\rho_i\sin(\varphi_i-\varphi_j)}{\cos^2(\varphi_i-\psi_3)\cos^3(\varphi_j-\psi_3)}=0$$

После подстановки найденных параметров в выражение для дискриминантной функции и соответствующих преобразований получим окончательную формулу для вычисления $L(k)$:

$$L(k)=s\ln\frac{\sigma_3^2}{\sigma_1\sigma_2}.$$

Можно ожидать, что вблизи точек переключения значения функции $L(k)$ будут большими, чем внутри параметризованного участка. Выберем величину порога h так, чтобы можно было отслеживать поведение дискриминантной функции при изменении параметра k по истинности неравенства $L(k)\ge h$. При этом должно наблюдаться скопление точек $NS=\{\varphi_k\mid L(k)\ge h\}$ вблизи точек φ_t, в которых происходит переход с одного параметризованного участка на другой, и напротив – разреженность таких точек внутри параметризованного участка. Будем для простоты называть такие точки нестационарными. Поэтому возникает проблема кластеризации нестационарных точек на нестационарные участки LN_j, $j=1,2,...,n$, в которых по определению наблюдается лишь один переход с одного параметризованного участка на другой.

Определение параметризованных участков с помощью дискриминантной функции

Нестационарные участки LN_j, $j=1,2,...,n$, задают разбиение множества NS. Обозначим через θ отношение эквивалентности, связанное с этим разбиением. Определим условие принадлежности двух произвольных точек $\varphi_{i1},\varphi_{i2}\in NS$, $\varphi_{i1}<\varphi_{i2}$, одному нестационарному участку. Для этого рассмотрим функцию $E[\varphi_{i1},\varphi_{i2}]=\frac{1}{i_2-i_1-1}\sum_{j=i_1}^{i_2}L(j)$, которая, очевидно, дает среднее значение функции L на отрезке $[\varphi_{i1},\varphi_{i2}]$.

Внутри нестационарного участка наблюдается сгущение нестационарных точек, поэтому следует ожидать, что значение $E[\varphi_{i1},\varphi_{i2}]$, при условии $(\varphi_{i1},\varphi_{i2})\in\theta$, будет больше выбранного порога. Если же точки φ_{i1} и φ_{i2} принадлежат различным нестационарным участкам, т.е. $(\varphi_{i1},\varphi_{i2})\notin\theta$, то отрезок $[i_1,i_2]$, вероятно, включает интервал, в точках которого значения функции L меньше h. Следуя этим рассуждениям, получаем следующее условие принадлежности произвольных $\varphi_{i1},\varphi_{i2}\in NS$ одному нестационарному участку: $(\varphi_{i1},\varphi_{i2})\in\theta$, когда $E[\varphi_{i1},\varphi_{i2}]\ge h$. Это условие, очевидно, определяет толерантное, но не транзитивное отношение $\widetilde{\theta}\subseteq\theta$.

Введем дискретную функцию $\pi(\varphi_i)$, характеризующую плотность распределения нестационарных точек в окрестности точки φ_i, необходимую при описании нестационарных участков. Пусть $[\varphi_{i1},\varphi_{i2}]$ – произвольный отрезок на $[\varphi_1,\varphi_N]$. Тогда

$$\pi(\varphi_i)=\max_{\substack{[\varphi_{i1},\varphi_{i2}]\\ \varphi_i\in[\varphi_{i1},\varphi_{i2}]}}E[\varphi_{i1},\varphi_{i2}],$$

где максимум берется по всем отрезкам $[\varphi_{l1},\varphi_{l2}]$, содержащим точку φ_i. В основе конструктивного метода определения нестационарных участков лежит следующая теорема, доказанная в [11].

Теорема. Пусть $\varphi_l,\varphi_m\in NS$, $\varphi_m>\varphi_l$. Тогда $(\varphi_l,\varphi_m)\in\theta$ в том и только том случае, если для любого $\varphi_k\in[\varphi_i,\varphi_m]$ имеет место неравенство $\pi(\varphi_k)\ge h$.

Покажем, как функцию $\pi(\varphi_j)$ можно использовать для определения нестационарных участков. Пусть нестационарные точки φ_i,φ_m принадлежат некоторому нестационарному участку LN_j, т.е. $(\varphi_i,\varphi_m)\in\theta$, тогда он включает отрезок, т.е. $[\varphi_i,\varphi_m]\subseteq LN_j$. По теореме функция π не меньше пороговой величины h в каждой точке этого отрезка, т.е. для любого $\varphi_k\in[\varphi_i,\varphi_m]$ $\pi(\varphi_k)\ge h$. Поэтому будем считать, что точка $\varphi_k\in[\varphi_1,\varphi_N]$ принадлежит некоторому нестационарному участку тогда и только тогда, когда $\pi(\varphi_k)\ge h$; точки $\varphi_i,\varphi_m\in[\varphi_1,\varphi_N]$ при $\pi(\varphi_i)\ge h$, $\pi(\varphi_m)\ge h$, $\varphi_i<\varphi_m$, принадлежат одному нестационарному участку в том и только том случае, если для любого $\varphi_k\in[\varphi_i,\varphi_m]$ $\pi(\varphi_k)\ge h$. Это определение нестационарного участка удовлетворяет всем необходимым условиям. Легко проверяется, что каждая нестационарная точка принадле-

жит некоторому нестационарному участку, причем каждый такой участок имеет непустое пересечение с множеством нестационарных точек, т.е. для любого $j \in [2, n]$ $LN_j \cap NS \neq \varnothing$, и определенное таким образом множество $\{LN_j\}$ индуцирует отношение эквивалентности θ на множестве нестационарных точек NS.

Оценка моментов нарушения стационарности

Оценка положения точек переключения случайной последовательности с одного параметризованного участка на другой существенно упрощается, так как эти моменты локализованы внутри нестационарных участков, и можно перейти к однопараметрическим задачам определения точек переключения внутри сегментов, каждый из которых включает лишь один нестационарный участок. При этом могут оказаться полезными предварительные оценки положения точек переключения, учитывающие поведение дискриминантной функции L. В частности, для точки переключения $\hat{\varphi}_k$ можно использовать первоначальное приближение вида

$$\hat{\varphi}_k = \frac{\sum_{\varphi_i \in LN_k} \varphi_i L(i)}{\sum_{\varphi_i \in LN_k} L(i)}$$

В качестве $\hat{\varphi}_k$ в последней формуле берется ближайшее целое значение.

Данный метод обнаружения точек, в которых меняются свойства случайных величин, был апробирован на данных моделирования зашумленного профиля. При этом оказалось, что последняя оценка дает достаточно точное приближение и не возникает необходимости ее улучшения каким-либо другим способом.

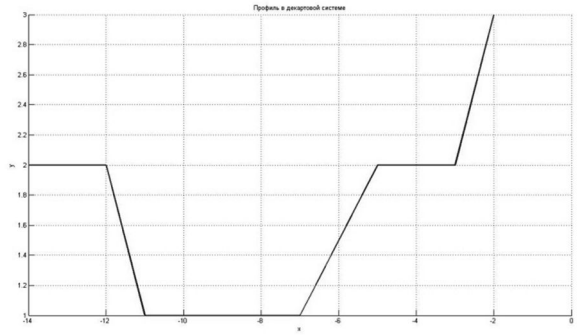
Пример численного моделирования

Рассмотрим численный пример обнаружения непараметрических участков в статистической последовательности. На рис. 5, а) показана форма идеального профиля, на рис. 5, б) - результаты численного моделирования случайной последовательности в полярной системе координат, на каждом параметризованном участке которой случайная последовательность описывается нормальным вероятностным распределением.

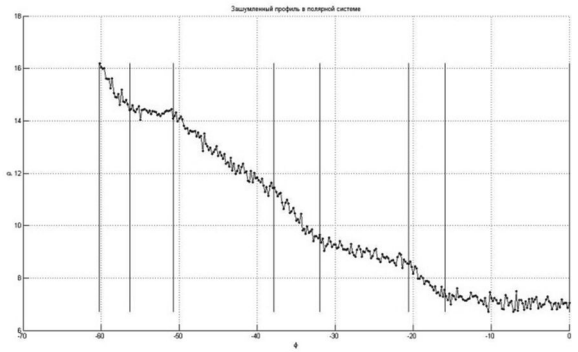
По результатам численного моделирования была построена дискриминантная функция и функция плотности нестационарных точек (рис. 5, в) и рис. 5, г)). Обратим внимание на то, что локальный метод обнаружения непараметрических участков в данном случае позволил осуществить правильную локализацию точек переключения случайной последовательности, если использовать порог $h = 22$. Восстановление геометрической формы профиля производилось по методу наименьших квадратов в декартовой системе координат. Восстановленный профиль изображен на рис. 5, д).

Заключение

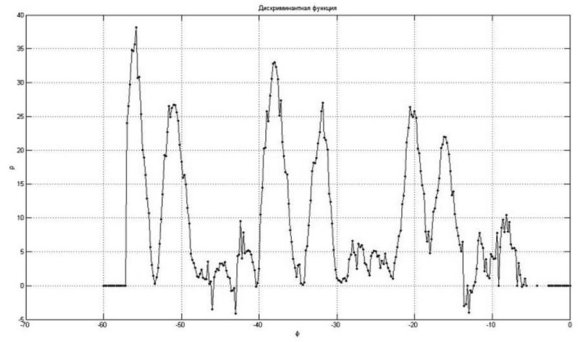
В работе показано, что автоматизация процедуры мониторинга поперечного профиля земляного полотна может быть сведена к задачам анализа случайной последовательности в различных постановках. Полезные результаты при этом открываются при использовании подходов, основанных на применении функций правдоподобия. В частности, на этом пути оказывается возможным решение задачи восстановления структурных линий, о чем упоминалось в начале статьи. Использование статистических методов накладывает и ряд ограничений, за рамками которых трудно рассчитывать на приемлемое по качеству решение. К числу таких ограничений следует отнести, прежде всего, требование достаточно представительной статистической выборки замеров. Методы вряд ли будут хорошо работать в условиях разрушений или сильных деформаций земляного полотна, когда не представляется возможным сделать правдоподобные предположения о форме профиля.



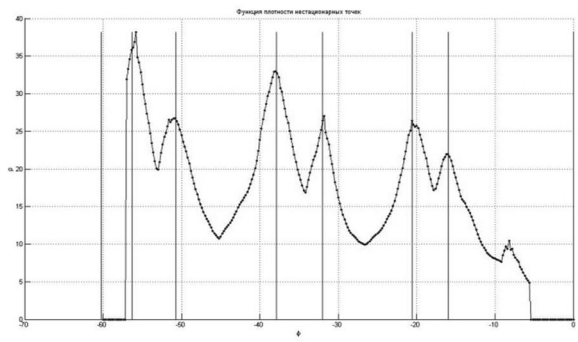
а)



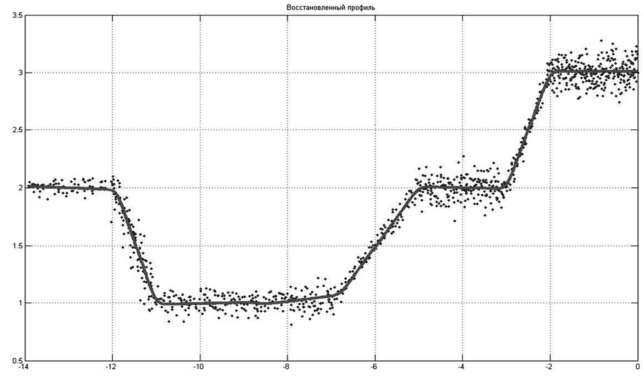
б)



в)



г)



д)

Рис. 5. Результаты моделирования:
а) исходный идеальный профиль в декартовой системе;
б) зашумленный профиль в полярной системе координат;
в) дискриминантная функция;
г) функция плотности нестационарных точек;
д) восстановленный профиль

Остается открытым вопрос об эффекте лазерных «теней», когда на профиле имеются места с крутизной, не позволяющей осуществить лазерное сканирование в принципе. Но это проблема самой технологии съема информации. По-видимому, в этом случае для восстановления профиля в «затененных» местах следует разрабатывать методы, основанные не только на чисто математических подходах, что могло бы стать темой отдельного исследования.

Литература

1. Сазонов В., Кармалин Б., Лебедев А. и др. Современные направления диагностики и мониторинга земляного полотна. Путь и путевое хозяйство. 2009, 6, с.34-37.

2. Уманский В.И. Технология построения трехмерных моделей железнодорожного полотна в высокоточном координатном пространстве. Сборник докладов 6-й Международной научно-практической конференции «Геопространственные технологии и сферы их применения». М., 2010.

3. Бриллинджер Д. Временные ряды. Обработка данных и теория. М.: Мир, 1980.

4. Гихман И.И., Скороход А.В. Введение в теорию случайных процессов. М.: Наука, 1965.

5. Королюк В.С., Портенко Н.И., Скороход А.В., Турбин А.Ф. Справочник по теории вероятностей и математической статистике. М.: Наука, 1985.

6. Клигене Н., Телькснис Л. Методы обнаружения моментов изменения свойств случайных процессов (обзор). Автоматика и телемеханика. 1983, № 10.

7. Торговицкий И.Ш. Методы определения момента изменения вероятностных характеристик случайных величин. Зарубежная радиоэлектроника. 1976, №1, с. 3-52.

8. Харин Ю.С. Классификация случайных серий неизвестной длины. Проблемы передачи информации. 1985, т. XXI, вып. 4, с. 64-65.

9. Дарховский Б.С. Непараметрический метод оценивания интервалов однородности случайной последовательности. Теория вероятностей и ее применение. 1985, т. 30, №4, с. 795-799.

10. Бродский Б.Е., Дарховский Б.С. Алгоритм апостериорного обнаружения разладок случайной последовательности. Автоматика и телемеханика. 1993, №1, с. 62-67.

11. Броневич А.Г., Каркищенко А.Н. Вероятностные и возможностные методы классификации случайных последовательностей. Таганрог: ТРТУ, 1996.

STATISTICAL METHOD OF RESTORING SUBGRADE FROM DATA OF LASER SCANNING

Bronevich A.G., Karkishchenko A.N., Umanskiy V.I.

The article is devoted to restoring the cross-section of subgrade from noisy data of laser scanning. A formal statement of this problem is given. It is shown that the problem can be reduced to the analysis of a random sequence of measurements. The paper presents a solution by using the Fisher discriminant function. It is also shown how to localize switching points in a piecewise para-meterized random sequence with the help of the constructed discriminant function. A computationally efficient way of estimating such points is proposed. Finally an example on simulated data is given.

10-ый Симпозиум по проектированию и тестированию

IEEE EAST-WEST DESIGN & TEST SYMPOSIUM - EWDTS 2012

пройдет с 14 по 17 сентября 2012 года в Харькове (Украина)

Подробнее: http://ewdtest.com/conf

УДК 621.372.54

МЕТОДЫ МНОГОСКОРОСТНОЙ ОБРАБОТКИ СИГНАЛОВ В МНОГОКАНАЛЬНЫХ СИСТЕМАХ С САМООРГАНИЗУЮЩЕЙСЯ СТРУКТУРОЙ: *ЭКВАЛАЙЗЕРЫ С ДИНАМИЧЕСКИМ ВЫБОРОМ ПОРЯДКОВ АДАПТИВНЫХ ФИЛЬТРОВ*

Линович А.Ю., *к.т.н., доцент, старший научный сотрудник Рязанского государственного радиотехнического университета, e-mail: rsrtu@yandex.ru*

Ключевые слова: эквалайзер, динамический выбор порядка, многоканальный фильтр, многоскоростная обработка сигналов.

Статья посвящена многоканальным адаптивным фильтрам, в которых используются принципы самоорганизации, и особенностям построения эквалайзеров на их основе. Приводятся результаты имитационного моделирования разработанных эквалайзеров, полученные для двухлучевой модели беспроводного канала связи, подтверждающие ряд преимуществ предложенного подхода.

Введение

Многоканальные адаптивные фильтры (МАФ), в которых внутриканальная обработка проводится на пониженной частоте дискретизации, имеют ряд преимуществ перед одноканальными адаптивными фильтрами [1–3]. С одной стороны, благодаря независимой обработке разных участков спектра разными каналами МАФ, удаётся существенно повысить скорость и точность настройки. С другой стороны, использование методов многоскоростной обработки сигналов даёт возможность значительно сократить объём вычислительных затрат, что особенно важно для повышения эффективности реализации адаптивных алгоритмов в реальном времени.

В предлагаемой вниманию читателя статье рассмотрен один из способов дальнейшего совершенствования МАФ, заключающийся в предоставлении каждому внутриканальному адаптивному фильтру самостоятельного выбора наилучшего (по некоторому критерию) порядка. Поскольку дальнейшее повествование ограниченно рассмотрением МАФ с настройкой внутри каналов по алгоритму НСКО с нормировкой шага адаптации[3] [4], в основе которого лежит нерекурсивный адаптивный фильтр, то порядок адаптивного фильтра равен числу его весовых коэффициентов и равен длине воспроизводимой им импульсной характеристики, выраженной в отсчётах, взятых с частотой дискретизации обрабатываемого этим фильтром сигнала. Известно, что алгоритм НСКО прост в реализации и нетребователен к вычислительным затратам. Вместе с тем его низкая скорость настройки во многих случаях оказывается серьёзным недостатком, побуждая разработчиков цифровых устройств обработки сигналов искать альтернативные пути решения. Одним из таких путей является переход к многоканальной структуре построения адаптивного фильтра [1–3]. Сам по себе переход к многоканальной адаптивной обработке при использовании методов многоскоростной обработки сигналов [5, 6] позволяет существенно снизить вычислительные затраты и значительно повысить качество работы адаптивных устройств. Но, как будет показано далее, введение специальной процедуры динамического выбора порядков внутриканальных адаптивных фильтров обеспечивает значительное улучшение характеристик МАФ. Несмотря на то, что основное внимание в статье уделено решению задачи обратного моделирования, решаемой в эквалайзерах аппаратуры передачи данных, многие положения нетрудно расширить на случай решения задачи прямого моделирования.

Структура МАФ в общем виде

Классическая структура МАФ описана во многих учебных и научных изданиях, посвящённых цифровым адаптивным фильтрам. В упрощённом виде она показана на рис. 1.

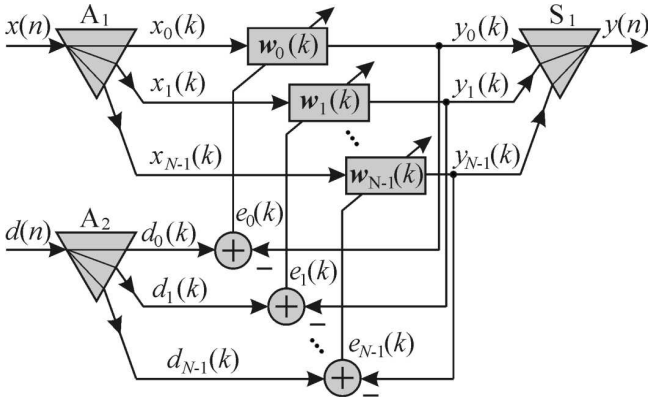


Рис. 1. Структурная схема МАФ

Подсистемы анализа А1 и А2 расщепляют входной сигнал $x(n)$ и обучающий сигнал $d(n)$ на N спектральных полос $x_0(k)$, $x_1(k)$, …, $x_{N-1}(k)$ и $d_0(k)$, $d_1(k)$, …, $d_{N-1}(k)$, соответственно. S1 – подсистема синтеза, выполняющая объединение выходных сигналов $y_0(k)$, $y_1(k)$, …, $y_{N-1}(k)$ в результирующий выходной сигнал МАФ $y(n)$. Поскольку спектры внутриканальных сигналов занимают более узкие полосы частот, чем спектры входного и выходного сигналов МАФ, в подсистемах анализа выполняется переход на более низкую частоту дискретизации, а в подсистеме синтеза исходная частота дискретизации восстанавливается. То есть, адаптивная обработка сиг-

[3] В англоязычной литературе этот алгоритм широко известен и носит название «Normalized least-mean square (NLMS) algorithm» [1, 4].

налов проводится на пониженной частоте. Набор внутриканальных адаптивных фильтров состоит из N трансверсальных фильтров: $\mathbf{w}_0(k)$, $\mathbf{w}_1(k)$, …, $\mathbf{w}_{N-1}(k)$. Здесь и далее в статье индексы n и k – индексы времени, предназначенные для нумерации отсчётов дискретных сигналов, определённых на высокой и на пониженной частотах дискретизации, соответственно. Для успешного функционирования МАФ подсистемы A1 и A2 по своим свойствам должны быть строго одинаковыми, а подсистема S1 должна быть надлежащим образом оптимизирована.

Алгоритм настройки МАФ с адаптивным выбором порядков внутриканальных фильтров

Из многочисленных работ по теории адаптивной фильтрации [1, 4, 7] известно, что при решении задачи обратного моделирования в классе нерекурсивных цифровых цепей предельная точность полученного решения, как правило, оказывается тем выше, чем выше порядок адаптивного фильтра (рис. 2), при условии отсутствия аддитивного шума в принимаемом сигнале.

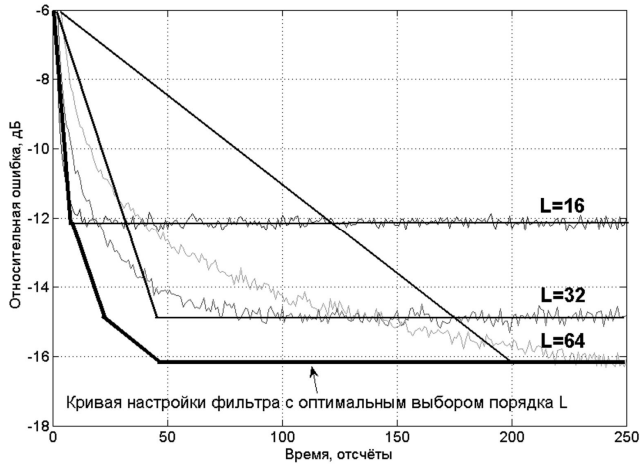


Рис. 2. Увеличение порядка фильтра позволяет повысить точность настройки, быстродействие при этом снижается

Однако при наличии в принимаемом сигнале аддитивного шума нельзя достичь нулевой среднеквадратической ошибки (СКО). Анализ формул, описывающих процедуру обновления вектора весовых коэффициентов алгоритма НСКО [4], показывает, что для адаптивных фильтров достаточно высокого порядка дисперсия сигнала ошибки содержит дополнительную составляющую, вызванную отклонением вектора весовых коэффициентов от оптимального вектора настройки под действием аддитивного шума (рис. 3).

Первую составляющую ошибки, вызванную усечением импульсной характеристики адаптивного фильтра в результате ограничения его порядка конечной величиной L, обозначим $\xi_y(L)$. Функция $\xi_y(L)$ является невозрастающей, на рис. 3 ей соответствует кривая, стремящаяся к нулю с ростом L. Вторую составляющую, наличие которой обусловлено действием шума, обозначим $\xi_ш(L)$. Функция $\xi_ш(L)$ увеличивается по мере роста L с линейной (или даже сверхлинейной) скоростью[4]. В результате, наиболее эффективное решение задачи обратного моделирования методами адаптивной фильтрации достигается при выборе порядка фильтра в пределах некоторого диапазона, ограниченного окрестностью точки минимума зависимости $\xi_\varepsilon(L) = \xi_y(L) + \xi_ш(L)$. В приведённом на рис. 3 примере оптимальная величина порядка фильтра лежит в диапазоне значений от 25 до 40.

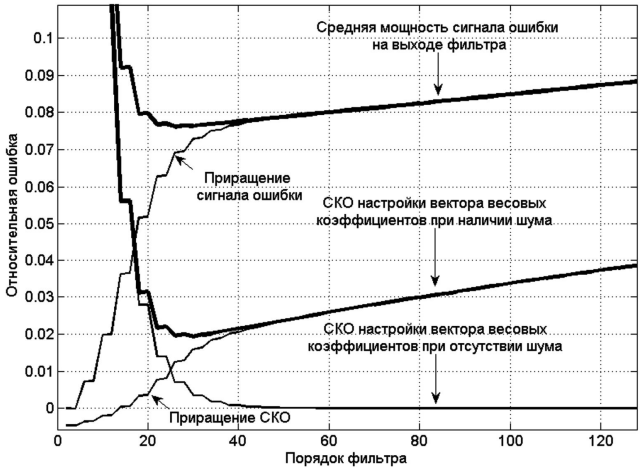


Рис. 3. Наличие шума снижает эффективность фильтров высокого порядка

Всё сказанное выше справедливо для стационарных систем. Если частотные характеристики моделируемой системы меняются с течением времени, то эквалайзер более высокого порядка медленнее реагирует на происходящие изменения. Поэтому эквалайзер меньшего порядка снова оказывается более предпочтительным. Однако вид функции $\xi_\varepsilon(L)$ при моделировании нестационарных систем изменяется со временем, а следовательно, изменяется и положение точки минимума. Функцию $\xi_\varepsilon(L)$ в данном случае можно рассматривать в каждый момент времени как сечение функции двух переменных $\xi_{\varepsilon t}(L, t)$.

Для организации динамического выбора порядка внутриканальных адаптивных фильтров, входящих в состав МАФ, воспользуемся процедурой, описанной в [3]. Критерием оптимальности в данной процедуре является степень приближения к такому равновесному состоянию, когда среднеквадратические ошибки основного и вложенного фильтров равны. Понятие вложенного фильтра поясняется на рис. 4. Множество весовых коэффициентов вложенного фильтра является подмножеством весовых коэффициентов основного фильтра, полученных исключением ΔL коэффициентов с наименьшими порядковыми номерами и ΔL коэффициентов с наибольшими порядковыми номерами. Поскольку весовые коэффициенты нерекурсивного фильтра, по сути, являются дискретными отсчётами воспроизводимой фильтром импульсной характеристики, то импульсную

[4] Правило обновления весовых коэффициентов алгоритма НСКО [1, 7] содержит линейный член поправки: $\boldsymbol{w}(n+1) := \boldsymbol{w}(n) + \mu\varepsilon\boldsymbol{x}(n)$. Если входной сигнал $x(n)$ является стационарным в широком смысле сигналом, то флуктуации всех весовых коэффициентов в установившемся режиме оказываются приблизительно одинаковыми и ошибка при вычислении выходного сигнала $y(n) = \boldsymbol{w}^{\mathrm{T}}(n)\,\boldsymbol{x}(n)$ растёт линейно по мере увеличения числа весовых коэффициентов – порядка фильтра.

характеристику вложенного фильтра можно рассматривать как результат усечения импульсной характеристики основного фильтра.

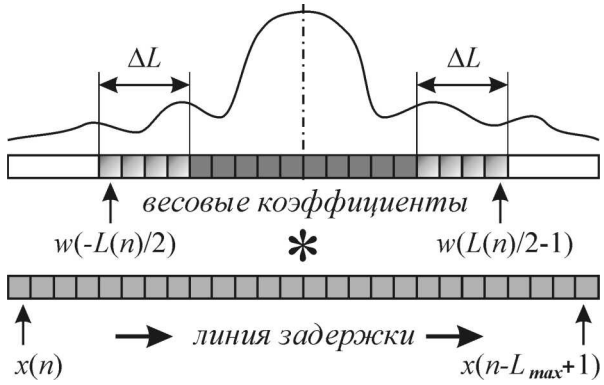


Рис. 4. Схема, поясняющая принцип выбора оптимального порядка

Выходной сигнал нерекурсивного фильтра $y(n)$ находится путём вычисления свёртки его импульсной характеристики $\mathbf{w}(n)$ (вектора его весовых коэффициентов) с последовательностью отсчётов входного сигнала $\mathbf{x}(n)$, продвигающихся по линии задержки фильтра. Причём векторы $\boldsymbol{w}(n)$ и $\boldsymbol{x}(n)$ одинаковы по длине, но поскольку рассматривается фильтр переменного порядка, то это означает, что их длина изменяется со временем. Сигнал ошибки настройки выражается разностью обучающего сигнала $d(n)$ и рассчитанного выходного сигнала $y(n)$:

$$e(n) = d(n) - \mathbf{w}^T(n)\mathbf{x}(n). \quad (1)$$

Обозначим вектор весовых коэффициентов и вектор отсчётов линии задержки вложенного фильтра, соответственно, как $\mathbf{w}_{вл}(n)$ и $\mathbf{x}_{вл}(n)$. Тогда сигнал ошибки вложенного фильтра:

$$e_{вл}(n) = d(n) - \mathbf{w}_{вл}^T(n)\mathbf{x}_{вл}(n). \quad (2)$$

Если использовать среднеквадратический критерий оценки, то из двух фильтров следует отдать предпочтение тому, который обеспечивает меньшую дисперсию ошибки. В большинстве практических задач, решаемых в реальном времени, опираются на приближённую оценку дисперсии, которую получают по ограниченному числу наблюдений K. Выполняя оценку дисперсии сигнала ошибки по ограниченному фрагменту, состоящему из K отсчётов сигнала ошибки, и применяя итеративный подход к выбору оптимального порядка адаптивного фильтра, запишем условие выбора оптимального порядка:

$$L(K\cdot(n+1)) = L(K\cdot n) + sign\left(\sum_{j=K\cdot n}^{K\cdot(n+1)-1}|e_{вл}(j)|^2 - \sum_{j=K\cdot n}^{K\cdot(n+1)-1}|e(j)|^2\right). \quad (3)$$

Здесь функция $sign(x)$ определяется следующим образом:

$$sign(x) = \begin{cases} 1 & при \quad x > 0, \\ 0 & при \quad x = 0, \\ -1 & при \quad x < 0. \end{cases} \quad (4)$$

Для N - канального МАФ, схема которого изображена на рис. 1, условие (3) превращается в систему, состоящую из N условий:

$$L_i(K\cdot(k+1)) = L_i(K\cdot k) + sign\left(\sum_{j=K\cdot k}^{K\cdot(k+1)-1}|e_{вл,i}(j)|^2 - \sum_{j=K\cdot k}^{K\cdot(k+1)-1}|e_i(j)|^2\right), \quad (5)$$

где $i = 0,\ldots,N-1$ – порядковый номер канала МАФ.

Описанный выше алгоритм настройки МАФ и вычислительные затраты на выполнение соответствующих действий, измеренные в операциях комплексного умножения, приводятся в табл. 1.

Разработка подсистемы анализа-синтеза

Вопросы оптимального проектирования подсистем анализа-синтеза с передискретизацией рассмотрены в [8–10]. Ограничимся рассмотрением класса комплексных подсистем анализа-синтеза с равномерным разбиением по частоте. Если входной сигнал разбивается по частоте на N полос (рис. 1) равной ширины $f_{кв}/N$, то частоту дискретизации, согласно теореме Котельникова, можно понизить в N раз. Однако при решении задач цифровой частотной селекции у каждого фильтра всегда существуют некоторые переходные зоны на границах полосы пропускания, из-за чего возникают нежелательные эффекты наложения спектров вблизи этих граничных частот. Поэтому при построении МАФ частоту дискретизации понижают не в N раз, а немного меньше – в M раз ($M < N$). Запас по частоте, выраженный соотношением

$$K_{ПД} = \sum_{i=0}^{N-1}\frac{1}{M_i}\cdot 100\% = \frac{N}{M}\cdot 100\%, \quad (6)$$

называют показателем передискретизации (ППД)[5]. Здесь i – порядковый номер канала МАФ. В общем случае коэффициенты децимации в разных каналах могут быть выбраны различными, но в данной статье будем рассматривать более простой случай, когда коэффициент децимации одинаков для всех каналов.

Расчёт подсистемы анализа-синтеза сводится к решению задачи оптимизации. Примерно десять лет назад наиболее известные методы расчёта предполагали использование оптимизации в классе нерекурсивных цифровых цепей при условии линейности фазы фильтра-прототипа. В последнее время стали широко известны более эффективные пути решения. В частности, в [9, 10] предлагается отказаться от требования линейности фазы и свести оптимизационную задачу к поиску минимума выпуклой целевой функции при наличии ограничений. Постановка задачи минимизации принимает следующий вид. Требуется найти минимум остаточной энергии зоны непрозрачности фильтра-прототипа:

$$\mathbf{g}^T\tilde{\mathbf{r}}_h \to \min_{\tilde{\mathbf{r}}_h}, \quad (7\text{ а})$$

при ограничениях $\tilde{\mathbf{v}}(\omega)^T\tilde{\mathbf{r}}_h \le \xi_2^2$ для всех

$$\omega \in [\pi/M, \pi], \quad (7\text{ б})$$

$$\tilde{\mathbf{v}}(\omega)^T\tilde{\mathbf{r}}_h \le \xi_3^2 \text{ для всех } \omega \in [0, \pi/M), \quad (7\text{ в})$$

$$\left\|\mathbf{S}\tilde{\mathbf{r}}_h\right\|_2 \le \frac{\xi_4}{\sqrt{2}}, \quad (7\text{ г})$$

[5] В англоязычной литературе этот термин носит устойчивое название «oversampling ratio (OSR)».

$$\left[\widetilde{\mathbf{r}}_h\right]_0 = \frac{M}{N}, \qquad (7\text{ д})$$

$$\widetilde{\mathbf{v}}(\omega)^T \widetilde{\mathbf{r}}_h \geq 0 \text{ для всех } \omega \in [0, \pi]. \qquad (7\text{ е})$$

Фильтр-прототип представляет собой низкочастотный фильтр. Условие (7 а) соответствует поиску фильтра с минимальной остаточной энергией в зоне непрозрачности. Вместо импульсной характеристики фильтра-прототипа оптимизируются отсчёты автокорреляционной функции (АКФ) $\mathbf{r}_h$, вернее её правой половины $\widetilde{\mathbf{r}}_h$, так как АКФ всегда симметрична. Элементы вектора $\mathbf{g}$ получаются следующим образом: $[\mathbf{g}]_0 = 1 - 1/M$, $[\mathbf{g}]_m = -2\sin(\pi m/M)/(\pi m)$ для $m \geq 1$. Если АКФ обладает чётной симметрией и содержит нечётное число коэффициентов, элементы вектора $\widetilde{\mathbf{v}}(\omega)$ находятся следующим образом: $\left[\widetilde{\mathbf{v}}(\omega)\right]_0 = 1$, $\left[\widetilde{\mathbf{v}}(\omega)\right]_m = 2\cos(m\omega)$ для $m \geq 1$.

Коэффициенты ξ_2, ξ_3 и ξ_4 вводят желаемые ограничения: ξ_2 ограничивает уровень лепестков АЧХ в полосе подавления фильтра, ξ_3 задаёт неравномерность АЧХ, ξ_4 определяет ортогональность банка фильтров и отвечает за степень перекрытия АЧХ смежных каналов[6]. Ограничивающее условие (7 г) можно переписать в форме

$$\sum_{i>0} r_h[iM]^2 \leq \frac{\xi_4^2}{2}. \qquad (8)$$

Как следует из (8), S может быть составлена из строк единичной матрицы размером $L \times L$ с номерами $iM + 1$, где $i > 0$.

Условие (7 д) использовано для нормировки энергии, теряемой в процессе децимации входного сигнала. Последнее условие (7 е) вытекает непосредственно из свойств АКФ.

Задача оптимизации (7) имеет линейную целевую функцию, три линейных ограничения-неравенства (7 б), (7 в) и (7 е), одно ограничение-неравенство в форме выпуклой квадратичной функции (7 г) и одно линейное ограничение-равенство (7 д). Следовательно, задача (7) относится к классу задач выпуклой оптимизации. Для решения такого рода задач в программе «MATLAB» имеется специальный пакет готовых функций «CVX», свободно доступный в сети «Интернет» [11]. Готовый пример программы, выполняющей расчёт АКФ приводится в табл. 2. Результатом расчёта является вектор «rtilde_h», содержащий центральный коэффициент и правую ветвь коэффициентов АКФ искомого фильтра-прототипа. Полученная с помощью данной программы АКФ (после дополнения левой ветвью) показана на рис. 5. Все коэффициенты АКФ должны быть вещественными по определению, в то же время коэффициенты импульсной характеристики нерекурсивного цифрового фильтра, соответствующего АКФ, могут быть комплексными величинами.

[6] Чтобы сделать условие (7 г) более ясным для восприятия, запишем его в виде $\sum_{i\neq 0}\left|\sum_n h(n)h(n+iM)\right|^2 \leq \xi_4^2$, где $h(n)$ – импульсная характеристика фильтра-прототипа. Чтобы представить (7 г) в форме линейного ограничения-неравенства, осуществляют переход к АКФ, вводя переменную $\widetilde{\mathbf{r}}_h$.

Таблица 1. Алгоритм настройки МАФ с динамическим выбором порядков внутриканальных адаптивных фильтров и вычислительные затраты

Выполняемые действия	*Затраты* (V/T_s)
Операции, выполняемые для отсчётов n = 0, 1, 2,… на частоте дискретизации $1/T_s$	
1) Разделение входного и обучающего сигналов на компоненты подсистемой анализа:	
$x_i(n) = \mathbf{h}_i^T \mathbf{x}(n)$, $i = 0,\ldots,N-1$,	$\frac{NL_H}{M}$
$d_i(n) = \mathbf{h}_i^T \mathbf{d}(n)$, $i = 0,\ldots,N-1$.	$\frac{NL_H}{M}$
2) Объединение компонентов выходного сигнала подсистемой синтеза:	
$y(n) = \sum_{i=0}^{N-1} \mathbf{g}_i^T \mathbf{y}_i(n)$.	$\frac{NL_G}{M}$
Операции, выполняемые для отсчётов k = 0, 1, 2,… на частоте дискретизации $1/(MT_s)$	
1) Уточнение оценки средней мощности входного сигнала:	
$\left\|\mathbf{x}_i(k)\right\|^2$, $i = 0,\ldots,N-1$.	N
2) Вычисление компонентов выходного сигнала и оценка сигналов ошибки:	
$y_i(k) = \mathbf{w}^T(k)\mathbf{x}_i(k)$, $e_i(k) = d_i(k) - y_i(k)$, $i = 0,\ldots,N-1$.	$\frac{NL_i(k)}{M^2}$
3) Обновление векторов весовых коэффициентов:	
$\mathbf{w}_i(k+1) = \mathbf{w}_i(k) + \mu_i \frac{\mathbf{x}_i^*(k)}{\left\|\mathbf{x}_i(k)\right\|^2} e_i(k)$, $i = 0,\ldots,N-1$.	$\frac{NL_i(k)}{M^2}$
Операции, выполняемые для отсчётов j = 0, 1, 2, на частоте дискретизации $1/(KMT_s)$	
1) Модификация порядков внутриканальных АФ в соответствии с (10):	
$L_i(K\cdot(k+1)) = L_i(K\cdot k) + sign\left(\sum_{j=K\cdot k}^{K\cdot(k+1)-1} \left\|e_{ea,i}(j)\right\|^2 - \sum_{j=K\cdot k}^{K\cdot(k+1)-1} \left\|e_i(j)\right\|^2\right)$, $i = 0,\ldots,N-1$.	$\frac{2N}{M}$

В таблице используются следующие обозначения. N – число каналов МАФ, μ_i – шаг адаптации фильтра i-го канала, $L = L_{eq}/M$ – число настраиваемых весовых коэффициентов в каждом из внутриканальных адаптивных фильтров, L_i – порядок эквивалентного одноканального адаптивного фильтра для канала МАФ с порядковым номером i, L_H – порядок фильтров подсистемы анализа, L_G – порядок фильтров подсистемы синтеза, M – коэффициент децимации. V/T_s – объём вычислительных затрат, определяемый как среднее число операций комплексного умножения, выполняемых за время T_s, равное шагу дискретизации входного сигнала МАФ.

Кроме того, в рассматриваемом методе расчёта снимается ограничение на линейность фазы фильтра-прототипа, что позволяет существенно повысить качественные показатели подсистемы анализа-синтеза при заданном порядке, составляющих её фильтров.

Таблица 2. Расчёт правой половины АКФ низкочастотного фильтра-прототипа в программе «MATLAB» с использованием пакета выпуклой оптимизации «CVX» [11]

```
L = 48;  % Порядок фильтра-прототипа
N = 8;   % Число каналов подсистемы анализа-синтеза
M = 6;   % Коэффициент децимации
level = -33;  % Подавление в зоне непрозрачности, дБ
alpha = 1e-3; % Показатель самоортогональности

xi_3_sqr = 10 ^ (0.1) * M;
xi_4 = alpha * M / N;
xi_2_sqr = xi_3_sqr * 10 ^ (level / 10);

sf = 15;
eps_sf = 1e-6;

omega_s = pi / M;
k_vecT = 1 : (L - 1);
Ns_pt = ceil(L * sf * omega_s / pi);
Ns_s = ceil(L * sf * (1 - omega_s / pi));
omega_ptb = linspace(0, omega_s, Ns_pt + 1)';
omega_ptb = omega_ptb(1 : end - 1);
omega_sb = linspace(omega_s, pi, Ns_s)';
VtT_ptb = [ones(Ns_pt, 1), 2*cos(omega_ptb*k_vecT)];
VtT_sb  = [ones(Ns_s,  1), 2*cos(omega_sb*k_vecT)];

N_isi = floor((L - 1) / N);
k_isi_vec = N * (1 : N_isi)';
S = eye(L);
S = S(k_isi_vec + 1, :);

gT = [1 - 1/M, -2*sin(pi*k_vecT/M) ./ (pi*k_vecT)];
cvx_begin
    cvx_solver sedumi;
    cvx_precision high;
    variable rtilde_h(L);
    variable X(L,L) symmetric;
    minimize (gT * rtilde_h)
    subject to
        VtT_sb * rtilde_h <= xi_2_sqr - eps_sf;
        VtT_ptb * rtilde_h <= xi_3_sqr - eps_sf;
        norm(S * rtilde_h) <= xi_4 / sqrt(2);
        [1, zeros(1,L-1)] * rtilde_h == M / N;
        X == semidefinite(L);
        trace(X) == rtilde_h(1);
        for n = 1 : L - 1,
            sum(diag(X, n)) == rtilde_h(n + 1);
        end;
cvx_end
```

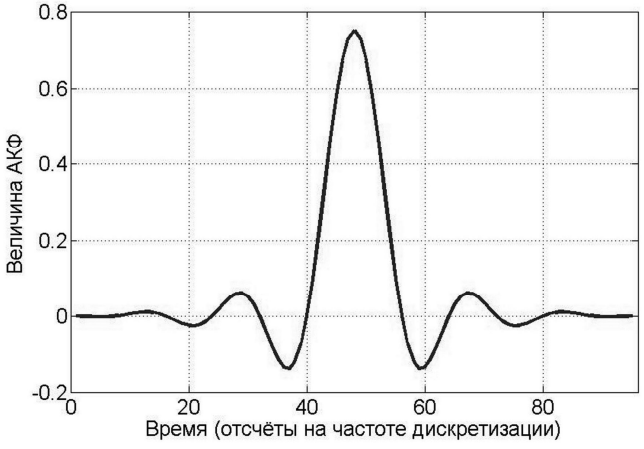


Рис. 5. АКФ низкочастотного фильтра-прототипа

Несмотря на то, что рассматриваемые по отдельности фильтры подсистем анализа и синтеза имеют несимметричные импульсные характеристики, а, следовательно, вносят в сигнал фазовые искажения, последовательное их соединение оказывается цифровой цепью, имеющей линейную фазу. Это значит, что подсистема синтеза полностью компенсирует фазовые искажения, внесённые фильтрами подсистемы анализа.

Для поиска импульсной характеристики обратимся к методам спектрального разложения (спектральной факторизации). Известно, что задача определения неизвестной импульсной характеристики некоторой системы по известной АКФ этой системы не даёт однозначного соответствия. Для определённости можно ввести ряд ограничений, например, потребовать, чтобы все коэффициенты найденной импульсной характеристики были вещественными числами.

Существует много численных методов решения данной задачи. Воспользуемся одним из методов, описанных в [12]. Метод[7] основан на использовании алгоритма быстрого преобразования Фурье (БПФ), который реализован в программе «MATLAB». Алгоритм БПФ хорошо известен специалистам в области цифровой обработки сигналов, поэтому данный метод спектральной факторизации легко осуществим при решении задачи в «MATLAB».

Метод основан на алгоритме непосредственного расчёта минимально-фазового спектрального множителя $S_{мф}(z)$ по заданной АКФ $r(n)$. Можно показать [12], что логарифм $S_{мф}(z)$ определён и является аналитической функцией во всех точках комплексной плоскости вне единичной окружности. Следовательно, $\ln S_{мф}(z)$ может быть разложен в бесконечный ряд по степеням z^{-1}:

$$\ln S_{мф}(z) = \sum_{i=0}^{\infty} a_i z^{-i} , \qquad (9)$$

где $|z| > 1$, $a_i \in \mathbf{R}$. В таком случае вещественная часть логарифма на единичной окружности определяется выражением:

$$\Re \ln S_{мф}\left(e^{j\omega}\right) = \ln\left|S_{мф}\left(e^{j\omega}\right)\right| = \frac{1}{2}\ln R(\omega) = \sum_{i=0}^{\infty} a_i \cos i\omega , \qquad (10)$$

где $R(\omega)$ – преобразование Фурье от АКФ. Коэффициенты a_i – коэффициенты Фурье функции $(1/2)\log R(\omega)$:

$$a_i = \frac{1}{2\pi}\int_0^{2\pi}\frac{1}{2}\ln R(\omega)e^{-ji\omega}d\omega,\ i = 0,1,\ldots , \qquad (11)$$

где j – мнимая единица.

После того как найдены коэффициенты a_i, $S_{мф}(e^{j\omega})$ можно найти следующим образом:

$$S_{мф}\left(e^{j\omega}\right) = \exp\sum_{i=0}^{\infty} a_i e^{-j\cdot i\omega} , \qquad (12)$$

Коэффициенты Фурье $S_{мф}(e^{j\omega})$ представляют собой искомые отсчёты импульсной характеристики:

$$h(n) = \frac{1}{2\pi}\int_0^{2\pi} e^{j\omega n}\exp\sum_{i=0}^{\infty} a_i e^{-j\cdot i\omega}d\omega . \qquad (13)$$

Далее приведём описание практической реализации в терминах цифровой обработки сигналов. На первом этапе по известной АКФ рассчитывается квадрат АЧХ $R(\omega)$, для чего вектор коэффициентов дополняется нулями и вычисляется БПФ. Для вычисления БПФ число отсчётов F (после дополнения нулями) должно быть кратно натуральной степени числа 2. Чем больше ну-

[7] Точно не известно, кто является автором данного метода. Вероятно, идея метода дана А.Н. Колмогоровым [12]. Следует также заметить, что метод не раз был предложен разными учёными, о чём свидетельствуют работы [13, 14].

лей, тем точнее будет выполнена спектральная факторизация.

Цель второго этапа состоит в поиске коэффициентов Фурье a_i для функции $(1/2)\ln R(\omega)$, который снова сводится к вычислению БПФ.

На третьем этапе вычисляется выражение (12), для чего в третий раз используется алгоритм БПФ и выполняется экспоненциальное взвешивание. Фактически на втором и третьем этапах происходит вычисление фазы функции $S_{мф}(e^{j\omega})$ с помощью преобразования Гильберта и последующей операции вычисления экспоненты.

Коэффициенты Фурье функции $S_{мф}(e^{j\omega})$ являются отсчётами искомой импульсной характеристики $h[n]$. Поэтому четвёртый этап связан с нахождением обратного преобразования Фурье, которое вновь сводится к вычислению БПФ.

Следует заметить, что при программной реализации описанного алгоритма в «MATLAB» в ряде преобразований следует принять дополнительные меры для гарантирования свойств симметрии, вещественности и ограничения размерностей получаемых результатов, как это сделано в примере, приведённом в табл. 3.

Таким образом, описанный здесь метод спектральной факторизации включает в себя четыре стадии вычислений, на каждой из которых применяется БПФ. Пример программы, позволяющий рассчитать импульсную характеристику низкочастотного фильтра-прототипа по известной АКФ этого фильтра приводится в табл. 3. Благодаря использованию стандартной операции вычисления БПФ, основная часть программы состоит всего из четырёх строчек. Вторая часть программы служит для отображения полученных результатов на графиках. Кроме того, вычисление БПФ позволяет быстро получать результат для фильтров высокого порядка (например, когда фильтр-прототип имеет несколько сотен весовых коэффициентов). Описанная процедура нахождения импульсной характеристики по известной АКФ отличается вычислительной эффективностью и устойчивостью вычислений. Полученная в результате спектральной факторизации импульсная характеристика показана на рис. 6. Она состоит из 48 коэффициентов и имеет несимметричный вид. Все коэффициенты полученного нерекурсивного фильтра принимают вещественные значения. Тем не менее, фильтры подсистемы анализа и синтеза получаются из фильтра-прототипа умножением на комплексные экспоненты соответствующих частот, поэтому МАФ, построенный на основе вещественного фильтра-прототипа, является комплексным.

О точности проведённой процедуры спектральной факторизации можно судить по рис. 7. Поиск АКФ фильтра по известной импульсной характеристике этого фильтра, в отличие от обратной задачи, описанной нами выше, даёт однозначное решение и может быть выполнен посредством вычисления операции свёртки. Рассчитав АКФ, соответствующую найденной импульсной характеристике, вычтем из неё первоначальную АКФ, для которой выполнялась спектральная факторизация. Найденная таким образом погрешность решения задачи поиска импульсной характеристики по известной АКФ показана на графике (рис. 7). Эта погрешность не превышает 10^{-10}.

Таблица 3. Расчёт импульсных характеристик фильтров подсистемы анализа-синтеза по АКФ низкочастотного фильтра прототипа

```
% Порядок ДПФ
F = 2^19;
% Формирование АКФ по её правой ветви (отражением)
r = rtilde_h .'; Lh = length(r);
r = [fliplr(r(2:end)) r];
% Расчёт спектра по известной АКФ
R = real(fft([r(Lh:end), ...
                    zeros(1,F-2*Lh+1) r(1:Lh-1)]));
% Преобразование Гильберта
a = 2 * 1/F * fft(1/2 * log(R)); a(1) = a(1) / 2;
a = [a(1:F/2) zeros(1,F/2)];
Smp = exp( fft(a) );
% Расчёт импульсной характеристики фильтра-прототипа
h = real( ifft(Smp) ); h = h(1:Lh);

% Отображение результатов на графиках
r0 = conv(h, fliplr(conj(h)));
figure(1); plot(h);
figure(2);
  plot([1:length(r)],r,'b',[1:length(r0)],r0,'r');
figure(3); plot(r-r0);
% ЛАЧХ фильтров анализа-синтеза
K = 4; Lh = length(h);
Freq = exp(j * pi * ([0:K-1]+1/2)'/4 * [1:Lh]);
H = h.' * ones(1,K) .* Freq.';
ACH = 20*log10(abs(fft(h.'*ones(1,K).*Freq.',512)));
figure(5); plot([0:512/2-1].'/512, ACH(1:end/2,:));
```

Рис. 6. Импульсная характеристика низкочастотного фильтра-прототипа

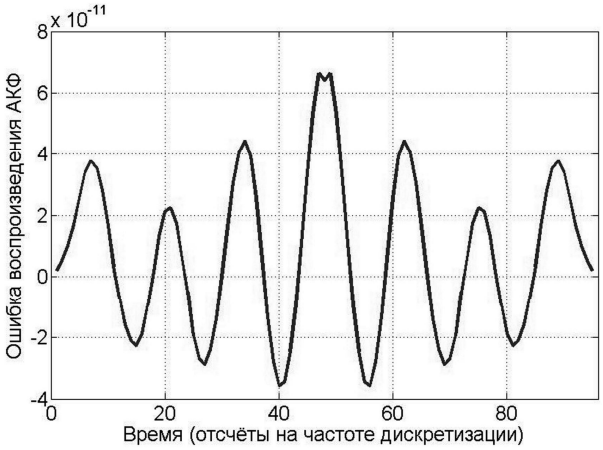


Рис. 7. Погрешность, полученная при решении задачи поиска импульсной характеристики по известной АКФ

На рис. 8 показаны АЧХ фильтров подсистемы анализа. Все фильтры получены на основе рассчитанного ранее фильтра-прототипа путём модуляции комплексными экспонентами соответствующих частот. При обработке вещественного входного сигнала достаточно четырёх фильтров, перекрывающих диапазон частот от нуля до половины частоты дискретизации в силу симметрии спектра. При обработке комплексного входного

сигнала фильтры для отрицательных частот получаются из фильтров положительных частот операцией комплексного сопряжения всех весовых коэффициентов. Фильтры подсистемы синтеза получаются из фильтров подсистемы анализа комплексным сопряжением всех весовых коэффициентов и перестановкой их в обратном порядке.

На рис. 9 показана зависимость остаточной энергии сигнала в зоне непрозрачности от уровня боковых лепестков для фильтра-прототипа 48-го порядка при выборе коэффициента ортогональности ξ_4 = $10^{-3}(M/N)$. Кружочком отмечено положение рассчитанного фильтра. Крестиком отражено положение фильтра-прототипа, найденного методом Хартенека и др. [12], имеющего линейную фазу.

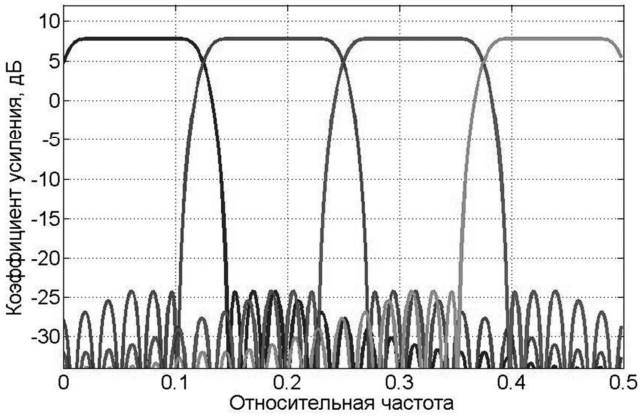


Рис. 8. Амплитудно-частотные характеристики фильтров подсистемы анализа

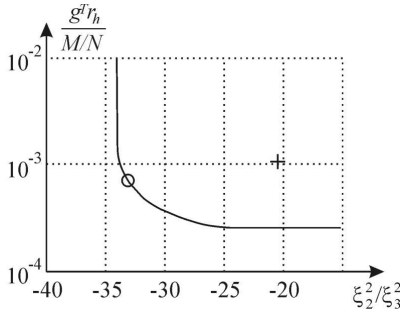


Рис. 9. Зависимость остаточной энергии сигнала в зоне непрозрачности от уровня боковых лепестков для фильтра-прототипа 48-го порядка при выборе коэффициента ортогональности ξ_4 = $10^{-3}(M/N)$

Результаты компьютерного моделирования

В данном разделе статьи представлены результаты двух экспериментов. Цель первого эксперимента состояла в проверке эффективности работы эквалайзеров с динамическим выбором порядка (ДВП) в стационарном канале связи с аддитивным белым гауссовским шумом (АБГШ). Во втором эксперименте исследовалась способность эквалайзеров с ДВП реагировать на изменения частотных характеристик нестационарного канала связи.

Для обоих экспериментов была выбрана модель, аналогичная предложенной в [13]: 4-позиционная относительная фазовая манипуляция (QPSK) с использованием кода Грея, скорость передачи равна 10 Мбит/с. Фильтрация импульсов осуществляется на приёмной и передающей сторонах цифровым фильтром с характеристикой типа «приподнятый косинус», нормированной на квадратный корень, и коэффициентом спада, равным 0,23. Как и в [15], при модуляции используется одна несущая частота, но обработка сигнала осуществляется во временной области (без вычисления спектра при помощи преобразования Фурье). Предполагается, что настройка происходит по известному в приёмнике эталонному сигналу, поэтому пилот-сигналы в модели отсутствуют.

Все эксперименты проводились для 8-канальных комплексных адаптивных фильтров (N = 8) с 6-кратным понижением частоты дискретизации (M = 6). Для адаптивных фильтров с фиксированным порядком (одноканального и многоканального) порядки были заданы таким образом, чтобы длина воспроизводимой импульсной характеристики получалась одинаковой. Порядок одноканального адаптивного фильтра равнялся 512, а порядок МАФ был в M раз меньше, то есть 86. Для адаптивных фильтров с ДВП эти значения являлись верхними ограничениями. Таким образом, для всех адаптивных фильтров был установлен один и тот же эквивалентный порядок: L_{eq} = 512 (для фильтров с адаптивным выбором порядка это наибольший допустимый порядок). Начальное значение порядка L_0 = 64 для одноканального адаптивного фильтра, начальные значения порядков внутриканальных фильтров МАФ $L_{0,i}$ = 10 для всех каналов (i = 0..7). Вложенные фильтры как в одноканальном, так и в многоканальном фильтрах получались исключением крайних ΔL = 5 весовых коэффициентов. Шаг адаптации был задан равным 0,1.

В первом эксперименте в качестве модели канала связи была выбрана стационарная модель двухлучевого канала. Оба луча имели равные коэффициенты передачи, задержка между лучами равнялась 1 мкс. При моделировании на выходе канала связи к полезному сигналу добавлялся АБГШ: отношение сигнал-шум равнялось 40 дБ.

Рис. 10 позволяет наглядно сравнить по скорости и точности настройки четыре алгоритма: классический алгоритм НСКО с нормировкой шага адаптации (алгоритм NLMS [4]), НСКО с ДВП (ДВП-НСКО), 8-канальный адаптивный фильтр с фиксированным порядком на основе алгоритма НСКО (МАФ) и 8-канальный адаптивный фильтр с ДВП (ДВП-МАФ). Под относительной ошибкой понимается отношение мощности сигнала ошибки на выходе адаптивного фильтра к мощности входного сигнала.

Рис. 11 отражает процесс изменения порядков адаптивных фильтров с ДВП. Разделение в семействе кривых для ДВП-МАФ объясняется наличием плавного спада вблизи частоты среза амплитудно-частотных характеристик согласованных фильтров с характеристикой типа «приподнятый косинус». Кроме того, на рис. 11 наблюдается эффект ограничения: кривые достигают наибольшего допустимого порядка, равного 86.

Во втором эксперименте в качестве модели канала связи была выбрана нестационарная модель двухлучевого канала. Отношение средних квадратов коэффициентов передачи по лучам было равным 5. Задержка между лучами равнялась 1 мкс. Отношение сигнал-шум было уменьшено до 80 дБ, чтобы влиянием шума можно

было пренебречь. Динамику изменений в канале связи позволяет оценить рис. 12. Быстродействие подсистемы выбора оптимального порядка отражено на рис. 13.

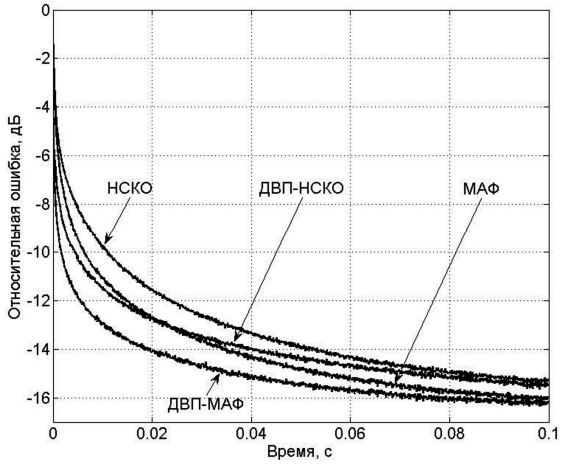


Рис. 10. Пример настройки четырёх эквалайзеров на основе алгоритма НСКО в условиях стационарного канала связи (эксперимент 1)

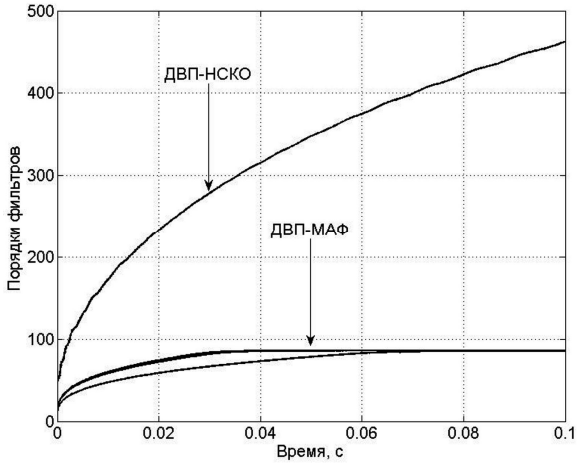


Рис. 11. Пример настройки четырёх эквалайзеров на основе алгоритма НСКО в условиях стационарного канала связи (эксперимент 1)

Рис. 14 позволяет наглядно сравнить по скорости и точности настройки четыре алгоритма: НСКО с нормировкой шага адаптации, ДВП-НСКО, МАФ и ДВП-МАФ. Некоторый провал на обучающих кривых в самом начале настройки объясняется выбранной моделью канала связи, в которой начальные значения коэффициентов передачи по лучам во всех экспериментах выбирались одинаковыми (1,0 и 0,2), что отражено также на рис. 12.

В первые моменты настройки, эквалайзеры с адаптивным выбором порядка снижают порядки АФ, благодаря чему настройка протекает значительно быстрее, чем у эквалайзеров с фиксированным порядком. В дальнейшем, порядки АФ растут, что позволяет повысить точность настройки в установившемся режиме. В данном примере отношение сигнал-шум удалось повысить с 10 до 14 дБ.

На рис. 15 представлен процесс изменения порядков адаптивных фильтров с ДВП. В отличие от рис. 13, на рис. 15 представлены результаты, полученные путём усреднения по более чем 10000 экспериментов. Оценим выигрыш по вычислительным затратам, достигаемый в эквалайзерах на основе ДВП-МАФ. Для определённости ограничимся результатами эксперимента с нестационарным двухлучевым каналом связи, представленными на рис. 15. Формулы для расчёта вычислительных затрат представлены в табл. 1.

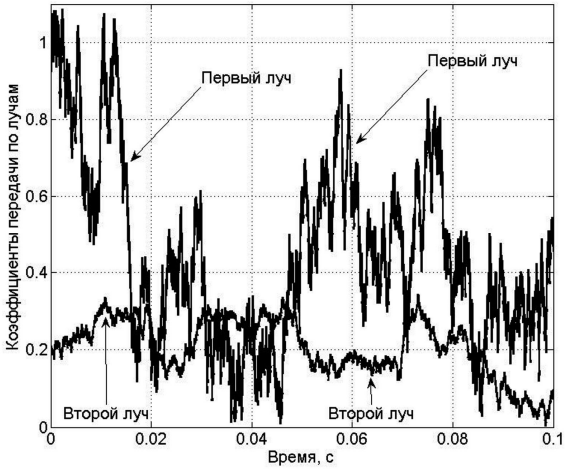


Рис. 12. Пример изменения коэффициентов передачи по лучам в двухлучевом нестационарном канале связи (эксперимент 2)

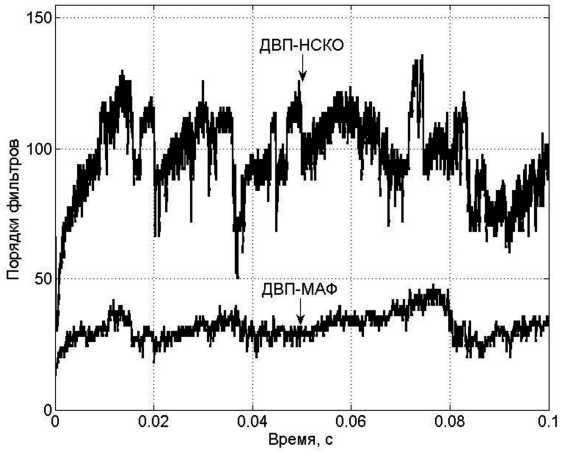


Рис. 13. Быстродействие подсистем выбора оптимального порядка в условиях нестационарного канала связи. Для МАФ для наглядности график построен для одного из 8 каналов (эксперимент 2)

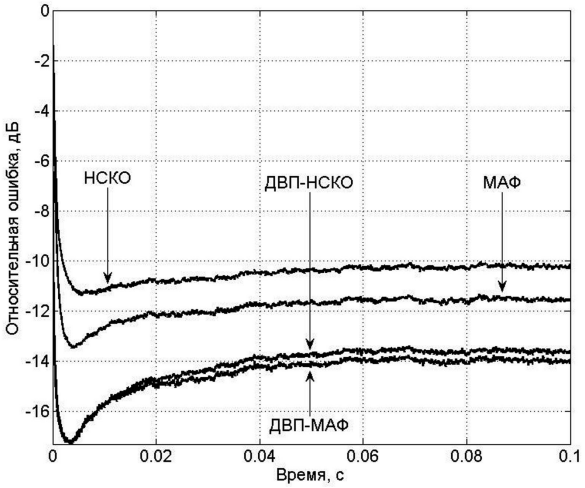


Рис. 14. Пример настройки четырёх эквалайзеров на основе алгоритма НСКО в условиях нестационарного канала связи (эксперимент 2)

При проведении компьютерного моделирования были выбраны следующие численные значения параметров. Использовался 8-канальный комплексный адаптивный фильтр (N = 8) с 6-кратным понижением частоты дискретизации (M = 6). Порядки фильтров анализа и

синтеза были выбраны одинаковыми L_H = L_G = 96, порядки адаптивных фильтров L_{eq} были равны 512.

Нетрудно оценить выигрыш по вычислительным затратам. Одноканальный эквалайзер с фиксированным порядком адаптивного фильтра и алгоритмом настройки НСКО требует $512 \cdot 2 + 1$ = **1025** комплексных умножений на каждый отсчёт дискретного входного сигнала (к.у./отс.). При переходе к многоканальной структуре эквалайзера с фиксированным порядком (МАФ) вычислительные затраты снижаются до $8 \cdot 96/6 + 1 \cdot 8 + 8 \cdot 512/6/6 \cdot 2$ = $136 + 8 \cdot 512/6/6 \cdot 2$ ≈ **366** к.у./отс. Предоставив эквалайзерам самостоятельный выбор наилучших порядков адаптивным фильтрам, можно достичь дополнительной экономии вычислительных затрат. В случае одноканального эквалайзера с адаптивным выбором порядка (ДВП-НСКО) потребуется $L_{eq} \cdot 2 + 1$ к.у./отс., а для многоканального эквалайзера с адаптивным выбором порядков (ДВП-МАФ) затраты составят $8 \cdot 96/6 + 3 \cdot 8 + 8 \cdot L_{eq}/6 \cdot 2 = 152 + 16/6 \cdot L_{eq}$ к.у./отс. Судя по рис. 15, в установившемся режиме работы затраты на ДВП-НСКО и ДВП-МАФ равны в среднем **313** и **259** к.у./отс., соответственно.

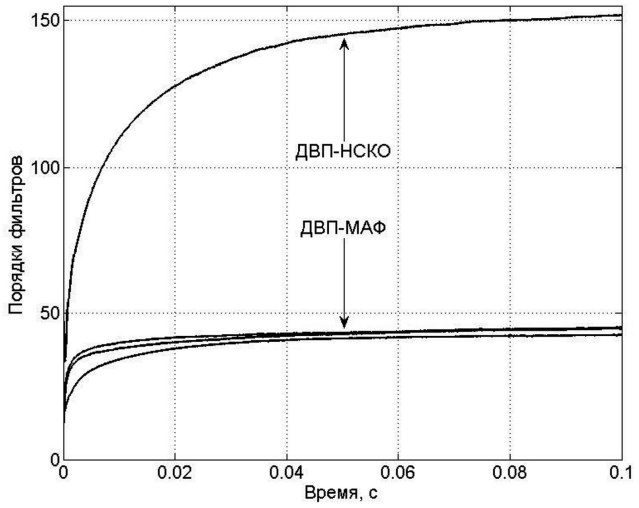


Рис. 15. Пример настройки четырёх эквалайзеров на основе алгоритма НСКО в условиях нестационарного канала связи (эксперимент 2)

Выводы

Введение процедуры динамического выбора порядков внутриканальных адаптивных фильтров в МАФ с независимой настройкой каналов значительно повышает быстродействие МАФ. При этом все основные преимущества МАФ перед одноканальным алгоритмом НСКО – повышение точности настройки и снижение вычислительных затрат на обработку в реальном времени – сохраняются.

Литература

1. Haykin S., Adaptive filter theory, 4th ed. – Upper Saddle River, NJ: Prentice-Hall, 2002. – 936 с.

2. Vityazev V.V., Linovich A.Y. A subband equalizer with the flexible structure of the analysis/synthesis subsystem // Proc. 2010 IEEE Region 8 International Conference on Computational Technologies in Electrical and Electronics Engineering “SIBIRCON-2010,” vol. 1, Irkutsk – Listvyanka, Russia. – July 11–15, 2010. – pp. 174–178.

3. Линович А.Ю. Многоскоростная обработка сигналов в задачах обратного моделирования – Германия: LAP LAMBERT academic publishing, 2011. – 219 с., (http://www.morebooks.de).

4. Sayed A.H., Fundamentals of adaptive filtering. – New York: Wiley, 2003.

5. Витязев В.В. Цифровая частотная селекция сигналов. – М.: Радио и связь, 1993. – 240 с.

6. Crochiere R.E., Rabiner L.R. Multirate digital signal processing. – NJ: Prentice-Hall, 1983. – 411 с.

7. Уидроу Б., Стирнз С. Адаптивная обработка сигналов / Пер с англ. – М.: Радио и связь, 1989. – 440 с.

8. Harteneck M., Weiss S., Stewart R.W. Design of near perfect reconstruction oversampled filter banks for subband adaptive filters // IEEE transactions on circuits and systems II. 1999. – № 8. – pp. 1081 – 1086.

9. Wilbur M.R., Davidson T.N., Reilly J.P. Efficient design of oversampled NPR GDFT filter banks // IEEE transactions on signal processing. 2004. – № 7. – pp. 1947 – 1963.

10. Davidson T.N. Enriching the art of FIR filter design via convex optimization // IEEE signal processing magazine. 2010. – № 3. – pp. 89 – 101.

11. Grant M., Boyd S. Graph implementations for nonsmooth convex programs. Recent Advances in Learning and Control // Blondel V., Boyd S., Kimura H., Eds. New York: Springer. 2008. – pp. 95 – 110. [Online]. Available: http://stanford.edu/~boyd/cvx

12. Wu S.-P., Boyd S., Vandenberghe L. FIR filter design via spectral factorization and convex optimization // Applied and computational control, signals, and circuits, vol 1, Datta B., Ed. Cambridge, MA: Birkhauser, May 1999, ch. 5, pp. 215 – 245.

13. R. Boite and H. Leich. A new procedure for the design of high order minimum phase FIR digital or CCD Filters. Signal Processing, 3:101 – 108, 1981.

14. G. A. Mian and A. P. Nainer. A fast procedure to design equiripple minimum-phase FIR Filters. IEEE Trans. Circuits Syst., 29(5):327 – 331, 1982.

15. Ng B., Lam C.T., Falconer D. Turbo frequency domain equalization for single-carrier broadband wireless systems // IEEE transactions on wireless communications. 2007 – № 2. – pp. 759 – 767.

METHODS OF MULTIRATE SIGNAL PROCESSING IN SUBBAND ADAPTIVE FILTERS WITH THE SELF-ORGANIZING STRUCTURES: EQUALIZERS WITH DYNAMIC ORDER SELECTION

Linovich A.Y.

Adaptive algorithms for the self-organizing subband equalizers are proposed. Simulation results demonstrating their performance in the wideband communication equipment are supplied. The series of advantages and drawbacks of all described multi-channel structures are considered. Many different details for this way of equalizing in wireless communications are discussed.

УДК 681.513.6

АДАПТИВНАЯ ЦИФРОВАЯ ЛИНЕАРИЗАЦИЯ ПЕРЕДАТОЧНЫХ ХАРАКТЕРИСТИК ЭЛЕКТРОННЫХ УСИЛИТЕЛЬНЫХ УСТРОЙСТВ

Гудкова Н.В., *к.т.н., доцент кафедры систем автоматического управления Таганрогского технологического института Южного Федерального Университета, e-mail: tala_gud@rambler.ru;*
Гудков В.И., *директор проектов Московского представительства ГК «Телрос», e-mail: v.gudkov62@gmail.com*

Ключевые слова: усилитель мощности, амплитудная характеристика, линеаризация, динамический диапазон, адаптивная обработка.

Предлагается один из возможных вариантов решения задачи линеаризации нелинейных статических характеристик электронных устройств, базирующийся на принципах адаптивного обратного моделирования объектов типа «черный ящик». В качестве примера рассмотрен алгоритм адаптивной компенсации нелинейности усилителя мощности (УМ), который может служить альтернативой или дополнением к традиционному цифровому предыстортеру, реализованному на основе LUT-таблиц. Приведены результаты компьютерного моделирования системы компенсации нелинейности амплитудной характеристики УМ, подтверждающие эффективность предложенных решений. Отмечается простота цифровой реализации адаптивного алгоритма средствами микросхемотехники.

Введение

Важнейшей интегральной характеристикой электронных устройств обработки информационных сигналов является динамический диапазон (ДД), который связан, с одной стороны, с чувствительностью устройства, определяемой его собственными шумами, а с другой стороны, с проявлением его нелинейных свойств.

На первых этапах развития техники приема и обработки информации требования к ДД были второстепенными, но в настоящее время их значение возрастает с каждым годом. Это объясняется в основном желанием использовать экономичные режимы работы выходных каскадов для повышения общего к.п.д. РЭА.

В результате относительное изменение уровней помех и полезного сигнала на входе устройства обработки в обычных условиях, т.е. ДД входных воздействий, может составлять 90…100 дБ. При работе радиосредств в условиях экстремальной электромагнитной обстановки он достигает 140…160 дБ и более, при этом помехи на входе радиоприемных устройств измеряются единицами и десятками вольт. В таких условиях прием сигналов с заданным качеством проблематичен.

Поэтому актуальной задачей проектирования устройств приема и обработки сигналов является повышение линейности передаточных характеристик блоков, составляющих данное устройство, и доведение диапазона линейной зависимости выходного отклика от входного воздействия до максимально возможной величины.

В частности, возрастает потребность в линейных усилителях мощности (УМ), сочетающих высокие энергетические характеристики с низким уровнем интермодуляционных искажений. Эта потребность определяется необходимостью передавать все большие объемы информации, что приводит к все более плотному размещению каналов в частотном диапазоне. Но линейные УМ, имеющие низкий уровень искажений, характеризуются низким КПД. Последнее оборачивается высоким уровнем потребляемой мощности, что нежелательно.

Таким образом, обеспечение качества передачи, определяющее высокие требования к линейности применяемого усилителя, вступает в противоречие с необходимостью снижения потребляемой усилителем мощности. Выходом из этого тупика является использование подходов, позволяющих создавать линейные усилительные устройства, имеющие высокий КПД, характерный для нелинейных УМ. В настоящее время существует значительное количество способов линеаризации усилителей мощности средствами аналоговой техники. Среди них такие методы, как введение цепи обратной связи, отдельное усиление радиочастотного сигнала и огибающей с последующей модуляцией радиосигнала, предварительное искажение входного сигнала перед непосредственным усилением и ряд других.

Наряду с аналоговыми методами весьма успешно развивается направление, связанное с разработкой численных методов линеаризации. Особенно это касается задачи компенсации нестабильности характеристик усилительных устройств, работающих в условиях неопределенности. Такого рода компенсация должна обеспечивать адаптацию устройств к случайным изменениям мощности входного сигнала УМ, температурным изменениям, колебаниям напряжения источника питания, старению транзисторов и т.п.

Об актуальности проблемы свидетельствуют многочисленные посвященные ей публикации, а также гранты, выделяемые компаниями связи на проведение работ в этой области. Причем, особое внимание уделяется разработкам цифровых устройств предыскажения, так называемым, «предыстортерам» (predistorter) [1, 2, 3].

Считается, что применение предыстортеров сулит большие перспективы, поскольку позволяет увеличить КПД усилителя при сохранении его габаритов и почти без увеличения цены. Например, значение КПД современных УМ базовых станций беспроводной связи в среднем составляет приблизительно 10%, тогда как при использовании недавно выпущенной фирмой Texas Instruments микросхемы GC5322 оно возрастает до 15%-40%. В большинстве случаев в основу алгорит-

мов предыстортера заложено представление модели нелинейного усилителя в виде LUT (Look Up Table)-таблицы добавлений, которая представляет собой набор коэффициентов многомерного ряда Вольтерра, описывающего УМ во временной области. Работа алгоритмов линеаризации заключается в анализе входного сигнала и последующем внесении в него «предыскажений» с тем, чтобы этим компенсировать последующие нелинейные искажения, вносимые усилителем.

Основной трудностью при построении табличной модели нелинейного усилителя является необходимость его предварительного тестирования, что обычно требует применения специального дорогостоящего оборудования, специальных методик и хорошо подготовленных для этой работы специалистов. Кроме того, существуют определенные сложности, связанные с поиском компромисса между требуемой памятью процессора для ее хранения, скоростью адаптации модели и быстродействием алгоритма расчёта в режиме реального времени. С учетом сказанного можно констатировать, что традиционные способы линейной коррекции характеристик УМ не свободны от недостатков. Ниже предлагается метод цифровой адаптивной линеаризации передаточной характеристикой усилителя мощности, который, по мнению авторов, может служить альтернативой или дополнением к общепринятым подходам.

Метод адаптивной линеаризации

Метод базируется на принципах управления, использующих адаптивное обратное моделирование неопределенных динамических объектов типа «черный ящик» [4, 5, 6]. Смысл этого вида моделирования состоит в создании цифровой адаптивной обратной модели (АОМ) некоторого объекта с неизвестными параметрами, которая должна служить наилучшим приближением дискретной передаточной функции, обратной передаточной функции этого объекта в процессе его функционирования. На рис. 1 показана одна из возможных структур системы управления «черным ящиком», который в рассматриваемом случае представляет собой усилитель мощности с неизвестной, в общем случае изменяющейся во времени, нелинейной амплитудной характеристикой $x = f(u)$.

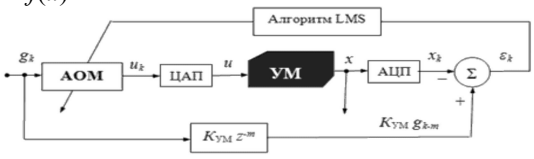


Рис. 1. Структурная схема системы управления амплитудной характеристикой УМ с АОМ

При этом предполагается, что с течением времени может также случайным образом изменяться мощность входного сигнала и положение «рабочей точки» усилителя. Роль управляющего устройства (адаптивного регулятора) в данной системе играет обратная модель УМ, реализованная в виде адаптивного трансверсального фильтра [4], дискретное уравнение которого имеет вид

$$u_k = \sum_{l=0}^{L} w_{lk}\, g_{k-l} \qquad (1)$$

где, g_k – дискретные отсчеты входного сигнала системы, u_k – отсчеты управляющего сигнала на выходе АОМ, L – длина фильтра, а w_{lk} – его весовые коэффициенты, которые в процессе адаптации перестраиваются по методу LMS в соответствии с рекуррентным соотношением

$$w_{l(k+1)} = w_{lk} + 2\mu\, g_{k-l}\, \varepsilon_k . \qquad (2)$$

В выражении (2) ошибка адаптации $\varepsilon_k = K_{УМ} \cdot g_{k-m} - x$, где $K_{УМ}$ – желаемое значение коэффициента усиления УМ; μ – параметр сходимости адаптивного алгоритма; m – число тактов задержки входного сигнала, которая вводится для учета инерционности усилителя.

После завершения процесса регулирования амплитудной характеристики $x = f(u)$ среднеквадратическая ошибка $CKO = E[\varepsilon_k^2]$ становится равной своему минимальному значению, которое зависит от заданной относительной средней CKO адаптации M ($0<M<1$). При этом скорость сходимости адаптивного процесса можно приближенно оценить по числу итераций [5]

$$T_{\text{СКО}} \cong \left[\frac{L+1}{M}\right] \qquad (3)$$

В тех случаях, когда инерционные свойства УМ выражены слабо, адаптивная процедура упрощается, так как для линеаризации амплитудной характеристики такого усилителя достаточно одного единственного весового коэффициента w_0 ($L = 0$) и нулевой временно́й задержки ($m = 0$). Анализ показал, что для устойчивой работы алгоритма компенсации нелинейности безынерционного усилителя должно выполняться условие

$$\mu \le \frac{M}{E[g_k^2]} \qquad (4)$$

где $E[g_k^2]$ – средняя мощность сигнала g_k.

Поскольку в общем случае величина $E[g_k^2]$ носит случайный характер, можно использовать для расчета параметра μ максимально допустимую мощность входного сигнала системы. Оптимальную настройку обеспечит автоматический расчет текущего значения μ в режиме реального времени. Желаемая величина коэффициента усиления должна быть задана из условия $K_{УМ} \le \frac{x_{\max}}{g_{\max}}$, где значения x_{max} и g_{max} определяются предельно допустимыми уровнями этих сигналов в системе.

Компьютерное моделирование

Принцип действия предлагаемых алгоритмов иллюстрируют результаты имитационного моделирования на ПК адаптивной системы линеаризации безынерционного УМ при подаче на ее вход случайной последовательности отсчетов белого шума со средней мощностью $E[g_k^2] \approx 0{,}33$. На рис. 2 показаны нормированные нелинейная и скомпенсированная амплитудные характеристики усилителя. Рис. 3 отражает процессы, происходящие в системе с некомпенсированной амплитудной характеристикой УМ ($u = g$), а рис. 4, 5 - процессы в адаптивной системе с компенсированной амплитудной характеристикой УМ ($u = w_0 g$) при двух относительных средних значениях СКО M = 0,5 и M = 0,1.

Из рисунков видно, что СКО в системе с компенсированной амплитудной характеристикой УМ при обоих значениях параметра M как минимум на порядок меньше СКО в системе с некомпенсированной характеристи-

кой. При этом число итераций адаптивных процессов $T_{СКО}$ приблизительно соответствует теоретическим значениям.

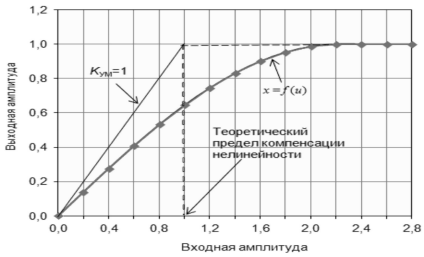


Рис. 2. Нормированные нелинейная и компенсированная амплитудные характеристики УМ

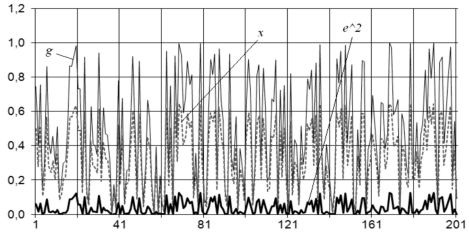


Рис. 3. Процессы в УМ с некомпенсированной амплитудной характеристикой ($u = g$; $СКО \approx 0{,}004$)

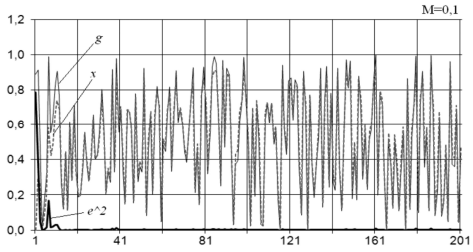


Рис. 4а. Процессы в системе с компенсированной амплитудной характеристикой УМ ($T_{СКО} \cong 10$; $СКО \approx 0{,}002$)

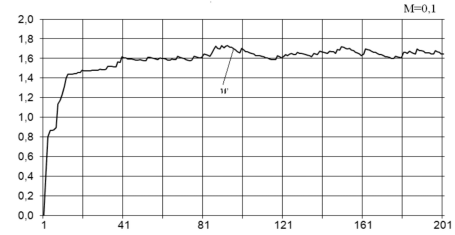


Рис. 4б. Весовой кэффициент в системе с компенсированной амплитудной характеристикой УМ ($u = w_0 g$)

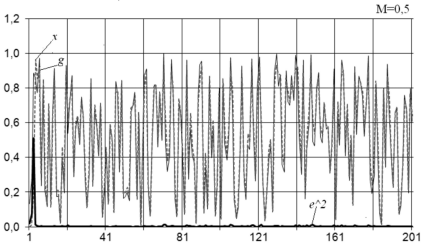


Рис. 5а. Процессы в системе с компенсированной амплитудной характеристикой УМ ($T_{СКО} \cong 2$; $СКО \approx 0{,}002$)

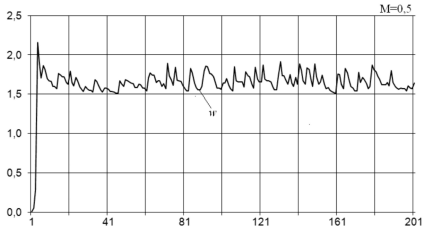


Рис. 5б. Весовой кэффициент в системе с компенсированной амплитудной характеристикой УМ ($u = w_0 g$)

Заключение

Выполненные в работе исследования позволяют сделать вывод о том, что использование предлагаемых решений может служить эффективным и экономичным способом цифровой линеаризации амплитудных характеристик безынерционных УМ в динамическом диапазоне, определяемом теоретическим пределом компенсации нелинейности, который в свою очередь зависит от выбранной величины $K_{УМ}$.

Как показала практика, при правильной настройке АОМ предложенный метод также обеспечивает хорошие результаты в задачах линеаризации характеристик УМ с инерционным запаздыванием.

Следует подчеркнуть, что алгоритмы метода достаточно легко реализуются средствами микросхемотехники и могут найти широкое применение в системах адаптивной линеаризации характеристик разнообразных технических устройств, например, датчиков физических величин технологических объектов и др.

Литература

1. D.-S. Han and T. Hwang, «An adaptive predistorter for the compensation of HPA nonlinearity», IEEE Transactions on Broadcasting, vol. 46, no. 2, 2000.

2. Chang, S. and Powers, E. J., «A simplified predistorter for compensatoin of nonlinear distortion in OFDM systems», in Proc. IEEE Global Telecommun. Conf., vol. 5, Nov. 2001.

3. P. L. Gilabert, G. Montoro and A. Cesari, «A recursive digital predistorter for linearizing RF power amplifiers with memory effects», in Proc. APMC'06, Pacifico Yokohama, Japan, Dec. 2006, vol. 2.

4. Уидроу Б., Стирнз С. Адаптивная обработка сигналов / М.: Радио и связь, 1989.

5. Гудкова Н.В. Цифровое управление техническими объектами с применением адаптивного обратного моделирования // Автоматизация и современные технологии. М.: Машиностроение, 2006. №4.

6. Widrow B., Walach E. Adaptive Inverse Control. A Signal Processing Approach, Wiley, Hoboken, NJ, 2008.

ADAPTIVE DIGITAL LINEARIZATION TRANSFER CHARACTERISTICS OF ELECTRONIC AMPLIFYING DEVICES

Gudkova N., Gudkov V.

Suggests possible solution of the the problem of linearization nonlinear static characteristics of electronic devices, based at the principle adaptive inverse modeling of objects such as «black box». As an example are considered the algorithms of adaptive compensation of nonlinearity of power amplifier (PA), that may be an alternative or complementation to conventional digital predystorter implemented in terms of LUT-tables. The results of computer simulation of the system of compensation of the nonlinearity of the amplitude characteristics of the PA, are confirming the effectiveness of the proposed solutions. The digital adaptive algorithms may be realized by means of micro circuitry.

УДК 620.178.53

МОНИТОРИНГ СПЕКТРАЛЬНЫХ ИЗМЕНЕНИЙ ВО ВРЕМЯ ИСПЫТАНИЙ НА ВИБРОУСТОЙЧИВОСТЬ

Бегишев С.В., *программист ЗАО «Электронные технологии и метрологические системы»,*
e-mail: begishev@zetms.ru

Одним из важных этапов производства технических изделий являются испытания на стойкость к воздействию механических факторов. Эти испытания проводятся для выявления конструкционных недостатков, производственных дефектов, определения работоспособности в сложных условиях эксплуатации. Сложность испытаний на воздействие механических факторов заключается в возможности частичного или полного разрушения испытуемого образца. При разрушении испытуемого образца информация о ходе испытаний может оказать помощь для выяснения причин разрушения, и чем детальнее и шире информация, тем лучше. Если разрушение испытуемого образца недопустимо, то необходимо определить момент остановки испытаний.

Во время испытаний на воздействие классического удара к испытуемому образцу прикладывается на короткий промежуток времени значительная сила. Большая часть энергии, передаваемой испытуемому образцу, идёт на перемещение, оставшаяся поглощается им и идёт на деформацию. Количество энергии, поглощённой телом в ходе виброиспытаний, может показать степень деформации испытуемого образца.

В эксперименте в качестве демпфера использовался кусок пластилина весом 100 грамм, которому была придана форма кубика со сторонами 41x41x41 мм. Кусок пластилина поместили в центр расширительного стола вибростенда. Первый датчик был установлен рядом с куском пластилина на расширительном столе, второй – непосредственно на пластилине. В ходе виброиспытаний выполнялся классический удар полусинусоидальной формы с амплитудой 1 g и длительностью 4 мс. На рис. 1 хорошо виден «провал» на пике сигнала, снятого со второго датчика.

При длительных испытаниях деформации в испытуемом образце будут накапливаться постепенно. Для определения величины деформации воспользуемся методом сравнения спектра ударного импульса с образцовым спектром. В качестве образцового спектра возьмём среднее арифметическое спектров первых ударных импульсов. В ходе виброиспытаний датчик, установленный на пластилине, отклонился от вертикального положения, что, несомненно, повлияло на передаточную характеристику.

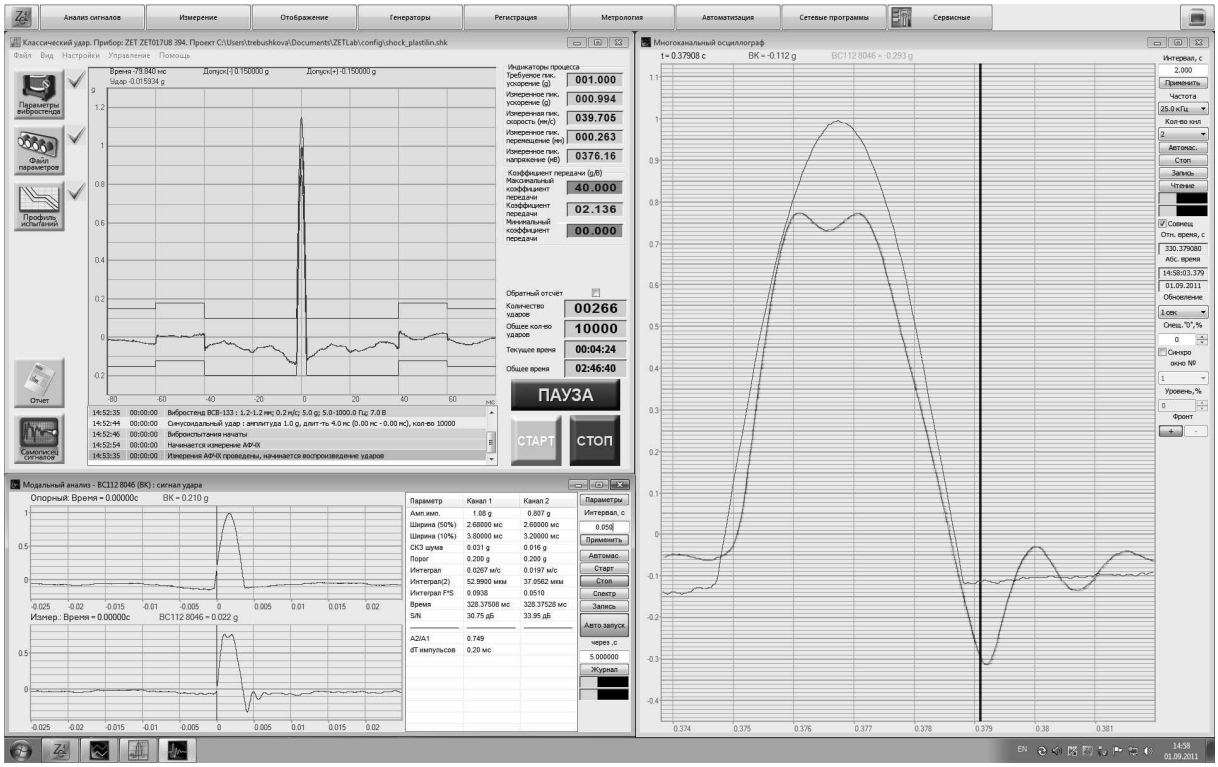


Рис. 1. Классический удар

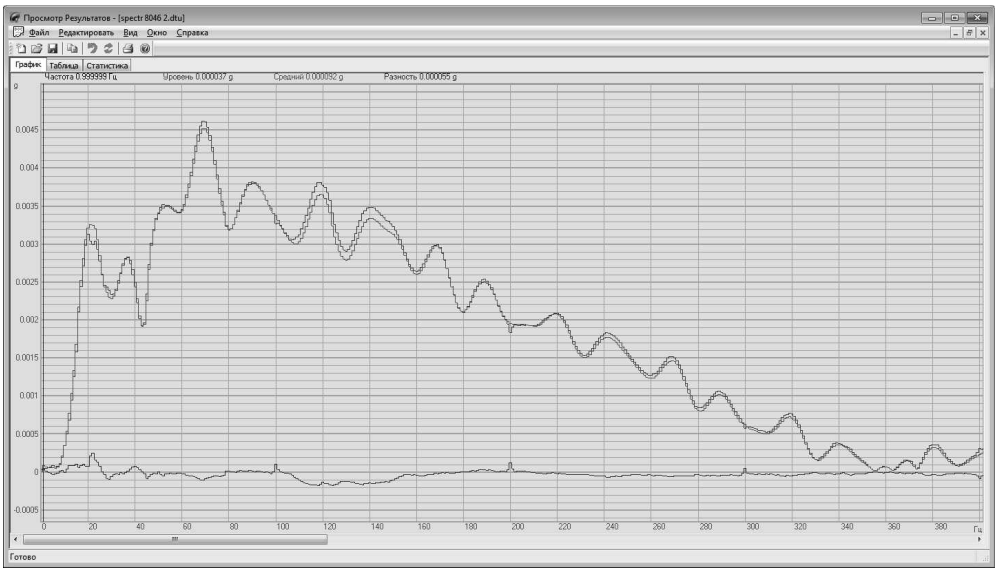


Рис. 2. Спектры удара: мгновенный, усреднённый и их разность

Разность между энергиями колебаний двух датчиков, вычисленная в программе «Модальный анализ» (рис. 1, «Интеграл F*s»), может дать определённую информацию о деформации испытуемого образца, но полезнее было бы получить информацию о происходящих изменениях в испытуемом образце. Такую информацию может дать спектр сигнала и определение его отклонения от значений в начале испытаний. На рис. 2 представлены графики мгновенного спектра в конце виброиспытаний, график усреднённого спектра за первые 100 секунд и график разности первых двух спектров.

СКЗ мгновенного спектра составило 0,001316 g, скз усреднённого спектра 0,001293 g, скз разности спектров 0,000075 g, что составляет 5.8%.

В ходе испытания требуется контроль над отклонениями от образцового спектра в реальном времени, а не расчёт отклонения по их завершении. Алгоритм работы программы контроля достаточно прост: в течение заданного времени в начале испытаний накапливаются измерения спектра для вычисления образцового спектра, а потом в течение всего времени испытаний сравниваются с ним, и строится график отклонений. При превышении заданного порогового уровня отклонений подаётся сигнал оператору.

Данная программа была реализована в SCADA-система ZETView. На рис. 3 приведён проект, реализующий описанный выше алгоритм. SCADA-система ZETView позволяет легко и быстро реализовывать алгоритмы любой сложности и представить результаты в удобном для работы оператора виде.

Задачи контроля спектральных характеристик возникают не только при виброиспытаниях, но и во многих других областях науки и техники. Приведённый выше алгоритм и его реализация в SCADA-системе ZETView (рис. 4) могут быть легко применены для решения широкого спектра подобных задач.

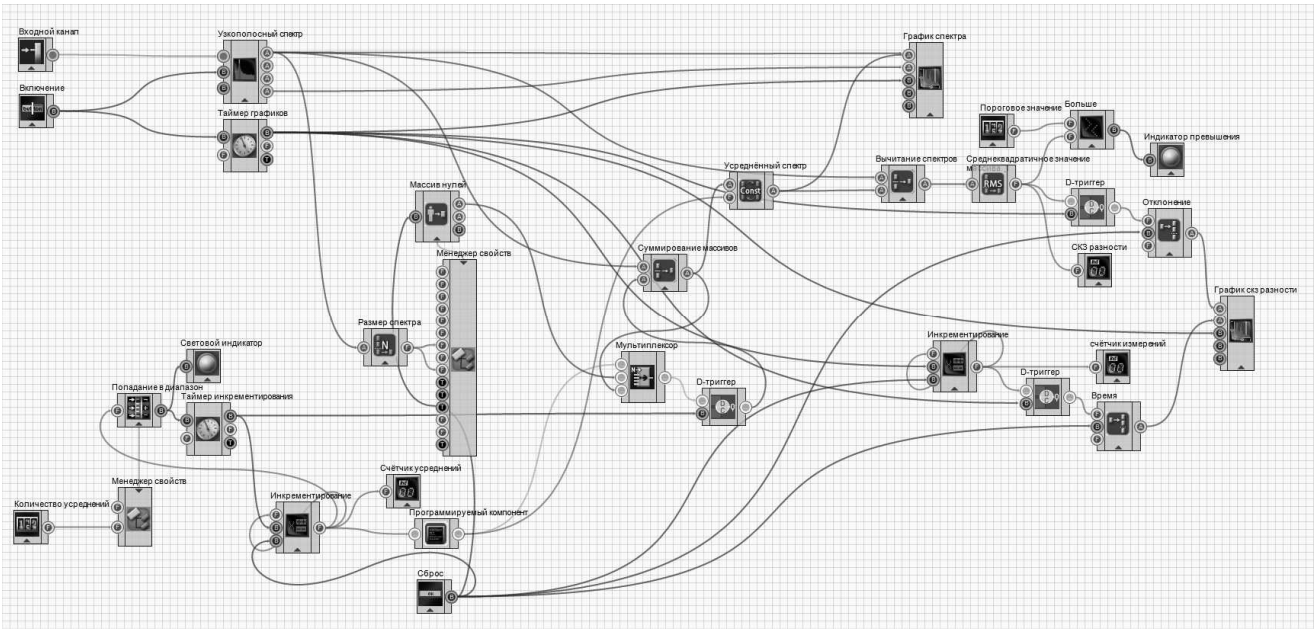


Рис. 3. SCADA-проект ZETView регистрации отклонений от усреднённого спектра

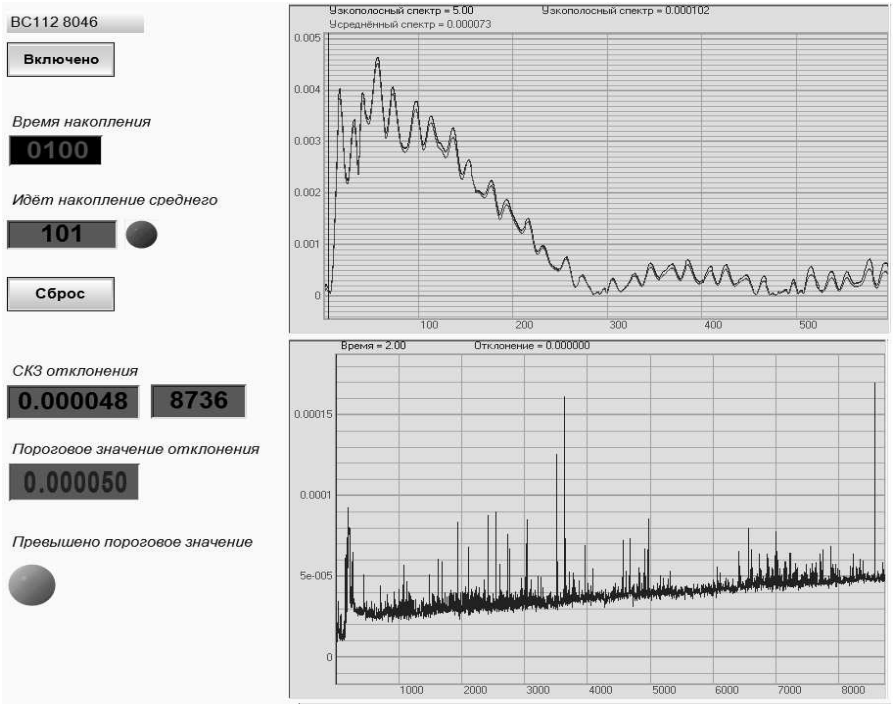


Рис. 4. SCADA-проект ZETView регистрации отклонений от усреднённого спектра, вид оператора

14-я Международная Конференция
ЦИФРОВАЯ ОБРАБОТКА СИГНАЛОВ И ЕЕ ПРИМЕНЕНИЕ

28 марта – 30 марта 2012 г., Москва, Россия

***The 14th International Conference* DIGITAL SIGNAL PROCESSING AND ITS APPLICATIONS**

ИНФОРМАЦИОННОЕ ПИСЬМО

Уважаемый коллега!
Приглашаем Вас принять участие в работе 14-й Международной конференции «Цифровая обработка сигналов и ее применение- DSPA'2012», которая состоится в ИПУ РАН 28 марта – 30 марта 2012 года

ОРГАНИЗАТОРЫ:

- Российское научно-техническое общество радиотехники, электроники и связи им. А.С. Попова
- IEEE Signal Processing Society
- Российская секция IEEE
- Институт радиотехники и электроники РАН
- Институт проблем управления РАН
- Институт проблем передачи информации РАН
- Московский научно-исследовательский телевизионный институт (ЗАО МНИТИ)
- Компания AUTEX Ltd. (ЗАО «АВТЭКС»

ПРИ УЧАСТИИ:

- Федеральное агентство по промышленности РФ
- Министерство образования и науки РФ
- Международный союз приборостроителей и специалистов по информационным и телекоммуникационным технологиям
- ФГУП ГКРЧ
- ГСКБ «АЛМАЗ-АНТЕЙ»
- ЗАО «Инструментальные системы»
- НТЦ «МОДУЛЬ»
- ЗАО «СКАН Инжиниринг Телеком»
- ГУП НПЦ «Элвис»
- Владимирский государственный университет
- Московский авиационный институт
- Московский государственный технический университет им. Н.Э. Баумана
- Московский институт радиотехники, электроники и автоматики
- Московский технический университет связи и информатики
- Московский энергетический институт
- Рязанский государственный радиотехнический университет
- Санкт-Петербургский государственный университет телекоммуникаций
- Санкт-Петербургский государственный электротехнический университет - ЛЭТИ
- Ульяновский государственный технический университет
- Ярославский Государственный Университет
- Московский Физико-технический институт (университет)