

ЦИФРОВАЯ ОБРАБОТКА СИГНАЛОВ

Научно-технический журнал
№ 1/2008

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. Бабаян, Ю.И. Борисов,
С.А. Грибачев, Г.В. Зайцев, Р.В. Зубарев,
А.П. Кирпичников, М.И. Кривошеев,
Н.А. Кузнецов, М.С. Куприянов, А.А. Ланнэ,
В.Г. Мистюков, С.Л. Мишенков,
А.А.Петровский, Ю.Н. Прохоров,
А.Н. Соловьев, Ю.Г. Сосулин,
В.В. Шахильдян, Ю.С. Шинаков

Адрес редакции:

107031 Москва, Рождественка, 6/9/20, стр. 1
Тел.: (495) 621-71-08, (495) 621-16-39
Факс: (495) 621-16-39
E-mail: nto.popov@mtu-net.ru
tor@rgta.ryazan.ru
http://www.dspsa.ru

Для писем:

129090 Москва, а/я 48.

Издатель:

Российское научно-техническое общество радиотехни-
ки, электроники и связи им. А.С. Попова
Ответственный редактор: П.И. Зудков
Компьютерная верстка: В.В. Андреехина
Дизайн: И.Е. Артюхина

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

Подписано в печать 11.03.08
Формат 60x90/8.

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

Заказ № 1348. 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:

Зубарев Ю.Б., Золотарев В.В., Овечкин Г.В.

**Обзор методов помехоустойчивого кодирования с
использованием многопороговых алгоритмов 2**

Витязев В.В., Витязев С.В., Зайцев А.А.

**Многоскоростная обработка сигналов: ретроспектива
и современное состояние (часть 1) 12**

Парфенюк М., Петровский А.А.

**Параунитарные банки фильтров на основе алгебры
кватернионов: теория и применение 22**

Дворкович В.П., Гильманшин А.В.

**Новый подход к использованию вейвлет-фильтров при
обработке изображений 37**

Моисеев А.А., Кобелев В.Ю., Волохов В.А.

**Курвлет-преобразование в задаче подавления шума в
изображениях 43**

Приоров А.Л.

Согласованные трехмерные вейвлет-фильтры 51

*Балдин М.В., Воронков Д.И., Руткевич А.В., Сенченко М.Н.,
Стешенко В.Б., Шишкин Г.В.*

**Эволюция применения ПЛИС в системах ЦОС: от
специализированного сопроцессора к аппаратной
платформе (часть 1) 58**

Витязев С.В.

**Новые разработки компаний Texas Instruments и Analog
Devices в 2007 году 64**

**ЦОС в университетах: университетская программа
компании Texas Instruments, часть 2 68**

Кирпичников А.П., Киришин Е.А.

**Микроконтроллер USIP® от INNOVA CARD: новая
платформа для систем с высокими требованиями к
безопасности 69**

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

ОБЗОР МЕТОДОВ ПОМЕХОУСТОЙЧИВОГО КОДИРОВАНИЯ С ИСПОЛЬЗОВАНИЕМ МНОГОПороГОВЫХ АЛГОРИТМОВ

Зубарев Ю.Б., Золотарёв В.В., Овечкин Г.В.

Введение

Быстрый рост объемов обработки данных, развитие цифровых систем вещания и вычислительных сетей предъявляют весьма высокие требования к минимизации ошибок в используемой дискретной информации. Переход всех видов создания, хранения, использования и передачи данных, а также средств вещания на цифровые методы, происходящий сейчас во всём мире, ещё более повышает важность высококачественной передачи цифровых потоков и особо надёжно-

Рассмотрены основные достижения теории помехоустойчивого кодирования и вытекающие из неё прикладные результаты, базирующиеся на важном классе простых мажоритарных итеративных методов исправления ошибок, получивших название многопороговых декодеров. Описаны как достижения последних десяти лет в области кодирования и последующего декодирования двоичных кодов, что наиболее применимо в различных высокоскоростных спутниковых каналах, так и принципиально новые возможности, которые появились у разработчиков после детального изучения характеристик недвоичных, так называемых символьных, декодеров этого же типа. Проводится краткое сопоставление характеристик многопороговых декодеров и других методов: декодера Витерби, декодера турбо и низкоплотностных кодов, а также некоторых методов исправления ошибок для кодов Рида-Соломона, которые оказываются намного сложнее мажоритарных алгоритмов.

го хранения сверхбольших объёмов данных. Успешная работа этих систем возможна только при наличии специальной эффективной аппаратуры, которая позволяет гарантировать достоверную передачу информации. Важнейший вклад в повышение достоверности обмена цифровыми данными вносит теория помехоустойчивого кодирования. На её основе разрабатываются всё новые методы защиты от ошибок, базирующиеся на использовании корректирующих кодов. Эффект от их применения может выражаться в том, что в системе связи они позволяют при прочих равных условиях многократно увеличивать скорость или дальность передачи, снижать размеры очень дорогих антенн или работать при существенно пониженном уровне полезного сигнала. Применение кодирования можно рассматривать и просто как способ многократного увеличения к.п.д. дорогих спутниковых и прочих цифровых каналов связи.

Ниже кратко охарактеризована сложившаяся к настоящему времени ситуация в технике декодирования. Ограничимся в анализе только системами передачи данных без обратной связи от приемника к передатчику, каналами с аддитивным белым гауссовским шумом (АБГШ) и линейными кодами, которые существенно облегчают реализацию декодеров. Основное внимание уделим методу коррекции ошибок, называемому многопороговым декодированием (МПД). Основными причинами выбора для глубоких разносторонних исследований и анализа в этом обзоре именно многопороговых алгоритмов являются следующие:

- способность мажоритарных методов исправлять большое число ошибок за пределами гарантированной корректирующей способности;
- крайне незначительная сложность пороговых процедур декодирования;

- свойство разработанных МПД алгоритмов почти всегда достигать оптимальных решений при весьма высоких уровнях шума в канале связи;

- предельная легкость реализации МПД даже для очень длинных кодов, когда только и возможно достижение максимально допустимых значений эффективности кодирования.

Одной из наиболее важных особенностей разрабатываемых многопороговых алгоритмов, обеспечивающей их высокую эффективность, оказывается идеология проводимых исследований. В то время как все усилия представителей современной западной науки направлены на повышение эффективности новых алгоритмов исправления ошибок при гораздо меньшем их внимании к вопросам сложности реализации, например, выраженной в числе операций декодирования, то обязательным исходным требованием к новым алгоритмам многопорогового декодирования, о которых, в основном, и будет идти далее речь, в течение всех 35 лет их развития всегда являлось предельно малое число операций декодера как при программной, так и при аппаратной их реализации. Столь своеобразный подход к проблеме простого и эффективного декодирования на базе МПД дал поразительные результаты. Оказалось, что практически всегда упрощение уже разработанных ранее алгоритмов становится новой практически неразрешимой проблемой, решением которой иногда оказываются более простые модификации исходных алгоритмов. Однако во многих случаях такие упрощения довольно заметно снижают эффективность декодирования по сравнению с исходными методами.

В то же время принцип «от простого к более эффективному», которому всегда следовали разработчики алгоритмов класса МПД, продемонстрировал за многие годы их исследования большие возможности

создания и последующей незначительной модификации исходных методов, которая значительно повышала их эффективность при большом уровне шума. Это и позволило сохранить минимальную сложность всех методов декодирования на основе МПД, которая, как будет показано далее, довольно часто оказывается при равной корректирующей способности существенно меньшей, чем у конкурирующих методов, в сотни, а иногда и в тысячи раз.

Принцип многопорогового декодирования

Многопороговый декодер самоортогональных кодов [1 – 5] является развитием простейшего порогового декодера Мессе и позволяет декодировать очень длинные коды с линейной от длины кода сложностью реализации. В основе работы МПД лежит итеративное декодирование, что позволяет вплотную приблизиться к решению оптимального декодера в достаточно широком диапазоне кодовых скоростей и уровней шума в канале. При этом МПД сохраняет простоту и быстродействие обычного порогового декодера, что делает его очень привлекательным для применения в существующих и вновь создаваемых высокоскоростных системах связи.

Рассмотрим принцип работы МПД. Пусть задан двоичный линейный систематический блочный или сверточный самоортогональный код (СОК), который используется для передачи сообщения из k двоичных символов. После кодирования общее число кодовых символов равно n , $n > k$. Пример схемы кодера блочного СОК и его многопорогового декодера представлен на рис. 1 и рис. 2. Данный код характеризуется следующими параметрами: длина кода $n=26$, длина информационной части кода $k=13$, кодовая скорость $R=1/2$, кодовое расстояние $d=5$, порождающий полином $g(x)=1+x+x^4+x^6$.

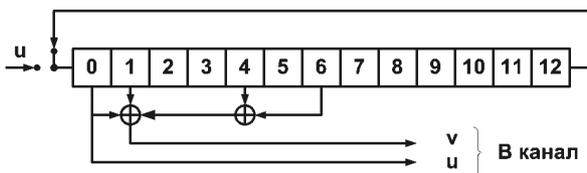


Рис. 1. Кодер блочного СОК с $R=1/2$, $d=5$ и $n=26$

Пусть необходимо передать информационный блок u . Устройство кодирования на основании информационного блока u формирует кодовое слово $c=(u,v)$. В результате передачи по двоичному симметричному каналу (ДСК) декодер получает вместо кодового слова c искаженное шумами сообщение $y=(u^*,v^*)$ длины n . Сначала, как и в обычном пороговом декодере, вычисляется синдром $s=yH^T$ (здесь H – проверочная матрица кода) принятого сообщения, и для каждого информационного символа u_j , $1 \leq j \leq k$, выделяется множество $\{s_{j_m}\}$ элементов синдрома с номерами $\{j_m\}$, называемых проверками относительно символа u_j и содержащих в качестве слагаемого ошибку e_j в этом символе.

Дополнительно вводится двоичный регистр d длины k , называемый разностным, первоначально заполненный нулями. Это – единственное отличие классического порогового алгоритма и МПД декодера. В регистре d будут отмечаться измененные информационные симво-

лы для того, чтобы декодер «помнил» принятое из канала сообщение и всегда мог вычислить разность между этим сообщением и кодовым словом, находящимся в информационном регистре.

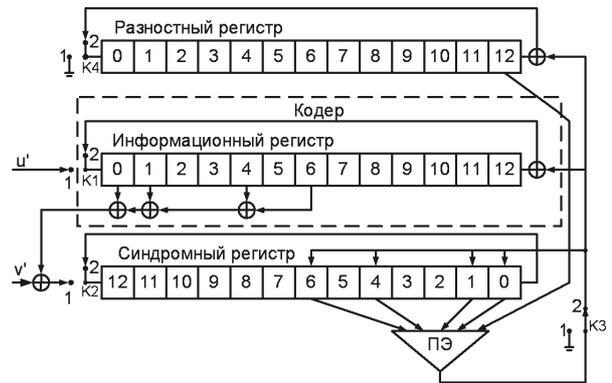


Рис. 2. Многопороговый декодер блочного СОК с $R=1/2$, $d=5$ и $n=26$

Основной шаг декодирования заключается в том, что для произвольно взятого символа u_j вычисляется функция правдоподобия L_j , зависящая от относящихся к нему проверок $\{s_{j_m}\}$ и j -го элемента регистра d :

$$L_j = \sum_{j_m} s_{j_m} + d_j \tag{1}$$

где $J=d-1$ – количество проверок (ненулевых элементов порождающего полинома $g(x)$); d_j – символ разностного регистра, относящийся к декодируемому символу u_j (равный 0 или 1); s_p – p -й элемент синдромного регистра, входящий в множество проверок относительно декодируемого символа u_j .

Общее число слагаемых в (1) равно минимальному кодовому расстоянию d . Если $L_j > T$, где $T=(d-1)/2$ – пороговое значение, то символ u_j , все проверки $\{s_{j_m}\}$ и символ d_j инвертируются, после чего выбирается другой символ u_m , $m \neq j$, для него снова вычисляется сумма L_m и т.д. Если же $L_j \leq T$, то сразу осуществляется переход к декодированию следующего символа u_m .

При выполнении основного шага декодирования все k информационных символов сообщения могут перебираться многократно и в любом порядке. При этом часть решений декодера на каких-то символах может быть ошибочной. Некоторые из этих ошибок будут исправлены при следующих попытках декодирования тех же символов (т.е. на следующих итерациях).

Основным свойством МПД является строгая сходимость его решения к решению оптимального (по максимуму правдоподобия) декодера (ОД), поскольку при каждом i -м изменении декодируемых символов суммарный вес синдрома s и разностного вектора d обязательно уменьшается, т.е. происходит переход к кодовому слову $c^{(i)}$, которое более правдоподобно, чем кодовое слово $c^{(i-1)}$, находившееся в МПД в предыдущий момент времени. Переходя от одного слова $c^{(i)}$ к другому, МПД может получить наиболее правдоподобное слово $c^{(ОД)}$, которое, собственно, и является решением оптимального декодера. Однако нельзя утверждать, что МПД обязательно достигнет решения оптимального декодера, так как во многих кодах, допускающих мажоритарное

декодирование, МПД на некоторых конфигурациях ошибок веса несколько большего, чем $d/2$, прекратит изменять информационные символы раньше, чем достигнет решения оптимального декодера.

Одной из основных причин этого является значительная подверженность пороговых декодеров, являющихся составной частью МПД, эффекту размножения ошибок [1, 4]. В результате вторая и последующие итерации декодирования вынуждены работать в основном с потоками пакетов ошибок от декодеров предыдущих итераций, что существенно уменьшает эффективность всего декодера. Следовательно, основным способом приближения решения МПД к решению оптимального декодера является уменьшение эффекта размножения ошибок.

Для уменьшения размножения ошибок необходим тщательный выбор кодов, характеризующихся малой степенью пересечения множеств ошибок, входящих в проверки относительно разных информационных символов, а также настройка параметров декодера (например, величины порогов на разных итерациях). Методика оценки качества кодов в соответствии с указанным выше критерием подробно изложена в [1, 4].

Отметим, что МПД, как и обычный пороговый декодер, легко модифицируется для суммирования проверок в (1) с некоторыми коэффициентами, в частности, при работе с квантованными на несколько уровней решениями мягкого модема, дополнительные выходные биты которого определяют надежность выносимого им решения. Использование мягких решений демодулятора позволяет достигать результатов на 1,4...1,7 дБ лучших, чем при использовании только жестких решений демодулятора. При этом выражение (1) для вычисления функции правдоподобия L_j принимает вид

$$L_j = \sum_{\{j_m\}} s_{j_m} w_{j_m} + d_j w_j, \quad (2)$$

где $\{w_{j_m}\}$ – коэффициенты, отражающие надежность проверок $\{s_{j_m}\}$; w_j – коэффициент, отражающий надежность принятого символа u_j .

Зависимость вероятности битовой ошибки $P_b(e)$ от отношения сигнал/шум E_b/N_0 в канале с АБГШ при использовании мягкого модема для многопорогового декодера самоортогональных кодов с кодовой скоростью $R=1/2$ и различным кодовым расстоянием d представлена на рис. 3. Пунктиром на рисунке показана вероятность ошибки декодирования тех же кодов с помощью оптимального декодера. Из представленных графиков видно, что МПД действительно обеспечивает близкое к оптимальному декодирование правильно выбранных кодов при достаточно высоком уровне шума в канале связи.

Сложность реализации многопороговых декодеров

Как уже отмечалось, сложность мажоритарных алгоритмов может быть формально оценена как линейно растущая с увеличением длины декодируемого кодового блока и кодового ограничения свёрточного кода.

Сложность программной реализации, как число операций на бит, в двоичном МПД свёрточного или блочного СОК равна $N_{МПД} \approx (I+1)(d+2)$ или $N_{МПД} \approx 4d+3I$ при пренебрежимо малых потерях в энергетике (менее 0,1 дБ) [5]. Здесь d – минимальное кодовое расстояние применяемого кода, I – число итераций декодирования.

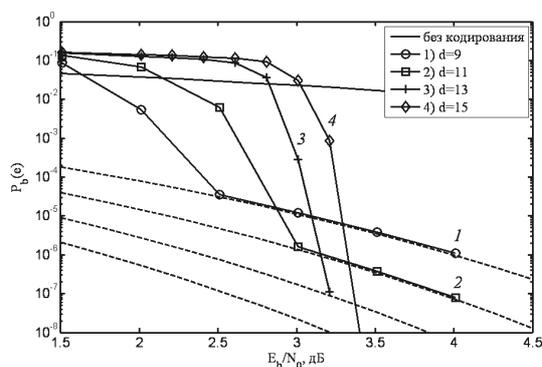


Рис. 3. Характеристики МПД для СОК с $R=1/2$ в канале с АБГШ

То, что при декодировании фактически каждая итерация в МПД требует всего несколько простейших операций типа сложение и сравнение, приводит к тому, что рост числа итераций почти не снижает реального быстродействия декодера, например, в программном варианте реализации. В качестве хорошей иллюстрации этого свойства алгоритма можно сослаться на демопрограмму МПД для свёрточного кода с $R=3/4$. Она используется в системе специального цифрового телевидения и обеспечивает на обычном ПК производительность порядка 4...10 Мбит/с, чего с огромным запасом хватает для обработки телевизионного цветного сигнала при весьма низком отношении сигнал/шум. Применяемые в ней коды проходят процедуру стандартизации. Демопрограмма и простая инструкция по её использованию может быть переписана с большого тематического веб-сайта ИКИ РАН www.mtdbest.iki.rssi.ru, где она размещена на учебной страничке. Действительно, сравнение МПД с другими методами показало, что скорость обработки на основе МПД оказывается примерно на 2 порядка более высокой, чем, например, у турбо декодеров с сопоставимыми параметрами энергетической эффективности

В случае аппаратной реализации МПД, например, на ПЛИС Xilinx или Altera, проведённые испытания подтвердили хорошие параметры их эффективности при одновременно очень высокой производительности, вплоть до 1,6 Гбит/с. Такая возможность появилась после реализации запатентованных технических решений для аппаратного МПД. Согласно этим решениям такой декодер превращается в одноктактовую решающую схему, причём на каждом такте работы он может принимать до 40 решений о декодируемых символах, а предельная частота тактов определяется максимально возможной скоростью сдвига принятых из канала данных по регистрам сдвига декодера, из которых он в основном и состоит. Типичная скорость продвижения информации по регистрам сдвига указанных ПЛИС находится в пределах от 100 до 250 Мбит/с, а число параллельно работающих регистров такого вида в МПД может превышать сотню. Это значит, что производительность МПД в аппаратном варианте реализации формально может существенно превышать даже 10 Гбит/с. По существу это снимает все ограничения по скорости обработки на такие устройства, что при обеспечиваемых этими алгоритмами

энергетических параметрах эффективности делает их единоличными лидерами среди всех методов высокоскоростной передачи цифровых потоков по дорогим спутниковым и прочим каналам. В частности, особенно полезны уже разработанные аппаратные версии МПД для систем дистанционного зондирования Земли, поскольку как раз их высокоскоростные потоки цифровых данных при ограниченной мощности передатчиков следует всемерно защищать методами помехоустойчивого кодирования.

Применение многопороговых декодеров в схемах с параллельным кодированием

Для приближения границы эффективной работы МПД к пропускной способности канала возможно его применение в схемах параллельного кодирования [4, 5]. В основе построения данных схем лежит выделение в СОК C_0 с кодовым расстоянием d_0 и кодовой скоростью R_0 некоторого составляющего кода C_1 с кодовой скоростью $R_1 > R_0$, тоже являющегося СОК. Кодовое расстояние d_1 выделенного кода выбирается значительно меньшим d_0 , и, как следует из рис. 3, область его эффективной работы будет ближе к границе Шеннона. При декодировании параллельного кода сначала выполняются несколько итераций декодирования составляющего кода C_1 , позволяющие примерно на порядок снизить вероятность ошибки в принятой из канала информационно последовательности, после чего в процесс декодирования включается оставшаяся часть кода C_0 . Отличительной особенностью данной схемы кодирования является то, что здесь внешний код работает с кодовой скоростью R_0 , в то время как в обычных каскадных кодах кодовая скорость внешнего кода близка к единице. Данное свойство обеспечивает существенное преимущество МПД перед другими каскадными конструкциями.

Для примера на рис. 4 представлены результаты моделирования схем с параллельным кодированием в канале с АБГШ для СОК с $R_0=6/12$, $d_0=13$ и $R_0=5/10$, $d_0=15$ (кривые «Параллельный»). В параллельном коде с $d_0=13$ в данном случае был выделен внешний код с $R_1=6/11$, $d_1=7$, а в коде с $d_0=15$ был выделен код с $R_1=5/9$, $d_1=9$. Кривые «Составляющий» на рисунках отражают вероятность ошибки на выходе выделенных кодов параллельной схемы. Пунктирными линиями без маркеров показаны вероятности ошибки оптимального декодирования кодов с $d=7$, 9, 13 и 15. Для сравнения на рис. 4 также представлены характеристики декодируемых с помощью МПД обычных СОК с аналогичными d и R (кривые «Обычный»). Отметим, что при получении данных характеристик использовались достаточно короткие коды, длиной до нескольких тысяч битов, и 10...20 итераций декодирования. При некотором увеличении объема памяти декодера и числа выполняемых итераций с помощью параллельного каскадирования уже удалось получить характеристики, представленные на рис. 4 кривой «Длинный». Как следует из анализа представленных графиков, применение параллельного кодирования позволяет МПД работать чуть больше чем в 1,5 дБ от пропускной способности канала.

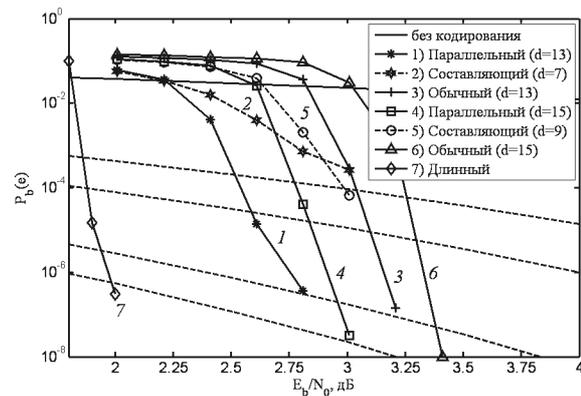


Рис. 4. Результаты моделирования параллельного кода на базе МПД в канале с АБГШ

Сложность МПД при параллельном кодировании (в смысле количества выполняемых операций) оказывается даже меньше сложности обычного МПД, поскольку в данном случае на первых итерациях декодирования некоторые элементы синдромного регистра просто не участвуют в процессе вычисления суммы на пороговом элементе.

Каскадные коды, декодируемые с использованием многопороговых декодеров

Высокие характеристики МПД способствуют его широкому применению в составе различных кодовых конструкций, поскольку эффективность последних непосредственно связана с эффективностью их составных элементов.

Особое место среди кодовых схем на базе МПД занимает его каскадирование с кодами контроля четности (ККЧ), использование которых позволяет существенно повысить эффективность применения кодирования. Особенность данной схемы заключается в том, что такое каскадирование практически не требует дополнительных затрат на оборудование (например, в схему кодирования требуется добавить лишь один сумматор по модулю 2), тогда как использование в каскадном коде, например, кодов Рида-Соломона несравненно труднее. Подробно принципы каскадирования МПД с ККЧ рассмотрены в [4, 5].

Эффективность работы каскадных схем, состоящих из СОК с $d=7$ и 9 и ККЧ с длиной 50, для канала с АБГШ отражена на рис. 5. Как видно из рисунков, каскадный код на основе МПД оказывается значительно лучше некаскадного. При этом использование простейшего ККЧ совместно с СОК позволяет получить дополнительный энергетический выигрыш около 1...1,5 дБ при вероятности ошибки на выходе декодера $P_b(e)=10^{-5}$. Также отметим, что каскадный код, состоящий из кода Рида-Соломона (255, 223, 33) и сверточного кода с кодовой скоростью 1/2 и длиной кодового ограничения $K=7$, декодируемый с помощью оптимального алгоритма Витерби, даже при меньшей общей кодовой скорости ($R=0,437$) уступает каскадной схеме на базе МПД при $P_b(e) > 10^{-6}$. Заметим, что при использовании каскадирования с ККЧ совместно с ранее рассмотренным параллельным кодированием, возможно получение значительного дополнительного энергетического выигрыша. Пример характеристик од-

ной из таких схем показан на рис. 5 кривой «МПДп+ККЧ». Использование малоизбыточного кода во внешнем каскаде данной схемы позволит получить сколь угодно малую вероятность ошибки при потерях в энергетике около 0,1...0,2 дБ.

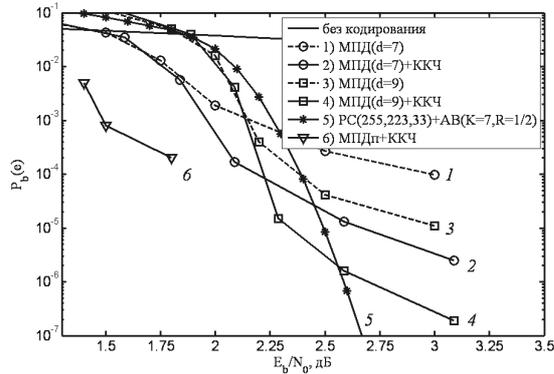


Рис. 5. Результаты моделирования каскадного кода на базе МПД в канале с АБГШ

Сложность декодирования рассмотренной каскадной схемы по сравнению со сложностью обычного МПД увеличивается на сложность реализации декодера ККЧ, примерно равную всего двум операциям на информационный бит.

Многopороговые алгоритмы в каналах с неравномерной энергетикой

Рассмотрим типичный пример очень простой модификации алгоритма МПД, которая с учётом конкретных свойств метода очень точно использует его свойства и фактически вообще без усложнения самого метода решающим образом улучшает его характеристики энергетической эффективности в полном соответствии с принципом «от простого к эффективному». Происходит просто согласование системы сигналов со свойствами кода и декодера.

Рассмотрим двухканальную схему передачи цифровых данных по спутниковым, космическим или иным каналам связи с достаточно большим уровнем гауссовского шума. Выберем для некоторого отношения сигнал/шум, первоначально одинакового для каждого из двух рассматриваемых каналов связи, такое перераспределение общей суммарной энергии, чтобы было обеспечено максимально возможное независимое последующее декодирование принятых информационных символов на основе многopорогового декодирования двоичных блоковых или свёрточных кодов. Иначе говоря, критерием наилучшего перераспределения энергии между каналами следует выбрать минимальный уровень проявления эффекта размножения ошибок при мажоритарном декодировании. В теории МПД эти вопросы достаточно полно проработаны [1, 4]. Снижение данного эффекта позволяет значительно улучшить сходимость решений МПД к оптимальным, что и создаёт условия для более эффективной работы алгоритмов МПД при больших уровнях шума.

При формировании такой достаточно простой новой сигнально-кодовой конструкции можно рассматривать различные способы балансировки энергетике. Например, каналы могут быть организованы так, что по одному из них передаются информационные символы кода, а по другому – проверочные. В этом случае анализ размножения ошибок упрощается в наибольшей степени, что позво-

ляет достаточно быстро и полно рассматривать применимость максимального числа кодов и соответствующих им МПД алгоритмов в подобных схемах кодирования. Такие модели получили название каналов с неравномерной энергетикой (НЭК) [4, 11]. Они могут быть просто реализованы в обычных трактах передачи цифровых данных.

Как показал детальный анализ ряда кодов и некоторых модификаций МПД алгоритмов для НЭК каналов с различными параметрами, перемещение границы области эффективной работы МПД в сторону более высокого уровня шума канала в диапазоне кодовых скоростей R от $1/4$ до $3/4$ может составлять до 1 дБ, что очень важно, поскольку уже исходная эффективность МПД в каналах обычного типа также оказывается весьма высокой. При этом отношение энергетики каналов должно находиться в диапазоне от 1,3 до 3,2.

Необходимость работы при более высоких уровнях шума требует увеличения числа итераций декодирования в МПД, но такое увеличение обычно оказывается не более чем двукратное, что сохраняет небольшую сложность реализации МПД как в программном, так и в аппаратном вариантах.

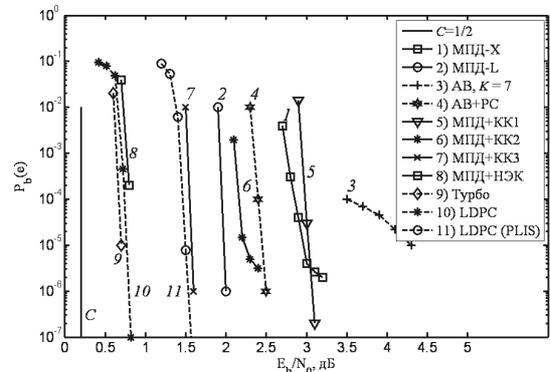


Рис. 6. Характеристики МПД, декодера Витерби, турбо и низкоплотностных кодов в гауссовском канале при $R=1/2$

Новые полученные результаты в этой области иллюстрируются графиками на рис. 6, на которых представлены возможности предложенных алгоритмов и уже известных методов. График «1) МПД-X» соответствует эффективности МПД декодера на ПЛИС Xilinx, кривые «5) МПД+КК1», «6) МПД+КК2» и «7) МПД+КК3» даны для применения МПД в простейших каскадных схемах с кодом контроля четности. Все они детально обсуждались в [9]. На рис. 6 также представлены кривые эффективности для алгоритма Витерби со стандартным кодом длины $K=7$ (кривая «3) АВ, $K=7$ ») и для каскадной схемы декодера Витерби с кодом Рида-Соломона (кривая «4) АВ+РС»), а также для турбо кода (кривая «9) Турбо») [12] и низкоплотностного кода, рекомендованного в стандарте DVB-S2 [13] (кривая «10) LDPC»). Отметим, что при реализации высокоскоростного декодера низкоплотностных кодов на ПЛИС потери в эффективности составляют более 0,5 дБ (кривая «11) LDPC (PLIS)»). Вертикаль $C=1/2$ определяет пропускную способность гауссовского канала, к которой стремятся разработчики при улучшении характеристик декодирования при кодовой скорости $R=1/2$. «МПД-L» – длинный код и МПД декодер с $L=40$ итерациями, реализованный в ИКИ РАН на ПЛИС Altera. Новый результат для МПД и канала НЭК – пункт «8) МПД+НЭК» – соответствует возможности декодирования кода при задержке принятия решения не бо-

лее 400000 битов, при котором сохраняется хорошо известная достаточно большая скорость работы МПД как в программном, так и особенно в аппаратном варианте.

С учетом уже достигнутой близости области эффективной работы МПД к пропускной способности канала связи, можно считать, что МПД имеет хорошие перспективы по дальнейшему приближению его характеристик к границе Шеннона. При этом значительное преимущество МПД перед другими алгоритмами по числу операций, составляющее один-два десятичных порядка для различных сочетаний параметров кодирования, дает основание полагать, что МПД можно активно использовать при разработках современной аппаратуры передачи цифровых данных для космических и спутниковых каналов связи.

Применение многопороговых декодеров с многопозиционными системами сигналов

В предыдущих разделах представлены результаты исследования МПД в гауссовских каналах передачи данных с двоичной фазовой модуляцией (ФМ2). Вместе с тем на каналы передачи данных часто накладываются значительные ограничения по занимаемой полосе частот и с каждым годом эти ограничения становятся все жестче. Одним из способов уменьшения занимаемой полосы частот является использование многопозиционных сигналов, для формирования которых обычно применяется многопозиционная фазовая (ФММ) или квадратурная амплитудная модуляция (КАММ). При переходе к таким сигналам все подходы к применению МПД остаются аналогичными двумерному случаю, что позволяет одновременно получить значительный энергетический выигрыш кодирования и существенно сэкономить полосу частот передаваемого сигнала.

Рассмотрим результаты моделирования МПД и других методов коррекции ошибок в канале с многопозиционными системами сигналов при использовании мягких решений демодулятора, когда демодулятор оценивает надежность выносимых им решений.

На рис. 7 кривыми «МПД, КАМ16», «МПД, КАМ32» и «МПД, КАМ64» представлены экспериментальные графики зависимости вероятности ошибки на бит $P_b(e)$ на выходе МПД от отношения сигнал/шум E_b/N_0 в канале с АБГШ и квадратурной амплитудной модуляцией при использовании 16, 32 и 64 символьных созвездий. При декодировании выполнялось от 10 до 20 итераций декодирования блочного СОК с кодовой скоростью $R=1/2$, кодовым расстоянием $d=11$ и длиной n порядка 10000. При этом использовался ранее рассмотренный код с параллельным кодированием.

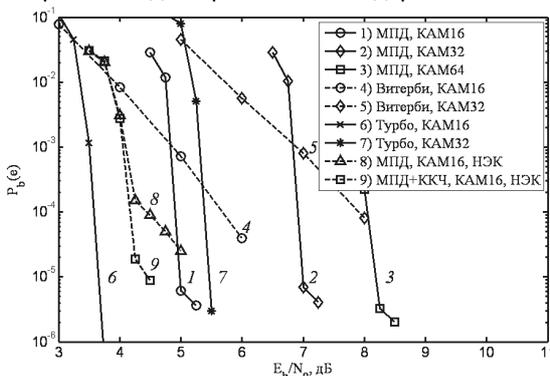


Рис. 7. Эффективность МПД в каналах с КАММ, мягкий модем

На этом же рисунке кривыми «АВ, КАМ16» и «АВ, КАМ32» представлены характеристики декодера Витерби для кода с длиной регистра $K=7$ при использовании КАМ16 и КАМ32 соответственно. Как видно, декодер Витерби в данных условиях при $P_b(e)=10^{-4}$ проигрывает МПД более 1 дБ. Кривыми «Турбо, КАМ16» и «Турбо, КАМ32» на рис. 7 показаны характеристики очень мощного турбо кода [5] с кодовой скоростью $R=1/2$, который образован путем параллельного каскадирования двух рекурсивных систематических сверточных кодов с конструктивной длиной $K=4$. В данном турбо коде применялся перемежитель типа S-random длиной $L=5000$ (общая длина турбо кода составляет $n=10000$). При декодировании турбо кода выполнялось 8 итераций, на каждой из которых для декодирования составляющих кодов применялся Max-Log-MAP алгоритм. Из сравнения характеристик декодера турбо кода и МПД видно, что эффективность последнего оказывается хуже примерно на 1,5 дБ, но МПД при этом почти на два порядка проще для практической реализации, чем данный турбо декодер.

На рис. 8 представлены характеристики МПД для того же блочного кода, что и на рис. 7, в канале с многопозиционной фазовой модуляцией (ФММ). Видно, что соотношения между характеристиками многопорогового декодера, декодера Витерби и декодера турбо кода сохраняются и при данном виде модуляции.

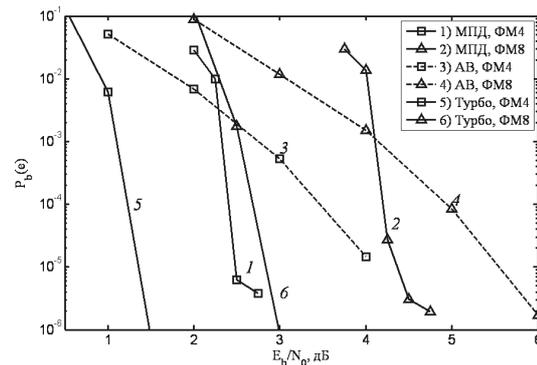
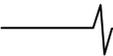


Рис. 8. Эффективность МПД в каналах с ФММ, мягкий модем

Рассмотрим один из возможных способов приближения области эффективной работы МПД к пропускной способности канала в каналах с многопозиционной модуляцией. Известно, что в таких каналах отдельные биты сигнального созвездия защищены по-разному. Например, для системы КАМ16 при использовании кода Грэя первый и третий биты могут быть защищены больше, а второй и четвертый – меньше. При этом вероятность ошибки в первом и третьем битах оказывается в два раза меньше вероятности ошибки во втором и четвертом битах. Похожая ситуация наблюдается и при использовании других типов модуляции. Таким образом, канал с многопозиционными системами сигналов можно рассматривать в качестве канала с неравномерной энергетикой. В результате для приближения области эффективной работы МПД к пропускной способности канала можно информационные символы сообщения располагать в более надежных битах созвездия, а проверочные – в менее надежных.



На рис. 7 кривой «МПД, КАМ16, НЭК» показаны характеристики МПД блокового СОК с кодовой скоростью $R=1/2$, кодовым расстоянием $d=11$ и длиной n порядка 10000 в канале с АБГШ при использовании КАМ16 и демодулятора, формирующего мягкие решения, при расположении информационных битов в более надежных позициях символа, а проверочных – в менее надежных. Заметим, что область эффективной работы МПД приблизилась к пропускной способности канала примерно на 0,5 дБ, но при этом область насыщения вероятности ошибки оказалась несколько выше. Вместе с тем для уменьшения вероятности ошибки в области эффективной работы, как показано в [26], возможно использование совместно с МПД простейшего кода с контролем четности (ККЧ). Характеристики каскадной схемы, состоящей из МПД и ККЧ длины 50, в случае расположения информационных битов в более надежных позициях символа для тех же условий показаны на рис. 7 кривой «МПД+ККЧ, КАМ16, НЭК». Из рисунка видно, что применение предложенного подхода позволило приблизить эффективность МПД к пропускной способности канала примерно на 0,7 дБ. В результате преимущество гораздо более сложного ранее рассмотренного декодера турбо кода (кривая «Турбо, КАМ16») перед МПД при вероятности ошибки порядка 10^{-4} оказалось даже меньшим, чем 0,5 дБ. Заметим, что похожего результата можно добиться и при использовании других многопозиционных систем сигналов.

Недвоичные многопороговые декодеры

Во многих случаях в реальных системах удобно работать с данными, имеющими байтовую структуру. Отметим, что кроме кодов Рида-Соломона (РС) в настоящее время вообще нет других сколько-нибудь эффективных и одновременно довольно простых методов декодирования недвоичных символьных данных, если выбранный код достаточно короткий.

В данном разделе рассмотрим обобщение МПД для q -ичных симметричных каналов (q СК). По аналогии с соответствующими декодерами для двоичных данных эти алгоритмы названы q -ичными многопороговыми декодерами (q МПД) [4, 5, 18]. Они также обладают свойством сходимости к решению ОД при сохранении линейной от длины кода сложности реализации, которая свойственна только пороговым процедурам. Ценность q МПД заключается в том, что в случае больших значений основания кода q , $q > 10$, практически невозможно создать эффективные ОД, поскольку их сложность в большинстве случаев будет пропорциональна q^k , где k – длина информационной части кода, выраженная числом q -ичных символов.

Изложим основные принципы работы q МПД. Пусть задан q СК ($q > 2$) с вероятностью ошибки $P_0 > 0$ такой, что при передаче любой исходный символ кода переходит в один из оставшихся $q-1$ символов случайно, независимо и равновероятно. Для этого канала решением ОД будет такое, может быть единственное, кодовое слово из q^k возможных, которое отличается от принятого из канала сообщения в минимальном числе кодовых символов.

Рассмотрим линейный недвоичный код, проверочная матрица которого имеет такой же вид, как и в двоичном

случае, т.е. состоит только из нулей и единиц, за исключением того, что вместо 1 в единичной подматрице будут -1 . Пусть эта матрица соответствует блоковому СОК. Поскольку проверочная (а значит, и порождающая) матрица кода содержит только нули и единицы, то для выполнения операций кодирования и декодирования принятого сообщения достаточно использовать операции сложения и вычитания по модулю q . Таким образом, для кодирования и декодирования не требуется наличие недвоичного поля, а достаточно создать только группу чисел, что существенно упрощает все процедуры кодирования и декодирования. Пусть декодер типа q МПД (рис. 9) устроен так, что после вычисления обычным образом вектора синдрома s принятого сообщения начинается процедура декодирования, заключающаяся в том, что для очередного контролируемого недвоичным пороговым элементом информационного символа кода u_j происходит подсчет количества и определение значений двух относящихся к нему и наиболее часто встречающихся проверок кода, например, b_1 и b_2 , причем b_1 встречается m_1 раз, $b_2 - m_2$ раз ($m_1 > m_2$), а остальные значения проверок для декодируемого символа u_j встречаются не более m_2 раз. Тогда на выходе q -ичного порогового элемента (q ПЭ) будет значение b_1 и q МПД при каждом изменении символа u_j будет переходить к все более правдоподобным решениям, так как при этом число различий между кодовым словом, соответствующим текущему содержимому информационного регистра, и принятой из канала последовательностью будет уменьшаться. Если окажется, что два наиболее часто встречающихся значения проверок таковы, что $m_1 = m_2$, то выход q ПЭ устанавливается равным нулю, т.е. символ u_j не изменяется, и делается попытка декодирования любого другого информационного символа кода. Наиболее существенным обстоятельством, повышающим корректирующие возможности описанного q МПД, является возможность принимать безошибочные решения при больших значениях q всего при двух правильных проверках относительно u_j из $d-1$ возможных, что происходит в том случае, когда все неправильные проверки относительно декодируемого символа u_j имеют различные значения.



Рис. 9. q МПД блокового СОК с $R=1/2$, $d=5$ и $n=26$. Верхние входы ячеек 0, 1, 4 и 6 регистра синдрома вычитаются по модулю q из содержимого соответствующих ячеек

Далее рассмотрены результаты моделирования работы q МПД в недвоичном симметричном канале q СК, характеристики каскадирования кодов для q МПД с кодами контроля

по mod q , а также возможности обычных декодеров кодов РС. Отметим, что в отличие от кодов Рида-Соломона для q МПД никаких ограничений по длине кода вообще нет, поскольку длина кода и величина его основания в недвоичных кодах с мажоритарным декодированием совершенно не зависят друг от друга. Объем моделирования в нижних точках графиков для q МПД составлял от $5 \cdot 10^{10}$ до $2 \cdot 10^{12}$ битов, что свидетельствует о крайней простоте метода.

На рис. 10 представлены зависимости вероятности ошибки на символ $P_s(e)$ от вероятности символьной ошибки P_0 в q СК для кодов Рида-Соломона, которые обозначены как RSn , где n – длина кода, выраженная в числе символов. Отметим, что реально сейчас используются только коды РС длины 255 или менее. Коды длины 4095 символов и, тем более, длины $n=65535$ (каждый символ – размером 16 битов), в обозримом будущем реализации не подлежат.

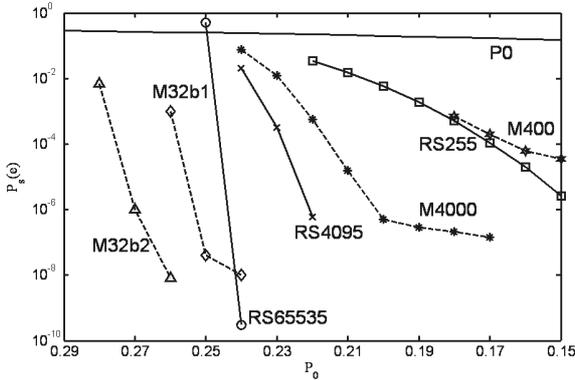


Рис. 10. Характеристики недвоичных блочных МПД и кодов РС с $R=1/2$

Здесь же пунктирными линиями показаны возможности кодов с мажоритарным декодированием при $R=1/2$ для случая $q=256$ (символ – один байт) для разных длин кодов, поскольку при любом $q>2$, как и в двоичном случае, для q МПД можно строить сколь угодно длинные коды с различными значениями кодового расстояния d и кодовой скорости R . Эти коды отмечены как M400 и M4000 с числами, обозначающими длины кодов, выраженные числом символов. Далее, обозначение M32b1 соответствует q МПД для кода длины 32000 символов. Как видно из рис. 10, возможности q МПД во всех случаях сопоставимы или лучше, чем у весьма сложных декодеров кодов РС. Более того, очень простой для реализации последний декодер для кода длины 32000 оказывается способным обеспечить простейшими мажоритарными методами помехоустойчивость, принципиально недостижимую даже для кода РС длины 65535 двухбайтовых символов, декодер для которого не будет создан никогда. А если перейти к двухбайтовым недвоичным кодам с мажоритарным декодированием, то его характеристики для длины кода 32000 символов будут соответствовать графику M32b2, ещё более показательному по уровню помехоустойчивости в области, где коды РС уже не работают. При этом q МПД для двухбайтовых символов практически ни в чём не сложнее однобайтового, так как даже обычные микропроцессоры одинаково просто и быстро работают и с однобайтовыми символами, и с 2-х и даже иногда с 8-байтовыми словами.

На рис. 11 показаны возможности q МПД и кодов РС в q СК при кодовой скорости $R=7/8$. Сплошными линиями с указанием длин блоков представлены вероятности ошибки на символ для кодов РС. Пунктирными линиями представ-

лены коды с q МПД декодированием и длиной 48000 символов: b1 – байтовых (символ – 8 битов) и b2 – двухбайтовых (символ – 16 битов). Подчеркнём, что возможность создания кодов РС длины 4095 при $R=7/8$ в ближайшее время останется очень проблематичной, в то время как даже для кодов длины 48000 байтов рассматриваемые недвоичные мажоритарные декодеры остаются очень простыми.

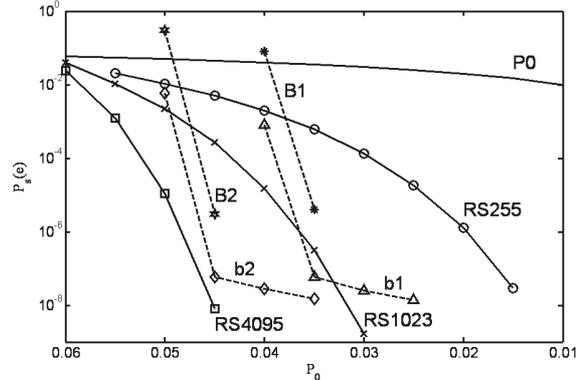


Рис. 11. Характеристики кодов РС и q МПД с $R=7/8$

На рис. 12 для кодов с малой избыточностью при $R=0,95$ представлены аналогичные характеристики для q МПД и кодов РС. Для сопоставления на рис. 12 приведён также график для кода РС с $n=255$ и $R=7/8$ с рис. 11. Пунктирами b1 и b2 указаны возможности двух q МПД для кодов длины $n \sim 80000$ и символов размером 1 и 2 байта.

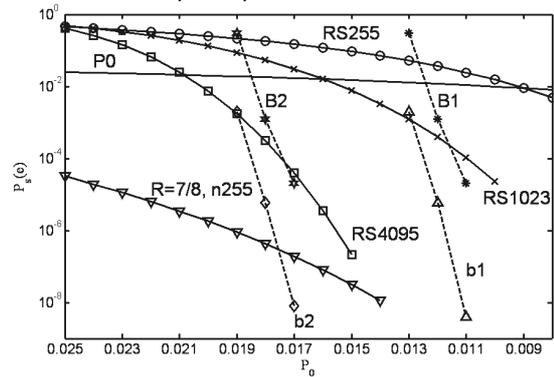
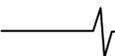


Рис. 12. Характеристики q МПД и кодов РС при $R=19/20$

Из сопоставления кодов РС длины $n=255$ при $R=7/8$ и $R=19/20$ видно, насколько последний менее эффективен первого и насколько труднее обеспечивать хорошую эффективность при уменьшении избыточности. Тем не менее характеристики малоизбыточных кодов с мажоритарным декодированием на основе q МПД оказываются весьма высокими и могут существенно поднять уровень помехоустойчивости, если выбранные коды имеют достаточно большие длины. Более того, q МПД при $R=0,95$, как следует из рис. 12, эффективнее кода РС с $R=7/8$, у которого избыточность в 2,25 раз больше. Этим снимаются все вопросы о применении любых новых самых сложных методов декодирования для кодов РС: они малоэффективны по сравнению с q МПД. В частности, например, алгоритмы Судана и прочие серьёзные усложнения декодеров для кодов РС приводят к росту их асимптотической от длины кода сложности с уровня n^2 до величины порядка n^3 . При этом в самом лучшем случае для этих алгоритмов вес ошибок канала, при котором возможно получение правильного решения, возрастает при использовании самого сложного алгоритма этого класса менее чем



на 4% для $R=7/8$ и менее чем на 1% для $R=0,95$, т.е. для тех значений скоростей, которые представлены на графиках. Стоит повторить, что рис. 12 показывает превосходство q МПД с $R=0,95$ перед кодом РС длины 255 даже с $R=7/8$, т.е. при его гарантированной корректирующей способности на 125% большей, чем при $R=0,95$. Поэтому рост корректирующей способности алгоритма Судана на 4% и 1% ничего не даст даже для длинных кодов РС, сопоставимых по длине с кодами, выбранными для q МПД.

Полезно отметить, что везде для графиков b1 и b2 нижние экспериментальные точки соответствуют оптимальному (совпадающему с переборным) декодированию с одиночными ошибками при частоте их появления, близкой к оценкам, представленным в [4 – 6]. Другие экспериментальные данные для q МПД можно найти в [4 – 6, 17].

Характеристики каскадирования недвоичных многопороговых декодеров

Согласно общим принципам теории кодирования, переход к каскадным принципам кодирования ещё более улучшает характеристики q МПД. При этом сложность декодирования увеличится по сравнению с исходным алгоритмом очень незначительно.

Возможности каскадирования q МПД с использованием кодов контроля по mod q также представлены для высокоскоростных кодов на рис. 11 и рис. 12. Вероятности ошибки на блок для каскадных кодов с внутренним СОК кодом при $R=7/8$ и внешнего кода длиной $L=190$ представлены кривой В1 для кода с $q=256$ (символ – 1 байт) и кривой В2 для кода с $q=65536$ (символ – 2 байта) на рис. 11.

Для каскадного кода с внутренним кодом при $R=0,95$ и внешнего кода с $L=190$ вероятность ошибки на блок на рис. 12 даётся кривыми В1 для $q=256$ и В2 для $q=65536=2^{16}$. Во всех случаях на графиках В1 и В2 нижние точки соответствуют величине N_B^{-1} , где N_B – число декодированных блоков, поскольку в экспериментах не было ни одного случая неправильного декодирования каскадного кода в этих точках. Как и ожидалось, применение каскадирования на много порядков снижает вероятность ошибки на блок по сравнению с обычным q МПД почти без сколько-нибудь заметного роста избыточности каскадного кода. При этом увеличение объёма вычислений в каскадном коде составляет менее 20% по сравнению с исходным алгоритмом q МПД.

Дальнейшее значительное улучшение эффективности декодирования методами q МПД возможно при переходе к свёрточным кодам, методам последовательного и параллельного каскадного кодирования, применению кодов с выделенными ветвями и другим мерам, некоторая часть из которых описана в [3 – 5].

Сложность реализации недвоичных многопороговых декодеров

Рассмотрение недвоичного (символьного) q МПД показывает, что линейная сложность декодирования сохраняется. При программной реализации такого алгоритма подпрограмма работы его порогового элемента, которая и составляет практически весь декодер, занимает менее десятка коротких строк на языке С++ и обеспечивает обработку одновременно такого количества байтов принятого или хранимого сообщения, которое допускает архитектура используемого процессора. Это ещё более увеличивает высочайшую

производительность декодеров класса МПД. Демопрограмма для символьного q МПД также выложена для всеобщего использования на веб-сайте www.mtdbest.iki.rssi.ru. Она показывает, что при крайне малой избыточности ($R=0,95$) и вероятности ошибки на символ в недвоичном канале более 10^{-2} достигается практически оптимальное декодирование очень длинных кодов со скоростью более 10 Мбит/с, а иногда, на относительно быстрых ПК общего назначения, – до 30 Мбит/с. При этом демопрограмма выполняет полную имитацию работы всего тракта передачи данных: формирование информационного потока, кодирование его в выбранных форматах, внесение шума, значительного по уровню искажений, и затем работу обсуждаемого алгоритма декодирования. Так что реальная производительность программного q МПД декодера может считаться большей примерно в два – три раза или даже более. Демопрограмма классического декодера кода РС также представлена в обзорной странице веб-сайта www.mtdbest.iki.rssi.ru.

Сложность декодеров для длинных недвоичных кодов естественно сравнивать с вычислительными затратами декодеров кодов РС, которые растут как n^2 . Более простые декодеры для недвоичных кодов, кроме МПД, неизвестны. Разнообразные методы повышения эффективности кодов РС, в том числе все вариации алгоритма Судана, приводят к сложности порядка n^3 . Для кодов длины 30000 символов это приводит к разнице в порядке сложности около $n^2=30000^2\approx 10^9$, т.е. миллиард раз. Однако при этом, как оказывается, улучшение помехоустойчивости благодаря более сложному декодированию, например, при $R=7/8$ весьма незначительно. Это определяется тем, что улучшение соответствует как бы увеличению минимального кодового расстояния кодов РС примерно в 1,04 раза. Конечно, реально этого не происходит, но доля случаев, когда это не так, и декодер ошибается при меньшем весе ошибок, невелика. Для этого случая оценки показывают, что код РС длины 4095 символов при $R=7/8$ обеспечит вероятность ошибки на символ $\sim 10^{-6}$ (при вероятности ошибки канала около $p_0=0,050$), алгоритм Судана будет столь же эффективен при $p_0=0,052$, тогда как q МПД обеспечивает эту же вероятность ошибки на блок при $p_0=0,047$ для кода длины 10^5 символов. При использовании с символьным МПД каскадирования с ККЧ характеристики МПД будут ещё на несколько порядков лучше почти без роста сложности декодера. Таким образом, характеристики всех трёх методов близки, но даже обычный декодер РС будет на несколько порядков, видимо в 10^4 раз, более сложен по числу операций, чем q МПД. Алгоритм Судана, в случае его реализации для этого же кода РС, потребует ещё примерно на три порядка большего числа операций, чем обычный декодер РС, что вообще снимает с обсуждения вопросы его практического использования. Отметим, что для $R=1/2$ и $R=0,95$ приведённые графики показывают абсолютное преимущество недвоичных МПД перед кодами РС любой разумной длины. Об этом же свидетельствуют и демопрограммы разных методов декодирования недвоичных кодов, представленные на веб-сайте [3].

Выводы

Применение МПД в спутниковых и прочих дорогих каналах позволяет реализовать произвольно высокие скорости обработки и существенно повысить к.п.д. их использования. Крайне простое устройство МПД по сравнению с прочими

сопоставимыми с ним по эффективности методами делает их предпочтительными в быстрых широкополосных каналах при аппаратной реализации. В достаточно медленных каналах связи даже программные реализации МПД очень эффективны и требуют написания лишь нескольких десятков команд программного кода для порогового элемента. Простые методы согласования кода и сигналов ещё более повышают возможности МПД и делают их реализацию особенно простой. Совершенно незначительная разница в эффективности МПД по сравнению с некоторыми особенно сложными декодерами других типов, как показывает динамика улучшения характеристик МПД в последние годы, будет преодолена, видимо, в самое ближайшее время.

Кроме естественных областей применения простых высокоэффективных методов кодирования в сетях связи следует отметить хорошие возможности применения q МПД для кодирования информации на дисках и других носителях больших объёмов информации, в сверхбольших базах аудио- и видеоданных с намного более высоким уровнем достоверности, чем это было доступно до недавнего времени, а также при обновлении, восстановлении, переносе и использовании хранимых там данных. При этом легко обеспечить оперативный постоянный контроль за качеством хранимой информации, а также своевременную корректировку и перенос данных вследствие старения и возникающих дефектов носителя. Все виды динамического контроля уровня достоверности, управления памятью и её резервированием самым очевидным образом реализуются на основе того, что q МПД алгоритмы непрерывно проводят различные простые, но очень информативные и удобные мажоритарные оценки надёжности записанных данных. Этим и определяются все дополнительные преимущества многопороговых алгоритмов в их приложениях по обеспечению принципиально нового, на много десятичных порядков более высокого уровня целостности и достоверности хранения информации в сверхбольших массивах данных практически любой структуры.

Таким образом, принципиально новый уровень помехоустойчивости, достигаемый с помощью алгоритмов МПД разных типов, позволяет решать задачи обеспечения высокой надёжности передачи и хранения данных без какой-либо дополнительной доработки этих алгоритмов или всего лишь при незначительной их адаптации к возможным дополнительным требованиям, возникающим в крупномасштабных цифровых системах.

Работа выполнена при финансовой поддержке РФФИ (грант №08-07-00078а)

Разработки алгоритмов МПД поддерживались Научным советом по комплексной проблеме «Кибернетика» АН СССР, НИИ «Квант» и ИКИ РАН. Дополнительная информация об МПД разных классов расположена на специализированном тематическом двуязычном веб-сайте ИКИ РАН www.mtdbest.iki.rssi.ru.

Литература

- Самойленко С.И., Давыдов А.А., Золотарёв В.В., Третьякова Е.И. Вычислительные сети. М.: Наука, 1981, 277 с.
- Золотарёв В.В., Овечкин Г.В. Эффективные алгоритмы помехоустойчивого кодирования для цифровых систем связи // Электросвязь. 2003, №9. С.34-37.
- Многопороговые декодеры. Веб-сайт ИКИ РАН www.mtdbest.iki.rssi.ru.
- Золотарёв В.В. Теория и алгоритмы многопорогового декодирования // Под научной редакцией члена-корреспондента РАН Ю.Б. Зубарева. М.: «Радио и связь», «Горячая линия – Телеком», 2006. 276 с.
- Золотарёв В.В., Овечкин Г.В. Помехоустойчивое кодирование. Методы и алгоритмы. Справочник. М.: «Горячая линия – Телеком», 2004. 124 с.
- Золотарёв В.В. Многопороговое декодирование для информационных потоков с байтовой структурой // Мобильные системы, 2006. №3. С. 25-27.
- Зубарев Ю.Б., Золотарёв В.В. Многопороговые декодеры: перспективы аппаратной реализации. // В сб.: «7-я Международная конференция и выставка «Цифровая обработка сигналов и её применение», М.: 2005. Вып. VII-1. С. 68-69.
- Зубарев Ю.Б., Золотарёв В.В., Овечкин Г.В., Строков В.В. Многопороговые декодеры для высокоскоростных спутниковых каналов связи: новые перспективы // Электросвязь, 2005. №2. С. 10-12.
- Зубарев Ю.Б., Золотарёв В.В., Овечкин Г.В., Дмитриева Т.А. Многопороговые алгоритмы для спутниковых сетей с оптимальными характеристиками // Электросвязь, 2006. №10. С. 9-11.
- Berrou C., Glavieux A., Thitimajshima P. Near Shannon Limit Error-Correcting Coding and Decoding: Turbo-Codes // Proceeding of ICC'93, Geneva, Switzerland, 1993. pp. 1064-1070.
- European Telecommunications Standards Institute. Digital video broadcasting (DVB) second generation framing structure, channel coding and modulation systems for broadcasting, interactive services, news gathering and other broadband satellite applications. DRAFT EN 302 307 (v. 1.1.2 06.2006).
- Zolotarev V.V., Averin S.V., Chulkov I.V. Optimum Decoding Characteristics Achievement on the Basis of Multithreshold Algorithms // 9-th ISCTA'07, July, UK, Ambleside, 2007.
- Zolotarev V.V., Nazirov R.R., Chulkov I.V. The Quick Almost optimal multithreshold decoders for Noisy Gaussian Channels // RCGSGO International Conference ESA in Moscow, Russia, June, 2007.
- Зубарев Ю.Б., Золотарёв В.В. Достижение характеристик оптимального декодирования на основе многопороговых алгоритмов // 9-я Международная конференция и выставка «Цифровая обработка сигналов и её применение», Доклады-1, Пленарный доклад, М.: 2007. С. 12–15.
- Zolotarev V.V., Averin S.V. Non-Binary Multithreshold Decoders with Almost Optimal Performance // 9-th ISCTA'07, July, UK, Ambleside, 2007.
- Золотарёв В.В. Обобщение алгоритма МПД на не двоичные коды // «Мобильные системы», №3, 2007, С. 39–42.
- Гринченко Н.Н., Овечкин Г.В. Помехоустойчивое кодирование для высокоскоростных спутниковых каналов связи // Цифровая обработка сигналов, 2006. №4, С. 29–33.
- Гринченко Н.Н., Золотарёв В.В., Овечкин Г.В., Овечкин П.В. Многопороговое декодирование в каналах с многопозиционной модуляцией // Вестник РГТУ, 2006 г. Вып. 19, С. 179–182.
- Гринченко Н.Н., Овечкин Г.В. Помехоустойчивое кодирование для цифровых систем связи // Известия ТРТУ, №15(70), Таганрог: Изд-во ТРТУ, 2006. С. 5–10.
- Золотарёв В.В., Овечкин Г.В. Многопороговые декодеры для каналов с предельно высоким уровнем шума // Телекоммуникации. М.: 2005. №9. С. 29–34.
- Золотарёв В.В., Овечкин Г.В. Использование многопорогового декодера в каскадных схемах // Вестник РГТА. 2003. Вып. 11. С. 112–115.

МНОГОСКОРОСТНАЯ ОБРАБОТКА СИГНАЛОВ: РЕТРОСПЕКТИВА И СОВРЕМЕННОЕ СОСТОЯНИЕ (ЧАСТЬ 1)

Витязев В.В., Витязев С.В., Зайцев А.А.

Введение

Многоскоростная обработка сигналов (multirate processing) предполагает, что в процессе преобразования цифровых сигналов возможно изменение частоты дискретизации в сторону уменьшения или увеличения и, как следствие, требуемой скорости обработки. Это приводит к более эффективной обработке сигналов, так как открывается возможность значительного уменьшения требуемой вычислительной производительности проектируемой цифровой системы.

В последние годы в области многоскоростной обработки сигналов достигнуты громадные успехи. Многоскоростная фильтрация и особенности ее применения стали предметом исследований многочисленных научных работ по цифровой обработке сигналов (ЦОС). Появились десятки монографий и учебных пособий, так или иначе связанных с научными и практическими достижениями в этой области [1–7]. Совершенно уникальные возможности дает использование многоскоростной обработки в системах адаптивной и нелинейной фильтрации, сжатия, анализа и восстановления речи, звука и изображений.

В предлагаемой вашему вниманию обзорной работе рассматривается эволюция теории и техники многоскоростной обработки сигналов в период с начала 70-х прошлого столетия до наших дней с позиции вклада, который внесли в эту область исследований работы российских ученых и специалистов. Но, прежде всего, - история вопроса!

Эволюция теории и техники многоскоростной обработки сигналов

Как показывает анализ работ, опубликованных по проблемам многоскоростной обработки сигналов за весь исторический период, в развитии данного направления исследований можно условно выделить три этапа. На первом этапе (1975 – 1985 гг.) были заложены теоретические основы многоскоростной фильтрации и обработки сигналов с применением децимации и интерполяции, эффектов прореживания по времени и по частоте [1, 2]. На втором этапе (1985 – 1995 гг.) получила развитие теория компенсации наложений и полного восстановления при синтезе банка фильтров с полной децимацией [3]. Была установлена связь с вейвлет-преобразованием [8-11]. Третий этап (1995 – 2005 гг. и

Рассматривается эволюция теории и техники многоскоростной обработки сигналов. Выделяются три взаимосвязанных этапа развития, в основе которых лежат методы цифровой частотной селекции сигналов с использованием эффектов прореживания по времени и по частоте, новые алгоритмы частотно-временной и адаптивной фильтрации, широкий спектр практического применения прежде всего в радиотехнических и телекоммуникационных системах. Изложение материала ведется с позиции вклада, который внесли в эту область исследований работы российских ученых и специалистов.

по настоящее время) – это этап широкого практического внедрения методов многоскоростной обработки сигналов и дальнейшего их развития и систематизации [6,7]. Важную роль на этом этапе играет создание эффективных инструментальных средств автоматизированного проектирования систем многоскоростной ЦОС, реализуемых на базе цифровых сигнальных процессорах (ЦСП) и программируемых логических интегральных схемах (ПЛИС).

Этап 1. Основы многоскоростной обработки сигналов

Исторически первыми работами по многоскоростной обработке сигналов являются исследования американских ученых Р. Шафера, Л. Рабинера [12], Г. Оуткина, Т. Паркса [13] по применению линейных цифровых фильтров к интерполяции сигналов. В то время как полиномиальная интерполяция недостающих данных оставалась классической проблемой численного анализа, новый подход вызвал повышенный интерес у специалистов в области ЦОС. Однако точкой отсчета следует считать 1974 год, когда появилась статья французских ученых М. Белланже, Ж. Догета, Г. Лепанола «Интерполяция, экстраполяция и уменьшение скорости вычислений в цифровых фильтрах» [14]. Год спустя выходит серия работ американских специалистов Л. Рабинера, Р. Крошье [15, 16], Р. Шивели [17], в которых рассматривается оптимизация многоскоростных структур узкополосных КИХ-фильтров по критерию минимизации общего числа умножений и ячеек памяти данных. Дальнейшие исследования этих авторов были продолжены в работах [18, 19]. В [20] впервые была предложена полифазная форма фильтра-дециматора и фильтра-интерполятора, позволяющая эффективно использовать простые математические модели односкоростных систем обработки сигналов для описания и программной реализации многоскоростных систем.

Начиная с 1975 года, проводятся активные исследования по многоскоростной фильтрации и в СССР [21-

35]. В работе [21] была представлена одноступенчатая структура узкополосного низкочастотного фильтра с применением вторичной дискретизации и последующей интерполяции, в которой фильтр-дециматор впервые был построен по параллельной форме. Это позволило для используемых приложений уменьшить в десятки раз не только объем вычислений, но и число регистров памяти данных. С тем чтобы расширить рассмотренный подход на более общий случай – синтез узкополосных фильтров с произвольной центральной частотой полосы пропускания и высокой прямоугольностью АЧХ, в [22] была предложена двухступенчатая реализация с использованием дополнительного формирующего фильтра. В последующие 10 лет идея многоскоростной обработки получила свое развитие применительно к построению набора цифровых полосовых фильтров на основе эффектов прореживания по времени [23-28] и по частоте [29-35]. Систематизированное изложение этих методов можно найти в монографии [2].

Одним из ярких событий в теории и применении многоскоростной обработки в эти годы было введение двухканального банка квадратурно-зеркальных фильтров (КЗФ) для компрессии речевого сигнала [36,37]. В соответствии с данным методом сигнал раскладывается на низкочастотную и высокочастотную субполосные составляющие с использованием банка фильтров анализа. Каждый субполосный сигнал децимируется с понижением частоты дискретизации в два раза и кодируется (квантуется). Проквантованные субполосные сигналы могут быть рекомбинированы с использованием банка фильтров синтеза для восстановления исходного сигнала. При этом возникает ошибка восстановления, которая обусловлена частично компрессией (субполосным квантованием) и частично погрешностями, возникающими на этапе реализации фильтров анализа и синтеза.

Ключевым источником ошибки восстановления являются наложения спектра цифрового сигнала при его децимации, поскольку фильтры анализа являются неидеальными в зоне их непрозрачности. Это требует принятия каких-то решений по компенсации наложений. Другим источником ошибки являются амплитудно-фазовые искажения, вносимые банком фильтров. Всё это предполагает, что банк фильтров полного восстановления должен обладать свойствами, которые устраняют все виды искажений. Поэтому дальнейшие усилия ученых и специалистов по многоскоростной обработке сигналов были направлены на разработку теории компенсации наложений и полного восстановления [38-40].

Этап 2. Банки фильтров с полным и квазиполным восстановлением

Задача синтеза систем полного восстановления была прежде всего решена для двухканального банка КЗФ с использованием полифазного представления структуры фильтров и условий параунитарности полифазной матрицы без потерь [36].

В последующих исследованиях [39-41] был выработан более общий подход к синтезу банка фильтров с максимальной децимацией и полным восстановлением, расширенный на M -полосные системы анализа-синтеза.

В случае равномерного разбиения исходного сигнала $x(nT)$ на субполосные составляющие ($v_i = M$ для всех

$i = \overline{1, M}$) Z -образ восстановленного сигнала принимает вид

$$\hat{X}(z) = \frac{1}{M} \sum_{k=0}^{M-1} \sum_{l=0}^{M-1} F_k(z) H_k(zW^l) X(zW^l). \quad (1)$$

Из (1) видно, что наряду с исходным сигналом восстановленный сигнал содержит $(M-1)$ составляющих наложения $X(zW^l), l > 0$. Поэтому главная задача на этапе синтеза – компенсация $(M-1)$ -кратных наложений.

Условие компенсации наложений в матричной форме принимает вид:

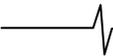
$$\begin{bmatrix} H_0(z) & H_1(z) & \cdots & H_{M-1}(z) \\ H_0(zW) & H_1(zW) & \cdots & H_{M-1}(zW) \\ \vdots & \vdots & \ddots & \vdots \\ H_0(zW^{M-1}) & H_1(zW^{M-1}) & \cdots & H_{M-1}(zW^{M-1}) \end{bmatrix} \begin{bmatrix} F_0(z) \\ F_1(z) \\ \vdots \\ F_{M-1}(z) \end{bmatrix} = \begin{bmatrix} MT(z) \\ 0 \\ \vdots \\ 0 \end{bmatrix} \quad (2)$$

Если матрица компонент наложений $[H_k(zW^l)]$ допускает обращение, то теоретически можно найти фильтры синтеза $F_k(z), k = 0, M-1$, компенсирующие наложения полностью. Однако при этом обратная матрица при обращении матрицы компонент может соответствовать БИХ-фильтрам и, как следствие, система становится неустойчивой или потребует использования фильтров очень высоких порядков. По этим причинам обращение матрицы компонент наложения не приемлемо, и гораздо более удачным является подход, опирающийся, как и для двухканального КЗФ, на полифазное представление фильтров анализа-синтеза, рассматриваемое в классе параунитарных матриц без потерь [42,46,47].

Во многих приложениях построение системы анализа-синтеза с теоретически полным восстановлением является необязательным. Достаточно синтезировать сигнал с наперед заданной точностью. Для построения банка фильтров с «квазиполным» восстановлением разными авторами предложено большое количество разнообразных методик. Многие из них имеют эвристический характер и предназначены для расчета ограниченного класса банка фильтров [48]. Другие подходы, претендующие на общность, можно разделить на класс банков фильтров, получаемых при помощи так называемой « H_∞ -оптимизации» [49,50], и на класс банков фильтров на основе косинусной модуляции фильтра-прототипа [51-57].

Особый класс банков фильтров, играющих для ряда приложений решающую роль, представляют банки фильтров с неравномерным разбиением по частоте [58 – 63]. Простым и надежным способом для их построения, с учетом требований полного восстановления, является использование древовидных структур на базе двух- и трехканальных КЗФ. Однако данный подход приемлем только для целочисленных значений коэффициентов децимации (интерполяции). Основы теории синтеза банка фильтров с рациональными коэффициентами децимации рассматриваются в [59, 60].

Многоскоростные банки фильтров, как показали параллельные исследования [8 – 11], очень тесно связаны с вейвлет-преобразованием. Вейвлет-преобразование выполняет декомпозицию сигнала по базисным функциям с различным «весом» последних



в различных частотно-временных участках. Каждая базисная функция является версией расширения и сдвига так называемой «материнской» (порождающей) вейвлет-функции. Вейвлеты – это очень полезное представление сигналов с переменными характеристиками.

Вейвлет-преобразование было изначально развито для аналоговых сигналов. Одним из наиболее важных достижений за прошедшее десятилетие является прежде всего установление математической зависимости между непрерывными по времени вейвлетами и банком дискретных фильтров, а также их связи с представлением сигналов, предполагающим множественное разрешение по частоте. Например, ортонормальное вейвлет-представление с базисными функциями конечной длительности относится к древовидному соединению двухканального ортонормального банка фильтров. Синтезируя один из фильтров анализа в двухканальном модуле таким образом, чтобы он имел достаточное число нулей в области высоких частот, можно добиться вейвлет-базиса, наиболее близкого к желаемому.

Этап 3. Оптимальное проектирование систем анализа-синтеза на сигнальных процессорах

В конце 90-х прошлого столетия и начале нового века продолжают дальнейшие интенсивные исследования и систематизация знаний в области общей теории многоскоростной обработки сигналов и синтеза банка фильтров [64]. Теория многоскоростного банка фильтров вызвала целое семейство теорем дискретизации, дополняющих классическую теорему Котельникова-Найквиста [65]. Вместе с тем основной акцент в эти годы смещается в сторону прикладных разработок, прежде всего применительно к системам телекоммуникаций и компрессии аудио- и видеосигналов [66,67]. Решающую роль играет разработка методики, алгоритмов и программ моделирования и оптимального автоматизированного проектирования систем анализа-синтеза сигналов на основе банка фильтров и адаптивной обработки [68-70].

В основе дальнейшего развития субполосного кодирования, эффективной реализации эквалайзеров и эквалайзеров лежит идея многоскоростной адаптивной фильтрации [71-73]. Суть идеи заключается в предварительном частотно-временном «расщеплении» входного сигнала с помощью банка фильтров и последующей адаптивной обработке каждой субполосной составляющей по заданному алгоритму. При необходимости на этапе синтеза выполняется дуальная процедура восстановления выходного сигнала, предварительно сжатого или «очищенного» от помех, шумов и искажений.

При проектировании систем адаптивной многоскоростной обработки сигналов встает целый комплекс проблем, связанных с поиском наиболее эффективных надежно работающих структур и оптимизацией параметров системы. Каким образом реализовать совместную адаптацию в субполосных каналах? Как будут влиять наложения при использовании децимации? Сколько субполосных каналов и с какими характеристиками обеспечат наивысшую производитель-

ность и качество работы всей системы адаптивной многоскоростной обработки сигналов? Все это и является предметом исследований в данной области на современном этапе.

Поиск эффективных структур системы анализа-синтеза с применением банка фильтров, требующий оптимизации «структурных» параметров (число фильтров в наборе, значение коэффициентов децимации, порядки фильтров, число ступеней преобразования), поставил вопрос о разработке методики оптимального проектирования. Поскольку в большинстве случаев аппаратно-программная реализация системы ЦОС ориентируется на применение ЦСП, то, как следствие, подобная методика должна была отражать все особенности архитектуры и внутренние ресурсы данного класса микропроцессорных средств обработки. Методика оптимального проектирования на сигнальных процессорах систем ЦОС была заложена в [2] и получила дальнейшее развитие и применение в последние годы [69,70].

Цифровая частотная селекция сигналов на основе эффектов прореживания по времени и по частоте

Вторичная дискретизация (эффект прореживания по времени) и интерполяция для построения узкополосных цифровых фильтров стали применяться в наших работах, начиная с середины 70-х прошлого столетия. В одной из первых статей, опубликованных в 1977 году в центральной печати [21], впервые была предложена параллельная форма построения фильтра-дециматора, выполняющего вторичную дискретизацию одноступенчато с последующей интерполяцией по обычной прямой форме (рис. 1).

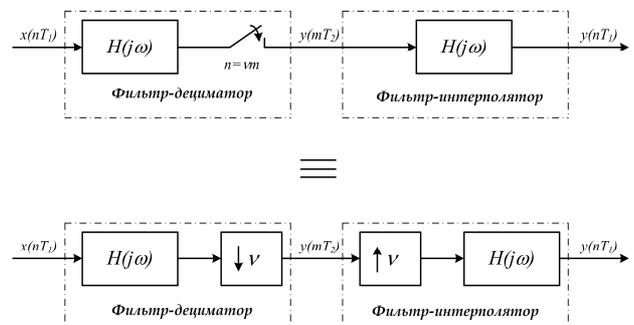


Рис. 1. Одноступенчатая структура узкополосного КИХ-фильтра на основе вторичной дискретизации (децимации).

На рис. 2,а представлена структурная схема КИХ-фильтра-дециматора, работающего по алгоритму параллельной свертки, а рис. 2,б иллюстрирует способ формирования весовых функций $h_i(n)$ для каждого i -го умножителя\накопителя. Из рис. 2,а видно, что вместо памяти входных данных $x(nT_1)$ используется параллельный набор регистров-аккумуляторов, число которых в V раз (V - коэффициент децимации) меньше порядка фильтра N . В этой же работе дан аналитический вывод и предложена структура цифрового полосового КИХ-фильтра с произвольной центральной частотой полосы пропускания, реализуемая с использованием квадратурной демодуляции\модуляции.

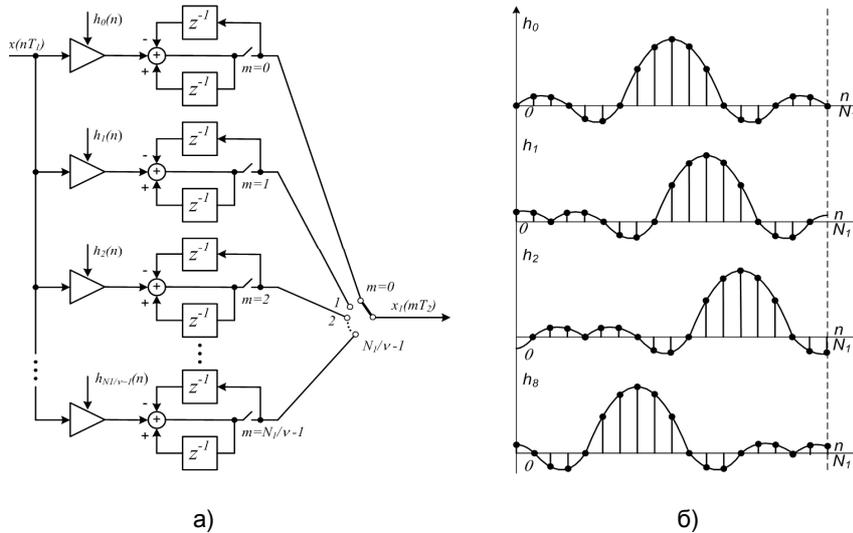


Рис. 2. Параллельная форма фильтра-дециматора.

Дальнейшее обобщение данного метода на класс БИХ-фильтров получено в [24]. В работе [74] проведен анализ шума квантования в цифровых полосовых фильтрах с прореживанием и интерполяцией отсчетов выходного сигнала. Получена оценка дисперсии шума квантования на выходе фильтра, позволяющая правильно выбрать разрядность входного и промежуточного сигналов.

Идея предварительной трансформации частот на основе квадратурной демодуляции/модуляции оказалась полезной при синтезе узкополосных фильтров методом частотной выборки. В работах [75,76] предложена структура интерполирующего фильтра частотной выборки с произвольно заданной частотой ω_k (рис. 3), не требующая высокоточного представления коэффициентов параллельного набора цифровых резонаторов, однако исключающая возможность округления результатов умножения на входах цифровых интеграторов. Попытка обойти эту проблему была предпринята в последующих работах [2,30].

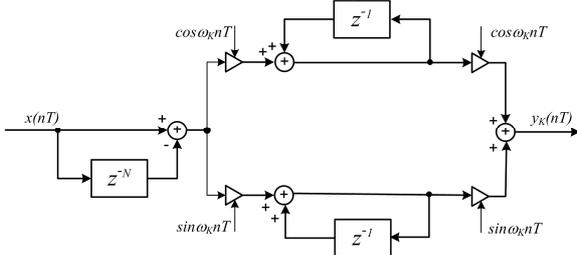


Рис. 3. Структурная схема интерполирующего фильтра k-ой частотной выборки.

В 1980 году в журнале «Известия вузов – Приборостроение» вышла статья «Синтез структуры цифрового узко-

полосного фильтра с использованием вторичной дискретизации» [22], обобщающая на этом этапе исследований применение эффекта прореживания по времени для построения одиночных полосовых фильтров. Приведена оптимизация двухступенчатой структуры фильтра, включающая на выходе фильтра-дециматора формирующий фильтр, который, работая на существенно пониженной частоте дискретизации, обеспечивает воспроизведение заданной высокой прямоугольности АЧХ проектируемого фильтра. Приводится методика оптимизации параметров двухступенчатой структуры, гарантирующая минимизацию приведенных вычислительных затрат (числа операций умножения) при заданном ограничении на выделенные ресурсы памяти данных.

Принимая во внимание, что эффекту прореживания по времени всегда сопутствует один общий недостаток – наличие шума вторичной дискретизации (децимации), особенно когда речь идет о проектировании полосовых фильтров с относительно небольшим подавлением в зоне непрозрачности (не более 60 дБ), в начале 80-х была выдвинута идея использования альтернативного подхода – прореживания импульсной характеристики фильтра вместо выходного сигнала. Поскольку такая процедура приводила автоматически к периодичности частотных характеристик фильтра с показателем периодичности, определяемым коэффициентом децимации импульсной характеристики V , то соответствующий фильтр был назван цифровым гребенчатым кратности V , а сам процесс преобразования сигнала при его прохождении через гребенчатый фильтр представлял собой эффект прореживания по частоте. На рис.4 представлена общая структурная схема цифрового гребенчатого КИХ-фильтра кратности V .

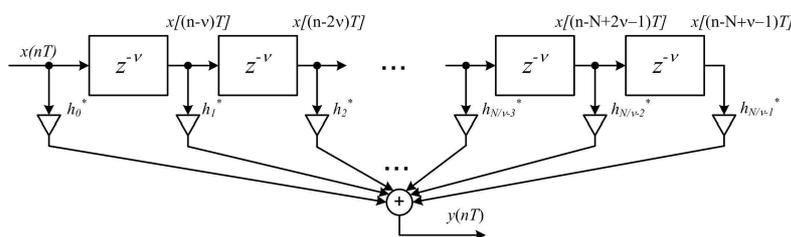
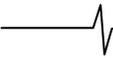


Рис. 4. Структурная схема цифрового гребенчатого КИХ-фильтра.



Из схемы видно, что структура гребенчатого КИХ-фильтра, реализуемая по прямой форме, отличается от структуры обычного фильтра только заменой элементов задержки z^{-1} на элементы задержки z^{-V} . Подобная структура цифрового гребенчатого фильтра (ЦГФ) впервые была использована для эффективной реализации набора цифровых полосовых фильтров (ЦПФ) с однотипными частотными характеристиками, отличающихся высокой прямоугольностью АЧХ [25]. Значительное уменьшение требуемых вычислительных затрат достигалось за счет того, что высокая прямоугольность АЧХ набора ЦПФ обеспечивалась одним гребенчатым фильтром с двумя параллельными выходами, разделяющим спектр входного сигнала на совокупность четных и нечетных каналов. К тому же, сам гребенчатый фильтр, отличающийся прореженностью импульсной характеристики, требует в V раз меньше вычислительных затрат по отношению к однополосному фильтру.

Идея последовательного соединения ЦГФ и последующего сглаживающего (маскирующего) фильтров (рис.5) для эффективной реализации одиночного узкополосного фильтра вначале использовалась только для решения проблемы расчета и оптимизации параметров КИХ-фильтра высокого порядка, а затем в работе [29] стала основой двухступенчатой, а впоследствии многоступенчатой реализаций.

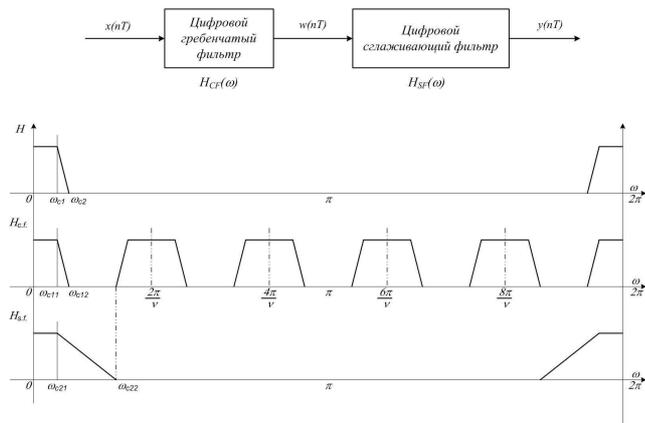


Рис. 5. Метод синтеза цифрового узкополосного фильтра на основе эффекта прореживания по частоте.

На рис. 6 представлена трехступенчатая структура узкополосного фильтра с использованием двух гребенчатых и одного сглаживающего фильтров. Многократное уменьшение общих вычислительных затрат и памяти коэффициентов связано с «прореживанием» (наличием $V-1$ нулей между каждой соседних отсчетов) импульсной характеристики ЦГФ и малым порядком широкополосного сглаживающего фильтра. Для сверх узкополосного фильтра число ступеней преобразования может пропорционально нарастать. При этом встает задача поиска оптимального значения коэффициентов децимации импульсных характеристик ЦГФ на каждой ступени преобразований по критерию минимизации памяти коэффициентов и, как следствие, минимизации вычислительных затрат [29].

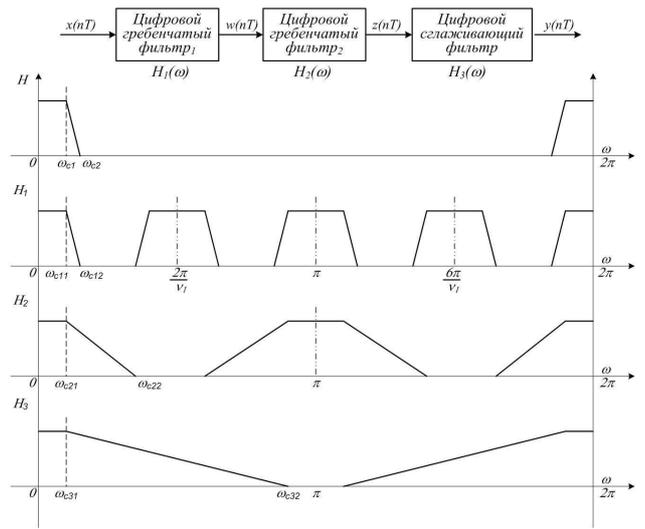


Рис. 6. Трехступенчатая реализация узкополосного фильтра на основе эффекта прореживания по частоте.

Естественным шагом дальнейшего развития теории многоскоростной обработки сигналов стало применение эффекта прореживания по времени и по частоте для построения набора (банка) ЦПФ с однотипными частотными характеристиками, равномерно перекрывающих заданную полосу рабочих частот, как правило, от нуля до частоты дискретизации.

В первой публикации на эту тему [23] фактически ставилась задача построения устройства преобразования ЧРК в ВРК, т.е. группового сигнала с частотным разделением каналов (ЧРК) в множество сигналов с временным разделением каналов (ВРК). Оригинальность нового подхода к решению известной проблемы заключалась в использовании алгоритма двойного БПФ с усечением боковых составляющих дискретного спектра за пределами полосы пропускания каждого канального сигнала. При этом множество обратных преобразований (из частотной области во временную), после соответствующей трансформации выделенного участка спектра частот группового сигнала, выполнялось по массиву малой размерности, что и обеспечило значительное уменьшение общих вычислительных затрат. В работе [25] проведен анализ влияния усечения дискретного спектра канальных сигналов на точность фильтрации и преобразования ЧРК в ВРК по ранее предложенному методу двойного БПФ и даны рекомендации по его применению. В дальнейшем была предложена двухступенчатая структура набора цифровых фильтров-демодуляторов (ЦФДМ) на основе двойного БПФ [27], которая позволила значительно уменьшить размерность прямого и обратного преобразований на каждой ступени и, как следствие, обеспечить уменьшение памяти данных и коэффициентов при высокой общей производительности всей системы, что особенно важно, когда используются специализированные СБИС процессора БПФ с относительно небольшой размерностью окна преобразования.

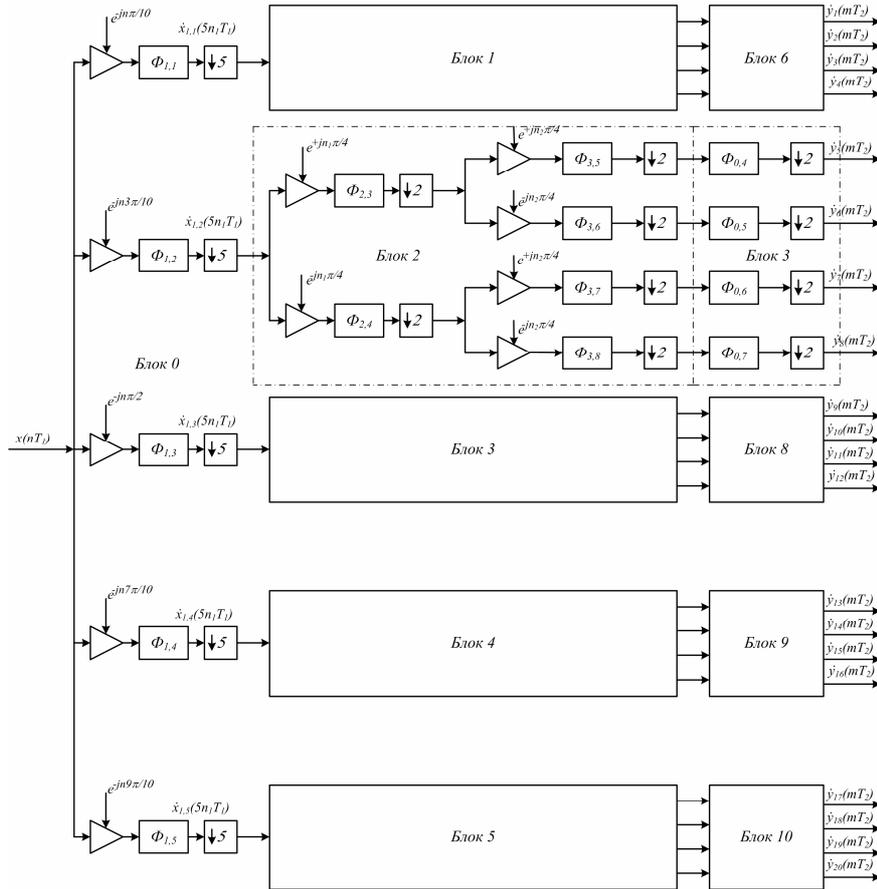


Рис. 7. Трехступенчатая пирамидальная форма набора ЦФДМ.

Идея последовательного понижения частоты дискретизации при построении набора ЦФДМ получила развитие в [28]. На рис.7 дан пример реализации трехступенчатой пирамидальной структуры набора ЦФДМ. Предполагается, что на первой ступени с помощью пяти полосовых фильтров-дециматоров спектр входного сигнала разбивается на пять поддиапазонов с трансформацией каждого выделенного поддиапазона в низкочастотную область и понижением частоты дискретизации в 5 раз. На второй ступени каждый из поддиапазонов разбивается на четыре более мелких поддиапазона с понижением частоты дискретизации в 4 раза и, наконец, на третьей ступени производится окончательное формирование высокопрямоугольных АЧХ каждого из 20 фильтров-демодуляторов с понижением частоты дискретизации в 2 раза (эта ступень может отсутствовать при невысокой прямоугольности АЧХ фильтров). Заметим, что на второй ступени преобразований в блоках 1-5, разбиение каждого поддиапазона на соответствующие полосы частот, выделяемые набором ЦФДМ, может производиться по двухступенчатой структуре с последовательным понижением частоты дискретизации в 2 раза. Более того, если число частотных каналов кратно степени двойки, т.е. $M = 2^m$, то в целом ряде приложений наиболее предпочтительной является m -ступенчатая пирамидальная структура, в рамках которой на каждой ступени преобразований число используемых фильтров-демодуляторов предварительной обработки удваивается с одновременным понижением частоты дискретизации в 2 раза (метод Цуды [77]).

Главное достоинство многоступенчатой пирамидальной структуры – частотно-временное представление входного сигнала с одновременным использованием временных скользящих окон разной длины и формы. Именно в этом и проявилась прямая связь теории синтеза банка фильтров-демодуляторов с вейвлет-преобразованием. Отметим также и другие положительные стороны нового метода синтеза по отношению к прямой параллельной форме. Это существенное уменьшение общих вычислительных затрат и памяти коэффициентов, возможность «адаптации» к частотно-временной структуре входного сигнала с выделением на выходе системы только информативных частотных каналов, снижение уровня собственных шумов и другие [2].

Как и для рассмотренной выше задачи проектирования узкополосного фильтра, для построения набора ЦФФ можно эффективно использовать альтернативный подход, базирующийся на последовательном соединении набора ЦГФ. Это особенно целесообразно и продуктивно, когда речь идет о разработке системы цифровой частотной селекции сигналов без понижения частоты дискретизации на выходах субполосных каналов. Такая задача была впервые связана с цифровой реализацией так называемого блока защиты от узкополосных помех. Предполагалось, что режекция (подавление) узкополосных помех на фоне широкополосного полезного сигнала может быть выполнена путем разбиения спектра входного сигнала на субполосы с последующим суммированием выходов только тех частотных каналов, которые не поражены помехой.

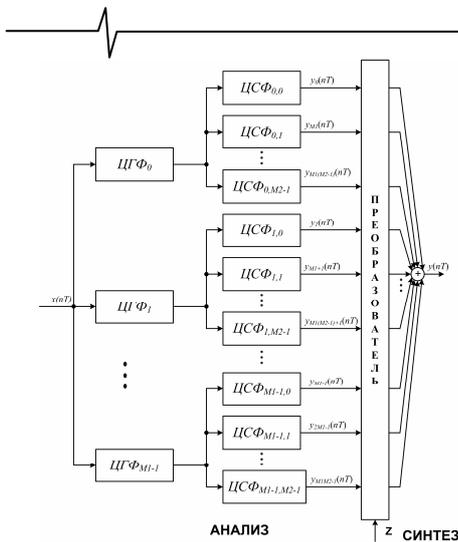


Рис. 8. Двухступенчатая реализация набора ЦФФ.

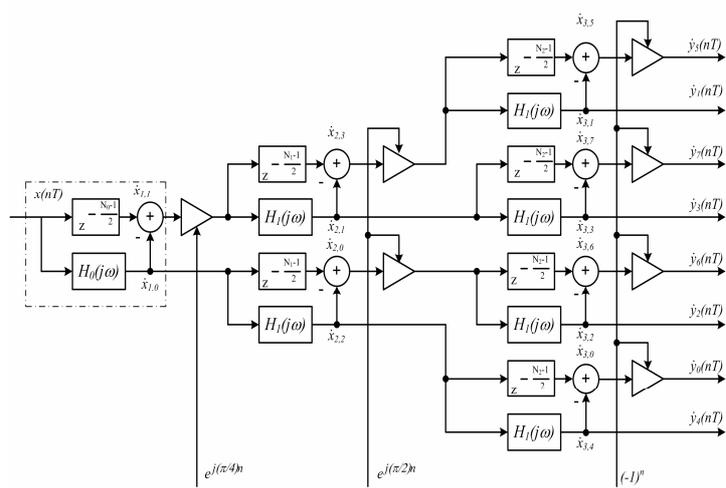


Рис. 9. Трехступенчатая пирамидальная форма набора ЦФФ с использованием полуполосных гребенчатых фильтров.

Еще в 1981 году в [25] была продемонстрирована эффективность применения ЦФФ для решения поставленной задачи. В дальнейшем было показано, что путем использования параллельного набора ЦФФ предварительной обработки входного сигнала и последующего параллельного набора цифровых сглаживающих полосовых фильтров (рис.8) при оптимальном сочетании числа фильтров можно добиться значительного уменьшения общих вычислительных затрат и памяти коэффициентов.

Естественным развитием нового подхода стал переход к многоступенчатой реализации набора ЦФФ с использованием эффекта прореживания по частоте. В работе [32] дана оценка вычислительной эффективности многоступенчатой пирамидальной структуры набора ЦФФ по отношению к другим известным методам. Показано, что многократное уменьшение приведенных вычислительных затрат и памяти коэффициентов связано как с прореживанием импульсных характеристик ЦФФ, так и с уменьшением их порядков от ступени к ступени последовательного преобразования. И, наконец, в 1988 году в [34] предложен метод синтеза цифровой частотной селекции сигналов на основе полуполосных гребенчатых фильтров, отличающийся наивысшей эффективностью решения поставленной задачи (рис. 9).

По логике преобразований и достигаемой вычислительной эффективности он подобен алгоритму БПФ, но в отличие от последнего дает в каждой частотной полосе полное временное представление сигнала на исходной частоте дискретизации. Предполагается, что на ка-

ждой ступени преобразований с помощью набора однотипных (в масштабе приведенных частот) полуполосных гребенчатых фильтров спектры входных сигналов разбиваются на совокупности четных и нечетных каналов. При этом коэффициенты децимации импульсных характеристик ЦФФ принимают максимально допустимые значения, а, следовательно, вычислительные затраты минимизируются. Если число частотных каналов M кратно степени двойки, т.е. $M = 2^m$, то общее число ступеней преобразования равно m , и достигается наивысшая эффективность предложенного метода синтеза структуры набора ЦФФ.

Идея «скрытой» децимации импульсной характеристики и применения ЦФФ, реализуемых в классе БИХ-цепей, не могла быть оставлена без внимания. В работе [31] впервые дан анализ чувствительности характеристик гребенчатого БИХ-фильтра к неточному представлению коэффициентов. Показано, что даже в рамках обычной прямой формы по структуре, представленной на рис.10, удается достигнуть значительного уменьшения чувствительности полюсов передаточной функции ЦФФ, а, следовательно, и характеристик фильтра по отношению к однополосной реализации. Таким образом, использованием последовательного соединения гребенчатого и сглаживающего (маскирующего) БИХ-фильтров, можно добиться более эффективного решения проблемы высокой чувствительности характеристик узкополосного БИХ-фильтра к неточному представлению коэффициентов как для одиночного фильтра, так и для набора ЦФФ, реализуемых по пирамидальной структуре (рис. 11) [33].

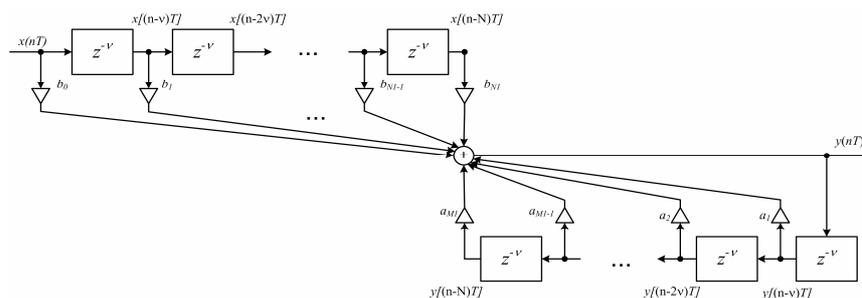


Рис. 10. Структурная схема цифрового гребенчатого БИХ-фильтра.

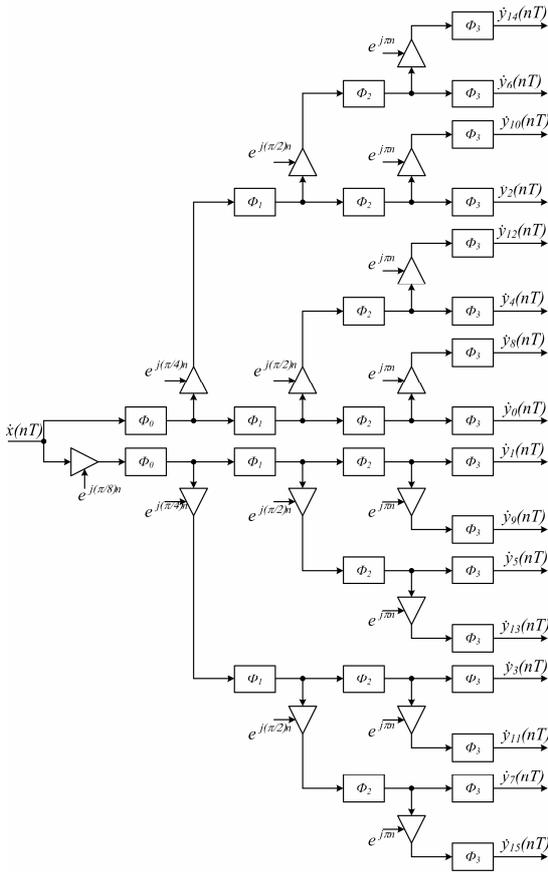


Рис. 11. Пирамидальная структура набора цифровых полосовых БИХ-фильтров.

Заключение

Подводя общий итог проведенных выше исследований, можно утверждать, что к концу 80-х прошедшего столетия была создана теория цифровой частотной селекции сигналов на основе эффектов прореживания по времени и по частоте, основы которой изложены в монографии [2].

Дальнейшие исследования, относящиеся к новому этапу эволюции теории и техники многоскоростной обработки сигналов, были связаны с построением адаптивных поисковых систем на основе частотно-временной обработки сигналов и разработки методики оптимального проектирования систем ЦОС на цифровых сигнальных процессорах. Но, это уже новая страница истории, к которой мы вернемся в очередной публикации на предложенную тему.

Литература

1. Crochiere R.E., Rabiner L. Multirate Digital Signal Processing. Prentice Hall. Englewood Cliffs.- NJ, 1983.
2. Витязев В.В. Цифровая частотная селекция сигналов. М.: Радио и связь, 1993. 240 с.
3. Vaidyanathan P.P. Multirate Systems and Filter Banks. Prentice Hall. Englewood Cliffs.- NJ, 1993.
4. Mitra S.K. Digital Signal Processing: a computer-based approach. McGraw-Hill. Comp. Inc., 1998.
5. Основы цифровой обработки сигналов: Курс лекций / А.И. Солонина, Д.А. Улахович, С.М. Арбузов и др. СПб.: БХВ-Петербург, 2003. 608 с.

6. Айфичер Э.С., Джервис Б.У. Цифровая обработка сигналов: практический курс : Пер. с англ. М.: Изддом «Вильямс», 2004. 992 с.
7. The Digital Signal Processing Handbook / Ed. Vijay K. Madisetti, Douglas B. Williams by CRC Press LLC, 1998.
8. Meyer Y. Wavelets and Operators // Cambridge University Press. Cambridge.- UK, 1990.
9. Meyer Y. Wavelets: Algorithms and Applications // SIAM. Philadelphia.- PA, 1993.
10. Vetterli M., Kovacevic J. Wavelets and Subband Coding. Prentice Hall. Englewood Cliffs.- NJ, 1995.
11. Strang G., Nguyen T. Wavelets and Filter Banks. Wellesley-Cambridge Press. Cambridge.- UK, 1996.
12. Shafer R.W., Rabiner L.R. A digital signal processing approach to interpolation // Proc. IEEE. V. 61. June, 1973. P. 692-702.
13. Oetken G., Parks T.W., Schussler W. New results in the design of interpolators // IEEE Trans. Acoust., Speech and Signal Processing. V. ASSP-23. June, 1975. P. 301-309.
14. Bellanger M.G., Daguët J.L., Hephagnol G.P. Interpolation, extrapolation and reduction of computation speed in digital filter // IEEE Trans. Acoust., Speech and Signal Processing. V. ASSP-22. Aug., 1974. P. 231-235.
15. Rabiner L.R., Crochiere R.E. A novel implementation for narrow-band FIR digital filters // IEEE Trans. Acoust., Speech and Signal Processing. V. ASSP-23. Oct., 1975. P. 457-464.
16. Crochiere R.E., Rabiner L.R. Optimum FIR digital filter implementations for decimation, interpolation and narrow band filtering // IEEE Trans. Acoust., Speech and Signal Processing. V. ASSP-23. Oct., 1975. P. 444-456.
17. Shively R.R. On multistage FIR filters with decimation // IEEE Trans. Acoust., Speech and Signal Processing. V. ASSP-23. Aug., 1975. P. 353-357.
18. Crochiere R.E., Rabiner L.R. Further considerations in the design of decimators and interpolators // IEEE Trans. Acoust., Speech and Signal Processing. V. ASSP-24. Aug., 1976. P. 296-311.
19. Crochiere R.E., Rabiner L.R. Interpolation and decimation of digital signals: a tutorial review // Proc. IEEE. V. 69. March, 1981. P. 300-331.
20. Bellanger M.G., Bonnerot G., Coudreuse M. Digital filtering poly-phase network: Application to sample rate alteration and filter banks // IEEE Trans. Acoust., Speech and Signal Processing. V. ASSP-24. Apr., 1976. P. 109-114.
21. Витязев В.В., Степашкин А.И. Метод синтеза цифровых узкополосных фильтров с усеченной импульсной характеристикой // Изв. вузов. Приборостроение. 1977. Т. 20. № 6. С. 25—29.
22. Витязев В.В., Степашкин А.И. Синтез структуры цифрового узкополосного фильтра с использованием вторичной дискретизации // Изв. вузов. Приборостроение. 1980. Т. 23. № 6. С. 32—38.
23. Витязев В.В., Алпатов Б.А., Степашкин А.И. Синтез цифровых узкополосных фильтров для частотного разделения каналов // Электросвязь. 1979. № 7. С. 47—51.
24. Витязев В.В., Степашкин А.И. Синтез цифровых полосовых фильтров // Радиотехника. 1978. Т. 33. № 3. С. 75—77.
25. Витязев В.В., Муравьев С.И., Степашкин А.И. Метод синтеза цифровых узкополосных КИХ-фильтров // Изв. вузов. Радиоэлектроника. 1981. Т. 24. № 7. С. 55—59.
26. Витязев В.В., Степашкин А.И. К синтезу цифрового фильтра-деמודулятора на основе двойного быстрого преобразования Фурье // Радиотехника 1981. Т. 36. № 7. С. 20—24.
27. Витязев В.В., Степашкин А.И. Метод синтеза цифровых фильтров-деמודуляторов на основе двойного быстрого преобразова-



- ния Фурье // Электросвязь. 1982. № 3. С. 45—47.
28. Витязев В.В. Синтез пирамидальной структуры набора цифровых фильтров-демодуляторов // Электросвязь. 1983. № 7. С. 45—49.
 29. Витязев В.В., Муравьев С.И., Степашкин А.И. Метод проектирования цифровых полосовых фильтров с конечной памятью // Изв. вузов. Радиоэлектроника. 1983. Т. 26. № 9. С. 3—9.
 30. Витязев В.В., Муравьев С.И. Модификации метода частотной выборки на основе структуры с квадратурной модуляцией // Радиотехника. 1984. № 2. С. 19—23.
 31. Витязев В.В., Муравьев С.И., Степашкин А.И. Метод проектирования цифровых полосовых фильтров с бесконечной памятью // Радиотехника. 1984. № 2, С. 19—23.
 32. Витязев В.В., Муравьев С.И., Степашкин А.И. Синтез пирамидальной структуры набора цифровых полосовых фильтров // Электросвязь. 1985. № 8. С. 52—56.
 33. Витязев В.В., Муравьев С.И. Пирамидальная структура цифровых полосовых фильтров с бесконечной памятью // Радиотехника. 1985. № 9. С. 45—49.
 34. Витязев В.В., Муравьев С.И. Синтез цифровой системы частотной селекции сигналов на основе полуполосовых гребенчатых фильтров // Электросвязь. 1988. № 3. С. 57—61.
 35. Витязев В. В. Оптимальное проектирование многоступенчатых структура цифровых фильтров на процессорах обработки сигналов // Электросвязь. 1992. № 4. С. 23—27.
 36. Esteban D, Galand C. Application of quadrature mirror filters to split band voice coding schemes // Proc. IEEE Int. Conf. Acoust. Speech and Signal Processing. May, 1977. P. 191-195.
 37. Crochiere R.E. Subband coding // Bell Syst. Tech. J. V. 60, Sept. 1981.
 38. Smith M.J.T., Barnwell III. Exact reconstruction techniques for tree-structured subband coders // IEEE Trans. Acoust., Speech and Signal Processing. V. ASSP-34. June, 1986. P. 434-441.
 39. Vetterli M. Filter banks allowing for perfect reconstruction // Signal Processing. V. 10. April, 1986. P. 219-244.
 40. Vaidyanathan P.P. Quadrature mirror filter banks, M-band extensions and perfect-reconstruction techniques // IEEE ASSP Mag. V. 4. July, 1987. P. 4-20.
 41. Vaidyanathan P.P. Multirate digital filters, filter banks, polyphase networks and applications: a tutorial // Proc. IEEE. V. 78. January, 1990. P. 56-93.
 42. Vaidyanathan P.P. Theory and design of M-channel maximally decimated quadrature mirror filters with arbitrary M having perfect reconstruction property // IEEE Trans. Acoust., Speech and Signal Processing. V. ASSP-35. Apr., 1987. P. 476-492.
 43. Vaidyanathan P.P., Vincent C. Liu. Classical sampling theorems in the context of multirate and polyphase digital filter bank structures // IEEE Trans. Acoust., Speech and Signal Processing. V. ASSP-36. Sept., 1988. P. 1480-1495.
 44. Viscito E., Allebach J. The design of tree-structured M-channel filter banks using perfect reconstruction filter blocks // In Proc. IEEE Int. Conf. on ASSP. NY: Apr., 1988. P. 1475-1478.
 45. Vaidyanathan P.P., Hoang P.-Q. Lattice structures for optimal design and robust implementation of two-channel perfect reconstruction QMF banks // IEEE Trans. Acoust., Speech and Signal Processing. V. ASSP-36. Jan., 1988. P. 81-94.
 46. Vaidyanathan P.P., Nguyen T.Q., Doganata Z., Saramaki T. Improved approach for design of perfect reconstruction FIR QMF banks with lossless polyphase structures // IEEE Trans. Acoust., Speech and Signal Processing. V. ASSP-37. July, 1989.
 47. Doganata Z., Vaidy P.P., Nguyen T.Q. General synthesis procedures for FIR lossless transfer matrices for perfect reconstruction multirate filter bank applications // IEEE Trans. Acoust., Speech and Signal Processing. V. ASSP-36. Oct., 1988. P. 1561-1574.
 48. Spurbeck M.S., Mullis C.T. Least squares approximation of perfect reconstruction filter banks // IEEE Trans. Acoust., Speech and Signal Processing. V. ASSP-46. Apr., 1998. P. 968.
 49. Huang J., Gu G. A direct approach to the design of QMF banks via frequency domain optimization // IEEE Trans. Acoust., Speech and Signal Processing. V. ASSP-46. Aug., 1998. P. 2131.
 50. Chen T., Francis B.A. Design of multirate filter banks by hinf optimization // IEEE Trans. Acoust., Speech and Signal Processing. V. ASSP-43. Dec., 1995. P. 2822-2830.
 51. Koilpillai R.D., Vaidyanathan P.P. Cosine-modulated FIR filter banks satisfying perfect reconstruction // IEEE Trans. on Signal Processing. V. SP-40. April, 1992. P. 770.
 52. Yuan-Pei Lin, Vaidyanathan P.P. Linear phase cosine modulated maximally decimated filter banks with perfect reconstruction // IEEE Trans. Signal Proc. V. SP-43. Nov., 1995. P. 2525.
 53. Nguen T.Q., R. D. Koilpillai R.D. The theory and design of arbitrary-length cosine-modulated filter banks and wavelets, satisfying perfect reconstruction // IEEE Trans. Signal Proc. V. SP-44. March, 1996. P. 473.
 54. Xu H., Lu W.S., Antoniou A. Efficient iterative design method for cosine-modulated QMF banks // IEEE Trans. Signal Proc. V. SP-44. July, 1996. P. 1657.
 55. Goh C.-K., Lim Y-C. An Efficient Algorithm for the Design of Weighted Minimax M-Channel Cosine-Modulated Filter Banks // IEEE Trans. Signal Proc. V. 46. May, 1998. P. 1426.
 56. Mertins A. Subspace Approach for the Design of Cosine-Modulated Filter Banks with Linear-Phase Prototype Filter // IEEE Trans. Signal Proc. V. 46. Oct., 1998. P. 2812.
 57. Argenti F., Del E. Re. Design of Biorthogonal M-Channel Cosine-Modulated FIR/IIR Filter Banks // IEEE Trans. Signal Proc. V. 48. March, 2000. P. 876.
 58. Kovacevic, Vetterli M. Perfect Reconstruction Filter Banks with Rational Sampling Factors // IEEE Trans. Signal Proc. V. 41. June, 1993. P. 2047.
 59. Nayebi K., Barnwell T.P., Spith J.T. Nonuniform filter banks: A reconstruction and design theory // IEEE Trans. Signal Proc. V. SP-41. March, 1993. P. 1114.
 60. Akkarakaran S., Vaidyanathan, P.P. New results and open problems on nonuniform filter-banks // Proc. IEEE Int. Conf. Acoust. Speech, and Signal Proc. Phoenix. March 1999.
 61. Chen T. L. Qiu, Bai E. General Multirate Building Structures with Application to Nonuniform Filter Banks, Special Issue on Multirate Systems, Filter Banks, Wavelets, and Applications // IEEE Trans. on Circuits and Systems II: Analog and Digital Signal Processing. V. 45. 1998. P. 948.
 62. Princen J. The Design of nonuniform modulated filterbanks // IEEE Trans. Signal Proc. V. SP-43. Nov., 1995. P. 2550.
 63. Chen T. Nonuniform Multirate Filter Banks: Analysis and Design with an Hinf Performance Measure // IEEE Trans. on Signal Processing. V. 45. March, 1997. P. 572.
 64. Зайцев А.А. Методы построения банков цифровых фильтров: тематический обзор // Цифровая обработка сигналов. 2003. № 1. С. 2-10.
 65. Djokovic I., Vaidyanathan P.P. Generalized sampling theorems in multiresolution subspaces // IEEE Trans. SP-45. March, 1997. P. 583-599.
 66. Wornell G.W. Emerging applications of multirate signal processing and wavelets in digital communications // Proc. IEEE. V. 84-1. April,

1996. Р. 586-603.
67. Зубарев Ю.Б., Витязев В.В., Дворкович В.П. Цифровая обработка сигналов – информатика реального времени // Цифровая обработка сигналов. 1999. № 1. С. 5-17.
68. Витязев В.В., Бодров К.А., Иванов С.И. Адаптивная многоскоростная фильтрация узкополосных процессов // Цифровая обработка сигналов и ее применение-DSPA'1999: Тр. 1-й Международной науч.-техн. конф. М., 1999. Т. 1. С. 150-160.
69. Витязев В.В., Бодров К.А., Сорокин Д.В. Оптимальный синтез многоскоростных систем фильтрации на цифровых сигнальных процессорах // Цифровая обработка сигналов и ее применение-DSPA'1999: Тр. 1-й Международной науч.-техн. конф. М., 1999. Т. 1. С. 50-54.
70. Витязев В.В., Зайцев А.А. Оптимальное проектирование многоступенчатых структур фильтров-дециматоров на сигнальных процессорах // Цифровая обработка сигналов. 2001. № 2. С. 2-9.
71. Линович А.Ю., Витязев В.В. Субполосная адаптивная фильтрация в задачах обратного моделирования // Цифровая обработка сигналов. 2004. № 1. С. 41-48.
72. Гусинская Е.И., Зайцев А.А. Оптимизация банка фильтров в задачах субполосного кодирования // Цифровая обработка сигналов. 2004. № 3. С. 18-28.
73. Линович А.Ю. Методы многоскоростной обработки сигналов в задачах обратного моделирования динамических систем // Цифровая обработка сигналов и ее применение-DSPA'2005: Тр. 7-й Международной науч.-техн. конф. М., 2005. Т. 1. С. 50-54.
74. Витязев В.В., Демашов В.С., Степашкин А.И. Шум квантования цифрового узкополосного фильтра с прореживанием и интерполяцией отсчетов выходного сигнала // Изв. вузов. Приборостроение. – 1979. – Т. 22, № 5.- с. 3-7.
75. Алпатов Б.А., Витязев В.В., Степашкин А.И. Синтез цифровых резонаторов для систем узкополосной фильтрации // Радиотехника. – 1979. – Т. 34, № 12.- с. 57-59.
76. Степашкин А.И., Алпатов Б.А. Цифровой резонатор с ограниченной импульсной характеристикой // Изв. вузов. Приборостроение. – 1980. – Т.23, № 2. – с.7-10.
77. Tsuda T., Morita S., Fujii Y. Digital TDM/FDM translator with multistage structure // IEEE Trans. – 1978, May. – Vol. COM-26.- p. 734-741.

Уважаемые коллеги!

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

В 2008 году планируется выпуск 4-х номеров журнала (тираж до 1000 экз.). Журнал будет распространяться по подписке через агентство «Роспечать» в России, СНГ и странах Балтии (индекс 82185), а также на Выставках: «Цифровая обработка сигналов и ее применение – DSPA'2008», «ExpoElectronica», «СвязьЭкспокомм», «ЭЛЕКТРОНИКА: компоненты, оборудование, технологии» (г. Москва) и др.

Размещение рекламы Вашей фирмы на страницах журнала «Цифровая обработка сигналов» на плановой основе (не менее 2-х полных или 4-х половинчатых рекламных полос в течение года) предоставит Вам следующие возможности и права:

1. Первоочередное право расположения рекламных материалов на всех обложках (кроме 1-й) и страницах журнала.
2. Публикация представленных Вами рабочих (рекламных) материалов (статей) объемом до 6 полос в каждом очередном номере (в счет оплаченной рекламы).
3. Установка баннера Вашего сайта (или логотипа вашей организации) на 1-й странице сайта журнала «Цифровая обработка сигналов» (www.dspra.ru) в течение всего года, что привлечет внимание к продукции (услугам) Вашей фирмы новых участников на рынке DSP-технологий (ежедневно фиксируется до 100 и более посещений сайта www.dspra.ru).
4. Предоставление до 10 экз. очередного выпуска журнала.

Ориентировочная стоимость рекламных услуг:

4-я (внешняя) страница цветной обложки - 20 тысяч рублей.

2-я и 3-я (внутренние) страницы цветной обложки - 13 тысяч рублей.

1\2 цветной внутренней полосы - 7 тысяч рублей.

1\2 черно-белой внутренней полосы – 1 тысяча рублей.

Ждем Ваших предложений.

С наилучшими пожеланиями, зам. главного редактора
д.т.н., профессор Витязев Владимир Викторович

Предложения прошу направлять по адресу: E-mail: tor@rgta.ryazan.ru или info@dspra.ru

ПАРАУНИТАРНЫЕ БАНКИ ФИЛЬТРОВ НА ОСНОВЕ АЛГЕБРЫ КВАТЕРНИОНОВ: ТЕОРИЯ И ПРИМЕНЕНИЕ

Парфенюк М., Петровский А.А.

1. Введение в банки фильтров

1.1. Определения

Банк фильтров – цифровая система, состоящая из секций анализа и синтеза, называемых банком фильтров анализа и банком фильтров синтеза (рис. 1.1). Входной сигнал $x(n)$, представленный последовательностью отсчетов, разбивается при помощи фильтров блока анализа $H_k(z)$ ($k = 0, 1, \dots, M-1$) на M субполосных составляющих, которые в идеальном случае в частотной области не перекрываются. Операции, выполняемые блоком синтеза, являются обратными операциями блока анализа. Подобранным соответствующим образом набор фильтров блока синтеза $F_k(z)$ ($k = 0, 1, \dots, M-1$), можно восстановить исходный сигнал $y(n)$ из его субполосных компонент.

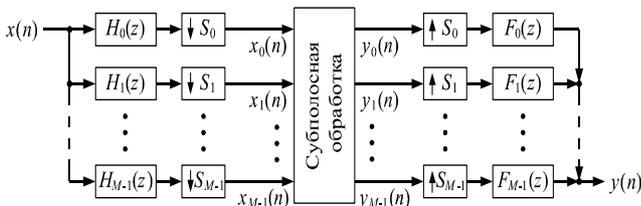


Рис. 1.1. Банк фильтров: система анализа/синтеза сигнала

Банк фильтров относится к классу многоскоростных систем цифровой обработки сигналов [2-4], в которых частота дискретизации различна в разных точках системы. Значение коэффициента темпа поступления отсчетов в канале (коэффициента децимации S_k) задается утверждением о дискретизации сигнала в зависимости от ширины частотной полосы канала B_k и его положения в банке фильтров. Оригинальная частота дискретизации f_s сигнала в k -ом канале теоретически может быть уменьшена в $S_k \leq f_s/B_k$ раз. Равенство, в данном случае, означает, что канал является максимально децимированным: $\sum_{k=0}^{M-1} 1/S_k = 1$. Таким образом, в частотно-временном представлении сигнала исключена избыточность, т.е. сумма отсчетов во всех каналах соответствует количеству отсчетов в исходном сигнале. Банк фильтров считается передецимированным, если

Параунитарные банки фильтров (ПУБФ) могут рассматриваться как наиболее значимое преобразование среди многоскоростных систем цифровой обработки сигналов [1]. Это обусловлено тем фактом, что подобные банки являются преобразованиями без потерь в дополнение к гарантированной перфективной реконструкции сигнала. Точное соотношение между энергиями во всем частотном диапазоне и в субполосах сильно упрощает теоретические выкладки и делает ПУБФ полезными в применении к компрессии (кодированию) изображений. Серьезной практической проблемой, связанной с реализацией ПУБФ на арифметике с фиксированной запятой, является потеря свойства перфективной реконструкции сигнала банком фильтров. Это обусловлено чувствительностью структуры ПУБФ к квантованию коэффициентов. В данной работе представлен новый подход к синтезу и конструированию решетчатых структур 4-х и 8-ми канальных ПУБФ на основе алгебры кватернионов. В отличие от известных решений, предложенные структуры являются структурами без потерь даже при выполнении вычислений с конечной точностью. Финансовая поддержка работы осуществлялась Министерством науки и высшей школы Республики Польша (грант № 519 030 32/3775).

$\sum_{k=0}^{M-1} 1/S_k > 1$, т.е. достаточно, чтобы хотя бы в одном канале коэффициент децимации не был равен числу каналов. Более сложные случаи позиционирования субполос каналов в банках фильтров рассмотрены в [5-7].

Соотношение между z - преобразованием сигналов на входе и выходе M -канального банка фильтров, изображенного на рис. 1.1, является следующим [2]:

$$Y(z) = \sum_{k=0}^{M-1} F_k(z) \frac{1}{S_k} \sum_{l=0}^{S_k-1} H_k(zW_{S_k}^l) X(zW_{S_k}^l), \quad (1.1)$$

где $W_{S_k} = e^{-j2\pi/S_k}$. Анализ данного выражения показывает, что в банке возможны искажения входного сигнала: амплитудные, фазовые и отражения частотных характеристик (элайзинг), обусловленные наличием дециматоров и интерполяторов. Величина данных искажений зависит как от частотных характеристик канальных фильтров, так и выбора коэффициентов децимации S_k . Анализ искажений, возникающих в структуре банка фильтров, проще анализировать в максимально децимированном банке фильтров, для которого соотношение (1.1) значительно упрощается:

$$Y(z) = \sum_{k=0}^{M-1} T_k(z) X(zW_M^k), \quad (1.2)$$

где $T_k(z)$ – передаточная функция k -го канала:

$$T_k(z) = \frac{1}{M} \sum_{l=0}^{M-1} F_l(z) H_l(zW_M^k). \quad (1.3)$$

Выходной сигнал $y(n)$ системы анализа-синтеза банка фильтров будет свободен от элайзинговой составляющей $X(zW_M^k)$, $k > 0$ в случае, если

$$T_k(z) = 0, \text{ для } 1 \leq k < M. \quad (1.4)$$

В банке фильтров, для которого справедливо условие (1.4), остаются только амплитудные и фазовые искажения, которые определяются согласно следующему выражению:

$$\frac{Y(z)}{X(z)} = T_0(z) = \frac{1}{M} \sum_{k=0}^{M-1} H_k(z) F_k(z). \quad (1.5)$$

Очевидно, что для получения перфективной реконструкции входного сигнала $x(n)$ банком фильтров, передаточная функция искажений $T_0(z)$ должна принять форму

$$T_0(z) = cz^{-\Delta}, c \neq 0, \Delta \in \mathbb{N} \quad (1.6)$$

простого звена задержки с некоторым масштабированием амплитуды.

Полное восстановление или перфективная реконструкция – свойство банка цифровых фильтров, заключающееся в том, что сигнал, прошедший через схему анализа-синтеза, идентичен входному с точностью до задержки. Для этого фильтры синтеза должны подавлять наложение частотных характеристик (элайзинг) и устранять амплитудные и фазовые искажения [1,2].

Параунитарный (ортогональный) банк фильтров (ПУБФ) – банк фильтров, у которого передаточные функции анализирующих и синтезирующих фильтров и их соответственно смещенные версии ортогональны друг другу. Фильтры синтеза в параунитарных банках являются транспонированными версиями фильтров анализа [1]:

$$F_k(z) = H_k^T(z^{-1}). \quad (1.7)$$

При соблюдении этого условия обеспечивается возможность перфективной реконструкции банком фильтров входного сигнала $x(n)$.

1.2. Полифазное представление

Пусть передаточные функции секций анализа и синтеза банка фильтров представлены в виде соответствующих векторов:

$$\mathbf{H}(z) = [H_0(z) \ H_1(z) \ \dots \ H_{M-1}(z)]^T, \\ \mathbf{F}(z) = [F_0(z) \ F_1(z) \ \dots \ F_{M-1}(z)]^T, \quad (1.8)$$

тогда можно выбрать такие полифазные матрицы анализа

$$\mathbf{E}(z) = \begin{bmatrix} E_{00}(z) & E_{01}(z) & \dots & E_{0,M-1}(z) \\ E_{10}(z) & E_{11}(z) & \dots & E_{1,M-1}(z) \\ \vdots & \vdots & \ddots & \vdots \\ E_{M-1,0}(z) & E_{M-1,1}(z) & \dots & E_{M-1,M-1}(z) \end{bmatrix}$$

и синтеза

$$\mathbf{D}(z) = \begin{bmatrix} D_{00}(z) & D_{01}(z) & \dots & D_{0,M-1}(z) \\ D_{10}(z) & D_{11}(z) & \dots & D_{1,M-1}(z) \\ \vdots & \vdots & \ddots & \vdots \\ D_{M-1,0}(z) & D_{M-1,1}(z) & \dots & D_{M-1,M-1}(z) \end{bmatrix},$$

что векторы передаточных функций секций анализа и синтеза можно представить следующим образом [2]:

$$\mathbf{H}(z) = \mathbf{E}(z^M) \begin{bmatrix} 1 & z^{-1} & \dots & z^{-(M-1)} \end{bmatrix}^T,$$

$$\mathbf{F}(z) = \begin{bmatrix} z^{-(M-1)} & z^{-(M-2)} & \dots & 1 \end{bmatrix} \mathbf{D}(z^M). \quad (1.9)$$

Для получения перфективной реконструкции на компоненты полифазных матриц накладывается дополнительное ограничение:

$$\mathbf{D}(z)\mathbf{E}(z) = cz^{-\Delta}\mathbf{I}, \quad c \neq 0, \Delta \geq 0, \quad (1.10)$$

где c – ненулевая константа; Δ – задержка, выраженная целым числом интервалов дискретизации, вносимая секциями анализа-синтеза; \mathbf{I} – единичная матрица.

1.3. Решетчатые структуры параунитарных банков фильтров

M -канальный параунитарный банк фильтров анализа в *общем виде* может быть описан следующей факторизацией полифазной передаточной матрицы [8]:

$$\mathbf{E}(z) = \mathbf{R}_N \mathbf{\Lambda}(z) \mathbf{R}_{N-1} \mathbf{\Lambda}(z) \dots \mathbf{R}_1 \mathbf{\Lambda}(z) \mathbf{E}_0, \quad (1.11)$$

которая содержит $\mathbf{\Lambda}(z) = \text{diag}(z^{-1}, \mathbf{I}_{M-1})$ - матрицу задержек, ортогональные матрицы: общую \mathbf{E}_0 с $M(M-1)/2$ степенями свободы и $\mathbf{R}_i, i = 1 \dots N-1$, ограниченные до $M-1$ степени свободы. Оба типа матриц параметризованы планарными оборотами Гивенса, в которых каждый из них соответствует одной степени свободы.

Банк фильтров анализа, определяемый выражением (1.11), может быть представлен следующей решетчатой структурой (рис. 1.2).

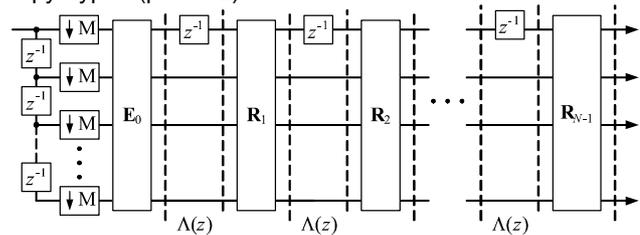


Рис. 1.2. Решетчатая структура банка фильтров анализа

На практике широкое распространение получили банки фильтров с линейными фазочастотными характеристиками (ФЧХ). Для ПУБФ анализа с линейной ФЧХ и четного числа каналов M хорошо известна факторизация передаточной полифазной матрицы $\mathbf{E}(z)$:

$$\mathbf{E}(z) = \mathbf{G}_{N-1}(z) \mathbf{G}_{N-2}(z) \dots \mathbf{G}_1(z) \mathbf{E}_0, \quad (1.12)$$

где $\mathbf{E}_0 = \frac{1}{\sqrt{2}} \mathbf{\Phi}_0 \mathbf{W} \text{diag}(\mathbf{I}_{M/2}, \mathbf{J}_{M/2})$, (1.13)

$$\mathbf{G}_i(z) = \frac{1}{2} \mathbf{\Phi}_i \mathbf{W} \mathbf{\Lambda}(z) \mathbf{W}, \quad i = 1, \dots, N-1, \quad (1.14)$$

$$\mathbf{W} = \begin{bmatrix} \mathbf{I}_{M/2} & \mathbf{I}_{M/2} \\ \mathbf{I}_{M/2} & -\mathbf{I}_{M/2} \end{bmatrix}, \quad (1.15)$$

$$\mathbf{\Lambda}(z) = \text{diag}(\mathbf{I}_{M/2}, z^{-1} \mathbf{I}_{M/2}), \quad (1.16)$$

$$\mathbf{\Phi}_i = \text{diag}(\mathbf{U}_i, \mathbf{V}_i). \quad (1.17)$$

\mathbf{U}_i и \mathbf{V}_i - произвольные ортогональные матрицы размера $M/2 \times M/2$. Кроме того, для упрощения вышеобозначенной факторизации для $i > 0$ матрица \mathbf{U}_i может быть представлена единичной матрицей [9], тогда

$$\mathbf{\Phi}_i = \text{diag}(\mathbf{I}, \mathbf{V}_i), \quad i > 0. \quad (1.30)$$

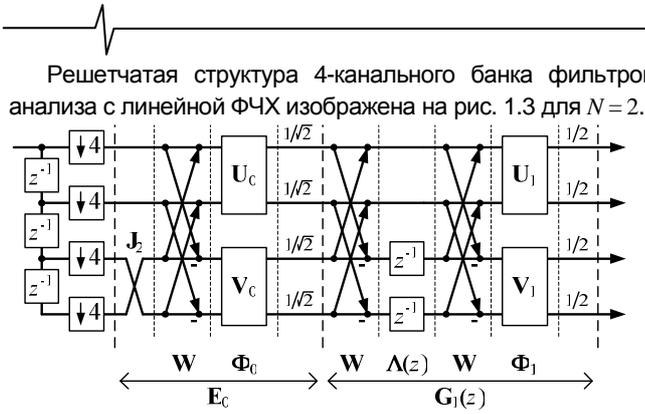


Рис. 1.3. Решетчатая структура 4-канального банка фильтров анализа с линейной ФЧХ

Среди ПУБФ с линейной ФЧХ выделяются банки фильтров, у которых наблюдается попарная зеркальная симметрия частотных характеристик фильтров относительно $\pi/2$ [10]. Для этого передаточные функции или импульсные характеристики анализирующих фильтров должны удовлетворять соответственно следующим условиям:

$$H_{M-1-k}(z) = \pm H_k(-z) \text{ или } h_{M-1-k}(n) = \pm (-1)^n h_k(n),$$

где $k=0, \dots, N-1$, $n=0, \dots, L-1$, L – длина импульсной характеристики фильтра. Для четного числа каналов M ПУБФ с попарной зеркальной симметрией частотных характеристик фильтров может быть легко получен, если незначительно модифицировать факторизацию решетчатой структуры ПУБФ с линейной ФЧХ. Действительно, достаточно ассоциировать U_i с V_i в (1.17) и представить матрицу U_i следующим образом [10]:

$$U_i = \Gamma V_i \Gamma, \quad i=0, \dots, N-2, \quad (1.20)$$

$$U_{N-1} = J_{M/2} V_{N-1} \Gamma,$$

где $J_{M/2}$ – обратная единичная матрица; Γ – диагональная матрица, элементы которой определяются как $\gamma_{mm} = (-1)^{m-1}$, $m=1, \dots, M/2$.

1.4. Регулярность и эффекты квантования коэффициентов

Для M -канального банка фильтров регулярность K -го порядка может быть определена как число нулей на частотах элайзинга $2\pi k/M$, $k=1, \dots, M-1$ фильтра нижних частот $H_0(z)$. Чтобы получить K -ый порядок регулярности, полифазная матрица $\mathbf{E}(z)$ должна удовлетворять условию [11]:

$$\frac{d^n}{dz^n} \left\{ \mathbf{E}(z^M) \left[1 \ z^{-1} \ \dots \ z^{-(M-1)} \right]^T \right\}_{z=1} = c_n \mathbf{e},$$

где $c_n \neq 0$ для $n=0, \dots, K-1$, $\mathbf{e} = [1 \ 0 \ 0 \ 0]^T$. Например, в 8-ми канальном банке фильтров ($M=8$) с регулярностью 1-го порядка ($K=1$) значения амплитудно-частотных характеристик всех канальных фильтров $H_k(z)$, $k=1, \dots, M-1$ равны нулю на нулевой частоте. Таким образом, постоянная составляющая на выходе банка фильтров локализуется фильтром нижних частот $H_0(z)$ и не «просачивается» в другие канальные фильтры, т.е. не проявляется так называемый эффект «DC leakage». В итоге, в обработанном изображении таким банком фильтров отсутствуют артефакты «шахматной доски» (checkerboard artifact).

К сожалению, описанные выше факторизации теряют свойство параунитарности, когда матрицы U_i и V_i представ-

ляются с ограниченной точностью [12]. Перфективная реконструкция не обеспечивается банком фильтров даже решетчатыми структурами с фиксированной запятой. Это обусловлено тем, что квантование компонент ортогональных матриц в общем случае приведет к изменению норм столбцов. Например, квантованная матрица оборотов Гивенса

$$\begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & Q(\cos \alpha) & -Q(\sin \alpha) & 0 \\ 0 & Q(\sin \alpha) & Q(\cos \alpha) & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

не ортогональна, так имеет две различные нормы столбцов: 1 и $\sqrt{Q^2(\cos \alpha) + Q^2(\sin \alpha)} \neq 1$, здесь Q – оператор квантования.

2. Кватернионы и ортогональные матрицы

2.1. Кватернионы

Кватернион – это гиперкомплексное число размерности 4 [13,14]

$$Q = q_1 + q_2 i + q_3 j + q_4 k, \quad q_1, q_2, q_3, q_4 \in \mathbb{R} \quad (2.1)$$

с одной реальной и тремя мнимыми частями. Мнимые единицы i, j, k соотносятся между собой согласно следующим уравнениям:

$$i^2 = j^2 = k^2 = ijk = -1; \quad ij = -ji = k; \quad jk = -kj = i;$$

$$ki = -ik = j. \quad (2.2)$$

Особенность умножения кватернионов в том, что оно не подчиняется переместительному закону умножения ($PQ \neq QP$, где P, Q – произвольные кватернионы), но подчиняется сочетательному. Сложение кватернионов осуществляется аналогично сложению комплексных чисел. Сопряженным для кватерниона вида (2.1) будет кватернион:

$$\bar{Q} = q_1 - q_2 i - q_3 j - q_4 k. \quad (2.3)$$

Модуль кватерниона определяется аналогично модулю комплексного числа:

$$|Q| = \sqrt{q_1^2 + q_2^2 + q_3^2 + q_4^2}. \quad (2.4)$$

Основные свойства модуля кватерниона также совпадают с подобными свойствами модуля комплексного числа, в частности:

$$Q \cdot \bar{Q} = |Q|^2; \quad |Q_1 \cdot Q_2|^2 = |Q_1|^2 \cdot |Q_2|^2. \quad (2.5)$$

Деление кватернионов определяется как умножение на обратную величину $Q^{-1} = \bar{Q}/|Q|^2$. При этом удовлетворяется тождество $QQ^{-1} = Q^{-1}Q = 1$. Полярная форма кватерниона определяется следующим образом [13]:

$$\begin{cases} q_1 = |Q| \cdot \cos \varphi \\ q_2 = |Q| \cdot \sin \varphi \cdot \cos \psi \\ q_3 = |Q| \cdot \sin \varphi \cdot \sin \psi \cdot \cos \chi \\ q_4 = |Q| \cdot \sin \varphi \cdot \sin \psi \cdot \sin \chi \end{cases}, \quad (2.6)$$

где φ, ψ, χ – произвольные значения углов ($0 \leq \varphi \leq \pi$, $0 \leq \psi \leq \pi$, $0 \leq \chi \leq 2\pi$). Полярное представление кватернионов сокращает число степеней свободы до трех и позволяет легко параметризовать кватернионы с фиксированными модулями, в частности, с единичными кватернионами, модуль которых равен 1.

2.2. Матрицы умножения

Операции сложения и умножения кватернионов также могут быть реализованы с использованием векторно-матричной арифметики. Для этого кватернионы представляются в виде четырехмерных векторов: $P=[p_1 \ p_2 \ p_3 \ p_4]^T$, $Q=[q_1 \ q_2 \ q_3 \ q_4]^T$. В этом случае сложение кватернионов будет производиться по правилу сложения векторов, а умножение кватернионов в векторной форме производится по правилу умножения вектора на матрицу:

$$R = PQ = \begin{bmatrix} r_1 \\ r_2 \\ r_3 \\ r_4 \end{bmatrix} = \begin{bmatrix} p_1 & -p_2 & -p_3 & -p_4 \\ p_2 & p_1 & -p_4 & p_3 \\ p_3 & p_4 & p_1 & -p_2 \\ p_4 & -p_3 & p_2 & p_1 \end{bmatrix} \times \begin{bmatrix} q_1 \\ q_2 \\ q_3 \\ q_4 \end{bmatrix} = \begin{bmatrix} p_1q_1 - p_2q_2 - p_3q_3 - p_4q_4 \\ p_2q_1 + p_1q_2 - p_4q_3 + p_3q_4 \\ p_3q_1 + p_4q_2 + p_1q_3 - p_2q_4 \\ p_4q_1 - p_3q_2 + p_2q_3 + p_1q_4 \end{bmatrix} = \begin{bmatrix} q_1p_1 - q_2p_2 - q_3p_3 - q_4p_4 \\ q_2p_1 + q_1p_2 + q_4p_3 - q_3p_4 \\ q_3p_1 - q_4p_2 + q_1p_3 + q_2p_4 \\ q_4p_1 + q_3p_2 - q_2p_3 + q_1p_4 \end{bmatrix} = \begin{bmatrix} q_1 & -q_2 & -q_3 & -q_4 \\ q_2 & q_1 & q_4 & -q_3 \\ q_3 & -q_4 & q_1 & q_2 \\ q_4 & q_3 & -q_2 & q_1 \end{bmatrix} \times \begin{bmatrix} p_1 \\ p_2 \\ p_3 \\ p_4 \end{bmatrix} \quad (2.7)$$

Данные выражения можно представить в более простом виде:

$$R = P \cdot Q = \mathbf{M}^+(P) \cdot Q = \mathbf{M}^-(Q) \cdot P, \quad (2.8)$$

где матрица $\mathbf{M}^+(P)$ – левый операнд умножения кватерниона P , а матрица $\mathbf{M}^-(Q)$ представляет собой правый операнд умножения кватерниона Q :

$$\mathbf{M}^+(P) = \begin{bmatrix} p_1 & -p_2 & -p_3 & -p_4 \\ p_2 & p_1 & -p_4 & p_3 \\ p_3 & p_4 & p_1 & -p_2 \\ p_4 & -p_3 & p_2 & p_1 \end{bmatrix},$$

$$\mathbf{M}^-(Q) = \begin{bmatrix} q_1 & -q_2 & -q_3 & -q_4 \\ q_2 & q_1 & q_4 & -q_3 \\ q_3 & -q_4 & q_1 & q_2 \\ q_4 & q_3 & -q_2 & q_1 \end{bmatrix} \quad (2.9)$$

Для изображения операндов умножения кватернионов на структурных схемах могут быть использованы следующие графические обозначения:

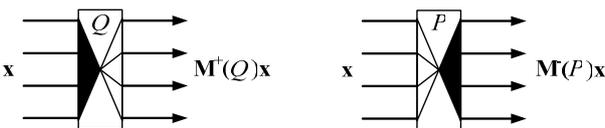


Рис. 2.1. Графические изображения операндов умножения кватернионов

Обе матрицы операндов умножения являются ортогональными, то есть справедливо условие:

$$\mathbf{M}^\pm(Q)^{-1} = \mathbf{M}^\pm(Q)^T \quad (2.10)$$

и имеют детерминант равный +1, что позволяет отнести их к группе специальных матриц размерностью 4x4, на-

зываемых SO(4) [17]. Для рассматриваемых матриц операторов умножения верны следующие равенства:

$$\mathbf{M}^+(Q_{N-1})\mathbf{M}^+(Q_{N-2})\dots\mathbf{M}^+(Q_0) = \mathbf{M}^+(Q_{N-1}Q_{N-2}\dots Q_0), \quad (2.11a)$$

$$\mathbf{M}^-(Q_{N-1})\mathbf{M}^-(Q_{N-2})\dots\mathbf{M}^-(Q_0) = \mathbf{M}^-(Q_{N-1}Q_{N-2}\dots Q_0). \quad (2.11b)$$

Существует другое интересное и полезное соотношение:

$$\mathbf{M}^\pm(\bar{Q}) = \mathbf{M}^\pm(Q)^T. \quad (2.12)$$

Соотношения между матрицами умножения кватернионов и произвольными 4 x 4 и 8 x 8 ортогональными матрицами является интригующим с точки зрения перспектив синтеза ПУБФ.

2.3. Параметризация ортогональных матриц

Для обоснования применения умножителей кватернионов в качестве составляющих блоков ПУБФ было выдвинуто и доказано несколько теорем, основными из которых являются следующие:

1) для каждой ортогональной матрицы A размерностью 4x4 существует уникальная (с точностью до знака) пара единичных кватернионов P и Q , таких что [16]:

$$A = \mathbf{M}^+(P) \cdot \mathbf{M}^-(Q) = \mathbf{M}^-(Q) \cdot \mathbf{M}^+(P), \quad (2.13)$$

2) когда блочно-диагональная матрица $\text{diag}(U, V)$ размерностью 8x8 образована из двух произвольных ортогональных матриц U и V размерностью 4x4, то существуют четыре единичных кватерниона P, Q, R, S , для которых справедливо соотношение [17,18]:

$$\text{diag}(U, V) = \text{diag}(\mathbf{M}^-(S), \mathbf{M}^-(S)) \cdot \text{diag}(\mathbf{M}^-(\bar{R}), \mathbf{M}^-(R)) \cdot \text{diag}(\mathbf{M}^+(\bar{Q}), \mathbf{M}^+(Q)) \cdot \text{diag}(\mathbf{M}^+(P), \mathbf{M}^+(P)). \quad (2.14)$$

В случае, когда оба блока диагональной матрицы равны ($U = V$), выражение (2.14) примет следующий вид:

$$\text{diag}(V, V) = \text{diag}(\mathbf{M}^-(Q), \mathbf{M}^-(Q)) \cdot \text{diag}(\mathbf{M}^+(P), \mathbf{M}^+(P)) \quad (2.15)$$

Если же блок U диагональной матрицы $\text{diag}(U, V)$ является единичной матрицей, то выражение (2.14) преобразуется к виду:

$$\text{diag}(I_4, V) = \text{diag}(\mathbf{M}^-(Q), \mathbf{M}^-(Q)) \cdot \text{diag}(\mathbf{M}^-(\bar{Q}), \mathbf{M}^-(Q)) \cdot \text{diag}(\mathbf{M}^+(\bar{P}), \mathbf{M}^+(P)) \cdot \text{diag}(\mathbf{M}^+(P), \mathbf{M}^+(P)). \quad (2.16)$$

2.4. Умножитель кватернионов как составной блок ПУБФ

Факторизация (2.13) всегда соответствуют ортогональному преобразованию сигналов, даже если ее компоненты квантованы. Это происходит потому, что столбцы каждой матрицы умножения кватернионов составлены из одних и тех же элементов с точностью до знака, например:

$$\begin{bmatrix} Q(q_1) & -Q(q_2) & -Q(q_3) & -Q(q_4) \\ Q(q_2) & Q(q_1) & -Q(q_4) & Q(q_3) \\ Q(q_3) & Q(q_4) & Q(q_1) & -Q(q_2) \\ Q(q_4) & -Q(q_3) & Q(q_2) & Q(q_1) \end{bmatrix}.$$

Видно, что как строчные, так и столбцовые нормы постоянны и равны $\sqrt{Q(q_1)^2 + Q(q_2)^2 + Q(q_3)^2 + Q(q_4)^2}$, но не обязательно единичны, невзирая на конечную точность представления матриц (или кватернионных коэффициентов). Таким образом, матрица A в (2.13) всегда представляет ортогональ-

ное преобразование, а умножение кватернионов может рассматриваться как основной строительный блок ПУБФ.

Произвольная 4 x 4 матрица требует хранения 16 действительных чисел и 32 чисел для блочно-диагональной матрицы 8 x 8. В данном же случае, необходимо запомнить в памяти 8 и 16 действительных чисел (2 и 4 кватерниона), соответственно. Один из недостатков кватернионной факторизации заключается в большом количестве вычислений, что является ценой за сохранение ортогональности. Вычислительная сложность факторизаций (2.13) в два раза, а (2.14) в четыре раза больше, чем простое умножение матрицы 4 x 4 на вектор. Однако, учитывая специфическую структуру матриц умножения кватернионов, можно выполнить данные операции только за 8 реальных умножений [19]. Реализация умножителей кватернионов на основе распределенной арифметики [20] или с использованием 4-х мерных алгоритмов CORDIC [21] хорошо согласуется со структурами FPGA.

3. Решетчатые структуры ПУБФ на основе кватернионов

3.1. Четырехканальный ПУБФ

Матрицы \mathbf{E}_0 и \mathbf{R}_i в факторизации (1.11) решетчатой структуры *общего параунитарного банка фильтров анализа*, как показано в [22], можно выразить через матрицы умножения кватернионов следующим образом:

$$\mathbf{E}_0 = \mathbf{M}^+(Q_0)\mathbf{M}^-(P_0), \quad (3.1)$$

$$\mathbf{R}_i = \mathbf{M}^\pm(Q_i), \quad i = 1, \dots, N-1, \quad (3.2)$$

где P_0 и все Q_i – некоторые единичные кватернионы, тогда передаточная функция решетчатой структуры *общего ПУБФ анализа* будет равна:

$$\mathbf{E}(z) = \mathbf{M}^\pm(Q_{N-1})\Lambda(z)\mathbf{M}^\pm(Q_{N-2})\Lambda(z)\dots \dots \mathbf{M}^\pm(Q_1)\Lambda(z)\mathbf{M}^\pm(Q_0)\mathbf{M}^-(P_0) \quad (3.3)$$

Как доказано в [22], если в (3.2) для всех $i = 1, \dots, N-1$ используются левые операнды умножения и значения кватернионов удовлетворяют равенству:

$$P_0 = \pm \frac{1}{2} \overline{O \cdot Q_{N-1} \cdot Q_{N-2} \cdot \dots \cdot Q_0}, \quad (3.4)$$

где O – кватернион, все части которого равны 1: $O = 1 + i + j + k$, то для решетчатой структуры ПУБФ (3.3) выполняется условие регулярности 1-го порядка.

Решетчатая структура 4-канального *общего ПУБФ анализа* (для $N = 3$), определяемая выражением (3.3), показана на рис. 3.1.

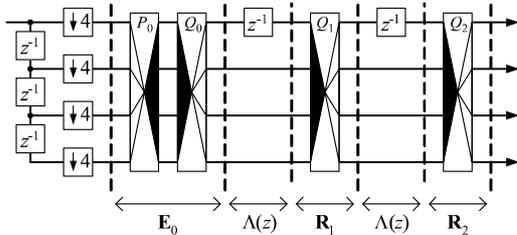


Рис. 3.1. Решетчатая структура 4-канального ПУБФ анализа на кватернионах

Выражение (1.10) задает связь между матрицами передаточных функций фильтров банков анализа и синтеза для обеспечения перфективной реконструкции.

Следовательно, требуемый банк фильтров синтеза может быть определен как:

$$\mathbf{D}(z) = \mathbf{E}(z)^{-1} c z^{-A} \mathbf{I}. \quad (3.5)$$

Подставив (3.3) в формулу (3.5) и учитывая соотношения (2.10) и (2.12) для матриц операндов умножения кватернионов, можно получить следующую факторизацию:

$$\mathbf{D}(z) = c z^{-A} \mathbf{M}^-(\overline{P_0})\mathbf{M}^+(\overline{Q_0})\Lambda(z)^{-1}\mathbf{M}^\pm(\overline{Q_1})\dots \dots \Lambda(z)^{-1}\mathbf{M}^\pm(\overline{Q_{N-1}})\Lambda(z)^{-1}\mathbf{M}^\pm(\overline{Q_N})$$

где $\overline{Q_i}, \overline{P_0}$ – кватернионы, комплексно-сопряженные к Q_i, P_0 . Приняв константу c равной 1 (для сохранения масштаба) и использовав следующие замены:

$$\hat{\Lambda}(z) = \Lambda(z)^{-1} z^{-1} \mathbf{I}_2, \quad (3.6)$$

$$\hat{\mathbf{E}}_0 = \mathbf{M}^-(\overline{P_0})\mathbf{M}^+(\overline{Q_0}), \quad (3.7)$$

$$\hat{\mathbf{R}}_i = \mathbf{M}^\pm(\overline{Q_i}), \quad i = 1, \dots, N-1, \quad (3.8)$$

окончательная факторизация для ПУБФ синтеза будет выглядеть следующим образом:

$$\mathbf{D}(z) = \mathbf{M}^-(\overline{P_0})\mathbf{M}^+(\overline{Q_0})\hat{\Lambda}(z)\mathbf{M}^\pm(\overline{Q_1})\dots \dots \hat{\Lambda}(z)\mathbf{M}^\pm(\overline{Q_{N-2}})\hat{\Lambda}(z)\mathbf{M}^\pm(\overline{Q_{N-1}}) = \hat{\mathbf{E}}_0 \hat{\Lambda}(z) \hat{\mathbf{R}}_1 \dots \hat{\Lambda}(z) \hat{\mathbf{R}}_{N-2} \hat{\Lambda}(z) \hat{\mathbf{R}}_{N-1}. \quad (3.9)$$

Факторизация (3.9) описывает решетчатую структуру ПУБФ синтеза (для $N = 3$), показанную на рис. 3.2.

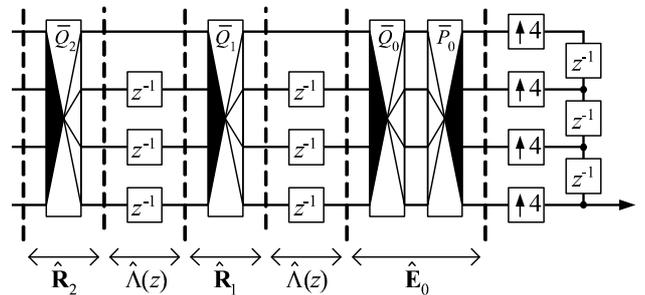


Рис. 3.2. Решетчатая структура 4-канального банка фильтров синтеза на кватернионах

Реализация *параунитарного банка фильтров анализа с линейной ФЧХ* (1.12) на основе умножителей кватернионов предполагает определение Φ_0 и Φ_i как [22]:

$$\Phi_0 = \mathbf{M}^-(P_0)\mathbf{M}^+(Q_0), \quad (3.10)$$

$$\Phi_i = \mathbf{M}^-(P_i), \quad i = 1, \dots, N-1, \quad (3.11)$$

где все P_i и Q_0 – некоторые единичные кватернионы. Если в (3.10) и (3.11)

$$Q_0 = \pm \frac{1}{\sqrt{2}} \overline{P_0 \cdot P_1 \cdot \dots \cdot P_{N-1} \cdot A}, \quad (3.12)$$

то факторизация (1.12), (3.10) – (3.12) соответствует регулярности 1-го порядка. Здесь A – специальный кватернион: $A = 1 + i$.

Решетчатая структура ПУБФ анализа с линейной ФЧХ на основе умножителей кватернионов (для $N = 3$) представлена на рис. 3.3.

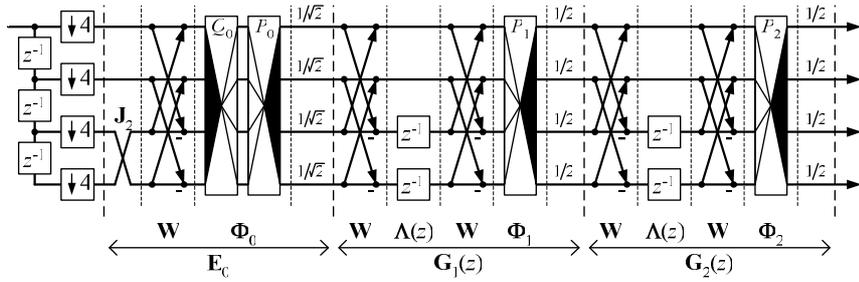


Рис. 3.3. Решетчатая структура 4-канального банка фильтров анализа с линейной ФЧХ на кватернионах

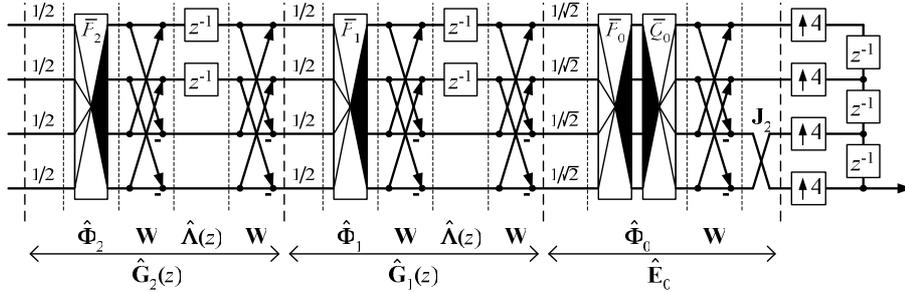


Рис. 3.4. Решетчатая структура 4-канального банка фильтров синтеза с линейной ФЧХ на кватернионах

Факторизация ПУБФ синтеза на кватернионах задается следующим образом:

$$\mathbf{D}(z) = \hat{\mathbf{E}}_0 \hat{\mathbf{G}}_1 \dots \hat{\mathbf{G}}_{N-2} \hat{\mathbf{G}}_{N-1}, \quad (3.13)$$

$$\hat{\mathbf{E}}_0 = \text{diag}(\mathbf{I}_2, \mathbf{J}_2) \mathbf{W} \hat{\Phi}_0 \frac{1}{\sqrt{2}}, \quad (3.14)$$

$$\hat{\mathbf{G}}_i(z) = \mathbf{W} \hat{\Lambda}(z) \mathbf{W} \hat{\Phi}_i \frac{1}{2}, \quad i = 1, \dots, N-1, \quad (3.15)$$

$$\hat{\Lambda}(z) = \text{diag}(z^{-1} \mathbf{I}_2, \mathbf{I}_2), \quad (3.16)$$

$$\hat{\Phi}_0 = \mathbf{M}^+(\bar{Q}_0) \mathbf{M}^-(\bar{P}_0), \quad (3.17)$$

$$\hat{\Phi}_i = \mathbf{M}^-(\bar{P}_i), \quad i = 1, \dots, N-1. \quad (3.18)$$

Решетчатая структура 4-канального ПУБФ синтеза с линейной ФЧХ, соответствующая факторизации (3.13) – (3.18) (для $N=3$), изображена на рис. 3.4.

Решетчатая структура на кватернионах ПУБФ с линейной ФЧХ и попарно зеркальной симметрией частотных характеристик фильтров получается в результате факторизации (1.12), (1.20) для соответственно определенных матриц Φ_i и Φ_{N-1} [22]:

$$\Phi_i = \mathbf{M}^-(P_i), \quad i = 0, \dots, N-2, \quad (3.19)$$

$$\Phi_{N-1} = \mathbf{M}^-(P_{N-1}) \text{diag}(\mathbf{J}_2 \Gamma, \mathbf{I}_2), \quad (3.20)$$

где $\Gamma = \text{diag}(1, -1)$ и P_i – некоторые единичные кватернионы, выбор которых в соответствии с

$$P_{N-1} = \pm \frac{1}{\sqrt{2}} \overline{A \cdot P_0 \cdot P_1 \dots P_{N-2}} \quad (3.21)$$

обеспечивает регулярность первого порядка банку фильтров.

Соответствующая факторизация передаточной функции банка фильтров синтеза задается следующим образом:

$$\mathbf{D}(z) = \hat{\mathbf{E}}_0 \hat{\mathbf{G}}_1 \dots \hat{\mathbf{G}}_{N-2} \hat{\mathbf{G}}_{N-1}, \quad (3.22)$$

$$\hat{\mathbf{E}}_0 = \text{diag}(\mathbf{I}_2, \mathbf{J}_2) \mathbf{W} \hat{\Phi}_0 \frac{1}{\sqrt{2}}, \quad (3.23)$$

$$\hat{\mathbf{G}}_i(z) = \mathbf{W} \hat{\Lambda}(z) \mathbf{W} \hat{\Phi}_i \frac{1}{2}, \quad i = 1, \dots, N-1, \quad (3.24)$$

$$\hat{\Lambda}(z) = \text{diag}(z^{-1} \mathbf{I}_2, \mathbf{I}_2), \quad (3.25)$$

$$\hat{\Phi}_{N-1} = \text{diag}(\Gamma \mathbf{J}_2, \mathbf{I}_2) \mathbf{M}^-(\bar{P}_{N-1}), \quad (3.26)$$

$$\hat{\Phi}_i = \mathbf{M}^-(\bar{P}_i), \quad i \text{ д\o} i = 0, \dots, N-2. \quad (3.27)$$

Решетчатые структуры ПУБФ анализа и синтеза с попарно зеркальной симметрией частотных характеристик фильтров приведены на рис. 3.5 и 3.6 соответственно ($N=3$).

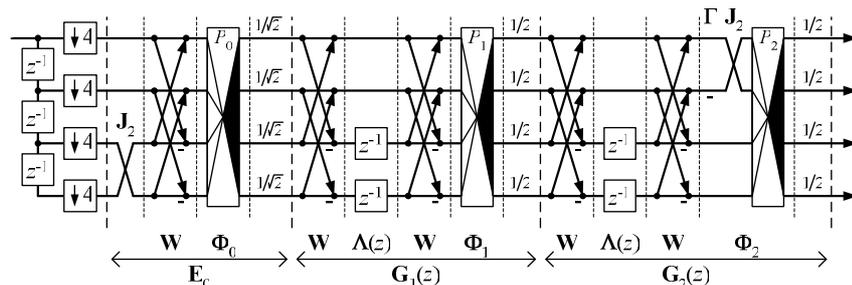


Рис. 3.5. Решетчатая структура 4-канального банка фильтров анализа с попарно зеркальной симметрией частотных характеристик фильтров на кватернионах

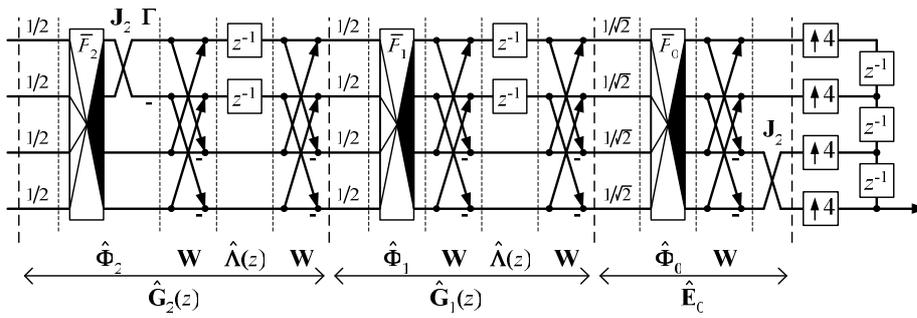


Рис. 3.6. Решетчатая структура 4-канального банка фильтров синтеза с попарно зеркальной симметрией частотных характеристик фильтров на кватернионах

3.2. Восьмиканальный ПУБФ

Вариант кватернионной факторизации (1.12) – (1.18) для 8-ми канального ПУБФ с линейной ФЧХ анализа определяется следующими подстановками [5,17]:

$$\Phi_i = \text{diag}(\mathbf{M}^-(\bar{R}_i), \mathbf{M}^-(R_i)) \cdot \text{diag}(\mathbf{M}^+(\bar{Q}_i), \mathbf{M}^+(Q_i)), \quad (3.28)$$

äëý $i=1, \dots, N-1$

$$\Phi_0 = \text{diag}(\mathbf{M}^-(S_0), \mathbf{M}^-(S_0)) \cdot \text{diag}(\mathbf{M}^-(\bar{R}_0), \mathbf{M}^-(R_0)) \cdot \text{diag}(\mathbf{M}^+(\bar{Q}_0), \mathbf{M}^+(Q_0)) \cdot \text{diag}(\mathbf{M}^+(P_0), \mathbf{M}^+(P_0)), \quad (3.29)$$

где S_0, P_0 и все Q_i, R_i – единичные кватернионы. Данный ПУБФ (рис. 3.7) будет банком фильтров с регулярностью первого порядка, если

$$R_{N-1} = \frac{1}{2} \bar{Q}_{N-1} \dots \bar{Q}_0 \cdot P_0 \cdot O \cdot \bar{R}_0 \cdot S_0 \cdot \bar{R}_1 \dots \bar{R}_{N-2}. \quad (3.30)$$

Для соответствующего 8-канального банка синтеза (рис. 3.8) в факторизацию (3.13) – (3.18) вводятся следующие подстановки:

$$\hat{\Phi} = \text{diag}(\mathbf{M}^+(Q_i), \mathbf{M}^+(Q_i)) \cdot \text{diag}(\mathbf{M}^-(R_i), \mathbf{M}^-(R_i)), \quad i=1, \dots, N-1. \quad (3.31)$$

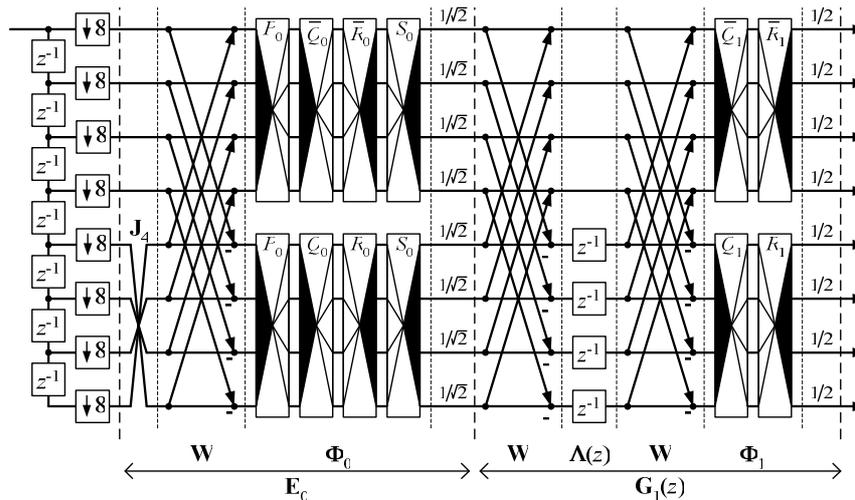


Рис. 3.7. Решетчатая структура 8-канального банка фильтров анализа с линейной ФЧХ на кватернионах (для $N=2$)

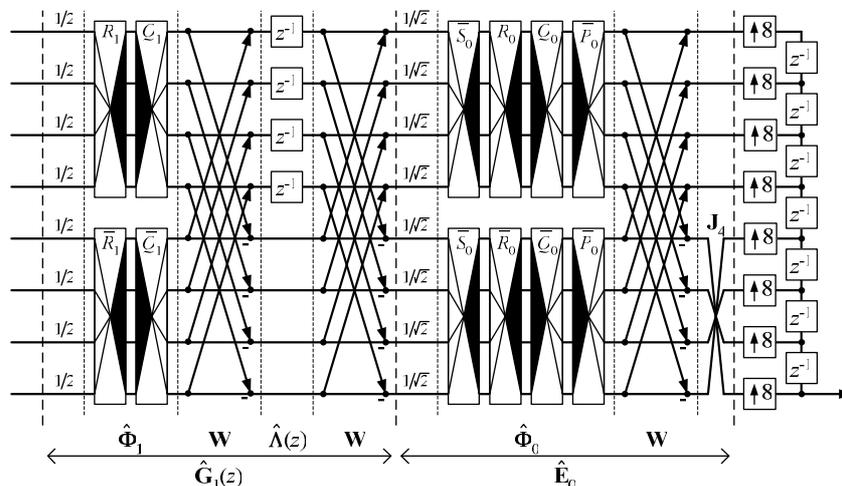


Рис. 3.8. Решетчатая структура 8-канального банка фильтров синтеза с линейной ФЧХ на кватернионах (для $N=2$)

$$\hat{\Phi}_0 = \text{diag}(\mathbf{M}^+(\bar{P}_0), \mathbf{M}^+(\bar{P}_0)) \cdot \text{diag}(\mathbf{M}^+(Q_0), \mathbf{M}^+(\bar{Q}_0)) \cdot \text{diag}(\mathbf{M}^-(R_0), \mathbf{M}^-(\bar{R}_0)) \cdot \text{diag}(\mathbf{M}^-(S_0), \mathbf{M}^-(\bar{S}_0)), \quad (3.32)$$

Решетчатые структуры 8-ми канальных ПУБФ с линейной ФЧХ и попарно зеркальной симметрией частотных характеристик фильтров (1.12) – (1.20) на квантаренионах получаются для соответственно определенных матриц Φ_i и Φ_{N-1} [5,17]:

для банка анализа (рис. 3.9):

$$\Phi_i = \text{diag}(\Gamma, \mathbf{I}_4) \cdot \text{diag}(\mathbf{M}^-(Q_i), \mathbf{M}^-(Q_i)) \cdot \text{diag}(\mathbf{M}^+(P_i), \mathbf{M}^+(P_i)) \cdot \text{diag}(\Gamma, \mathbf{I}_4), \quad \text{для } i = 0, \dots, N-2, \quad (3.33)$$

$$\Phi_{N-1} = \text{diag}(\mathbf{J}_4, \mathbf{I}_4) \cdot \text{diag}(\mathbf{M}^-(Q_{N-1}), \mathbf{M}^-(Q_{N-1})) \cdot \text{diag}(\mathbf{M}^+(P_{N-1}), \mathbf{M}^+(P_{N-1})) \cdot \text{diag}(\Gamma, \mathbf{I}_4) \quad (3.34)$$

где все Q_i, P_i – единичные кватернионы, выбор которых в соответствии с соотношением

$$P_{N-1} = \frac{1}{2} \hat{E} \bar{Q}_{N-1} \dots \bar{Q}_0 \cdot P_0 \cdot \bar{Q}_0 \cdot \bar{P}_0 \cdot \bar{P}_1 \dots \bar{P}_{N-2} \quad (3.35)$$

обеспечивает регулярность первого порядка банку фильтров;

для банка синтеза (рис. 3.10):

$$\hat{\Phi}_i = \text{diag}(\Gamma, \mathbf{I}_4) \cdot \text{diag}(\mathbf{M}^+(\bar{P}_i), \mathbf{M}^+(\bar{P}_i)) \cdot \text{diag}(\mathbf{M}^-(\bar{Q}_i), \mathbf{M}^-(\bar{Q}_i)) \cdot \text{diag}(\Gamma, \mathbf{I}_4), \quad \text{для } i = 0, \dots, N-2 \quad (3.36)$$

$$\hat{\Phi}_{N-1} = \text{diag}(\Gamma, \mathbf{I}_4) \cdot \text{diag}(\mathbf{M}^+(\bar{P}_{N-1}), \mathbf{M}^+(\bar{P}_{N-1})) \cdot \text{diag}(\mathbf{M}^-(\bar{Q}_{N-1}), \mathbf{M}^-(\bar{Q}_{N-1})) \cdot \text{diag}(\mathbf{J}_4, \mathbf{I}_4) \quad (3.37)$$

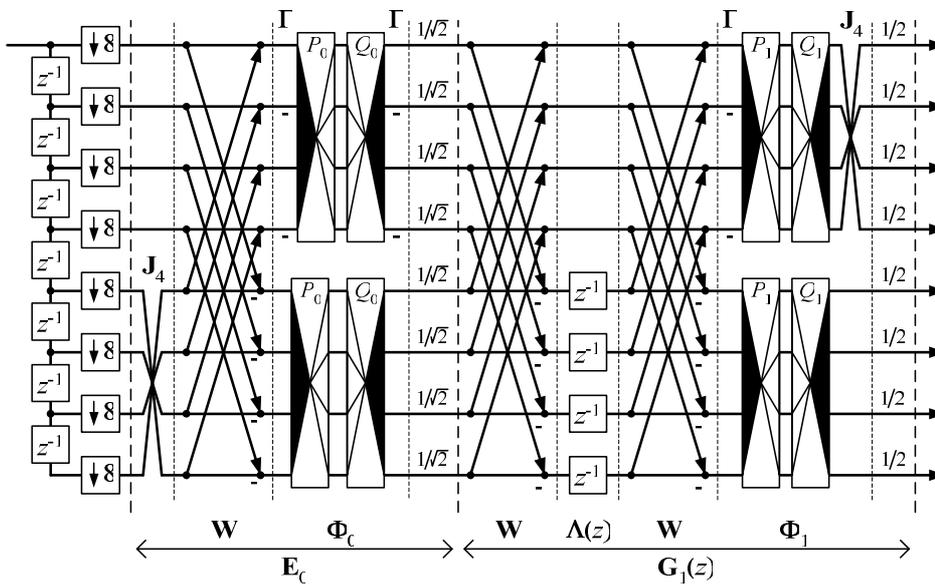


Рис. 3.9. Решетчатая структура 8-канального банка фильтров анализа с попарно зеркальной симметрией частотных характеристик фильтров на кватернионах (для $N = 2$)

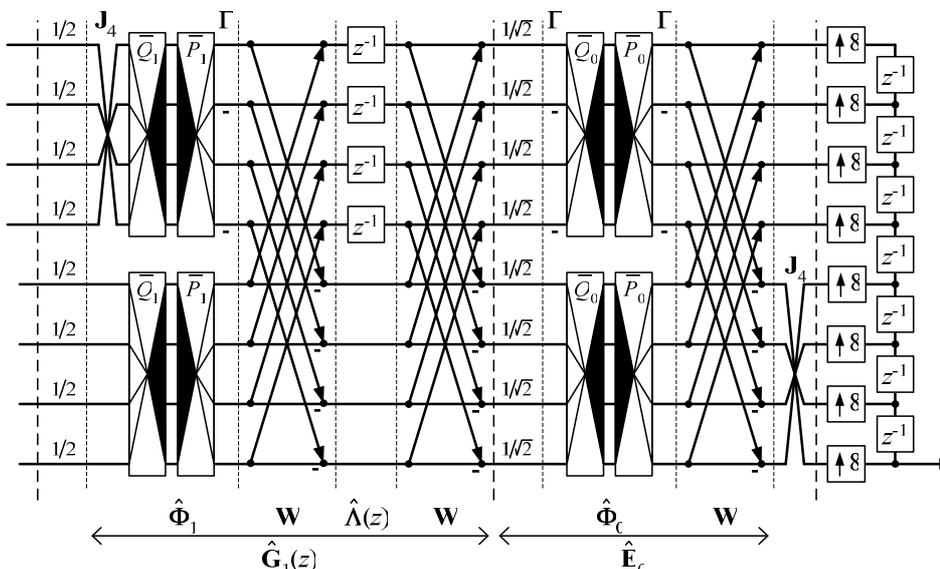


Рис. 3.10. Решетчатая структура 8-канального банка фильтров синтеза с попарно зеркальной симметрией частотных характеристик фильтров на кватернионах (для $N = 2$)

4. Примеры синтеза

4.1. Процедура расчета

Для получения ПУБФ с высоким коэффициентом эффективности кодирования (coding gain) и хорошей частотной избирательностью кватернионные коэффициенты оптимизируются путем минимизации двух показателей:

- степень ослабления в полосе задержки каналов, выраженная в терминах энергии:

$$\mathcal{E}_{SBE} = \sum_{k=0}^{M-1} \int_{\omega \in \Omega_k} |H_k(e^{j\omega})|^2 d\omega,$$

где Ω_k определяют полосу задержки k -го фильтра банка фильтров;

- коэффициент эффективности кодирования:

$$CG = 10 \lg \frac{\frac{1}{M} \sum_{k=0}^{M-1} \sigma_{x_k}^2}{\left(\prod_{k=0}^{M-1} \sigma_{x_k}^2 \right)^{\frac{1}{M}}},$$

где $\sigma_{x_k}^2$ - дисперсия сигнала в k -ом канале банка фильтров, значение которой для всех M каналов определяются через диагональные элементы автокорреляционной матрицы \mathbf{R}_{yy} выходного сигнала $y(n)$: $\sigma_{x_k}^2 = [\mathbf{R}_{yy}]_{kk}$. Автокорреляционная матрица $\mathbf{R}_{yy} = \mathbf{H} \mathbf{R}_{xx} \mathbf{H}^T$. Здесь \mathbf{R}_{xx} - автокорреляционная матрица входного сигнала $x(n)$, а \mathbf{H} - матрица преобразования, сформированная из импульсных характеристик фильтров банка следующим образом: $[\mathbf{H}]_{kn} = h_k(L-1-n)$, где $k = 0, \dots, M-1$, $n = 0, \dots, L-1$. В рассматриваемых ниже примерах синтеза ПУБФ матрица \mathbf{R}_{xx} конструировалась для входного сигнала $x(n)$, который формировался с помощью авторегрессионной

модели первого порядка AR(1) с единичной дисперсией и коэффициентом корреляции 0,95. Такая модель сигнала соответствует только обработке естественных изображений, для других типов сигналов потребуются другие модели.

Кватернионные коэффициенты решетчатых структур ПУБФ, приведенных выше, для процедуры синтеза представляются в полярной форме (2.6). Для решения задачи оптимизации (минимизация \mathcal{E}_{SBE} и CG) используются программы `fminsearch` и `fminunc` из пакета MATLAB. Далее, полученные коэффициенты конвертируются в рациональные числа с целью ликвидации операции умножения при аппаратной или программной реализации банка фильтров.

4.2. Четырехканальные ПУБФ

Результаты синтеза *общего 4-х канального ПУБФ* (3.3) с длиной фильтра равной 8 показаны ниже. Коэффициенты решетчатой структуры банка фильтров приведены в табл. 4.1, а соответствующая АЧХ - на рис. 4.1. Для данного банка фильтров коэффициент эффективности кодирования $CG = 8.1227$ дБ, и минимальное ослабление в полосе задержки составляет -20дБ, что вполне согласуется с результатами из [23]. Из анализа АЧХ (рис. 4.1) следует, что полученный ПУБФ действительно имеет регулярность 1-го порядка: кроме фильтра нижних частот все фильтры имеют нулевое значение АЧХ на частоте DC. С другой стороны, на комплексной плоскости корней (рис. 4.2) видно, что фильтр нижних частот имеет один ноль в каждой точке на единичной окружности, соответствующей зеркально элайзинговой частоте.

Таблица 4.1

Коэффициенты	Re(·)	Im _i (·)	Im _j (·)	Im _k (·)	Длина слова (бит)
r_0	-45/128	9/16	31/128	-5/8	8
s_0	-11/16	-1/2	1/16	7/16	5
s_1	3/8	1/8	3/4	-1/2	4

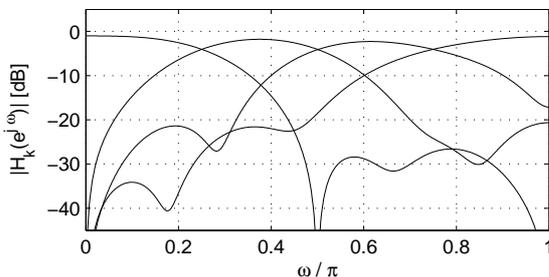


Рис. 4.1. АЧХ 4-х канальный ПУБФ

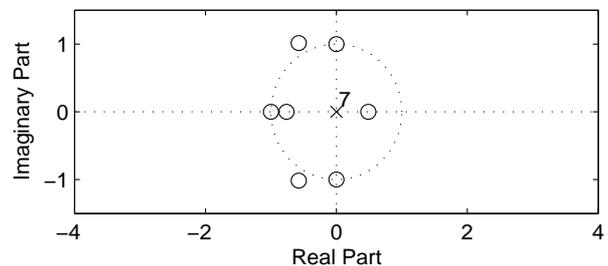


Рис. 4.2. Плоскость корней ПУБФ

Таблица 4.2

Коэффициенты	Re(·)	Im _i (·)	Im _j (·)	Im _k (·)	Длина слова (бит)
s_0	-231/512	459/1024	0	0	11
r_0	-7/8	-3/8	0	0	4
r_1	-3/16	15/16	0	0	5
r_2	-9/16	-13/16	0	0	5

Другой пример расчета показывает эффективность предложенного подхода построения ПУБФ на кватернионах с линейной ФЧХ. В табл. 4.2 приведены коэффициенты банка фильтров (длина фильтра – 12), а на рис. 4.3 иллюстрируется АЧХ данного ПУБФ, из которого видно, что банк также имеет регулярность первого рода. Коэффициент эффективности кодирования и минимальное ослабление в полосе затухания соответственно равны: 8.1845 дБ и -20дБ.

В табл. 4.3 и на рис. 4.5 и рис. 4.6 иллюстрируются результаты расчета 4-х канального ПУБФ с попарно зеркальной симметрией частотных характеристик фильтров на кватернионах (длина фильтра – 12), которые очень близки к результатам предыдущего примера: $CG= 8.1699$ дБ, минимальное ослабление в полосе затухания – 20дБ. Это объясняется схожестью расположения нулей на плоскости корней (сравни рис 4.4 и рис. 4.6)

Таблица 4.3

Коэффициенты.	$Re(\cdot)$	$Im_i(\cdot)$	$Im_j(\cdot)$	$Im_k(\cdot)$	Длина слова (бит)
r_0	7/8	3/8	0	0	4
r_1	3/16	-1	0	0	5
r_2	-17/128	43/64	0	0	8

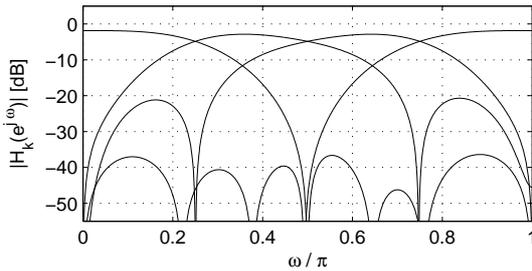


Рис. 4.3. ПУБФ с линейной ФЧХ

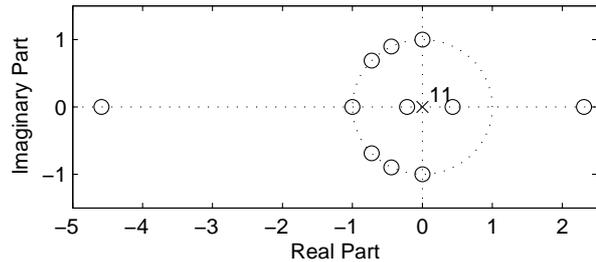


Рис. 4.4. Плоскость корней ПУБФ

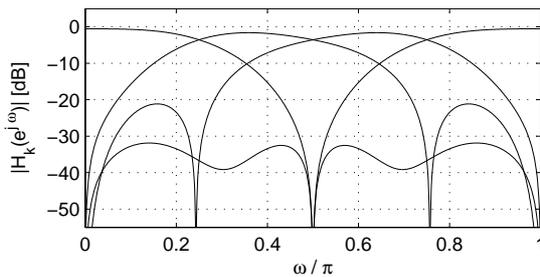


Рис. 4.5. АЧХ ПУБФ с попарно зеркальной симметрией частотных характеристик фильтров

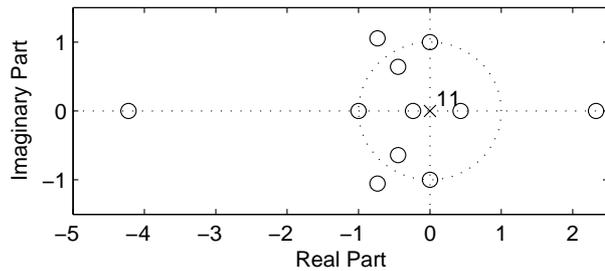


Рис. 4.6. Плоскость корней ПУБФ

4.3. Восьми канальные ПУБФ

В данном параграфе рассматриваются результаты двух примеров расчета 8-канальных ПУБФ как для арифметики с неограниченной точностью, так и для арифметики с фиксированной запятой: с линейной ФЧХ и с попарно зеркальной симметрией частотных характеристик фильтров. В обоих случаях факторизация содержит три ступени при длине фильтра 24.

Для первого банка фильтров с линейной ФЧХ и коэффициентами (см. Табл 4.4), представленными в арифметике с неограниченной точностью, коэффициент эффективности кодирования $CG= 9.3747$ дБ, а минимальное ослабление в полосе затухания равно – 21дБ, что видно из рис. 4.7.

В табл. 4.5 приведены коэффициенты второго банка фильтров, которые являются квантованными коэффициентами первого банка фильтров. Как видно из табл. 4.5, реализация данного ПУБФ не потребует умножителей. При этом, $CG= 9.3040$ дБ практически остался прежним, однако минимальное ослабление в полосе затухания уменьшилось до -13дБ, как показано на рис. 4.9.

Следует отметить, оба банка являются банками с регулярностью первого порядка, потому что их кватернионные коэффициенты (см. Табл. 4.4 и табл. 4.5) соответствуют условию (3.30). Анализ АЧХ (рис. 4.9 и рис 4.7) показывает, что в данных ПУБФ не проявляется эффект «DC leakage» (см. Раз-

дел 1.4). На комплексной плоскости корней (рис. 4.8) видно, что фильтр нижних частот имеет один ноль в каждой точке на единичной окружности, соответствующей зеркально элайзинговым частотам.

Синтез второго банка фильтров с попарно зеркальной симметрией частотных характеристик фильтров приводится сначала в арифметике с неограниченной точностью (табл. 4.6): коэффициент эффективности кодирования $CG= 9.3655$ дБ, минимальное ослабление в полосе затухания равно – 22дБ (рис. 4.10), а затем коэффициенты данного банка представляются в арифметике с фиксированной запятой – в рациональных числах (табл. 4.7), что позволяет избежать умножителей при практической реализации банка.

Квантование кватернионных коэффициентов (табл. 4.7), в данном случае, привело к незначительному уменьшению коэффициента эффективности кодирования $CG= 9.2549$ дБ и значения минимального ослабления в полосе затухания - -19 дБ (рис. 4.12). Учет условия (3.35) при синтезе банка фильтров с попарно зеркальной симметрией частотных характеристик фильтров гарантирует отсутствие искажений, обусловленных постоянной составляющей, а также анализ расположения нулей на комплексной плоскости корней показывает их корректность локализации (рис. 4.11).

Таблица 4.4.

Коэффициенты	Re(\cdot)	Im _i (\cdot)	Im _j (\cdot)	Im _k (\cdot)
s_0	-0.0297993	0.0669165	0.8901262	-0.449788
p_0	-0.2784728	-0.5164164	0.6659689	-0.4607087
q_0	-0.9687160	-0.0713704	-0.2281805	0.0665523
q_1	0.2571736	-0.2085346	-0.8265446	0.4551912
q_2	0.8835067	-0.2654561	0.0588609	-0.3814242
r_0	-0.1264641	0.6558758	-0.1001648	-0.7374285
r_1	0.9928326	0.0690029	-0.0687024	-0.0692971
r_2	0.9069558	0.3701135	-0.1523974	-0.1312336

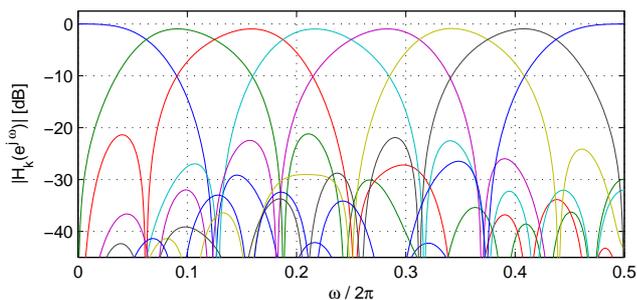


Рис. 4.7. АЧХ 8-канальных ПУБФ с линейной ФЧХ

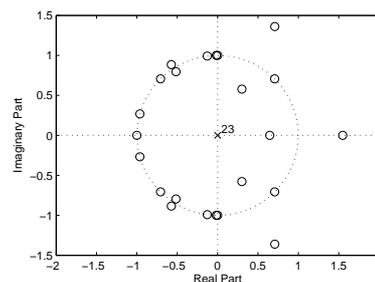


Рис. 4.8. Плоскость корней ПУБФ

Таблица 4.5

Коэффициенты	Re(\cdot)	Im _i (\cdot)	Im _j (\cdot)	Im _k (\cdot)	Длина слова (бит)
s_0	1/2	0	7/8	-1/8	4
p_0	-1/4	-1/2	5/8	-1/2	4
q_0	-15/16	-1/16	-1/4	0	5
q_1	1/4	-1/4	-7/8	1/2	4
q_2	7/8	-1/4	1/8	-3/8	4
r_0	-1/16	15/16	1/4	-1/4	5
r_1	15/16	1/16	-1/16	-1/16	5
r_2	$3820719/2^{22}$	$611699/2^{21}$	$-54443/2^{19}$	$-435031/2^{22}$	23

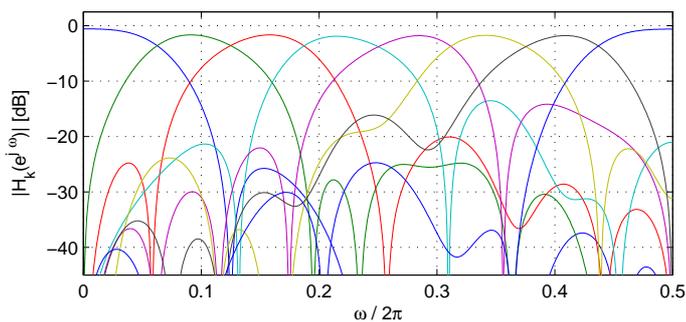


Рис. 4.9 АЧХ 8-канальных ПУБФ с линейной ФЧХ с квантованными коэффициентами

Таблица 4.6.

Коэффициенты	Re(\cdot)	Im _i (\cdot)	Im _j (\cdot)	Im _k (\cdot)
p_0	-0.0546530	-0.8381704	-0.2287760	0.4920823
p_1	0.9557711	0.0614376	0.1436822	0.2491636
p_2	-0.6224147	0.0345565	-0.5578179	-0.5479461
q_0	0.9916213	0.0606809	-0.0625301	-0.0953675
q_1	-0.0144669	0.6629532	0.1228110	0.7383774
q_2	-0.9805812	-0.0261340	0.0941266	0.1700519

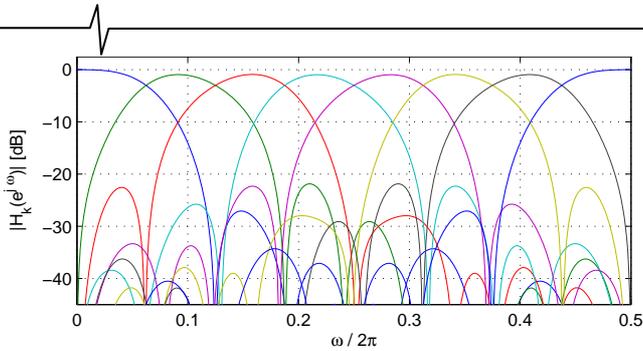


Рис. 4.10. АЧХ 8-канальных ПУБФ с попарно зеркальной симметрией частотных характеристик фильтров

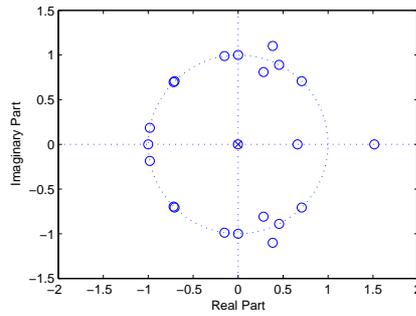


Рис. 4.11. Комплексной плоскости корней ПУБФ

Таблица 4.7

Коэффициенты	Re(·)	Im _i (·)	Im _j (·)	Im _k (·)	Длина слова (бит)
p_0	-1/8	-7/8	-1/4	1/2	4
p_1	7/8	0	1/8	1/4	4
p_2	$-4307/2^{13}$	$1349/2^{14}$	$-8563/2^{14}$	$-3261/2^{13}$	15
q_0	7/8	0	0	-1/8	4
q_1	0	5/8	1/8	3/4	4
q_2	-1	-1/16	1/16	1/8	5

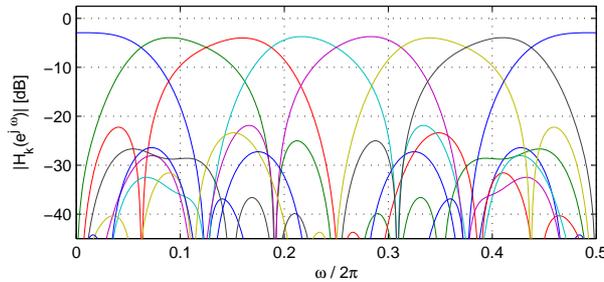


Рис. 4.12. АЧХ 8-канальных ПУБФ с попарно зеркальной симметрией частотных характеристик фильтров

5. Функциональная реализация параунитарного банка фильтров

5.1. Умножитель кватернионов на распределенной арифметике

В последнее время стремительно развиваются технологии проектирования специализированных встраиваемых процессоров реального времени на основе FPGA [24], главным преимуществом которых является возможность создавать встраиваемые процессоры реального времени, где возможен компромисс между производительностью и потребляемой мощностью. В данной работе рассматривается задача построения FPGA-ориентированных поточных процессоров ПУБФ как анализа, так и синтеза на умножителях кватернионов, которые реализованы на распределенной арифметике [20,25]. Исследуются характеристики данных процессоров: производительность, потребляемая мощность в зависимости от степени параллелизма в умножителях кватернионов.

Базовой операцией умножения кватернионов на правый операнд $M(Q)$ является вычисление суммы вида:

$$r_i = \sum_{k=1}^4 q_{ik} p_k = q_{i1} p_1 + q_{i2} p_2 + q_{i3} p_3 + q_{i4} p_4, i = 1, \dots, 4,$$

где q_{ik} – постоянные коэффициенты. Исходя из того, что обрабатываемые числа являются нормированными (выполняется условие $|p_k| \leq 1$) и представлены в дополнительном коде, то p_k можно записать как:

$$p_k = -b_{k0} + \sum_{n=1}^{B-1} b_{kn} 2^{-n},$$

где b_{kn} – значения битов (0 или 1); b_{k0} – знаковый разряд; B – разрядность входных чисел; $b_{k,B-1}$ – младший бит (LSB). Введем новые переменные:

$$c_{kn} = b_{kn} - \bar{b}_{kn}, c_{k0} = -(b_{k0} - \bar{b}_{k0}),$$

где $c_{kn}, c_{k0} \in \{-1, +1\}$; $\bar{b}_{kn}, \bar{b}_{k0}$ – инверсные значения битов b_{kn}, b_{k0} , тогда используя методику, изложенную в [18], продукт умножения кватернионов приводится к виду:

$$r_i = \sum_{n=0}^{(B/L)-1} F_i(b_{kn}) 2^{-nL} + 2^{-(B-L)} F_i(0),$$

где L – количество одновременно обрабатываемых битов;

$$F_i(b_{kn}) = \sum_{k=1}^4 \sum_{l=0}^{L-1} \frac{1}{2} q_{ik} 2^{-l} c_{k,nL+l}, F_i(0) = -2^{-L} \sum_{k=1}^4 q_{ik} -$$

суммы коэффициентов q_{ik} , рассчитываемые предварительно. Данное выражение получения продукта умножения кватернионов на основе распределенной арифметики описывает возможность одновременной обработки L битов входных данных (при этом L должно быть кратно B). Подобная реализация арифметических операций на распределенной арифметике получила название L-BAAT (L-bit-at-a-time). Структурная схема умножителя кватернионов по принципу L-BAAT приведена на рис. 5.1. Наличие блока коррекции переполнения вызвано

наличием дополнительного разряда в частичной сумме (разрядность $B+1$), что требует коррекции результата к разрядности B . Блок контроля нулевых данных предназначен для анализа входных данных на равенство нулю,

так как в таком случае результат будет нулевым. Усредненные результаты синтеза структур умножителей кватернионов по схеме L-BAAT на основе FPGA Xilinx Virtex v400-4 приведены в табл. 5.1.

Таблица 5.1

Результаты синтеза структур умножителей кватернионов по схеме L-BAAT

Параметры	Реализация L-BAAT				
	1-BAAT	2-BAAT	4-BAAT	8-BAAT	16-BAAT
Function Generators (LUTs)	219	339	561	958	1637
CLB Slices	110	170	281	479	819
Dffs or Latches	205	207	212	223	246
Макс. частота работы умножителя, МГц	58,1	56,0	40,6	29,8	18,6

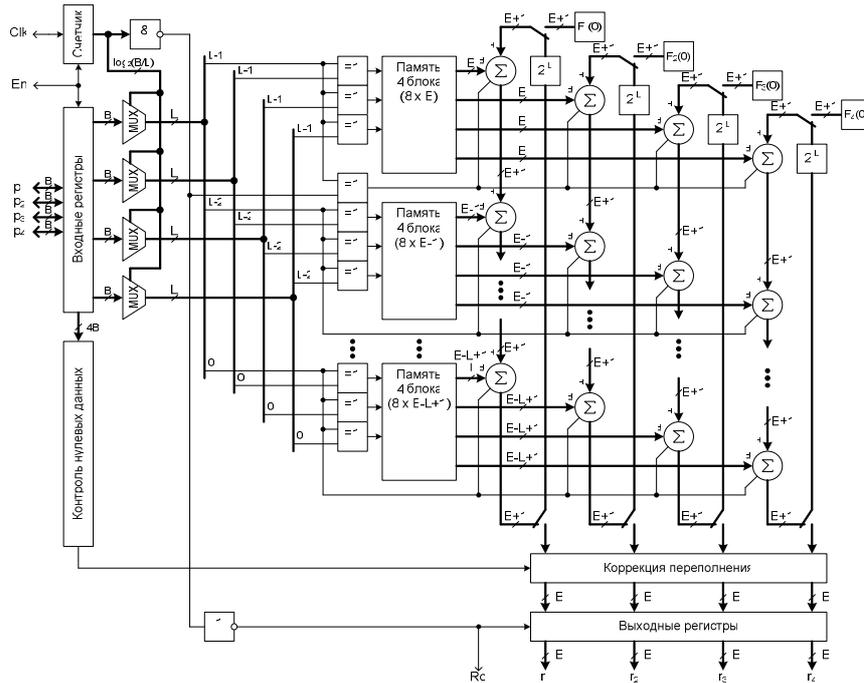


Рис. 5.1. Структурная схема умножителя кватернионов по схеме L-BAAT

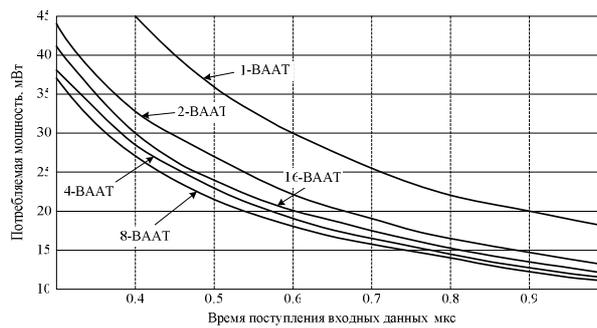


Рис. 5.2. Потребляемая мощность умножителем кватернионов на FPGA Xilinx Virtex v400-4

Таблица 5.2

Результаты синтеза поточных процессоров ПУБФ

Параметр	Значения					
	Банк анализа			Банк синтеза		
	1-BAAT	2-BAAT	4-BAAT	1-BAAT	2-BAAT	4-BAAT
Function Generators	1228	1863	2969	1298	1938	3064
CLB Slices	614	932	1485	702	969	1532
Dffs or Latches	1190	1200	1225	1404	1414	1439
Частота (ClkMul), МГц	58,1	46,3	40,4	58,1	46,5	40,0

Исследование реализации в FPGA схемы умножителя кватернионов по схеме L-BAAT для $B = 16$ показало нелинейную зависимость между потребляемой умножителем мощностью и скоростью поступления входных отсчетов (рис. 5.2). Как можно видеть из приведенных данных, наиболее эффективной по потребляемой мощности является схема 8-BAAT. Функция времени умножения от схемы построения умножителя кватернионов носит нелинейный характер: 1-BAAT: 0,292598 мкс; 2-BAAT: 0,1607142 мкс; 4-BAAT: 0,123152 мкс; 8-BAAT: 0,100671 мкс; 16-BAAT: 0,107526 мкс. Данный результат обусловлен двумя факторами: 1) «разрастанием» дерева сумматоров при увеличении значения L (см. рис. 5.1), что снижает максимальную частоту работы схемы (см. данные табл. 5.1); 2) наличием дополнительного такта инициализации умножителя, что с учетом необходимости B/L тактов для выполнения умножения дает следующее выражение для частоты работы умножителя:

$$f_{L-BAAT} = \left(\frac{B}{L} + 1 \right) / \Delta t,$$

где f_{L-BAAT} – частота работы умножителя, МГц; Δt – время поступления входных данных на умножитель, мкс. Таким образом, для построения умножителя кватернионов с минимально занимаемой площадью кристалла целесообразно использовать схему 1-BAAT, а для обеспечения максимальной пропускной способности умножителя и минимально потребляемой мощности – схему 8-BAAT. Реализация умножителя кватернионов на левый операнд будет отличаться от реализации на правый лишь содержанием блоков памяти.

5.2. Поточный процессор ПУБФ

Аппаратные затраты реализаций поточного процессора в зависимости от схемы умножителя кватернионов для $L=1,2,4$ приведены в табл. 5.2. Откуда видно, что площадь (количество CLB блоков), занимаемая банками, практически полностью определяется умножителями кватернионов, а частота работы процессора также ограничивается возможностями используемых умножителей (практически соответствует значениям из табл. 5.1). Латентность данных поточных процессоров определяется как $Lat = 2 \cdot (N+2) + (N+1)$, где $N+2$ – количество умножителей кватернионов в каждом банке фильтров; $N+1$ – задержка, вносимая алгоритмом, равна количеству блоков задержки.

Оценка ошибки реконструкции сигнала пропорциональна отклонению АЧХ структуры банк анализа – банк синтеза от идеальной характеристики (отклик схемы на дельта-импульс). Как можно видеть из рис. 5.3, наибольшие искажения наблюдаются для варианта 16-BAAT и составляют около 0,04 дБ. При этом отношение сигнал-шум для поточных процессоров ПУБФ в зависимости от схемы умножителя кватернионов составляет следующие значения: 1-BAAT: -74,3 дБ; 2-BAAT: -73,2 дБ; 4-BAAT: -70,5 дБ; 8-BAAT: -64,8 дБ; 16-BAAT: -63,0 дБ. Ухудшение результатов

при использовании схемы 16-BAAT по сравнению со схемой 1-BAAT находится примерно на уровне 10 дБ. Это объясняется сокращением разрядности предварительно рассчитываемых функций $F_i(b_{kn})$ и $F_i(0)$, хранимых в блоках памяти данных (см. рис. 5.1), что и вызывает нелинейное возрастание уровня шумов с увеличением уровня параллелизма на принципе L-BAAT. Улучшение данных характеристик можно получить, если использовать специальные методики синтеза ПУБФ, когда значения функций $F_i(b_{kn})$ и $F_i(0)$ не требуют округления при загрузке в память (например, 19/128, 1/16, 5/32 и т.д.), что позволит повысить точность выполнения умножения и, за счет больших возможностей методики синтеза ПУБФ, сократить занимаемую площадь кристалла и потребляемую мощность.

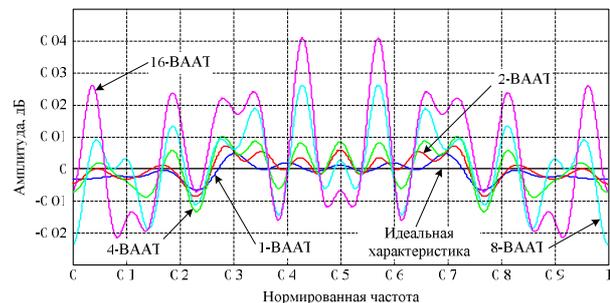


Рис. 5.3. АЧХ банка фильтров анализа-синтеза для различных реализаций умножителя кватернионов

6. Пример обработки изображения. Эффект регулярности первого рода

С целью продемонстрировать обработку изображения системой анализ-синтез на ПУБФ были спроектированы два 8-ми канальных ПУБФ с линейной ФЧХ ($N=1$): первый банк фильтров – банк, в котором отсутствует регулярность первого рода, а второй банк фильтров – ПУБФ с регулярностью первого порядка. Кватернионные коэффициенты обоих банков представлены 8-ми битными словами. При этом, коэффициенты эффективности кодирования были приблизительно одинаковые: $CG = 9,39$ дБ и $CG = 9,37$ дБ соответственно для первого и второго банков.

Далее осуществлялась обработка образа «Лена» на системе анализ-синтез двух данных ПУБФ. Объективные показатели обработки примерно равные: $PSNR = 34,82$ дБ, $MSE = 21,45$ для первого банка фильтров и $PSNR = 35,08$ дБ, $MSE = 20,21$ для второго банка с регулярностью первого порядка. На рис. 6.1а и рис. 6.1б представлены изображения реконструированные обоими ПУБФ. Изображение на рис. 6.1а обработано банком фильтров на кватернионах, коэффициенты которого не удовлетворяют условию (3.30). Здесь четко видна сетка поверх образа, представляющая собой артефакт «шахматной доски», обусловленный эффектом DC leakage. С другой стороны, изображение на рис. 6.1б чисто, потому что синтез второго ПУБФ на кватернионах осуществлялся с учетом требования (3.30), накладываемого на коэффициенты банка фильтров.



Рис. 6.1 а



Рис. 6.1 906

Заключение

В статье рассмотрен новый подход проектирования и реализации 4-х и 8-ми канальных ПУБФ на основе теории гиперкомплексных чисел, в частности, кватернионов, которая еще не применялась в области синтеза и анализа цифровых банков фильтров. Здесь умножитель кватернионов представлен как альтернативный модуль построения ПУБФ и может рассматриваться как обобщение планарных оборотов Гивенса. Представленные решетчатые структуры ПУБФ на кватернионах являются системами без потерь, несмотря на квантование коэффициентов, и могут рассматриваться как расширение классических двухканальных решетчатых структур, разработанных Вайдиянасаном и Хоангом. Более того, предложенный подход дает возможность непосредственного выражения условия регулярности первого порядка, которое зафиксировано в значениях коэффициентов решетчатой структуры ПУБФ и, следовательно, может легко удовлетворяться даже для арифметики с конечной точностью вычислений. Таким образом, данная теория синтеза решетчатых структур ПУБФ на кватернионах представляет большой интерес с практической точки зрения обработки сигналов и является хорошей альтернативой стандартной технике построения ПУБФ.

Литература

1. P.P. Vaidyanathan, Z. Doganata, "The role of lossless systems in modern digital signal processing: a tutorial", IEEE Trans. on Education, vol. 32, no. 3, pp. 181-197, 1989.
2. P.P. Vaidyanathan, Multirate systems and filter banks, Prentice-Hall, Englewood Cliffs, NJ, USA, 1993.
3. R.E. Crochiere, L.Rabiner, Multirate digital signal processing, Prentice-Hall, Englewood Cliffs, NJ, USA, 1983.
4. В.В.Витязев, Цифровая частотная селекция сигналов, Радио и связь, М., 1993, - 240с.
5. A.Plotrowski, M.Parfieniuk, Cyfrowe banki filtrów: analiza, synteza i implementacja dla systemów multimedialnych, Politechnika Białostocka, Białystok, 2006, - 389p.
6. R.G. Vaughan, N.L. Scott, D.R. White, „The theory of bandpass sampling”, IEEE Trans. Signal processing, vol. 39, no. 9, pp. 1973-1984, 1991.
7. M. Parfieniuk, A.Petrovsky, „Simple rule of selection of subsampling ratios for warped filter banks”, in Proc. VIII Int. conf. "Modern communication systems", Naroch, Belarus, 2003, pp. 130-134, Special Issue of Trans. Belarussian Engineer Academy, no. 1(15)/3.

8. Z. Doganata, P.P. Vaidyanathan, T.Q. Nguyen, "General synthesis procedures for FIR lossless transfer matrices for perfect-reconstruction multirate filter bank applications", IEEE Trans. On Acoustics, speech and signal processing, vol. 36, n0. 10, pp. 1561-1574, 1988.
9. L. Gan, K. -K. Ma, "A simplified lattice factorization for linear-phase perfect reconstruction filter bank", IEEE Signal processing letters, vol. 8, n0. 7, pp. 207-209, 2001.
10. T.Q. Nguyen, P.P. Vaidyanathan, "Maximally decimated perfect-reconstruction FIR filter banks with pairwise mirror-image analysis (and synthesis) frequency responses", IEEE Trans. On Acoustics, speech and signal processing, vol. 36, no. 5, pp. 693-706, 1988.
11. S. Orantara, T.D. Tran, P.N. Heller, T.Q. Nguyen, "Lattice structure for regular paraunitary linear-phase filterbanks and M-band orthogonal symmetric wavelets", IEEE Trans. On Signal processing, vol. 49, no. 11, pp. 2659-2672, 2001.
12. P.P.Vaidyanathan, "On coefficient-quantization and computational roundoff effects in lossless multirate filter banks", IEEE Trans. On Signal processing, vol. 39, no. 4, pp. 1006-1008, 1991.
13. W.R. Hamilton, "On quaternions; or on a new system of imaginaries in algebra", The London, Edinburgh and Dublin Philosophical Magazine and Journal of Science, vol. 25, pp.489-495, 1844.
14. I.L. Kantor, A.S.Solodovnikov, Hypercomplex numbers: an elementary introduction to algebra, Springer, New York, NY, USA, 1989.
15. A. Baker, Matrix groups: an introduction to Lie group theory, Springer, London, UK, 2002.
16. H.G. Baker, "Quaternions and orthogonal 4x4 real matrices", Tech. Rep., June 1996, <http://www.gamedev.net/reference/articles/article428.asp>.
17. M. Parfieniuk, A. Petrovsky, "Hypercomplex factorizations for 8-channel linear phase paraunitary filter banks", in Proc. 7th Int. Conf. and exhibition "Digital signal processing and its applications" (DSPA), vol. 2, Moscow, Russia, 2005, pp. 509-513.
18. M. Parfieniuk, A. Petrovsky, "Quaternionic approach to 8-channel general paraunitary filter banks", in Proc. 13th European Signal processing conference (EUSIPCO), Antalya, Turkey, 2005, CD.
19. T.D.Howell, J.C. Lafon, "The complexity of the quaternion product", Tech. Rep. TR 75-245, Cornell University, Ithaca, NY, USA, June 1975, <http://citeseer.ist.psu.edu/howell75complexity.html>.
20. A. Verenik, M. Parfieniuk, A. Petrovsky, "An FPGA implementation of the distributed arithmetic based quaternionic multipliers for paraunitary filter banks", in Proc. of the MIXDES'2007, Ciechocinek, Poland, 2001, pp. 605-610.

УДК 621.397

НОВЫЙ ПОДХОД К ИСПОЛЬЗОВАНИЮ ВЕЙВЛЕТ-ФИЛЬТРОВ ПРИ ОБРАБОТКЕ ИЗОБРАЖЕНИЙ

Дворкович В.П., Гильманшин А.В.

Введение

Обработка изображения при помощи одномерного вейвлет-преобразования основана на последовательном разложении изображения по столбцам и строкам. Такой вид обработки соответствует использованию разделимых двумерных фильтров, импульсная характеристика которых есть тензорное произведение импульсных характеристик соответствующих одномерных фильтров.

Стандартное вейвлет-преобразование подразумевает использование двумерных вейвлет-фильтров, при этом обычно сначала вейвлет-преобразование применяется ко всем строкам изображения, в результате чего формируется 2 области (левая – низкочастотная и правая – высокочастотная). Затем вейвлет-преобразование применяется ко всем столбцам полученных коэффициентов, в результате получается уже 4 области (НН – низкочастотная область, левый верхний угол, и три высокочастотных области; НВ – нижний левый угол, ВН – верхний правый угол, ВВ – нижний правый угол). Ниже рассматривается применение двух- и трехполосных вейвлет-фильтров для решения задачи кодирования изображений.

Применение двухполосных вейвлет-фильтров

Если необходимо, данная процедура может быть снова применена к НН компоненте изображения, в результате чего вместо нее будет сформировано еще 4

Приводятся результаты применения для кодирования изображений двухполосных и трехполосных вейвлет-преобразований. Рассмотрены возможности разложения изображений на 6, 9, 11, 13 и 16 частотных компонент, а также достижимый выигрыш в изменении объема информации при использовании соответствующих разложений.

частотных компоненты. Для повышения эффективности преобразования изображений предлагается дополнительное вейвлет-преобразование, применяемое к НВ и ВН компонентам изображения. В результате все изображение раскладывается на 6 частотных компонент: 1 низкочастотную и 5 высокочастотных. Схема разложения изображения представлена на рис. 1.

Для исследования энергетических характеристик вейвлет-преобразования от вида вейвлет-базиса было выбрано 4 тестовых монохромных изображения с разрешением 512x512 пикселей. Отличительные особенности этих изображений приведены в таблице 1.

Для исследования характеристик вейвлет-преобразования использовался набор различных вейвлет-фильтров с разным количеством отсчетов в импульсных характеристиках (ИХ) низкочастотных и высокочастотных фильтров. При этом стоит отметить, что с ростом числа отсчетов в ИХ уменьшается площадь под АЧХ ВЧ фильтра и наблюдается уменьшение энергии в ВЧ компонентах.

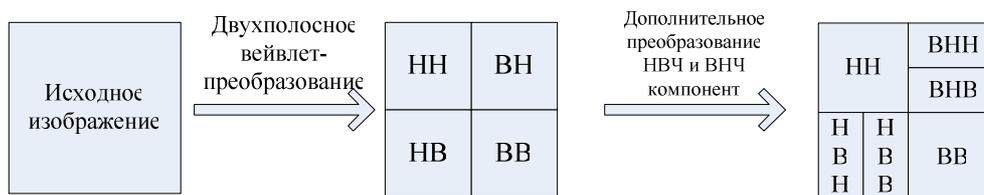
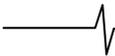


Рис. 1. Схема вейвлет-преобразования, с дополнительным разложением НВЧ и ВНЧ компонент.

Таблица 1

Отличительные особенности тестовых изображений

Название изображения	Особенности изображений
Lenna	Классическое тестовое изображение. Плавные цветовые переходы, наклонные границы.
Barbara	Черно-белое изображение. Полоски на скатерти, платке и брюках, здесь часто возникает муар и другие артефакты
Goldhill	Черно-белое изображение, содержащее множество мелких деталей, как на переднем, так и на заднем плане
Harbour	Классическое тестовое изображение. Содержащее большое число вертикально ориентированных элементов



С другой стороны, по мере увеличения числа отсчетов в ИХ требуется увеличение мощности вычислительных процедур на выполнение операции дискретного вейвлет-преобразования, поэтому в реальных системах сжатия изображений при помощи вейвлет- преобразования необходим компромисс между размером вейвлет-базиса и требуемым объемом вычислений.

В данной работе использовались фильтры, частотные характеристики которых определяются соотношениями:

$$\text{для НЧ фильтра } H(x) = h_0 + 2 \sum_{n=1}^4 h_n \cos(\pi n x), \quad (1)$$

$$\text{для ВЧ фильтра } G(x) = g_0 + 2 \sum_{n=1}^5 g_n \cos(\pi n x). \quad (2)$$

Восстановление изображений производится с применением следующих фильтров:

$$\text{Для НЧ-составляющих } \bar{H}(x) = g_0 + 2 \sum_{n=1}^5 (-1)^n g_n \cos(\pi n x), \quad (3)$$

$$\text{Для ВЧ-составляющих } \bar{G}(x) = h_0 + 2 \sum_{n=1}^4 (-1)^n h_n \cos(\pi n x). \quad (4)$$

В соотношениях (1) и (2) $0 \leq x \leq 1$, $x = \frac{2f}{f_d}$, f_d –

частота дискретизации. Коэффициенты h_n и g_n приведены в таблице 2.

На рис. 2 изображены соответствующие АЧХ вейвлет-фильтров.

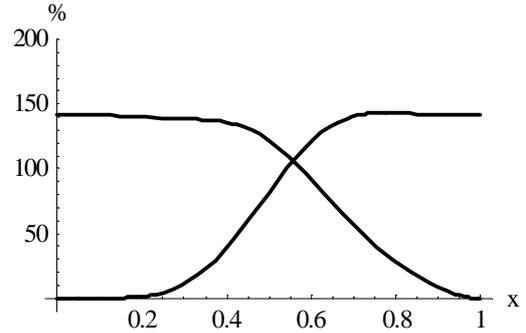


Рис 2. АЧХ двухполосных вейвлет-фильтров.

В качестве примера на следующем рисунке даны результаты разложения тестового изображения “Lenna” на 4 и 6 компонент.

Из этого рисунка видно, что разбиение на 6 компонент позволяет сконцентрировать энергию в НЧ компонентах, что обеспечивает квантование ВЧ компонент с меньшим количеством уровней.

Основной особенностью описанного выше вейвлет-преобразования является, то, что в общем случае, в сигнале выделяется лишь две компоненты – низкочастотная и высокочастотная составляющие.

Таблица 2

Коэффициенты двухполосного вейвлет-фильтра.

N	0	1	2	3	4	5
h_n	$\frac{306}{256}$	$\frac{75}{256}$	$-\frac{65}{256}$	$\frac{15}{256}$	$\frac{3}{256}$	
g_n	$\frac{168}{256}$	$-\frac{103}{256}$	$\frac{8}{256}$	$\frac{12}{256}$	$\frac{2}{256}$	$-\frac{1}{256}$



Рис. 3. Распределение вейвлет-коэффициентов по изображению Lenna, разбиение на 4 и 6 компонент соответственно. Величина НЧ-коэффициентов уменьшена в 2 раза, величина ВЧ-коэффициентов, расположенных на уровне серого, увеличена в 4 раза.

Применение трехполосных вейвлет-фильтров

Для повышения эффективности преобразования изображений предлагается производить разделение сигналов при помощи трехполосных фильтров, которые делят сигнал на три составляющих (низкочастотную, среднечастотную и высокочастотную).

При этом вейвлет преобразование выполняется по стандартной схеме: сначала преобразованию подвергаются все строки изображения, в результате чего формируются 3 области (низкочастотная, среднечастотная и высокочастотная). Затем вейвлет преобразование применяется ко всем столбцам полученных коэффициентов, в результате можно выделить уже 9 областей. Для повышения эффективности преобразования изображений предлагается дополнительное двухполосное вейвлет-преобразование, применяемое к низкочастотным и среднечастотным компонентам изображения, при этом в зависимости от глубины преобразования возможно разложение на 11, 13 или 16 компонент.

Структурная схема вейвлет-разложения изображения и обозначение полученных областей (при разложении на 16 компонент) представлены на рис. 4.

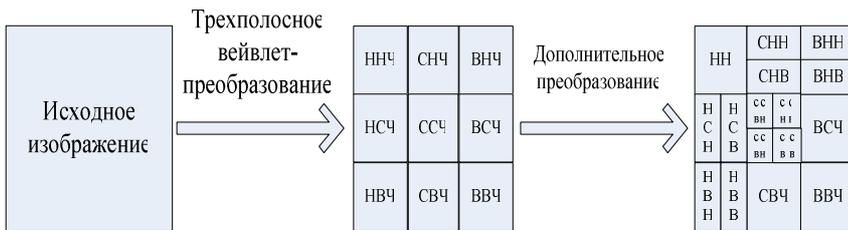


Рис. 4. Структурная схема трехполосного вейвлет-преобразования изображения.

Для вейвлет-преобразования по вышеуказанной схеме может использоваться большой набор банков вейвлет-фильтров с разным количеством отсчетов в импульсных характеристиках (ИХ) низкочастотных, среднечастотных и высокочастотных фильтров.

Частотные характеристики фильтров, используемых в данной работе, определяются следующими соотношениями:

для НЧ фильтра $H(x) = h_0 + 2 \sum_{n=1}^4 h_n \cos(\pi n x)$, (5)

для СЧ фильтра $V(x) = 2j \sum_{n=1}^4 b_n \sin(\pi n x)$, (6)

для ВЧ фильтра $G(x) = g_0 + 2 \sum_{n=1}^5 g_n \cos(\pi n x)$. (7)

В соотношениях (5), (6) и (7) $0 \leq x \leq 1$, $x = 2f/f_d$, f_d – частота дискретизации. Коэффициенты h_n , b_n и g_n – прямого преобразования приведены в таблице 2.

На рис. 5 приведены соответствующие АЧХ вейвлет-фильтров.

В таблице 3 даны коэффициенты обратного преобразования \bar{h}_n , \bar{b}_n и \bar{g}_n .

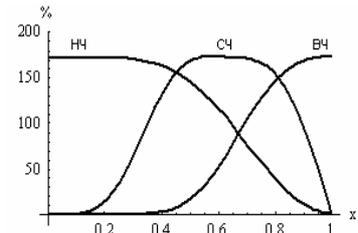


Рис. 5. АЧХ трехполосных вейвлет-фильтров.

Таблица 2

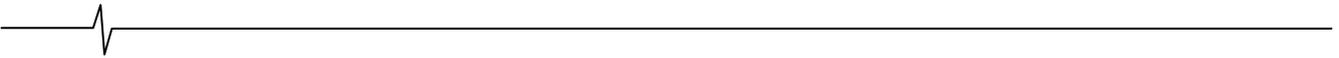
Коэффициенты прямого преобразования.

N	0	1	2	3	4	5	6
h_n	$\frac{295}{256}$	$\frac{107}{256}$	$-\frac{38}{256}$	$\frac{3}{256}$	$\frac{1}{256}$		
b_n	0	$\frac{200}{256}$	$-\frac{88}{256}$	$-\frac{10}{256}$	$-\frac{18}{256}$	$\frac{7}{256}$	$\frac{4}{256}$
g_n	$\frac{149}{256}$	$-\frac{114}{256}$	$\frac{45}{256}$	0	$-\frac{8}{256}$	$\frac{3}{256}$	

Таблица 3

Коэффициенты обратного преобразования.

N	0	1	2	3	4	5	6	7	8
\bar{h}_n	$\frac{149}{256}$	$\frac{118}{256}$	$\frac{51}{256}$	$-\frac{1}{256}$	$-\frac{16}{256}$	$-\frac{8}{256}$	0	$\frac{1}{256}$	$\frac{1}{256}$
\bar{b}_n	0	$-\frac{194}{256}$	$-\frac{73}{256}$	$\frac{4}{256}$	$-\frac{11}{256}$	$-\frac{4}{256}$			
\bar{g}_n	$\frac{143}{256}$	$-\frac{230}{256}$	$-\frac{91}{256}$	$-\frac{2}{256}$	$\frac{14}{256}$	$\frac{4}{256}$	0	$\frac{2}{256}$	$\frac{1}{256}$



На рис. 6-9 в качестве примера приведены результаты разложения тестового изображения "Harbour" на 9, 11, 13 и 16 компонент соответственно.

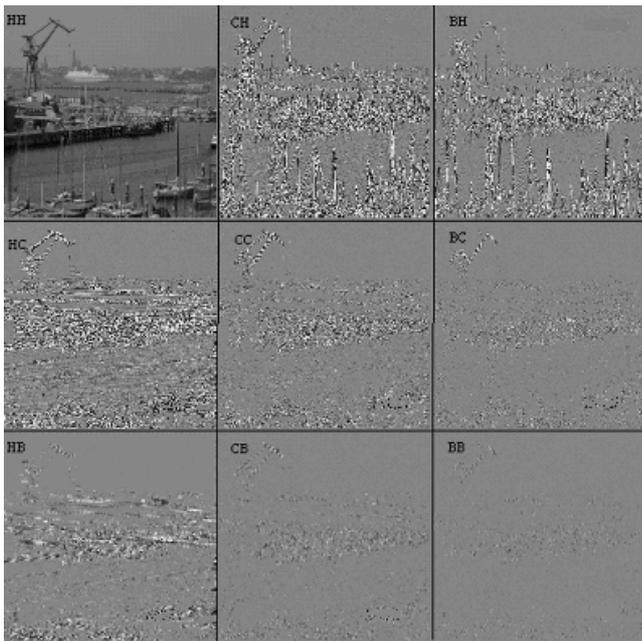


Рис. 6. Распределение вейвлет-коэффициентов по изображению Harbour, разбиение на 9 компонент. Величина НЧ-коэффициентов уменьшена в 3 раза, величина ВЧ-коэффициентов, расположенных на уровне серого, увеличена в 5 раза.

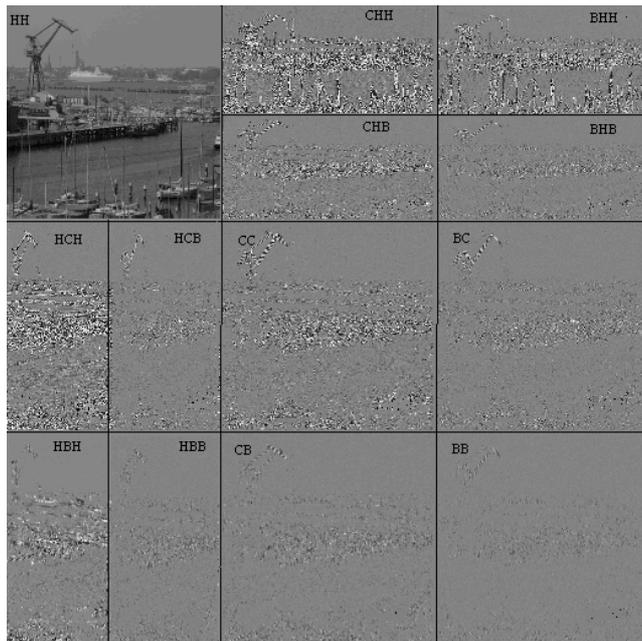


Рис. 8. Распределение вейвлет коэффициентов по изображению Harbour, разбиение на 13 компонент. Величина НЧ-коэффициентов уменьшена в 3 раза, величина ВЧ-коэффициентов, расположенных на уровне серого, увеличена в 5 раза.

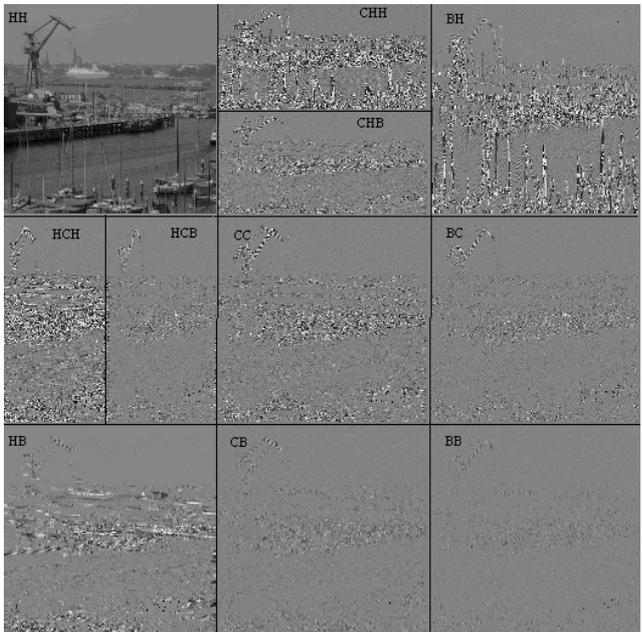


Рис. 7. Распределение вейвлет-коэффициентов по изображению Harbour, разбиение на 11 компонент. Величина НЧ-коэффициентов уменьшена в 3 раза, величина ВЧ-коэффициентов, расположенных на уровне серого, увеличена в 5 раза.

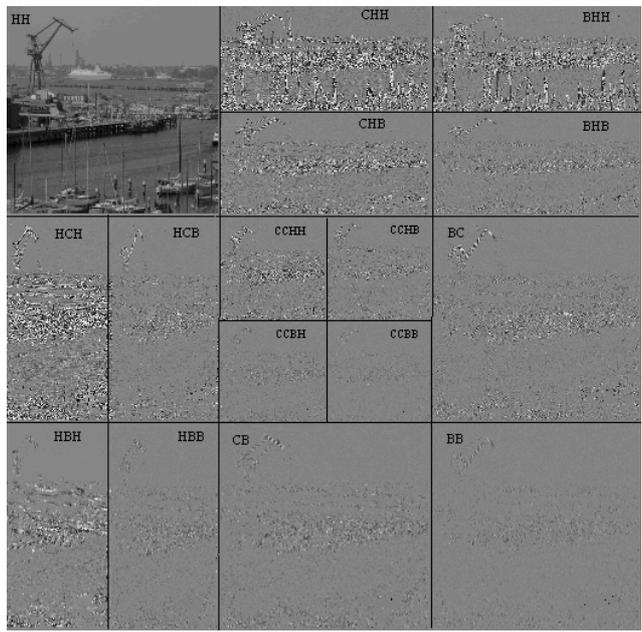


Рис. 9. Распределение вейвлет-коэффициентов по изображению Harbour, разбиение на 16 компонент. Величина НЧ-коэффициентов уменьшена в 3 раза, величина ВЧ-коэффициентов, расположенных на уровне серого, увеличена в 5 раза.

Из приведенных рисунков видно, что разбиение на большее число компонент позволяет сконцентрировать энергию в НЧ и СЧ компонентах, что обеспечивает квантование ВЧ компонент со значительно меньшим количеством уровней.

После вейвлет-преобразования в каждой из высокочастотных компонент изображения содержится различное количество энергии. Поэтому к каждой из компонент изображения возможно применение своей маски квантования, не значительно ухудшающей качество изображения.

Соответственно для описания случайной величины – яркости пикселей в каждой из ВЧ-компонент вейвлет-преобразования можно использовать одномерную плотность распределения вероятностей, с достаточной степенью точности подчиняются закону Лапласа:

$$w(x) = \frac{\lambda}{2} e^{-\lambda|x|}. \tag{8}$$

Величину λ можно найти, зная СКО данного распределения, т.к. по определению для непрерывного случайного процесса с нулевым математическим ожиданием:

$$СКО \cong \sqrt{\frac{1}{MN} \sum_{n=0}^{N-1} \sum_{m=0}^{M-1} x_{nm}^2}. \tag{9}$$

Откуда, используя формулы (8) и (9), не трудно показать, что:

$$\lambda = \frac{\sqrt{2}}{СКО}. \tag{10}$$

Выбор количества уровней квантования проводился экспериментальным путем на основе:

- приемлемого визуального ухудшения качества изображения при переходе к меньшему числу уровней квантования;
- уменьшению величин сигнал-шум (SNR) и пиковый сигнал-шум (PSNR) не более чем на 0.8дБ для каждой ВЧ компоненты при общем уменьшении SNR и PSNR от переквантования всех компонент не более чем на 3дБ.

При этом стоит отметить, что число уровней квантования выбиралось из ряда 0, 3, 7, 15, 31, 63. В этом случае каждому уровню можно сопоставить конечное число бит.

В результате исследования большого числа различных изображений была сформирована приведенная в таблице 4 маска выбора числа уровней квантования в зависимости от параметра λ .

Последнее значение λ свойственно в основном текстурам (искусственным изображениям) и по этому редко встречается при воспроизведении телевизионных изображений.

Стоит отметить и другие возможности определения параметра λ , например, имеет смысл для оценки распределения воспользоваться параметром модульной медианы N:

$$\text{mod} \cong \frac{1}{MN} \sum_{n=0}^{N-1} \sum_{m=0}^{M-1} |x_{nm}|. \tag{11}$$

В этом случае с помощью формул (5) и (8) не трудно показать что:

$$\lambda = \frac{1}{\text{mod}}. \tag{12}$$

Преимущество такого способа определения параметра λ – меньшие вычислительные затраты.

В таблице 5 приведены средние значения параметра λ для компонент по четырем тестовым изображениям при разложении на 4 и 6 компонент.

Таблица 4

Определение числа уровней квантования.

Параметр λ	$\lambda > 1$	$1 > \lambda > 0.4$	$0.4 > \lambda > 0.15$	$0.15 > \lambda > 0.05$	$0.05 > \lambda > 0.02$	$0.02 > \lambda$
Число уровней	0	3	7	15	31	63

Таблица 5

Средние значения параметра λ по компонентам.

Количество компонент	Среднее значение λ для всех компонент		
	НВ	ВН	ВВ
4	0,13	0,14	0,42
6	НВН	НВВ	0,42
	0,09	0,20	
	ВНН	ВНВ	
	0,10	0,22	

В следующей таблице приведены средние значения параметра λ для компонент по четырем тестовым изображениям при разложении на 9, 11, 13 и 16 компонент соответственно.

Таблица 6.

Средние значения параметра λ по компонентам.

Кол-во областей	Среднее значение λ для всех компонент														
	НС		СН		НВ	ВН	СС		СВ	ВС	ВВ				
9	0,09		0,08		0,15	0,13	0,18		0,36	0,32	0,7				
11	НСН	НСВ	НСН	НСВ	0,15	0,13	0,18		0,36	0,32	0,7				
	0,06	0,16										0,05	0,15		
13	0,06	0,16	0,05	0,15	НВН	НВВ	ВНН	ВНВ	0,18		0,36	0,32	0,7		
					0,13	0,25								0,12	0,24
16	0,06	0,16	0,05	0,15	0,13	0,25	0,12	0,24	НН	ВН	НВ	ВВ	0,36	0,32	0,7
									0,14	0,35	0,33	0,62			

Уменьшение объема информации.

Название изображения	Уменьшение объема информации в зависимости от количества компонент в разбиении.					
	4	6	9	11	13	16
Lenna	1,68	2,05	2,32	2,51	2,61	2,84
Barbara	1,60	1,94	2,16	2,34	2,46	2,63
Goldhill	1,68	2,05	2,16	2,29	2,42	2,67
Harbour	1,60	1,94	2,32	2,49	2,58	2,81

В таблице 7 приведены результаты по уменьшению объема информации, содержащейся в выходном потоке при применении разбиения на 4, 6, 9, 11, 13 и 15 компонент, соответственно, при условии $PSNR \approx 37$ дБ.

При этом под уменьшением информации понимается отношение количества бит, требуемое для передачи исходных тестовых монохромных изображений, к количеству бит после вейвлет-преобразования и последующего квантования высокочастотных областей по методике, указанной выше.

Заключение

Разложение изображения на большое количество компонент приводит к некоторому увеличению числа выполняемых операций в кодирующем устройстве по сравнению со стандартным вейвлет-разложением на 4 компоненты, число же операций, выполняемых декодером, в силу малого числа ненулевых отсчетов в СЧ и ВЧ компонентах, возрастает незначительно.

При стандартном вейвлет-разложении изображений на 4 компоненты и последующем квантовании высокочастотных областей объем информации в среднем может быть уменьшен в 1.64 раза. Дополнительная обработка двух ВЧ-областей этих вейвлет-преобразований позволяет допол-

нительно уменьшить объем информации на 25%.

При вейвлет-разложении изображений на 9, 11, 13 и 16 областей объем информации в среднем уменьшается по сравнению с разложением на 4 компоненты на 35%, 45%, 50% и 60% соответственно.

Литература

1. S. Mallat. A theory for multiresolution signal decomposition: the wavelet representation. IEEE Pattern Anal. And Machine Intel. 1989. vol.11, no 7, pp. 674-693
2. Дворкович В.П., Дворкович А.В. Расчет банков фильтров дискретного вейвлет-преобразования и анализ их характеристик // Цифровая обработка сигналов, 2006, №2
3. Дворкович В.П., Гильманшин А.В. Новый подход к использованию двумерных вейвлет-фильтров при обработке изображений // Доклад на 10-й Международной конференции «Цифровая обработка сигналов и ее применение», Москва, март 2008.
4. Дворкович В.П., Гильманшин А.В. Использование трехполосных фильтров вейвлет-фильтров для обработки изображений // Доклад на 10-й Международной конференции «Цифровая обработка сигналов и ее применение», Москва, март 2008.

Уважаемые коллеги!

Со второго полугодия 2008 года вы можете подписаться на наш журнал в любом отделении связи. Подписной индекс издания по каталогу ОАО «Роспечать» - 82185.

Для тех, кто не успел оформить подписку на первое полугодие 2008 года через ОАО «Роспечать», сохраняется возможность приобретения журналов непосредственно в редакции по адресу: 107031, г. Москва, Рождественка, 6\9\20, стр. 1, Российское научно-техническое общество радиотехники, электроники и связи им. А.С. Попова, или оформить Заказ в соответствии с требованиями, выставленными на сайте журнала: www.dsra.ru.

Справки по телефонам: (495) 621-71-08, 621-06-10.

Факс: (495) 621-16-39.

E-mail: nto.popov@mtu-net.ru

tor@rgta.ryazan.ru

УДК 621. 397

КУРВЛЕТ-ПРЕОБРАЗОВАНИЕ В ЗАДАЧЕ ПОДАВЛЕНИЯ ШУМА В ИЗОБРАЖЕНИЯХ

*Моисеев А.А., Кобелев В.Ю., Волохов В.А.***Введение**

На сегодняшний день существует множество различных методов цифровой обработки изображений. Они могут быть применены для решения различных задач, в частности, для анализа, восстановления и сжатия изображений. Довольно широкое применение в указанных областях в последние годы нашло вейвлет-преобразование, которое является эффективным инструментом для работы как с одномерными, так и с двумерными сигналами. В основу вейвлет-преобразования положено разложение произвольного сигнала (функции) по некоторому ортонормированному базису, построенному на основе вейвлет-функций, обладающих определенными свойствами [1], наиболее важными из которых являются частотно-временная локализация и масштабируемость. Одной из наиболее простых реализаций дискретного вейвлет-преобразования является банк фильтров. Как правило, для обработки изображений применяются двумерные разделимые банки фильтров (на практике обработка банка фильтров последовательно к строкам, а затем к столбцам изображения), что соответствует разделимому двумерному вейвлет-преобразованию.

Изначально разработанное для анализа точечных особенностей изображения, разделимое вейвлет-преобразование, обладающее свойством изотропного масштабирования, не позволяет анализировать более сложные элементы изображения. Значительно больше возможностей для анализа дает использование двумерного неразделимого вейвлет-преобразования [2, 3]. Соответствующие банки фильтров имеют более высокую сложность, и синтез фильтров для них является более сложной задачей. Наиболее известным примером двумерных неразделимых банков фильтров являются банки ромбовидной формы [4]. На рис. 1а, б приведены диаграммы частотного разбиения разделимого и неразделимого банков фильтров. Как можно заметить из рисунка, разделимые и неразделимые банки фильтров ромбовидной формы ограничены в выборе направлений. Более эффективным является применение двумерных неразделимых согласованных вейвлет-фильтров. Методика, описанная в [5], позволяет синтезировать вейвлет-фильтры для анализа изображений с особенностями заданного вида [6].

Однако, применение как разделимого, так и неразделимого вейвлет-преобразований не позволяет выполнять обработку элементов изображения в виде криволинейных контуров, границ и т.п. На рис. 2а приведен при-

Описывается метод кратномасштабного представления изображений, называемый курвлет-преобразованием. В настоящей работе данный тип преобразования используется для задачи восстановления изображений из зашумленных данных. Приведены результаты моделирования, показывающие основные особенности данного метода преобразования. Проведено сравнение полученных результатов с результатами, полученными на основе классических алгоритмов фильтрации.

мер аппроксимации кривой вейвлет-функциями различных масштабов. Видно, что для точного представления данной кривой потребуется достаточно большое количество вейвлет-коэффициентов. В связи с этим возникает необходимость создания методов, позволяющих более компактно представлять особенности изображений подобного вида. Рядом авторов разработаны преобразования, которые можно объединить в класс контурных преобразований.

На рис. 3 приведена общая классификация наиболее известных типов контурных преобразований. Далее кратко рассмотрим особенности каждого из них.

В 1992 году Р. Бамбергер и М. Смит исследовали методы синтеза двумерных направленных банков фильтров [7], которые могут обладать свойством максимальной децимации наряду с полным восстановлением сигнала. В то время внимание было сосредоточено на использовании направленной декомпозиции для целей анализа изображения, а проблемы, связанные с конструкцией базисов для представления изображения и связи банков фильтров с базисными функциями непрерывной области, не были рассмотрены. Они были изучены в 2001 году М. Ду, который ввел в рассмотрение новый тип кратномасштабного преобразования, названного контурлет-преобразованием [8]. Основой его стали направленные банки фильтров Бамбергера и Смита, которые придали контурлет-преобразованию свойства, отличные от свойств вейвлет-преобразования, а именно, возможность более точной аппроксимации криволинейных участков на изображениях.

Ключевое понимание конструкций направленного базиса на основе преобразования Радона было сделано ранее в 1998 году Э. Кандэ и Д. Донохо с использованием риджлет-преобразования [9, 10]. Идея последнего заключается в отображении линейных особенностей изображения в точечные с использованием преобразования Радона и последующем применении вейвлет-преобразования. Результатом данной операции является эффективное представление двумерных функций с кусочно-гладкими областями, разделенными линейными участками.

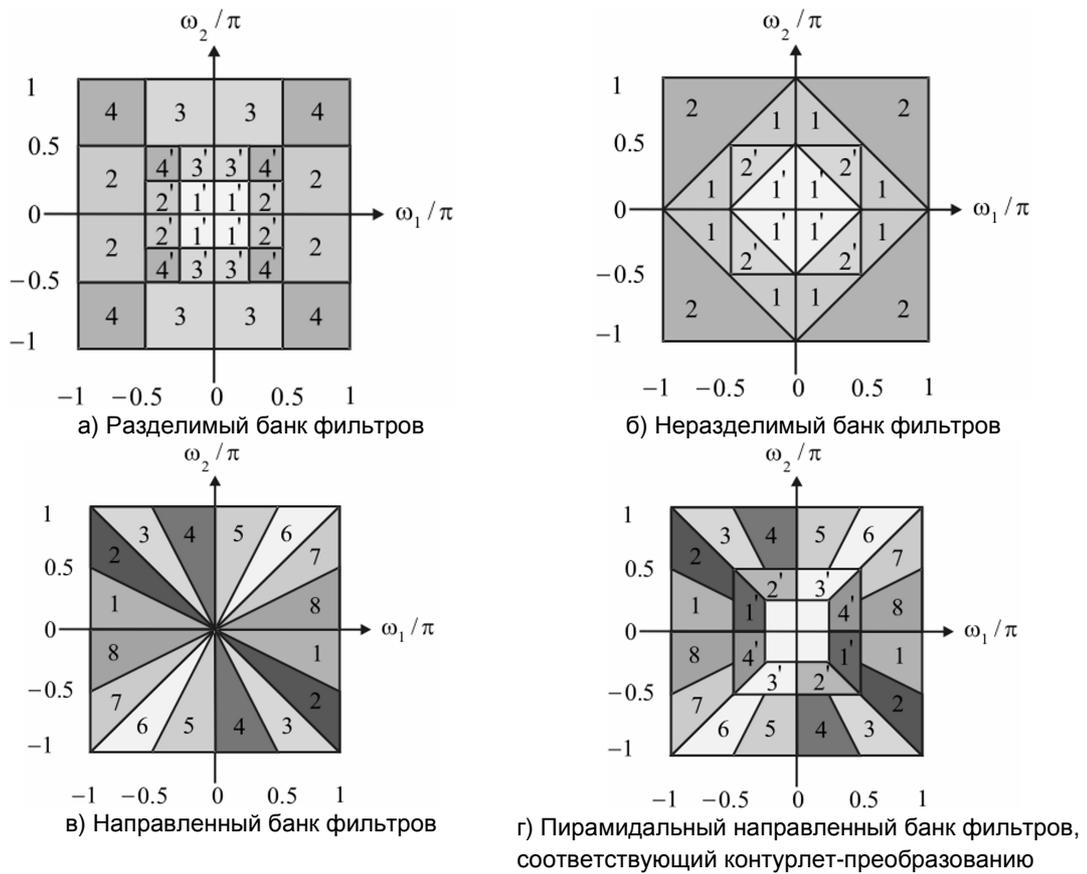


Рис. 1. Диаграммы частотного разбиения

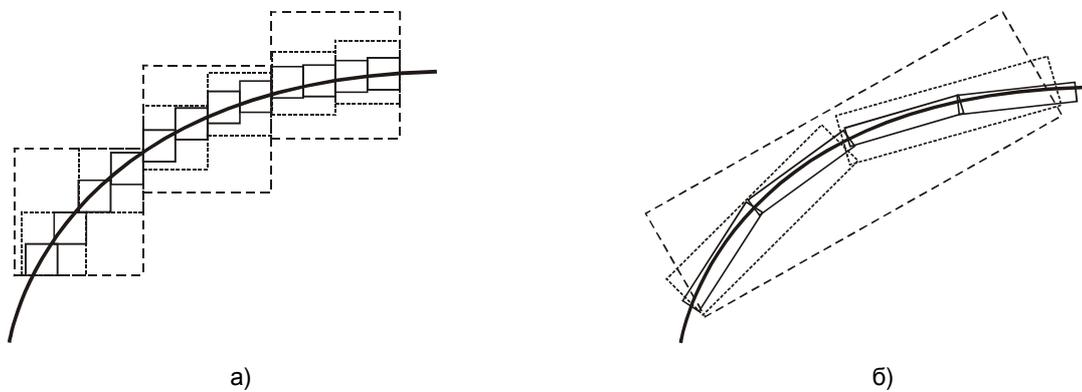


Рис. 2. Иллюстрация последовательного уточнения кривой, разделяющей две гладкие области с использованием: а) базиса вейвлет-функций; б) базиса курвлет-функций

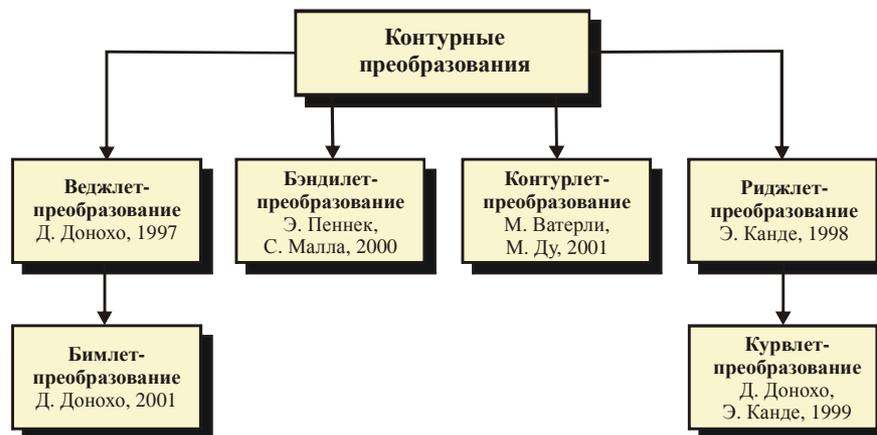


Рис. 3. Типы основных контурных преобразований

В 1999 году Э. Кандэ и Д. Донохо расширили риджлет-преобразование до курвлет-преобразования [11, 12], которое хорошо аппроксимирует особенности изображения вдоль гладких кривых (рис. 2б). По этим свойствам курвлет-преобразование схоже с контурлет-преобразованием, упомянутым выше. Однако общая идеология его выполнения отлична от контурлет-преобразования и основана на разбиении криволинейного участка изображения на линейные участки с последующим применением локального риджлет-преобразования. Помимо данных типов преобразований в 1997 и 2001 годах Д. Донохо предложил вейвлет- [13] и бимлет-преобразования [14]. Первое из них может быть применено для задач удаления шума и сжатия изображений, а второе – к детектированию границ на них.

В 2000 году Э. Пеннек и С. Малла предложили другой подход обработки геометрической структуры изображения, основанный на бэндилет-преобразовании [15]. В этом преобразовании изображение разбивается на области, для каждой из которых строится векторное поле – «геометрический поток», такое, что дисперсия в направлении, заданном векторами, является постоянной. Затем вдоль полученных направлений «геометрического потока» выполняется вейвлет-преобразование.

Теория риджлет и курвлет-преобразований

Основное отличие риджлет-функций от вейвлет-функций заключается в том, что риджлеты являются двумерными неразделимыми функциями и определяются не только параметрами масштаба и сдвига, но и параметром их ориентации в пространстве.

Впервые термин риджлет был введен Э. Кандэ [9]. В общем случае риджлет-функция может быть определена в виде:

$$\Psi_{\gamma}(\mathbf{x}) = a^{-1/2} \Psi\left(\frac{u \cdot \mathbf{x} - b}{a}\right), \tag{1}$$

где a – параметр масштаба, b – параметр смещения и u – параметр, определяющий ориентацию риджлета. Для того, чтобы функция (1) являлась риджлетом необходимо, чтобы функция Ψ удовлетворяла условию

$$C_{\Psi} = \int_{-\infty}^{\infty} \frac{|\widehat{\Psi}(\xi)|^2}{|\xi|^d} d\xi < \infty, \tag{2}$$

Из (2) следует, что функция Ψ должна удовлетворять условию нулевых моментов

$$\int_{-\infty}^{\infty} t^k \Psi(t) dt = 0, \tag{3}$$

где $k = 0, 1, \dots, \left[\frac{d+1}{2} - 1\right]$.

Условие (3) аналогично требованию, предъявляемому к вейвлет-функции [1].

Для $a > 0, b \in \mathbf{R}$ и $\theta \in [0; 2\pi)$ определим двумерную риджлет-функцию в виде

$$\Psi_{a,b,\theta}(\mathbf{x}) = a^{-1/2} \cdot \Psi\left(\frac{x_1 \cos \theta + x_2 \sin \theta - b}{a}\right). \tag{4}$$

Функция $\Psi_{a,b,\theta}(\mathbf{x})$ имеет постоянное значение вдоль прямой $x_1 \cos \theta + x_2 \sin \theta - b = const$. В направлении, перпендикулярном данной прямой, она является вейвлетом. Отсюда видна основная особенность риджлет-функции – зависимость от направления. Примеры двумерной разделимой вейвлет-функции и риджлет-функции приведены на рис. 4.

Для интегрируемой функции $f(\mathbf{x})$ прямое непрерывное риджлет-преобразование можно записать в виде криволинейного интеграла [16, 17]:

$$R_f(a, b, \theta) = \int f(\mathbf{x}) \Psi_{a,b,\theta}(\mathbf{x}) d\mathbf{x}. \tag{5}$$

Формула обратного риджлет-преобразования имеет вид:

$$f(\mathbf{x}) = \int_0^{2\pi} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} R_f(a, b, \theta) \Psi_{a,b,\theta}(\mathbf{x}) \frac{da}{a^3} db \frac{d\theta}{4\pi}. \tag{6}$$

Кроме того, для риджлет-преобразования (5), (6) справедливо равенство Парсеваля. Следовательно, аналогично преобразованию Фурье или вейвлет-преобразованию, из равенства (5) следует, что произвольную интегрируемую функцию можно представить в виде непрерывной суперпозиции риджлетов.

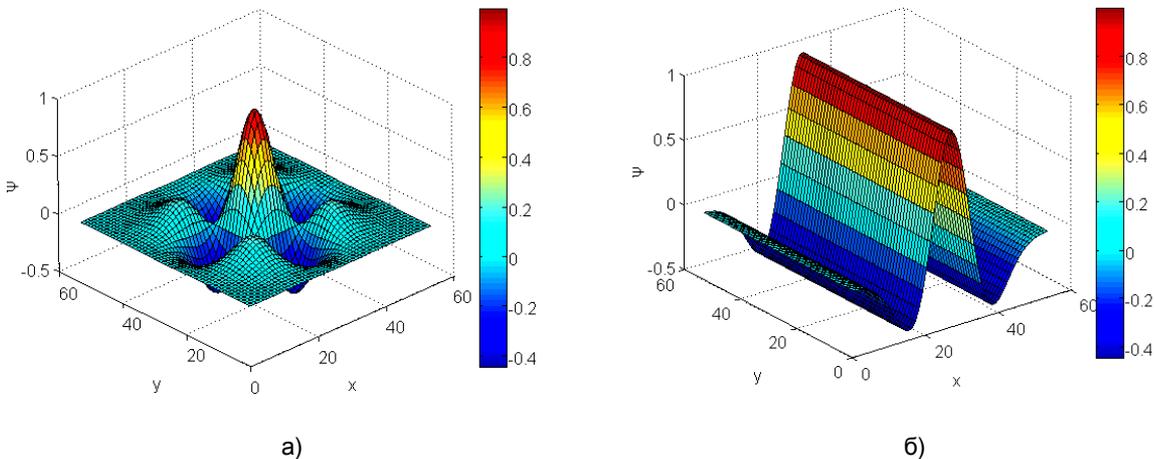
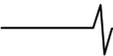


Рис. 4. Примеры двумерных базисных функций: а) вейвлет-функция; б) риджлет-функция



Основным средством для вычисления риджлет-коэффициентов является представление риджлет-анализа в виде вейвлет-преобразования в области коэффициентов преобразования Радона. Преобразованием Радона от функции f является совокупность линейных интегралов [18]

$$Rf(\theta, t) = \int f(x_1, x_2) \delta(x_1 \cos \theta + x_2 \sin \theta - t) dx_1 dx_2, \quad (7)$$

где δ – дельта-функция Дирака. Из (7) следует, что риджлет-коэффициенты $R_f(a, b, \theta)$ функции f с использованием преобразования Радона запишутся в виде

$$R_f(a, b, \theta) = \int Rf(\theta, t) \cdot a^{-1/2} \psi((b-t)/a) dt. \quad (8)$$

Как видно из выражения (8), риджлет-преобразование может быть выполнено посредством применения одномерного вейвлет-преобразования к «срезах» преобразования Радона (угловая переменная θ постоянна, t – изменяется).

Реализация дискретного риджлет-преобразования использует связь преобразования Радона с преобразованием Фурье, которая представлена выражением:

$$\hat{f}(\omega \cos(\theta), \omega \sin(\theta)) = \int Rf(\theta, t) \cdot e^{-i\omega t} dt. \quad (9)$$

Отсюда следует, что преобразование Радона может быть получено посредством применения обратного одномерного преобразования Фурье к коэффициентам двумерного преобразования Фурье, взятым вдоль радиальных прямых, проходящих через начало координат частотной сетки. Необходимо отметить, что дискретная реализация преобразования Радона, основанная на свойстве (9), выполняется с применением быстрого преобразования Фурье с последующей интерполяцией Фурье коэффициентов (для получения коэффициентов, соответствующих угловой частоте) и является приближенной. В то же время возможны и другие методы реализации дискретного преобразования Радона, дающие более точный результат. Применение данной схемы к осуществлению риджлет-преобразования позволяет достаточно эффективно реализовать последнее в случае дискретного двумерного сигнала (изображения).

Схема риджлет-преобразования для дискретного случая приведена на рис. 5. Основные этапы его осуществления следующие:

1. Вычисление прямого двумерного преобразования Фурье (FFT2D).

2. Переход в области преобразования Фурье от прямоугольной сетки координат к полярной сетке с использованием операции интерполяции соответствующих коэффициентов преобразования Фурье. Результатом последней операции является набор линий полярной сетки.

3. Применение обратного одномерного преобразования Фурье (IFFT1D) к каждой линии полученной полярной сетки. Результатом этой операции являются коэффициенты преобразования Радона.

4. Применение к плоскости преобразования Радона одномерного вейвлет-преобразования (WT1D) вдоль переменной, определяющей угол наклона линии. Результатом является область риджлет-коэффициентов.

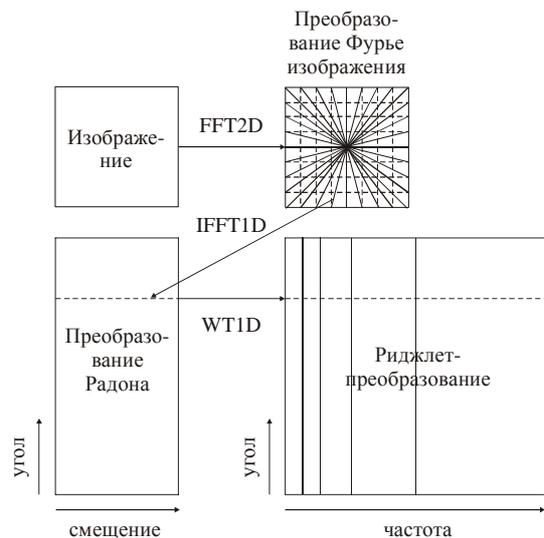


Рис. 5. Схема двумерного дискретного риджлет-преобразования

Возможны два подхода к анализу изображения при помощи риджлет-преобразования: а) анализ на одном масштабе, при котором изображение разбивается на блоки одинакового размера с последующим риджлет-анализом каждого блока; б) кратномасштабный анализ, в данном случае изображение многократно разбивается на блоки (аналогично варианту а) с уменьшением размера блока вдвое на каждом этапе разбиения и с последующим риджлет-анализом.

В заключение рассмотрения риджлет-преобразования необходимо отметить, что оно оптимально подходит для использования в задаче фильтрации прямых линий на изображениях, так как в основе него лежит преобразование Радона или, что то же самое, использование базисных функций вида (4). Но в то же время риджлет-преобразование является избыточным, так как выполняется полный анализ изображения на всех масштабах, независимо от масштаба деталей, содержащихся в изображении.

Как указывалось выше, курвлет-преобразование является более общим по отношению к риджлет-преобразованию и позволяет анализировать не только отрезки прямых, но и кривые линии, контуры и другие детали в изображениях. Курвлет-преобразование основано на применении кратномасштабного риджлет-преобразования к изображениям, восстановленным по отдельным частотным полосам, полученным в результате субполосной фильтрации исходного изображения. Далее кратко рассмотрим основные этапы реализации дискретного курвлет-преобразования, приведенные на рис. 6.

1. Разбиение исходного изображения на субполосы посредством дискретного вейвлет-преобразования без децимации с прореживанием импульсных характеристик вейвлет-фильтров на каждом последующем уровне разложения и с последующим восстановлением каждой из субполос в отдельности [19]. Применение вейвлет-преобразования позволяет разделить изображение в пространственной области на совокупность изображений деталей и одного аппроксимирующего изображения. Сумма этих изображений есть исходное изображение.

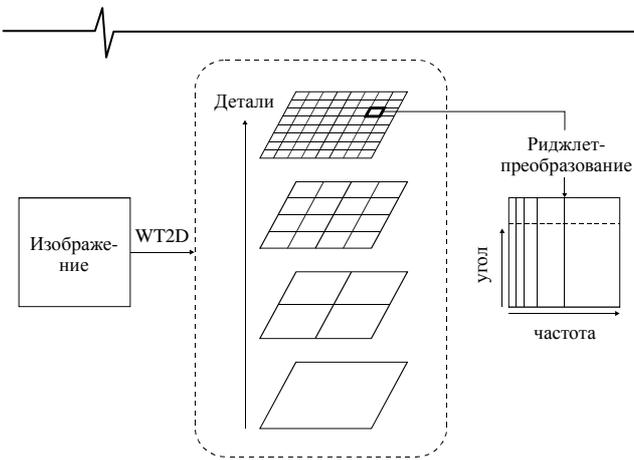


Рис. 6. Схема двумерного дискретного курвлет-преобразования

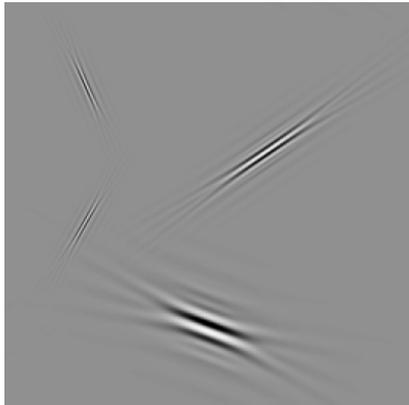


Рис. 7. Примеры курвлетов для различных масштабов, позиций и ориентаций

2. Полученные изображения деталей разбиваются на перекрывающиеся квадранты одинакового размера (во избежание артефактов блочности) с последующим взвешиванием некоторой функцией окна. При переходе к изображениям, содержащим более мелкие детали исходного изображения, выполняется уменьшение размера квадрантов в четыре раза так, как показано на рис. 6.

Разбиение на квадранты необходимо, так как последним этапом преобразования является риджлет-преобразование. В пределах квадрантов отдельные участки кривых контуров могут быть близки к прямым линиям.

3. Вычисление риджлет-преобразования по схеме, представленной на рис. 5.

Необходимо отметить, что курвлет-функция, в отличие от риджлет-функции, локализована в пространственной области и обладает следующим свойством: геометрические размеры курвлета связаны между собой соотношением – длина \approx ширина². На рис. 7 приведены несколько курвлет-функций для различных масштабов, ориентаций и позиций.

Исходя из указанного, можно отметить, что в отличие от риджлет-преобразования, курвлет-преобразование оптимально подходит для использования в задаче фильтрации кривых контуров на изображениях.

Результаты моделирования

Описанный выше алгоритм курвлет-преобразования может быть применен для решения задачи удаления шума из изображения. Рассмотрен метод пороговой фильтрации изображений на основе курвлет-преобразования, реализованный в среде Matlab. Для тестирования алгоритма использованы семь полутоновых изображений с разрешением 512*512 пикселей, различной степени детализации. В качестве искажения использован аддитивный белый гауссов шум. Для сравнения с рассматриваемым алгоритмом фильтрации выбраны следующие алгоритмы:

1. Пороговый алгоритм фильтрации на основе вейвлет-преобразования [20].
2. Пороговый алгоритм фильтрации на основе вейвлет-преобразования без децимации [19, 20].
3. Оптимальная винеровская фильтрация.

Общая схема алгоритма пороговой фильтрации изображений на основе курвлет-преобразования и его аналогов представлена на рис. 8.

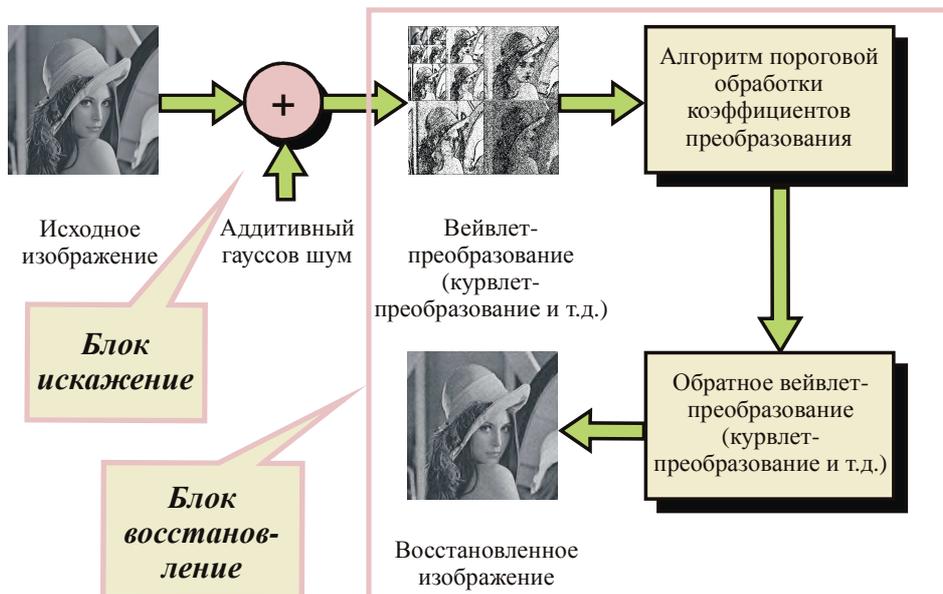
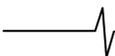


Рис. 8. Нелинейный алгоритм фильтрации, основанный на принципе кратномасштабного представления цифровых изображений



Основу схемы фильтрации (рис. 8) составляют: блок искажения, в который входит только аддитивный гауссов шум, и блок восстановления, состоящий из схемы прямого и обратного кодирования с преобразованием, а также блок обработки коэффициентов, полученных в результате преобразования.

Схема работы блока восстановления рассматриваемого алгоритма фильтрации следующая (рис. 8):

1. Вычисление прямого преобразования изображения (вейвлет-преобразования, риджлет-преобразования, курвлет-преобразования или их аналогов).

2. Изменение полученных коэффициентов преобразования по определенному правилу (в работе использо-

валась методика мягкой и жесткой пороговой обработки коэффициентов преобразования [20]).

3. Вычисление обратного преобразования от области измененных коэффициентов.

Результаты работы алгоритма для тестовых изображений «Лена», «Барбара» и «Отпечаток пальца» представлены на рис. 9.

Здесь ПОСШ – пиковое отношение сигнал/шум, используемое в настоящей работе для оценки степени искажения цифровых изображений, σ – среднеквадратическое отклонение рассматриваемой модели шума. Отметим, что зависимости носят монотонно спадающий характер, так как с увеличением σ растет степень искажения изображений.

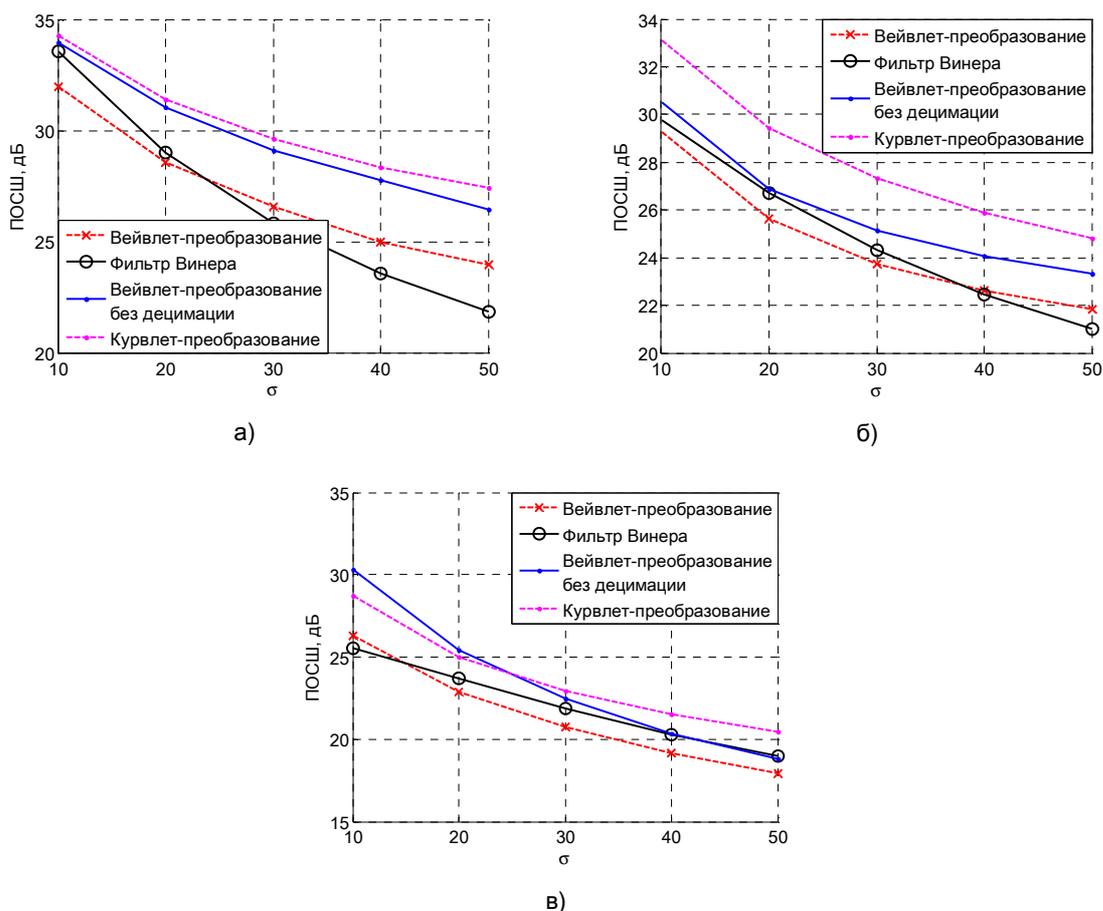


Рис. 9. Результат обработки тестовых изображений: а) «Лена»; б) «Барбара»; в) «Отпечаток пальца»

В табл. 1 приведены численные результаты работы алгоритмов для среднеквадратического отклонения $\sigma = 25$.

Таблица 1

Результаты работы алгоритмов

Тестовое изображение	Вейвлет-преобразование, ПОСШ, дБ	Вейвлет-преобразование без децимации, ПОСШ, дБ	Курвлет-преобразование, ПОСШ, дБ	Оптимальный фильтр Винера, ПОСШ, дБ
«Бабуин»	21.80	23.40	23.77	23.95
«Лена»	27.50	30.23	30.39	27.28
«Перцы»	27.73	30.41	30.23	27.27
«Корабли»	25.71	28.18	28.18	26.71
«Барбара»	24.55	26.48	28.28	25.45
«Сатурн»	28.20	28.38	37.53	26.19
«Отпечаток пальца»	21.70	23.77	23.83	22.73

На рис. 10 и рис. 11 представлены фрагменты восстановленных тестовых изображений, полученных с использованием различных алгоритмов цифровой фильтрации.

Из представленных на рис. 10 фрагментов восстановленных тестовых изображений видно, что наилучшие визуальные результаты получены в случае применения алгоритмов фильтрации на основе вейвлет-преобразования

без децимации (рис. 10г) и курвлет-преобразования (рис. 10д). Однако в первом случае на восстановленном изображении отсутствуют мелкие детали. Аналогичная ситуация наблюдается при использовании курвлет-преобразования, но контуры элементов изображения здесь обработаны лучше, хотя и наблюдаются специфические для этого преобразования артефакты.

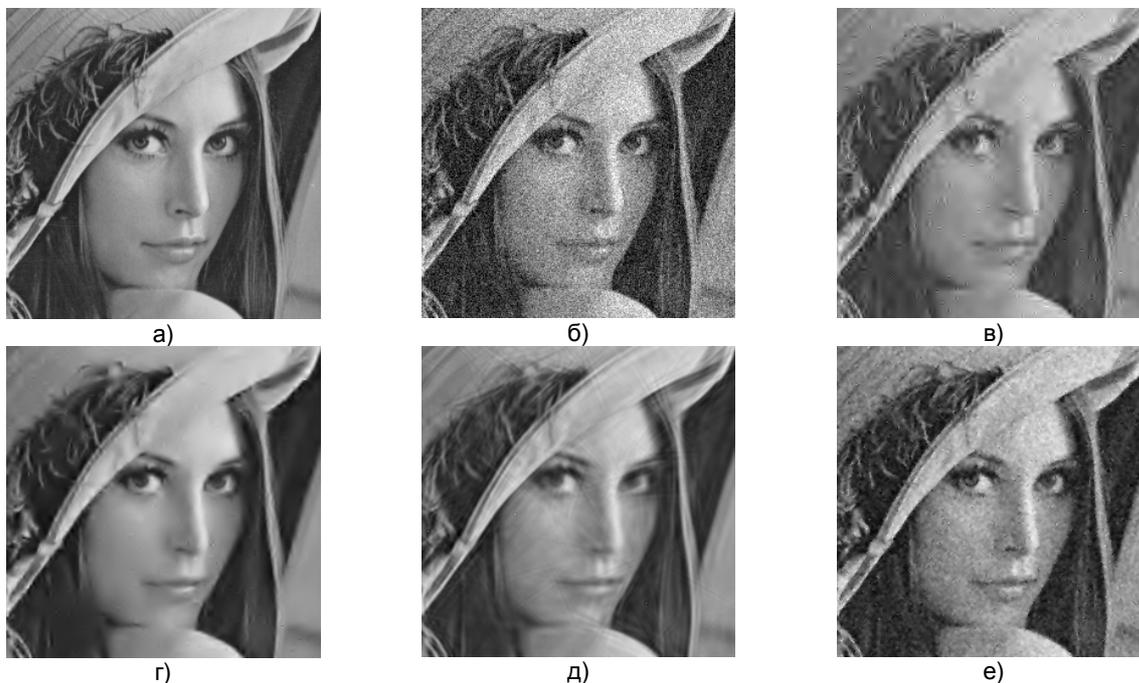


Рис. 10. Фрагменты изображения «Лена»: а) исходного; б) зашумленного, ПОСШ=20.15 дБ; в) восстановленного на основе схемы вейвлет-преобразования, ПОСШ=27.50 дБ; г) восстановленного на основе схемы вейвлет-преобразования без децимации, ПОСШ=30.23 дБ; д) восстановленного на основе курвлет-преобразования, ПОСШ=30.39 дБ; е) восстановленного на основе винеровской фильтрации, ПОСШ=27.28 дБ

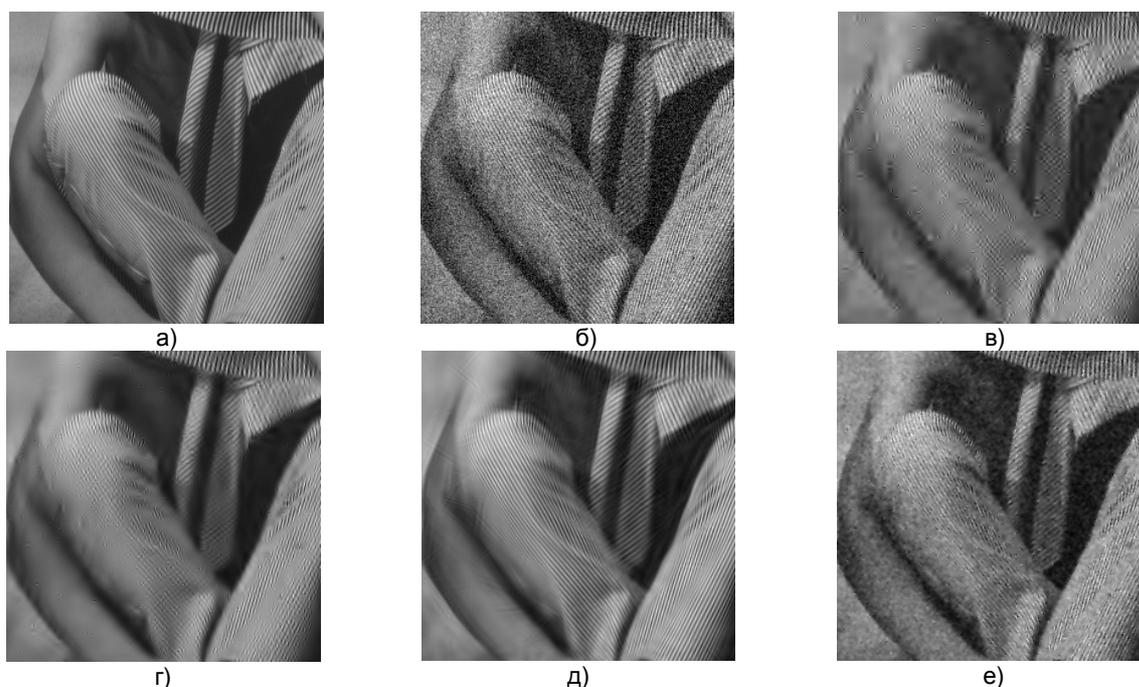


Рис. 11. Фрагменты изображения «Барбара»: а) исходного; б) зашумленного, ПОСШ=20.25 дБ; в) восстановленного на основе схемы вейвлет-преобразования, ПОСШ=24.55 дБ; г) восстановленного на основе схемы вейвлет-преобразования без децимации, ПОСШ=26.48 дБ; д) восстановленного на основе курвлет-преобразования, ПОСШ=28.28 дБ; е) восстановленного на основе винеровской фильтрации, ПОСШ=25.45 дБ



Из фрагментов восстановленных тестовых изображений, представленных на рис. 11, видно, что текстура элементов одежды в виде полос наилучшим образом восстановлена в случае использования алгоритма обработки на основе курвлет-преобразования (рис. 11д).

Заключение

В работе рассмотрено курвлет-преобразование и его применение в задаче удаления шума. Из полученных результатов обработки тестовых изображений следует, что использование алгоритма удаления шума на основе курвлет-преобразования позволяет более эффективно обрабатывать элементы изображения, содержащие контуры, детали в виде линий, кривых и т.п. Последнее связано с особой формой базисных функций данного преобразования. Отметим, что в обработанных изображениях присутствуют артефакты по форме, соответствующей форме базисных функций. Для фрагмента тестового изображения «Барбара», характеризуемого большим количеством мелких деталей, получено улучшение качества восстановленного изображения в среднем на 2 дБ по сравнению с другими алгоритмами восстановления.

Литература

1. Добеши И. Десять лекций по вейвлетам. – Москва-Ижевск: НИЦ «Регулярная и хаотическая динамика», 2004.
2. Kovacevic J., Vetterli M. Nonseparable multidimensional perfect reconstruction filter banks and wavelet bases for R^n // *IEEE Trans. on Information Theory, Special Issue on Wavelet Transforms and Multiresolution Signal Analysis*. 1992. V. 38, № 2. P. 533-555.
3. Cohen A., Daubechies I. Nonseparable bidimensional wavelet bases // *Rev. Mat. Iberoamericana*. 1993. V. 9, № 1. P. 51-137.
4. Karlsson G., Vetterli M. Theory of two-dimensional multirate filter banks // *IEEE Trans. Signal Proc.* 1990. V. 38, № 6. P. 925-973.
5. Кобелев В.Ю., Приоров А.Л. Применение неразделимых вейвлет-фильтров в задачах сжатия изображений // *Цифровая обработка сигналов*. 2006. № 2. С. 21-26.
6. Кобелев В.Ю., Приоров А.Л. Анализ изображений при помощи согласованных двумерных вейвлет-фильтров // *Радиотехника*. 2008. № 1. С. 12-19.
7. Bamberger R.H., Smith M.J.T. A filter bank for the directional decomposition of images theory and design // *IEEE Trans. Signal Proc.* 1992. V. 40, № 4. P. 882-893.
8. Do M.N. Directional multiresolution image representations // PhD dissertation, School Comput. Commun. Sci., Swiss Fed. Inst. Technol., Lausanne, Switzerland, 2001.
9. Candes E.J. Ridgelets: Theory and applications. PhD thesis, Department of Statistics, Stanford University. 1998.
10. Candes E.J., Donoho D.L. Ridgelets: a key to higher-dimensional intermittency // *Phil. Trans. R. Soc. Lond. A*. 1999. P. 2495-2509.
11. Candes E.J., Donoho D.L. Curvelets – a suprisingly effective nonadaptive representation for objects with edges / In Cohen A., Rabut C., Schumaker L.L., editors, *Curve and Surface Fitting*. Vanderbilt University Press. Saint-Malo. 1999.
12. Candes E.J., Donoho D.L. Curvelets, multiresolution representation, and scaling laws / In Aldroubi A., Laine A.F., Unser M.A., editors, *SPIE Wavelet Applications in Signal and Image Processing VIII*. 2000. V. 4119.
13. Donoho D.L. Wedgelets: Nearly-minimax estimation of edges // *Annals of Stat.* 1999. V. 27, P. 856-897.
14. Donoho D.L., Huo X. Beamlets and multiscale image analysis / In *Multiscale and Multiresolution Methods, Lecture Notes in Computational Science and Engineering*. Springer, 2001.
15. Pennec E., Mallat S. Image compression with geometric wavelets // In *Proc. IEEE Int. Conf. on Image Proc.*, Vancouver, Canada. 2000. P. 661-664.
16. Starck J. L., Emmanuel J., Candes E., Donoho D. The curvelet transform for image denoising // *IEEE Trans. Image Proc.* 2002. V. 11, № 6. P. 670-684.
17. Starck J.L., Murtagh F., Candes E., Donoho D. Gray and color image contrast enhancement by the curvelet transform // *IEEE Trans. Image Proc.* 2003. V. 12, № 6. P. 706-717.
18. Кравчук А.С. Основы компьютерной томографии. – М.: Дрофа, 2001.
19. Guo H., Burrus C.S. Convolution using the undecimated discrete wavelet transform // *Int. Conf. Acoustics, Speech, Signal Processing (ICASSP-96)*, 1996. V. 3. P. 1291.
20. Malla S. A wavelet tour of signal processing. Academic Press, 1999.

УДК 621. 397

СОГЛАСОВАННЫЕ ТРЕХМЕРНЫЕ ВЕЙВЛЕТ-ФИЛЬТРЫ

Приоров А.Л.

Введение

Использование вейвлет-теории прочно вошло в основные направления обработки сигналов: удаление шума, сокращение избыточности представления сигналов и изображений (сжатие) [1,2]. Одной из актуальных задач цифровой обработки является сжатие видеопотоков. Эволюция развития методов сжатия видеопотоков проходит путь от методов покадровой обработки (использующих прореживание кадров по времени, кодирование опорных кадров и вычисление векторов движения с целью предсказания прореженных кадров) до методов кодирования целостных видеофрагментов в трехмерном (3D) пространстве, где используются две пространственные координаты, а третья координата - временная [3]. Последние методы обладают лучшими возможностями для кодирования видеопотоков с высокими коэффициентами сжатия при минимальных потерях. Один из вариантов добиться этого – кодирование движения во фрагменте в глобальном смысле (т.е. не на уровне оценки смещения блока пикселей, а на уровне описания движения произвольных однородно движущихся областей фрагмента). Решение этой задачи в общем случае требует значительных вычислительных затрат. Одним из способов приблизиться к ее решению является разложение фрагмента по системе ортогональных функций (например 3D вейвлет-функций). В этом случае каждый трехмерный блок коэффициентов дерева разложения нужно оценивать, учитывая параметры движения (вектор скорости и т.п.). Это позволяет несколько уменьшить эффект блочности восстановленного фрагмента по сравнению со стандартными методами сжатия (например, JPEG2000) [3]. Согласованные 3D вейвлет-фильтры позволяют подойти к решению этой задачи с другой стороны, а именно, на первой стадии обработки оценить характер движения и в последующем заложить характер движения фрагмента в структуру используемого вейвлет-фильтра.

Другое направление многомерной цифровой обработки сигналов связано с обработкой 3D объектов (распознавание, классификация). Это обусловлено значительным развитием средств трехмерной визуализации. Разработаны многочисленные критерии оценки 3D объектов: в пространственной области (по форме поверхности тела, по осям симметрии и т.д.), в области трансформант (например, 3D Фурье-разложение, разложения по сферическим функциям [4] и т.д.), комбинированные методы, использующие вычисление 3D моментов во

Работа посвящена изучению согласованных трехмерных цифровых вейвлет-фильтров. Рассматриваются фильтры для работы с трехмерными объектами двух классов. Первый класс – видеопоследовательности при наличии движения, второй – тела, синтезируемые искусственно с использованием стандартных средств визуализации. Получено математическое выражение для расчета амплитудно-частотной характеристики (АЧХ) согласованных 3D вейвлет-фильтров произвольной кратности. Анализ свойств данной АЧХ показал возможность применения согласованных 3D вейвлет-фильтров в различных практических задачах, одна из которых – оценка параметров движения в видеопоследовательности, вторая – классификация тел, что может быть полезно в задаче распознавания как объемных тел, так и их проекций.

временной области и в области трансформант [5]. Рассматриваемые в работе трехмерные согласованные вейвлет-фильтры предлагается использовать как альтернативный инструмент оценки и дальнейшей классификации объемных тел. По способу работы предлагаемый метод относится к классу методов, работающих в области трансформант. Он направлен на выявление линейных областей (плоскостей) рассматриваемого тела и его дальнейшего описания посредством задания положения образующих плоскостей.

Согласованные 2D вейвлет-фильтры

Понятие “согласованного двумерного вейвлет-фильтра” введено и рассмотрено в работах [6-9] применительно к задаче сжатия изображений с минимальными потерями. В основе расчета такого фильтра лежит следующая задача: рассчитать такой двумерный вейвлет-фильтр, чтобы при многополосном вейвлет-разложении осуществить восстановление первоначального изображения, используя только отсчеты импульсной характеристики вейвлет-фильтра и отсчеты одной из полученных полос разложения (например, из низкочастотной области). Одним из результатов проведенных исследований стало получение алгоритма, позволяющего для заданного количества полос разложения получить соответствующий вейвлет-фильтр. Для обозначения такого фильтра и был предложен термин “согласованный двумерный вейвлет-фильтр”. Свойство “согласованности” такого фильтра обусловлено тем, что его импульсная характеристика согласована с отсчетами обрабатываемого изображения, а амплитудно-частотная характеристика получена из отношения амплитудных спектров исходного и искаженного сжатием изображений.

Структура многополосного двумерного вейвлет-разложения показана на рис. 1. Параметр D – есть кратность вейвлет-разложения, например, при параметре $D = 2$ реализуется вейвлет-разложение на 4 частотные полосы

(одна низкочастотная, три – высокочастотные). Это стандартная кратность разложения, принятая во многих кодеках. При значениях $D = 4, 8, 16$ реализуются разложения соответственно на 16, 64, 256 частотных плоскостей.



Рис. 1. Схема двумерного многополосного вейвлет-разложения кратности D

Опуская сопутствующие математические расчеты, рассмотрим выражение для расчета согласованного вейвлет-фильтра для кратности многополосного разложения, равной D . Пусть $f(n_1, n_2)$ – исходное изображение, $F^2(\omega_1, \omega_2)$ – квадрат модуля его Фурье-спектра, $H^2(\omega_1, \omega_2)$ – квадрат АЧХ низкочастотного согласованного вейвлет-фильтра, который вычисляется в соответствии с выражением:

$$H^2(\omega_1, \omega_2) = \frac{F^2(\omega_1, \omega_2)}{\sum_{k=0}^{D-1} \sum_{n=0}^{D-1} F^2\left(\omega_1 + \frac{2k\pi}{D}, \omega_2 + \frac{2n\pi}{D}\right)}. \quad (1)$$

Поиск соответствующей импульсной характеристики осуществляется путем решения классической фазовой задачи, более детально описанной в работе [10].

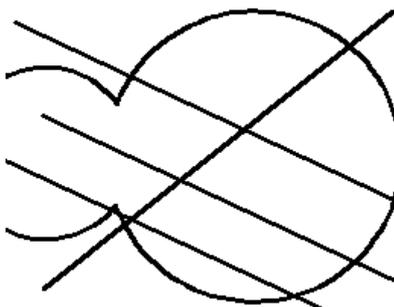
Далее в работе для краткости термин “квадрат АЧХ согласованного вейвлет-фильтра, соответствующего некоторому тестовому изображению” заменен на термин “Н-образ тестового изображения”.

О принадлежности полученного фильтра к классу вейвлет-фильтров говорит, прежде всего, ортогональность его импульсной характеристики $h_{n,m}$ к ее целочисленным сдвигам [1, 2]:

$$\sum_n \sum_m h_{n,m} h_{n+2k, m+2p} = \delta_{k,p}.$$

В частотной области это свойство соответствует свойству квадратурно-зеркальности квадрата АЧХ соответствующего вейвлет-фильтра

$$H^2(\omega_1, \omega_2) + H^2(\omega_1 + \pi, \omega_2) + H^2(\omega_1, \omega_2 + \pi) + H^2(\omega_1 + \pi, \omega_2 + \pi) = 1.$$



а)

В случае многополосного разложения произвольной кратности это свойство запишем в обобщенном виде:

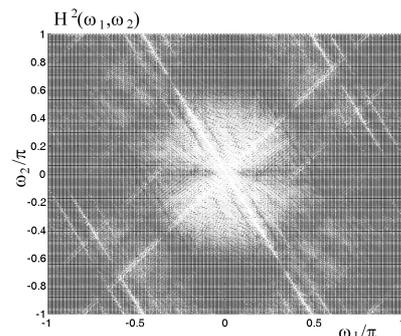
$$\sum_{k=0}^{D-1} \sum_{n=0}^{D-1} H^2\left(\omega_1 + \frac{2k\pi}{D}, \omega_2 + \frac{2n\pi}{D}\right) = 1. \quad (2)$$

Отметим, что функция $H^2(\omega_1, \omega_2)$, заданная выражением (1), удовлетворяет равенству (2).

Несмотря на то, что выражение (1) было получено при решении задачи синтеза как согласованных вейвлет-фильтров, так и близких к ним фильтров (обозначенных в [6, 7] как “оптимизированные вейвлет-фильтры”) с целью минимизации ошибки восстановленного изображения, дальнейший анализ полученной двумерной функции (1) показал, что это выражение может быть использовано и для оценки свойств анизотропности обрабатываемых изображений, и для распознавания на изображении различных прямых. Под анизотропностью изображения понимается наличие в текстуре изображения одного или нескольких преобладающих направлений. Например, для тестового изображения “Барбара” преобладающее направление задают полосы на одежде, для изображения “Фотограф” преобладающее направление определяет тренога фотоаппарата, а на изображении “Корабли” – мачты кораблей.

Свойства двумерных согласованных вейвлет-фильтров изучены в работе [6], где приведен их обзор. Основным свойством АЧХ согласованного вейвлет-фильтра является хорошая чувствительность к линейным объектам различной природы. Это явно выраженные линии, полученные при оцифровке чертежа, и линии, образованные путем пересечения плоскостей (например, ребра куба, показанного в изометрической проекции, линия горизонта и т.п.), а также линейные объекты с плавным перепадом яркости. Главной особенностью согласованных 2D вейвлет-фильтров применительно к линейным объектам является то, что Н-образом линейных объектов является набор прямых линий, приведенных к началу координат с точным сохранением угла наклона, причем их длина и интенсивность отображения пропорциональна совокупной длине линий заданного угла наклона на первоначальном изображении.

Второй особенностью согласованных 2D вейвлет-фильтров является то, что при увеличении кратности вейвлет-фильтра (т.е. параметра D в выражении (1)) уменьшается чувствительность к линейным объектам малой протяженности. Это позволяет выделять линейные объекты на общем фоне изображения (рис. 2), в том числе и при его высокой зашумленности. Видно, что при увеличении значения параметра D Н-образ, соответствующий криволинейным объектам, вырождается.



б)

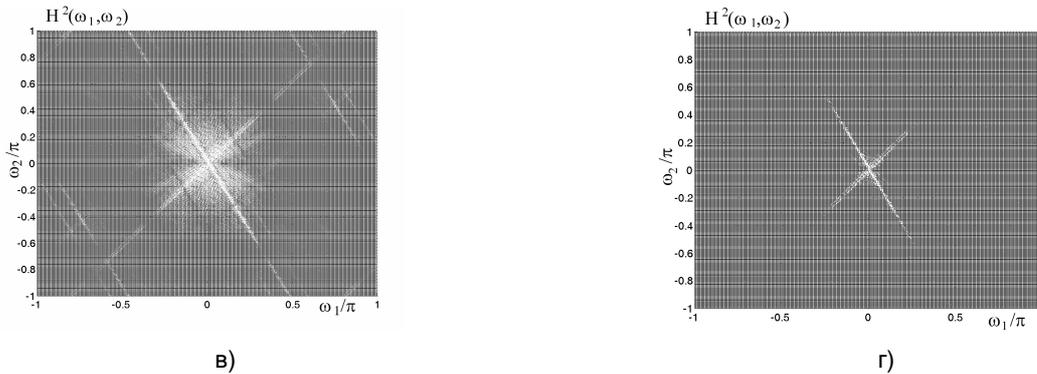


Рис. 2. Квадраты АЧХ согласованных вейвлет-фильтров для различных кратностей разложения: а) тестовое изображение; б) H-образ при кратности $D = 2$; в) H-образ при кратности $D = 4$; г) H-образ при кратности $D = 8$

Следует отметить, что хорошей чувствительностью к выделению линейных объектов обладают все методы, основанные на спектральном анализе изображения (это преобразование Радона, методы риджлет- и курвлет-анализа изображений, и др.). Несмотря на это, метод, основанный на использовании согласованных вейвлет-фильтров, имеет некоторые преимущества:

- для выявления линейных объектов синтезируется фильтр, который уже на этапе формирования относится к классу ортогональных вейвлет-фильтров, что позволяет выполнять дальнейшую обработку изображения с использованием классической теории вейвлет-анализа;

- обеспечивается различная чувствительность к линиям разной длины, что достигается путем изменения кратности разложения.

По этой причине и проводится работа по исследованию трехмерных согласованных вейвлет-фильтров. Основная цель проводимой работы – разработка нового инструмента для оценки параметров движения в видеопоследовательности и для классификации объемных тел. Предполагается, что с использованием согласованных трехмерных вейвлет-фильтров удастся приблизиться к решению задачи оптимального распознавания трехмерных объектов по их проекциям.

Согласованные 3D вейвлет-фильтры

В основе расчета такого фильтра лежит задача, аналогичная двумерному разложению: рассчитать такой трехмерный вейвлет-фильтр, чтобы при многополосном вейвлет-разложении осуществить восстановление первоначального объема, используя только отсчеты импульсной характеристики вейвлет-фильтра и отсчеты одной из полученных областей разложения (например, из низкочастотной области). На рис. 3 показана структура многополосного трехмерного разложения кратности D .

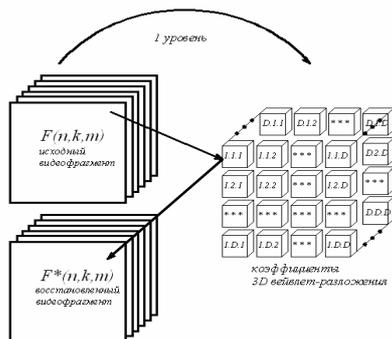


Рис. 3. Схема трехмерного многополосного прямого и обратного вейвлет-разложения кратности D

Аналитически выполнить расчет выражения для квадрата АЧХ согласованного трехмерного вейвлет-фильтра оказалось достаточно трудно. Поэтому соответствующее выражение получено путем обобщения одномерного и двумерного случая на трехмерную область (3а, 3б, 3в):

$$H^2(\omega) = \frac{F^2(\omega)}{\sum_{k=0}^{D-1} F^2\left(\omega + \frac{2k\pi}{D}\right)}, \tag{3а}$$

$$H^2(\omega_1, \omega_2) = \frac{F^2(\omega_1, \omega_2)}{\sum_{k=0}^{D-1} \sum_{n=0}^{D-1} F^2\left(\omega_1 + \frac{2k\pi}{D}, \omega_2 + \frac{2n\pi}{D}\right)}, \tag{3б}$$

$$H^2(\omega_1, \omega_2, \omega_3) = \frac{F^2(\omega_1, \omega_2, \omega_3)}{\sum_{k=0}^{D-1} \sum_{n=0}^{D-1} \sum_{m=0}^{D-1} F^2\left(\omega_1 + \frac{2k\pi}{D}, \omega_2 + \frac{2n\pi}{D}, \omega_3 + \frac{2m\pi}{D}\right)}. \tag{3в}$$

Конечно, для того чтобы практически с нулевой ошибкой выполнить сжатие видеопоследовательности в самом общем случае (с высоким уровнем детализации и с большим числом движущихся объектов) в 8 раз, порядок такого вейвлет-фильтра $N \times M \times K$ должен быть сопоставимым с порядком видеопоследовательности. Однако выражение (3в) может быть использовано для расчета оптимизированных трехмерных вейвлет-фильтров, и в перспективе оно может использоваться для оценки движения в глобальном смысле, т.е. в рамках всей сцены наблюдения. Под оптимизированным трехмерным вейвлет-фильтром понимается фильтр $H_{opt}(j\omega_1, j\omega_2, j\omega_3)$ относительно небольшого размера $n \times m \times k$ (например, $20 \times 20 \times 20$), структура которого приближенно описывает структуру соответствующего согласованного фильтра $H(j\omega_1, j\omega_2, j\omega_3)$:

$$\lim_{(n \times m \times k) \rightarrow (N \times M \times K)} \iiint (H_{opt}^2(\omega_1, \omega_2, \omega_3) - H^2(\omega_1, \omega_2, \omega_3)) d\omega_1 d\omega_2 d\omega_3 \rightarrow 0.$$

Аналогично двумерному случаю, вводится понятие “квадрат АЧХ согласованного 3D вейвлет-фильтра, соответствующего некоему объему”, которое предлагается обозначить как H-образ этого объема.

В ходе исследования трехмерных согласованных вейвлет-фильтров установлено их основное свойство: H-образом плоскости является плоскость, отнесенная к началу координат при сохранении ориентации в пространстве. H-образом исследуемого объема из нескольких

плоскостей с одинаковой ориентацией является плоскость с интенсивностью, пропорциональной удельному весу совокупности этих плоскостей относительно прочих элементов объема. И по аналогии с двумерным случаем, при увеличении кратности вейвлет-фильтра (параметр D (3)) уменьшается чувствительность к небольшим криволинейным объектам. На основании этого свойства формируются частные выводы, в том числе по оценке параметров движения, что будет рассмотрено ниже. На рис. 4а показан пример согласованного 3D вейвлет-фильтра кратности $D = 4$, полученного для тетраэдра (в объеме размером $128 \times 128 \times 128$ пикселей).

Из-за сложности представления 3D объема H-образа на плоскости листа, на рис. 4б, 4в показано только по одной проекции H-образа $H^2(\omega_1, \omega_2, \omega_3)$. На рис. 4г представлена объемная интерпретация полученного H-образа, построенная средствами так называемой воксельной графики ("voxel" – элемент объемного изображения). Для уменьшения вычислительной сложности при отображении H-образа соответствующий объем переведен в бинарный формат:

$$\begin{cases} \text{если } H^2(\omega_1, \omega_2, \omega_3) > T, \text{ то } H^2(\omega_1, \omega_2, \omega_3) = 1 \\ \text{иначе } H^2(\omega_1, \omega_2, \omega_3) = 0 \end{cases}$$

где порог T выбирается из условия $T = 0.25 \cdot \max(H^2(\omega_1, \omega_2, \omega_3))$.

Установлено, что H-образом тетраэдра является пересечение шести плоскостей, что соответствует теории о приведении каждой плоскости к началу координат (рис. 4).

С целью получения численной характеристики полученного квадрата АЧХ согласованного 3D вейвлет-фильтра, по аналогии с двумерным случаем, в работе используется функция углового распределения интенсивностей $\Theta(\alpha, \beta)$ значений квадрата АЧХ, полученная путем перевода в полярную систему координат и интегрированием по переменной ρ :

$$H^2(\omega_1, \omega_2, \omega_3) \rightarrow H^2(\alpha, \beta, \rho), \quad \Theta(\alpha, \beta) = \int H(\alpha, \beta, \rho) d\rho,$$

где $\omega_1 = \rho \cos(\alpha) \sin(\beta)$, $\omega_2 = \rho \sin(\alpha) \sin(\beta)$, $\omega_3 = \rho \cos(\beta)$.

Функция углового распределения для объекта на рис. 4 показана на рис. 5. На рис. 5б максимумы функции распределения специально выделены.

Для подтверждения полученного вывода на рис. 6а рассмотрена объемная фигура, состоящая из девяти граней (верхняя, нижняя и семь боковых). H-образом этой фигуры является набор пересекающихся плоскостей. Это хорошо видно на рис. 6в.

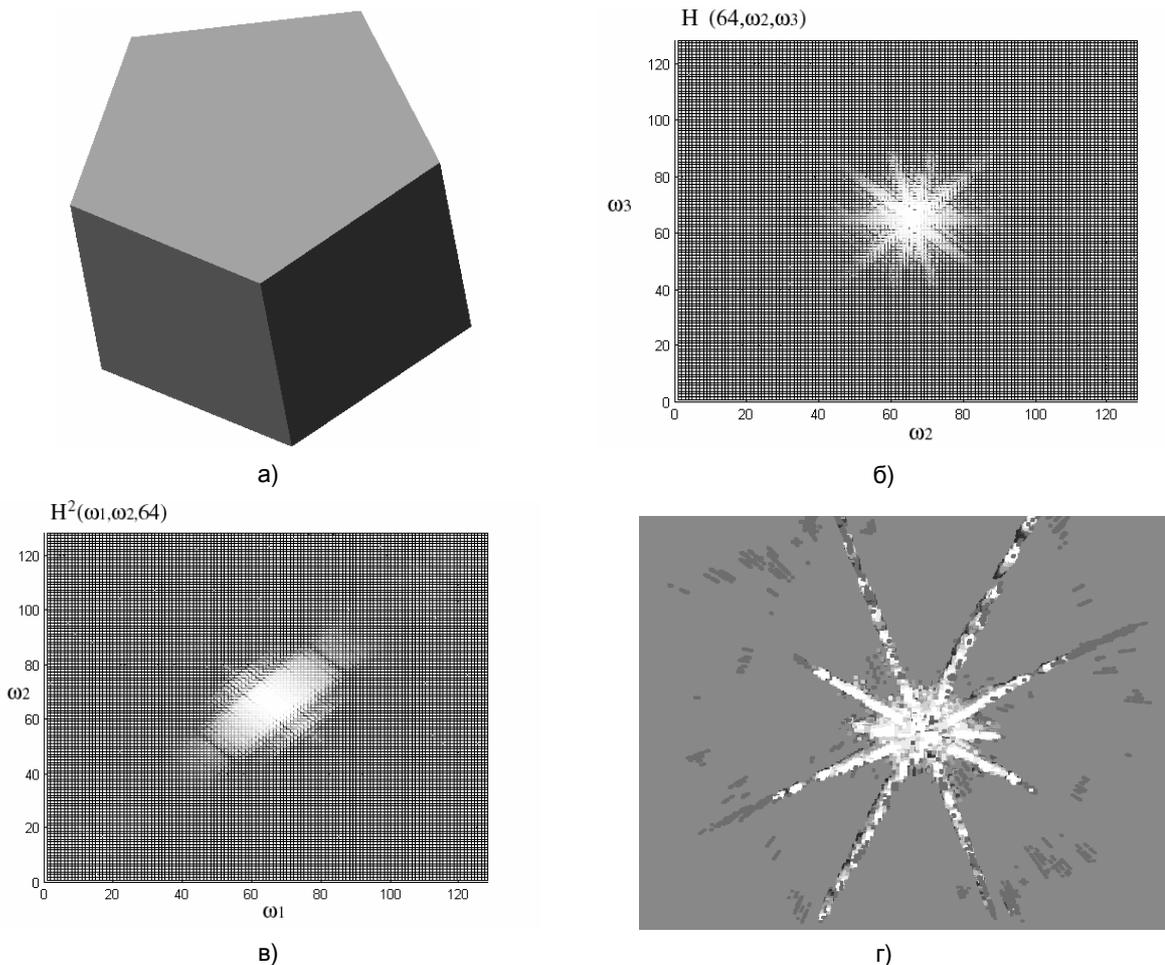


Рис. 4. Пример расчета согласованного 3D вейвлет-фильтра для объемного тела: а) тетраэдр; б), в) проекции квадрата АЧХ согласованного 3D вейвлет-фильтра; г) визуализация квадрата АЧХ согласованного вейвлет-фильтра (бинаризованный H-образ кратности $D=4$)

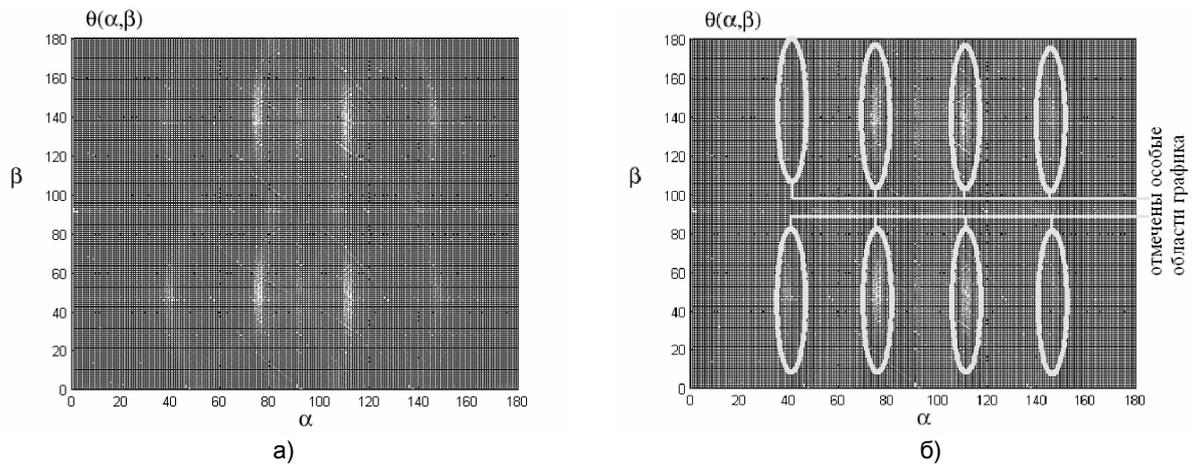


Рис. 5. Угловое распределение квадрата АЧХ согласованного 3D вейвлет-фильтра

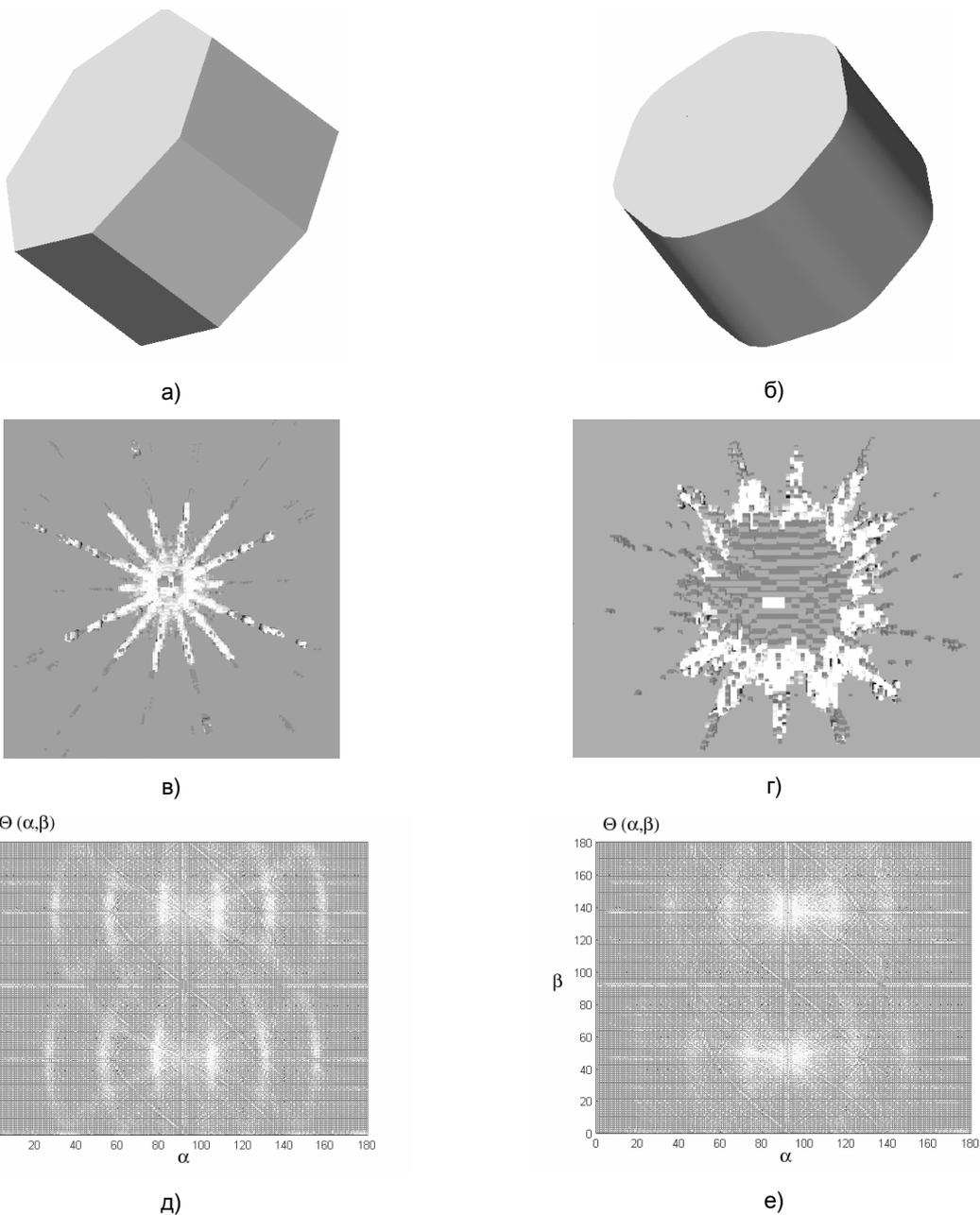
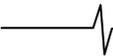


Рис. 6. Примеры расчета согласованных вейвлет-фильтров: а), б) примеры тестовых объемных тел; в), г) бинаризованные Н-образы (кратность $D=4$); д), е) функции углового распределения



Отметим, что основное свойство согласованных 3D вейвлет-фильтров состоит в том, что H -образом произвольной плоскости объемного тела является плоскость. Это распространяется и на объемные тела сложной формы со сглаженными участками поверхности. В этом случае любая поверхность может быть рассмотрена как совокупность плоскостей меньшей площади. Соответствующий пример приведен на рис. 6б. Здесь рассматривается объемное тело, аналогичное телу, представленному на рис. 6а, но со скругленными гранями. Видно, что общая площадь поверхности тела практически не изменилась, однако удельный вес отдельных плоскостей уменьшился, и, как следствие, уменьшается возможность выделения этих плоскостей с использованием H -образа.

Один из вариантов применения предлагаемого инструмента анализа – использование H -образов для классификации объемных тел путем выделения различных плоскостей поверхности тела и комплексной оценки их взаимной ориентации, что предлагается выполнять на базе вычисления 3D моментов полученных H -образов [5].

Как уже отмечалось, другой вариант использования инструмента H -образов – выявление движения в видеопоследовательности и оценка его параметров.

На рис. 7 показан пример функции $H^2(\omega_1, \omega_2, \omega_3)$ для простого движения точки в плоскости листа. Видно, что прямолинейному движению соответствует на H -образе некоторая плоскость, которая в проекции на рис. 7б вырождается в прямую линию с углом наклона, соответствующим вектору движения объекта. Соответственно, при прямолинейном движении нескольких объектов с различными направлениями движения H -образом будет являться пересечение нескольких

плоскостей, что в проекции вырождается в пересечение прямых линий. Отметим, что криволинейному движению соответствует конус с вершиной в точке начала координат.

Сформулируем некоторые частные свойства согласованных трехмерных фильтров, касающиеся анализа движения в видеопоследовательности:

1. Прямолинейному движению объекта произвольной формы соответствует H -образ, представляющий собой плоскость с углами наклона, однозначно определяющими направление и скорость движения.

2. Движению нескольких объектов с различными векторами движения соответствует H -образ, представляющий несколько наклонных плоскостей.

3. Движению нескольких объектов с одинаковыми векторами движения соответствует H -образ, представляющий собой одну наклонную плоскость.

4. Криволинейному движению соответствует конус с вершиной в точке $(0,0,0)$.

5. H -образ неподвижных объектов тождественно равен нулю.

6. При увеличении кратности разложения (параметр D в разложении (3)) уменьшается чувствительность H -образа к небольшим движениям в кадре.

Приведенные свойства согласованных вейвлет-фильтров позволяют использовать их для решения задачи обнаружения однородно движущихся областей в видеофрагменте. Рассмотрим небольшой пример, показывающий перспективы использования согласованных 3D вейвлет-фильтров. Съемка осуществляется камерой, почти равномерно двигающейся вдоль наблюдаемой сцены (например, скоростная купольная видеокамера работает в режиме автопанарамирования).

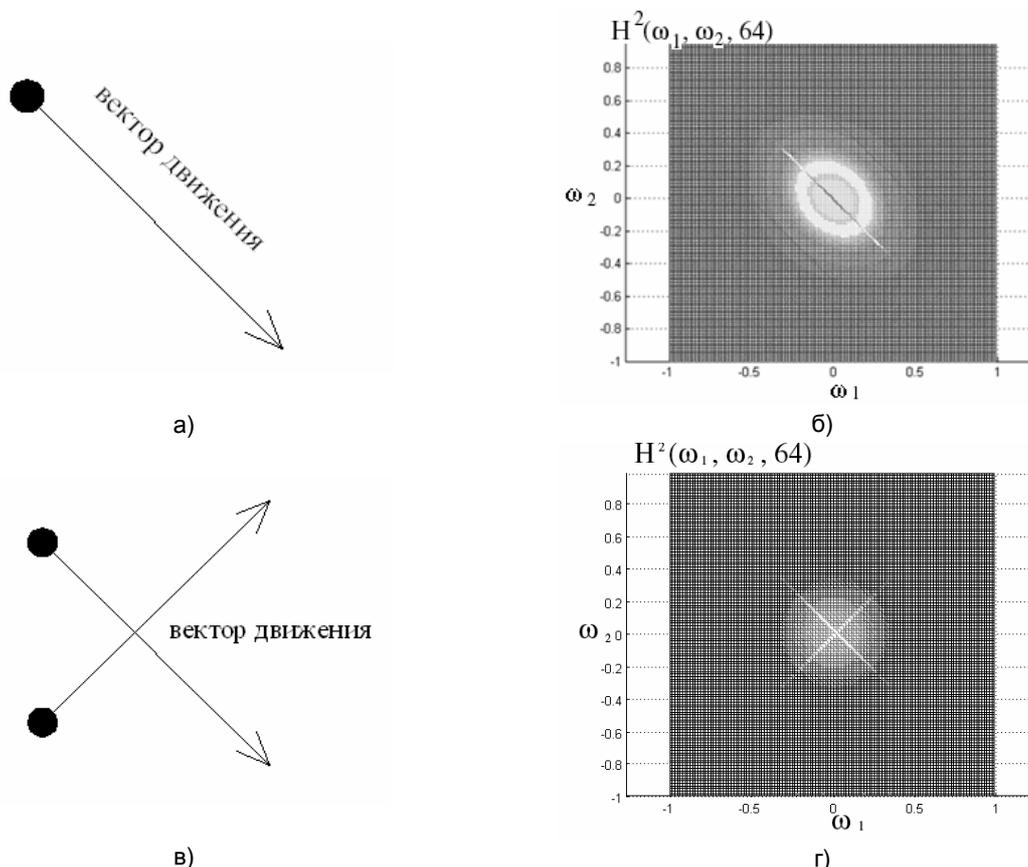


Рис. 7. Применение H -образов для оценки параметров движения: а), в) примеры движущегося объекта (разрешение кадра 128×128 пикселей, количество кадров – 128); б), г) проекции трехмерного H -образа

Полученный видеофрагмент будет содержать движения двух типов: равномерное перемещение фона и произвольное движение объектов наблюдения. Использование согласованных 3D вейвлет-фильтров высокой кратности позволит в каждый момент времени выполнять оценку перемещение фона, что позволит выделить движение объектов наблюдения на общем подвижном фоне.

Заключение

Работа посвящена изучению согласованных вейвлет-фильтров. В этой связи рассматриваются трехмерные объекты двух классов. Первый класс – видеопоследовательности при наличии движения. Второй класс – тела, синтезируемые искусственно с использованием стандартных средств визуализации (например, AutoCAD, 3D Studio). Получено математическое выражение для расчета амплитудно-частотной характеристики согласованных 3D вейвлет-фильтров произвольной кратности. Анализ свойств данной АЧХ показал возможность применения согласованных 3D вейвлет-фильтров в различных практических задачах. Одна из них – оценка параметров движения в видеопоследовательности (вычисление преобладающих направлений и скоростей движущихся объектов с возможностью установки порога отсеивания незначительных движений кадра). Вторая задача – классификация тел, что может быть полезно в задаче распознавания как объемных тел, так и их проекций.

Литература

1. Добеши И. Десять лекций по вейвлетам. – Москва-Ижевск: НИЦ "Регулярная и хаотическая динамика", 2004. – 464 с.
2. Воробьев В.П., Грибунин В.Г. Теория и практика вейвлет-преобразования. – СПб.: Военный университет связи, 1999. – 204 с.
3. Gonzalez J., Bernabe G., Garcia J.M. Reducing 3D fast wavelet transform execution time using blocking and streaming SIMD extensions // Journal of VLSI Signal Processing. 2005. № 41. P. 209-223.
4. Kazhdan M., Funkhouser T., Rusinkiewicz S. Rotation invariant spherical harmonic representation of 3D shape descriptors // Proc. of symposium on geometry processing. Aachen, Germany. 2003. P. 156-164.
5. Li Cui, Ying Li, Don Xu, Hua Li. Fractal scale descriptors base on wavelet moments for 3D objects // Proc. of 2007 int. conf. on wavelet analysis and pattern recognition. 2007. P. 1197-1202.
6. Кобелев В.Ю., Приоров А.Л. Анализ изображений при помощи согласованных двумерных вейвлет-фильтров // Радиотехника. 2008. № 1. С. 12-19.
7. Кобелев В.Ю., Приоров А.Л. Применение неразделимых вейвлет-фильтров в задачах сжатия изображений // Цифровая обработка сигналов. 2006. № 2. С. 21-26.
8. Кобелев В.Ю., Приоров А.Л. Синтез оптимизированных двумерных неразделимых вейвлет-фильтров для сжатия изображений // Телекоммуникации. 2006. № 9. С. 7-12.
9. Кобелев В.Ю., Моисеев А.А., Волохов В.А., Смоляков А.В. Синтез двумерных неразделимых вейвлет-фильтров с перестраиваемыми коэффициентами // Докл. 8-ой межд. конф. "Цифровая обработка сигналов и ее применение" (DSPA'06). Москва. 2006. Т. 2, С. 389-392.
10. Прэрт У. Цифровая обработка изображений. – М.: Мир, 1982. Т. 1, – 312 с.

У в а ж а е м ы е а в т о р ы !

Редакция научно-технического журнала "Цифровая обработка сигналов" просит Вас соблюдать следующие требования к материалам, направляемым на публикацию:

1) Требования к текстовым материалам и сопроводительным документам:

- *Текст - текстовый редактор Microsoft Word.*
- *Таблицы и рисунки должны быть пронумерованы. На все рисунки, таблицы и библиографические данные указываются ссылки в тексте статьи.*
- *Объем статьи до 12 стр. (шрифт 12). Для заказных обзорных работ объем может быть увеличен до 20 стр.*
- *Название статьи на русском и английском языках.*
- *Рукопись статьи сопровождается:*
 - *краткой аннотацией на русском и английском языках;*
 - *номером УДК;*
 - *сведениями об авторах (Ф.И.О., организация, телефоны, электронная почта).*

2) Требования к иллюстрациям:

- *Векторные (схемы, графики) - желательно использование графических редакторов Adobe Illustrator или Corel DRAW.*
- *Растровые (фотографии, рисунки) - М 1:1, разрешение не менее 300dpi, формат tiff.*

ЭВОЛЮЦИЯ ПРИМЕНЕНИЯ ПЛИС В СИСТЕМАХ ЦОС: ОТ СПЕЦИАЛИЗИРОВАННОГО СОПРОЦЕССОРА К АППАРАТНОЙ ПЛАТФОРМЕ (Часть 1)

Балдин М.В., Воронков Д.И., Руткевич А.В., Сенченко М.Н., Стещенко В.Б., Шишкин Г.В.

Введение

Десять лет назад были опубликованы работы [1,2], в которых рассматривались тенденции в реализации тех или иных классов алгоритмов цифровой фильтрации на существующей элементной базе. Следует отметить, что направление, заложенное в этих работах, оказалось востребованным – сейчас, когда основным критерием успешности проекта стало время выхода на рынок и окупаемость, вопрос о подходах к практической реализации алгоритмов в конкретных изделиях встал как нельзя остро.

Чем же характеризуется современное состояние в проектировании аппаратуры ЦОС? Выделим основные изменения с точки зрения элементной базы и подходов к проектированию систем ЦОС:

- Значительное возрастание вычислительной мощности сигнальных процессоров, появление многоядерных архитектур, снижение стоимости и потребляемой мощности.
- Появление ПЛИС эквивалентной емкостью более 1 млн. вентилях, встроенными процессорными ядрами, специализированными аппаратными блоками вычислителей, развитыми интерфейсами [3,4,5].
- Возрастание доступности современных микроэлектронных технологий для разработчика систем. Если еще 6 – 8 лет назад имелись только единичные попытки отечественных компаний выхода на современное субмикронное кремниевое производство, то на данный момент существует огромное число разработок под нормы 0,18 – 0,25 мкм, появились первые отечественные наработки с уровнем 0,13 – 0,09 мкм. Вводятся современные мощности и на отечественных предприятиях (ОАО «НИИ МЭ и завод «Микрон», ОАО «Ангстрем», ФГУП «НИИ ИС им. Седаква» и др.). Такой крен в сторону доступности кремниевых технологий порождает некоторые специфические методы в проектировании, о которых будет рассказано ниже.
- Возрастание сложности и возможностей отладочных средств сигнальных процессоров и средств САПР ПЛИС и СБИС. Практически повсеместным стал отказ от использования ассемблера как основного языка разработки ПМО. Возрастает роль использования современных маршрутов в проектировании аппаратуры [6,7,8,9]. Эти маршруты основаны на использовании языков описания аппаратуры высокого уровня, многокомпонентных наращиваемых сетевых средствах проектирования. Появилась необходимость интеграции разработки программного и аппаратного обеспечения в одном цикле (сопряженное проектирование), средства разработки алгоритмов стали интегрироваться со средствами разработки аппаратуры.

Дан анализ тенденций в развитии методов проектирования систем и аппаратуры ЦОС с учетом современных тенденций развития как массовой, так и специальной электроники с использованием технологий аппаратных платформ. Рассмотрены особенности процесса проектирования с учетом рыночных механизмов. Приведены примеры существующих зарубежных и отечественных аппаратных платформ для верификации СБИС и обработки сигналов, показаны особенности их проектирования.

- Изменилась номенклатура интерфейсов, возросла скорость информационного обмена. Современные последовательные интерфейсы обеспечивают как высокую скорость передачи информации, так и стандартизацию и унификацию этого процесса. Изменились подходы к схемотехническому проектированию и конструированию аппаратуры – возросла роль моделирования целостности сигналов, тепловых режимов, как при проектировании узлов (плат), так и кристаллов СБИС. Налицо значительное увеличение влияния технологии изготовления на процесс проектирования.

Вышеперечисленные особенности являются следствием того, что полупроводниковая промышленность занимает в рыночной мировой экономике уникальное положение: она развивается по детально разработанному плану, который, не только не препятствует конкурентоспособности участников, но даже всемерно помогает ей. Этот план известен как International Technology Roadmap for Semiconductors (ITRS) и представляет собой план-прогноз, ежегодно обновляемый и публикуемый международной организацией Semiconductor Industry Association (SIA). [10,11] В основе ITRS лежат несколько простых принципов, в том числе знаменитый закон Мура об удвоении числа элементов СБИС каждые 1,5-2 года. Закон Мура не отражает никаких фундаментальных законов природы, а лишь описывает ситуацию, складывающуюся на рынке в результате конкуренции между производителями, а также вследствие взаимного стимулирования радиоэлектронной и полупроводниковой отраслей. Огромную роль играет чисто психологический фактор: разработчики и производители стараются придерживаться прогнозных сроков закона Мура и ITRS, потому что знают, что так же действуют конкуренты. Более того, стремление обогнать конкурентов часто приводит к тому, что разработчики создают, а производители выводят на рынок новые изделия раньше, чем это предусмотрено ITRS. В результате в наши дни закон Мура, строго говоря, должен формулироваться иначе, поскольку из экспоненциального он превратился в суперэкспоненциальный. Главное значение ITRS состоит в том, что этот документ не просто прогнозирует динамику параметров, но и содержит точные указания относительно того, какими конструкторскими и технологическими средствами новые параметры могут быть достигнуты, когда и какие

технические средства должны быть разработаны и освоены производством.

Фундаментальным понятием ITRS является "масштабирование" (scaling) - пропорциональное уменьшение всех геометрических размеров ИС в каждом новом поколении, благодаря чему обеспечивается снижение массогабаритных параметров, повышение рабочих и тактовых частот, уменьшение потребляемой мощности и т.д. Каждому "узлу" сетевого графика ITRS соответствует очередное поколение ИС, характеризующееся некоторым линейным параметром, который в ITRS именуется просто "узлом". За величину характеристического размера (XP) данного поколения ИС обычно принимают половину шага между дорожками сигнала в схемах памяти; для микропроцессоров XP определяется как половина шага между поликремниевыми затворами МОП-транзисторов.

С переходом к "узлам" менее 130 нм в конструировании ИС возникли принципиально новые проблемы. Помимо проблем технологического свойства, связанных с тем, что традиционная конструкция МОП-транзистора перестает работать из-за различных паразитных эффектов, проявляющихся в малоразмерных конструкциях, возникли проблемы, связанные с программно-аппаратным и методологическим обеспечением процесса проектирования. Дело в том, что с улучшением проектных норм микроэлектронных изделий возрастает стоимость подготовки производства и, соответственно, цена риска технической и идеологической ошибки.

Так, в настоящее время, средняя стоимость подготовки производства (изготовление фотошаблонов) и выпуск опытной партии в количестве 10 -12 пластин (так называемый «инженерный лот») на фабриках Юго-Восточной Азии (X-fab Sarawak, Siltera и т.п.) составляет по технологии 0.18 мкм – 120 000 \$, то стоимость фотошаблонов при проектных нормах 0.13 мкм составляет 350 000 \$, а по технологии 0.09 мкм - около 1 000 000 \$. Очевидно, что при такой динамике, цена ошибки возрастает многократно. Тенденция уменьшения проектных норм в первую очередь связана со стремлением получить как можно больше кристаллов с одной пластины, поскольку стоимость пластины составляет 800 – 1800 долларов, при этом с уменьшением размера кристалла в два раза выход увеличивается в четыре раза. При этом известно, что выход годных кристаллов не зависит от минимального размера, а с уменьшением размеров элементов увеличивается съём кристаллов с пластины, значит, они становятся дешевле.

Таким образом, разработчик массовой аппаратуры оказывается под дамокловым мечом ответственности за принимаемые проектные решения.

Не легче ситуация и в высокотехнологичном, но малосерийном секторе аппаратуры для космической и военной техники. Для этого сегмента рынка характерны:

- Широкая функциональная номенклатура (по данным ESCIES (www.escies.org) – координационного центра Европейского космического агентства по применению электронных компонентов для космических программ порядка 1500 типоминималов).
- Крайне малая серийность (не более 15-100 тыс. шт.), не характерная для магистрального направления развития электронной промышленности.
- Высокие требования к надежности (безотказность, ресурс, сохраняемость).

- Стойкость к воздействию ионизирующих излучений космического пространства и других факторов.
- Расширенный температурный диапазон.
- Необходимость обеспечения длительных сроков безотказной работы (15 лет и более).

Как для массовой, так и для специальной электроники тенденцией последних лет стало жесткое требование по уменьшению срока выхода продукции на рынок. Это накладывает особенно жесткие требования на методологию и средства проектирования, которые должны обеспечить минимум итераций при освоении производства. Значительно возрастает сложность процесса проектирования. При использовании традиционных методов проектирования хороший дизайнер может выполнять проект со средней скоростью порядка 100 вентиля в день или 30 строк RTL кода. В этом случае, чтобы спроектировать СБИС сложностью 100 тыс. вентиля потребуется 1000 человеко-дней, т.е. команда из 5-ти человек сможет разработать такую СБИС в течение года. Следуя данной логике, чтобы разработать сложную СБИС порядка 10 млн. вентиля в течение одного года потребуется команда из 500 человек, что неприемлемо с точки зрения стоимости разработки.

По более точным прогнозам (рис. 1), если при переходе на глубокие субмикронные технологии (0,18 -0,13 мкм) пользоваться существующей методологией проектирования, то стоимость проекта увеличивается до 250 человеко-лет, что неприемлемо для заказчика.

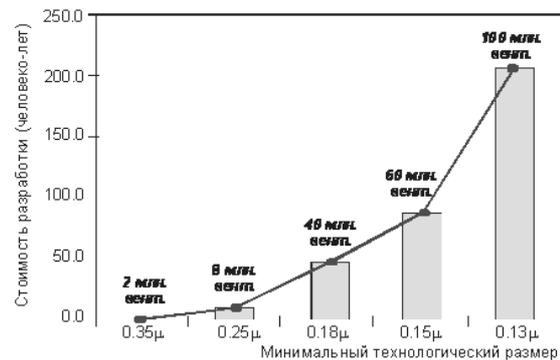


Рис.1. Зависимость стоимости разработки СБИС от технологии (по материалам фирмы Cadence Design Systems)

В последнее время сложилась тенденция постоянного роста доли затрат на разработку программного обеспечения (ПО) РЭА. Если вести разработку ПО и СБИС отдельно, то увеличивается вероятность выявления ошибок на этапах тестирования или эксплуатации всего комплекса аппаратуры.

Можно выделить дополнительно ряд причин, по которым необходимо переходить на новую методологию проектирования:

- в условиях рынка прибыль в значительной степени зависит от времени проектирования;
- такие технические параметры СБИС, как производительность, площадь кристалла и потребляемая мощность являются ключевыми элементами в продвижении товара на рынок;
- увеличение степени интеграции делает задачу верификации качественно более сложной;
- из-за новых особенностей технологии глубокого субмикрона (DSM – Deep Submicron) все труднее удовлетво-

ритель всем требованиям по временным ограничениям (timing);

- команды разработчиков высокоинтегрированных СБИС имеют различный уровень знаний и опыта в области проектирования, и часто при выполнении проектов СБИС расположены в различных частях мира.

Выход из создавшейся ситуации очевиден – необходимо изменить методологию проектирования СБИС. Наиболее перспективным направлением в настоящий момент представляется методология проектирования СБИС типа «система на кристалле» с использованием платформенного принципа организации [12].

Таким образом, становится очевидным общее место двух, казалось бы противоречивых направлений, – необходимость единого методического и программно-аппаратного обеспечения процесса проектирования. Такой подход к разработке мы назовем платформенным принципом проектирования (рис.2)



Рис. 2. Платформенный принцип проектирования СБИС и аппаратуры

Как обычно, в современной российской электронике нет четкого определения слову «платформа». В общем случае, в платформах СБИС класса «Система на кристалле» (СНК) должны содержаться, как минимум: процессорное ядро, контроллер памяти, универсальный асинхронный приемопередатчик, таймер, схема обеспечения безопасности, универсальные средства ввода/вывода с открытой шинной архитектурой для связывания всех элементов между собой, например, на базе шины AMBA или Wishbone (рис.3).

Платформенный подход к проектированию предполагает использование единой среды проектирования. Единая среда проектирования «комплекс – аппаратура – компоненты» -это технология создания систем и комплексов на

основе перспективных микроэлектронных технологий с учетом специализации компонентов под решение конкретных целевых задач аппаратуры и комплексов.

Данная технология предусматривает применение методов систематического повторного использования стандартных высокоинтегрированных сертифицированных СФ-блоков и платформенного подхода к проектированию и производству аппаратуры. Характеризуется высокой степенью унификации проектных решений, нормативной базой и возможностью контроля качества на любом этапе разработки, производства и эксплуатации.

Основными преимуществами такого подхода являются:

- значительное сокращение сроков и стоимости разработки сложных систем при сохранении высокой функциональности, за счет использования методологии «систематического повторного использования» IP-блоков и унифицированных аппаратных платформ;
- гибкость при решении нестандартных задач, за счет возможности наращивания физических интерфейсов и программного обеспечения;
- улучшение массогабаритных характеристик аппаратуры, построенной на базе СБИС аппаратных платформ;
- аппаратная верификация алгоритмов на всех стадиях разработки проекта.

Предлагаемая методика проектирования предусматривает инвариантность к используемым библиотекам, обеспечивая возможность миграции проекта. Включение платформ СНК в методологию конструирования имеет несколько преимуществ. Уменьшится риск интеграции, поскольку все СФ-блоки работают вместе, уменьшится время на лицензирование и на составление контракта, поскольку платформа ограничивается одной лицензией и значительно уменьшится стоимость, благодаря повторному использованию содержимого платформы в многочисленных последующих конструкциях.

Важным преимуществом СНК платформ является функциональная изоляция. Компоненты могут быть изолированы друг от друга и проверены. Сама платформа может быть также изолирована от своих компонентов и проверена независимо. Использование платформ помогает вводу конструкции в различные изделия, ввиду встроенной гибкости. При этом требуется минимальный объем «перепроектирования» и меньший объем повторной верификации. Если конструкторский коллектив не имеет ресурсов, времени, кадров для разработки гибкой

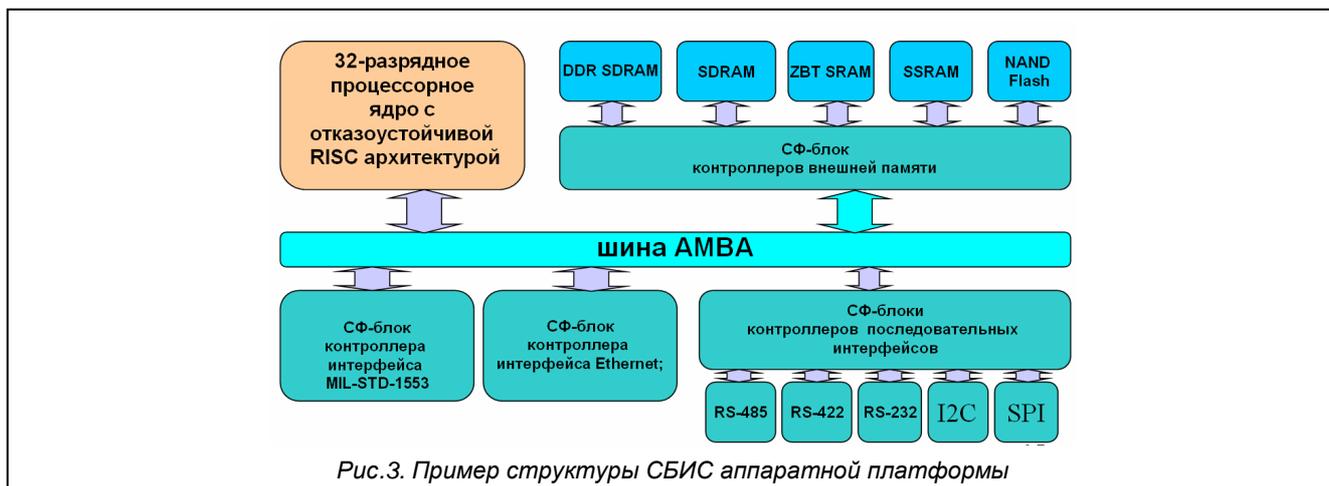


Рис.3. Пример структуры СБИС аппаратной платформы

платформы, он может приобрести необходимый СФ-блок у сторонней фирмы.

Платформенный принцип проектирования СнК является «прорывной» технологией, которая может сократить цикл конструирования на десять – двенадцать месяцев. Как и в случае инструментальных средств САПР, коэффициент окупаемости инвестиций в платформы СнК быстро становится очевидным, и реализация методологии платформ при конструировании представляется конкурентоспособным решением.

Разумеется, при работе с СнК платформами имеются определенные ограничения, связанные со случаями, когда требуются интенсивные модификации или когда требуются чрезмерные усилия по обеспечению конфигурируемости. Сюда могут относиться максимальная и минимальная ширина шины на кристалле, отображение системной памяти, максимальные размеры кристалла и частота, системная задержка. Совмещая концепцию СФ-блока с реконфигурируемой архитектурой и с базовой платформой, становится возможным говорить о реконфигурируемых аппаратных платформах для малосерийных и настраиваемых изделий.

При создании реконфигурируемых платформ также целесообразно использование совместного конструирования аппаратных и программных средств. Платформы могут быть поставлены с уже конфигурируемыми средами разработки, состоящими из сборочных файлов проекта, программ на макроязыке и возможно даже инструментальных средств разработки, таких как компиляторы и устройства отладки. Коды также могут быть предоставлены, начиная от простых процедурных интерфейсов до аппаратных средств, конфигурируемых программ верификации и до полностью отработанной операционной системы реального времени. Это позволяет осуществить разработку структуры программных средств системы и тестирование до интеграции дополнительных СФ-блоков, предназначенных для намеченного рынка.

На мировом рынке аппаратных платформ для верификации уже наметилось несколько лидеров. Одной из первых разработок является персональный эмулятор ZeBu-ZV компании EVE (рис.4) Архитектура ZeBu-ZV позволяет использовать эту платформу как разработчикам аппаратуры, так и разработчикам ПО.

Ресурсы эмулятора представлены логической емкостью до 1.5 миллионов ASIC эквивалентных вентилях, системной памятью объемом 128Мбит и производительностью до 12 MHz. Данная модель является идеальной для верификации отдельных логических блоков, IP блоков, небольших проектов, реализуемых в FPGA или ASIC. Эмулятор реализован в виде стандартной PCI платы, встав-

ляемой непосредственно в настольный PC, обеспечивая возможность интеграции с популярными логическими симуляторами и средствами синтеза. Интерфейс внутрисхемной эмуляции ICEPod позволяет подключить эмулятор к целевой системе или IP ядру с использованием 744 вх/вых контактов. Кроме того, эмулятор имеет возможность подключения популярных программных отладчиков через 16-контактный интерфейс SmartICE или виртуальный JTAG транзактор.

В последнее время одной из актуальных тенденций в проектировании стало построение платформ на базе открытых решений. Примером такого подхода является подход компании SPARC, впоследствии развитый при поддержке Европейского космического агентства в процессорном ядре LEON3 компании Gaisler Research.

Процессор LEON3 компании Gaisler Research

Процессор LEON3 является моделью 32-разрядного процессора, использующего SPARC V8 архитектуру. Ядро процессора хорошо структурировано и хорошо подходит для проектирования "систем на кристалле" (SOC). Ядро связано с помощью интерфейса, использующего AMBA-2.0 APB шину и поддерживающего IP plug&play, обеспеченный в Библиотеке IP Gaisler Research (GRLIB). Процессор может быть эффективно изготовлен и по FPGA, и по ASIC технологиям (технологиям специализированных интегральных схем). Leon3 использует стандартные синхронные ячейки оперативной памяти для обеих кэш-систем и регистрации файла. Архитектура SPARC упрощает раннюю оценку и макетирование процессора. Реализация ядра LEON3 также возможна и в радиационно-устойчивом исполнении для использования в космосе и других высокотехнологичных приложениях.

Архитектура полностью открыта, исходные VHDL коды свободно доступны, возможно их использование без каких либо лицензионных ограничений.

В таблице 1 даны сравнительные характеристики быстродействия процессора и занимаемой площади в зависимости от технологии, по которой произведен синтез.

LEON3 - 32-разрядный процессор, основанный на SPARC V8 архитектуре. Он использует расширенный пятитактный конвейер (архитектура Гарварда). Процессор поддерживает полную SPARC V8 систему команд, включая команды перемножения, деления и перемножения с накоплением. Дополнительный IEEE-754 модуль с плавающей запятой обеспечивает поддержку операций с плавающей запятой с одинарной и двойной точностью. Кэш-система поддерживает мультимножества с 4 наборами по 256 Кбайт в наборе.



Рис.4. Персональный эмулятор ZeBu-ZV компании EVE

Таблица 1.

Реализация	Ресурсы (Площадь, число лог. элементов)	Быстродействие
Atmel 0.18 CMOS std-cell	35K gates + RAM	165 MHz
Atmel 0.25 CMOS std-cell	33K gates + RAM	140 MHz
UMC 0.25 CMOS std-cell	35K gates + RAM	130 MHz
Atmel 0.35 CMOS std-cell	2 mm ² + RAM	65 MHz
Xilinx XC2V3000-6	5,000 LUT + block RAM	80 MHz
Altera 20K200C-7	5,700 LCELLs + EAB RAM	49 MHz
Actel AX1000-3	7,600 cells + RAM	48 MHz
Siltera 0.18 mkm	~0.5мм ²	239.5 MHz
ПЛИС XC2V1000-4	99%	50МГц

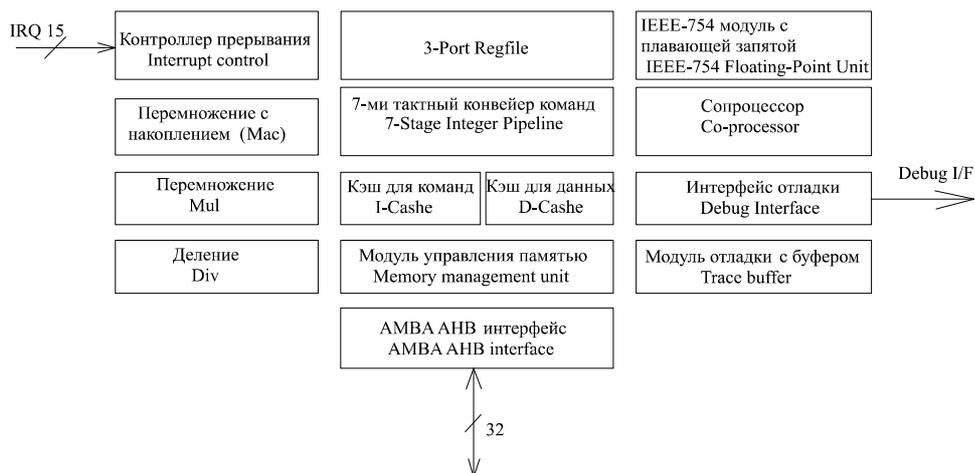


Рис.5. Структура базового процессора.

Конструктивно LEON3 включает в себя: отдельный командный Кэш и Кэш данных, аппаратный перемножитель и делитель, контроллер прерываний, модуль отладки с буфером, два 24-битных таймера, два универсальных асинхронных приемопередатчика, систему безопасности, 16-битный порт I/O (ввода/вывода), гибкий контроллер памяти, поддержку Ethernet порта, интерфейс PCI (32-разрядная шина с возможностью расширения до 64-разрядов, взаимодействие через которую происходит без участия CPU), поддержку шины AMBA.

Функционально LEON3 состоит из следующих модулей (рис. 5): модуля целочисленных вычислений, модуля вычислений с плавающей точкой, системы Кэш-памяти, модуля управления памятью, модуля поддержки отладки, интерфейса памяти (ПЗУ-PROM, статическая ОЗУ-SRAM, синхронная динамическая ОЗУ-SDRAM), двух 24-битных таймеров, контроллера прерываний, параллельного 32-битного порта ввода/вывода, шины AMBA (AHB и APB), Ethernet.

Ядро базового процессора (конвейер, контроллеры кэш-памяти и интерфейс AHB) использует приблизительно 20 000 вентилях и может быть реализован как на специализированных интегральных схемах, так и по технологии FPGA. На технологии в 0.13 мкм может быть достигнута тактовая частота, превышающая 400МГц.

Для аппаратной верификации систем на базе процессора LEON существуют аппаратные платформы компаний Surrey Space Centre и Astrium SAS.

Начинает формироваться и отечественный сегмент средств аппаратной верификации. Компанией НПП «Цифровые решения» разработано семейство цифро-

вых аппаратных платформ для решения задач обработки высокоскоростных сигналов (включая видео), задач передачи данных и макетирования различного рода процессов, в том числе и «прототипирования» СБИС.

Платформа DS-M-1000

Цифровая вычислительная платформа DS-M-1000 выполнена на шестислойной печатной плате, с размещенной на ней ПЛИС XILINX VIRTEX2 1000 (1 млн. вент.). На плате установлены два банка быстродействующей SSRAM памяти суммарным размером 4 Мбайта, работающие на частоте 165 МГц. Разрядность шины данных 36 бит.

Для высокоскоростной передачи данных на плате предусмотрены два интерфейса: 64-разрядная параллельная шина данных и приёмо-передатчик ETHERNET. Параллельная шина данных может быть использована как одна или как несколько параллельных шин данных нужной разрядности; каждый сигнал шины может также использоваться как линия с последовательной передачей данных. Порт ETHERNET реализован на трансивере AM79C874VI фирмы AMD и обеспечивает скорость приёма и передачи данных 100 Мбит/с (FAST ETHERNET).

Структурная схема вычислительной платформы DS-M-1000 представлена на рис.6, а фотография внешнего вида – на рис.7

Отличительной особенностью, платформы DS-M-1000 является возможность работы с цифровым видеосигналом, для чего на плате установлены две микросхемы видеокодеков (ADV202, фирмы ANALOG

DEVICES), обеспечивающие аппаратную компрессию или декомпрессию видео информации. Каждая микросхема может выполнять как функции сжатия (с регулируемым коэффициентом), так и функцию декомпрессии.

Функции контроля и функции отладки позволяет выполнять микросхема ПЛИС CPLD XILINX Cool Runner XCR3256-XL, установленная на плату. С помощью неё можно осуществлять перезагрузку основной ПЛИС, выполнять чтение и анализ контрольных сигналов основной ПЛИС и выводить индикационную информацию на двухразрядный светодиодный дисплей. На плате имеется контроллер датчика температуры, встроенный в ПЛИС VIRTEX, опрашивая который, можно получать текущую температуру кристалла ПЛИС.

На базе платформы DS-M-1000 проведена аппаратная верификация СБИС контроллера FLASH памяти с возможностью аппаратной реализации алгоритмов защиты информации. Данная СпК разрабатывалась с целью создания контроллера для поддержки больших массивов FLASH памяти (от 256Мбайт до 8Гбайт), а также с возможностью передачи данных из или в FLASH с высокой скоростью. В контроллере поддержан командный интерфейс SCSI, предназначенный для использования в промышленных системах сбора и хранения данных.

Дальнейшее развитие аппаратных платформ для верификации СБИС и обработки сигнала будет рассмотрена во второй части статьи, публикация которой планируется в очередном номере журнала.

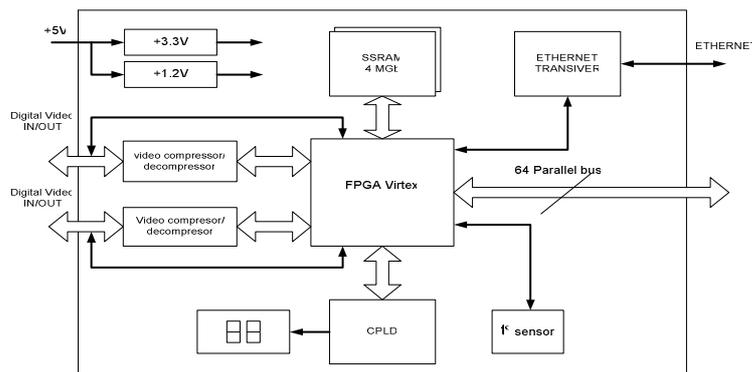


Рис.6. Структурная схема цифровой вычислительной платформы DS-M-1000



Рис.7. Внешний вид цифровой вычислительной платформы DS-M-1000.

Литература

1. В.Б.Стещенко, Губанов Д.А. «Методология реализации алгоритмов цифровой фильтрации на основе программируемых логических интегральных схем Сборник докладов 1-й Международной конференции «Цифровая обработка сигналов и ее применения» 30.06-3.07.1998, Москва, МЦНТИ, том 4
2. Губанов Д.А., Стещенко В.Б., Храпов В.Ю., Шипулин С.Н. Перспективы реализации алгоритмов цифровой фильтрации на основе ПЛИС фирмы ALTERA. // Chip News, № 9-10, 1997, с. 26–33.
3. www.dsol.ru
4. Стещенко В.Б. Оценка производительности новых семейств ПЛИС Altera для применения в системах связи с ОФДМ, Цифровая обработка сигналов №3, 2005
5. Стещенко В.Б. Алгоритмы цифровой обработки сигналов: реализация на ПЛИС, Электронные компоненты за 2006 г., №6
6. В.Б.Стещенко. EDA. Практика автоматизированного проектирования

7. В.Б.Стещенко. «Проектирование аппаратуры цифровой обработки сигналов на ПЛИС с использованием языка описания аппаратуры VHDL» Цифровая обработка сигналов, №1, 2001
8. А. А. Стемпковский, В. А. Шепелев, А. В. Власов, «Системная среда САПР БИС», Наука, 1994 г.
9. А. В. Бухтеев, «Методы и средства проектирования систем на кристалле», Chip news, 2003 г., №4, стр. 4-14.
10. R. Doering, Y. Nishi, «Limits of integrated circuit manufacturing», Proceedings of the IEEE, v. 89, №3, p. 375-393, 2001.
11. R. E. Bryant, K.-T. Cheng, A. B. Kahng, et al., «Limitations and challenges of computer-aided design technology for CMOS VLSI», Proceedings of the IEEE, v. 89, №3, p. 341-362, 2001.
12. H. Chang, L. Cooke, M. Hunt, et al., «Surviving the SOC revolution: A guide to platform-based design», Norwell, MA: Kluwer, 1999.

НОВЫЕ РАЗРАБОТКИ DSP КОМПАНИЙ TEXAS INSTRUMENTS И ANALOG DEVICES В 2007 ГОДУ

Витязев С.В.

Рынок цифровых сигнальных процессоров (DSP – Digital Signal Processor) продолжает стремительно развиваться. Растут скорости обработки информации, повышается степень интеграции устройств. Производители ежегодно расширяют предлагаемые ими линейки процессоров новыми устройствами, стремясь обеспечить потребителя более мощными и функциональными вычислительными элементами или микросхемами, оказывающимися более экономичными, чем альтернативные решения. На российском рынке основную роль играют две фирмы-производителя DSP: Analog Devices Inc. и Texas Instruments Inc. Статья посвящена обзору новых разработок этих компаний, появившихся в 2007 году.

ANALOG DEVICES INC.

Новое семейство процессоров Blackfin – ADSP-BF52x

Семейство процессоров Blackfin компании Analog Devices является на сегодняшний день одним из самых популярных семейств цифровых сигнальных процессоров. Ориентированные на стремительно развивающийся рынок мобильных коммуникационных устройств с поддержкой мультимедийных функций процессоры Blackfin находят применение в самых различных технических сферах. Стремясь дать потребителю (разработчику оборудования на базе Blackfin) возможность выбора модели процессора, наиболее точно удовлетворяющего требованиям конкретного приложения, компания Analog Devices предлагает широкое разнообразие процессоров внутри одного семейства (что обеспечивает их совместимость и простоту перехода от одной модели к другой), отличающихся стоимостью, энергопотреблением, производительностью, функциональностью и так далее. В 2007 году компания Analog Devices расширила семейство Blackfin новой линейкой устройств ADSP-BF52x.

Процессоры ADSP-BF52x характеризуются наименьшим в семействе Blackfin энергопотреблением. Производителем указывается следующий показатель эффективности: потребление ядра составляет 0.16 мВт/МГц при частоте 250 МГц. Малые энергозатраты обеспечиваются наличием пятой режимов функционирования: от режима «оперирования в полную мощность» до режимов «глубокого сна» и «режима бездействия» (Hibernate state); программной регулировкой тактовых частот и напряжения питания; возможностью независимой регулировки тактовых частот различных наборов устройств периферии. В частности, «режим бездействия» позволяет отключать питание ядра процессора и все тактовые сигналы ядра и синхронных периферийных устройств. То есть сигналы CCLK и SCLK отключаются, а внутреннее напряжение питания V_{DDINT} устанавливается равным 0 В. Перед переходом в данный режим

Приведен обзор новых разработок компаний Analog Devices и Texas Instruments в сфере цифровых сигнальных процессоров и средств их поддержки. Обзор сделан на основе материалов, представленных компаниями в сети Internet, и охватывает период 2007 года.

необходимо всю требуемую информацию (содержимое внутренней памяти, внутренних регистров и т.п.) скопировать на внешние устройства. Процессор выходит из состояния бездействия по одному из событий, например, прерыванию от модулей Ethernet или USB или от часов реального времени (RTC).

Другой отличительной особенностью новых процессоров является использование технологии безопасности Lockbox (Lockbox Secure Technology™). Ряд аппаратных и программных средств позволяет обеспечить защиту кодов, размещенных на процессоре. Ядром здесь является однократно программируемая память, в которую при изготовлении конечного устройства делается однократная прошивка набора данных: криптографических ключей, идентификатора устройства и другое. В дальнейшем, обеспечивается невозможность перезаписи «секретной» области памяти и ее «защита от взлома». Появляется, с одной стороны, возможность охраны интеллектуальной собственности (исходных кодов и алгоритмов, загруженных на процессор), а, с другой стороны, – возможность реализации приложений электронной коммерции и других. Дополнительно процессором поддерживается безопасный режим функционирования (Secure mode), при котором разрешается выполнение только кодов «с подтвержденной подлинностью». Отметим, что описанная технология безопасности уже была использована в ранее выпущенных моделях ADSP-BF54x, ориентированных на применение в автомобильной электронике.

Семейство ADSP-BF52x ориентировано на применение в портативных мультимедиа-устройствах, таких как различного рода медиаплееры, IP-телефоны, IP-камеры, мобильные ТВ-устройства. В семейство входит целый набор моделей, отличающихся соотношениями рабочих частот, цены и периферией. Наиболее производительным является процессор ADSP-BF527, работающий на частоте 600 МГц и оцениваемый в \$14.50*. Наиболее экономичная модель ADSP-BF522L работает на частоте 300 МГц и стоит \$5.50*.

Объем внутренней памяти для всех устройств составляет 132 Кбайта. Из них память программ имеет общий объем 64 Кбайта, память данных SRAM – 32 Кбайта, память данных SRAM/кэш – 32 Кбайта и сверхоперативная память (scratchpad) – 4 Кбайта. Периферия процессоров включает интерфейсы USB On-The-Go и Ethernet, хост-порт в режиме DMA, контроллер flash-

памяти типа NAND, до 48 выводов общего назначения и другие типовые в целом для семейства устройства.

Дополнительно, компания Analog Devices выпускает линейку процессоров ADSP-BF52xC. Они относятся уже не к самым низкопотребляющим процессорам Blackfin, а к классу средней производительности, и отличаются наличием встроенного малопотребляющего аудиокодека. Кодек строится на базе 24-битных стерео (двухканальных) ЦАП и АЦП с частотами выборки от 8 до 96 кГц.

Процессоры ADSP-BF52x выпускаются в корпусах CSP_BGA 189 шариков, 12×12 мм и 208 шариков, 17×17 мм. Для начала работы с новым семейством доступен стартовый набор разработчика ADSP-BF527 EZ-KIT Lite с размещенными на плате дополнительной микросхемой памяти, аудиокодеком, ЖК-дисплеем, клавиатурой и др.

Подключение системы MATLAB к среде VisualDSP++

Компании MathWorks и Analog Devices сообщают о возможности подключения системы MATLAB и пакета Simulink к среде разработки и отладки программ для процессоров компании Analog Devices VisualDSP++. Средства подключения MATLAB к VisualDSP++ поставляются компанией MathWorks. Их основная задача – упростить разработку и отладку программного обеспечения DSP за счет возможности проводить верификацию кодов, исполняемых на процессоре в рамках среды MATLAB и возможности создания проектов для среды VisualDSP++ автоматически на основе модели, разработанной в пакете Simulink. Новый инструмент поддерживает работу с семействами Blackfin, SHARC и TigerSHARC компании Analog Devices. Он позволяет при программировании DSP использовать принцип разработки, ориентированной на модель: разрабатывается структура рабочей модели, в пакете Simulink она собирается из набора типовых блоков, проводится моделирование и здесь же на базе той же модели осуществляется генерация программных кодов и их верификация. Команды анализа данных, отладки, контроля исполнения, действующие в рамках среды VisualDSP++, выполняются непосредственно из MATLAB. Становится возможным автоматический запуск выполнения программы на процессоре с различными параметрами, что оказывается очень эффективно с позиции верификации кодов.

Средство подключения системы MATLAB к среде VisualDSP++, работающее под операционной средой Microsoft Windows, уже является доступным для приобретения и стоит от \$2000*.

Среда разработки и отладки VisualDSP++ 5.0

Компания Analog Devices выпустила новую версию среды разработки и отладки программных кодов для процессоров этой фирмы VisualDSP++ 5.0. Усовершенствования повышают функциональность и надежность среды. Внесен ряд нововведений, детально описываемых в документе: "VisualDSP++ Release 5.0 Product Description". Перечислим ряд из них.

Файл ядра (Core File). В новой версии среды VisualDSP++ имеется возможность записать в файл всю информацию (содержимое регистров и памяти) о текущем состоянии системы. Впоследствии эта информация может быть загружена из файла на процессор. Это оказы-

вается полезно при отладке работы системы. Выполнение программы останавливается в том месте, где работа системы оказывается неудовлетворительной, состояние системы записывается в файл, поведение системы анализируется впоследствии без необходимости подключения всего оборудования. Проблему можно обсудить с коллегами, воспроизвести на другой системе. При отладке некоторой подпрограммы можно записать в файл полное состояние системы до обращения к ней и в другой файл – состояние системы после выхода из подпрограммы. Файлы могут быть представлены в текстовом формате и давать наглядную информацию о всех изменениях, которые произвел данный фрагмент кода.

Панель пользователя (Custom Board). В версии VisualDSP++ 5.0 пользователю предоставляется возможность самому создавать окна регистров, то есть формировать списки отображаемых регистров и задавать значения этих регистров, устанавливаемые при перезапуске или начальной инициализации. Это может существенно повысить удобство работы с системой, позволяя выводить на экран только необходимую в данный момент информацию.

Проверка на соответствие стандарту MISRA-C. В настройках компилятора может быть задана опция проверки программного кода на соответствие правилам MISRA-C:2004, дающим рекомендации по использованию языка Си для программирования особо важных систем. Эти рекомендации разработаны Ассоциацией надежности программного обеспечения промышленных двигателей (Motor Industry Software Reliability Association). Они оказываются очень полезны при разработке программного обеспечения, связанного, например, с электронными системами для автомобилей или системами промышленного управления.

Внесен еще целый ряд изменений в среду VisualDSP++, явившихся откликом на многочисленные пожелания пользователей. В частности, изменен механизм лицензирования и регистрации: вместо HOST ID (идентификация номера персонального компьютера) теперь может использоваться его сетевой MAC-адрес; лицензия может временно переноситься с одного компьютера на другой (например, если вы хотите продолжать работать в поездке, используя ноутбук). Эти и другие многочисленные нововведения существенно усовершенствовали среду VisualDSP++ и позволили говорить о выпуске ее новой версии VisualDSP++ 5.0.

TEXAS INSTRUMENTS INC.

Новые скорости процессоров TMS320C64x – тактовая частота достигла 1.2 ГГц

Семейство процессоров TMS320C64x является наиболее высокопроизводительной линейкой DSP компании Texas Instruments, ориентированных на проводные и беспроводные системы связи, обработку видео и изображений. Это процессоры с фиксированной точкой, впервые вышедшие на уровень тактовых частот в один гигагерц. В 2007 году компания Texas Instruments объявила о выпуске наиболее мощного устройства данного класса – процессора TMS320C6455 с тактовой частотой 1.2 ГГц. Это новый уровень скоростей! Процессор получил от независимого эксперта в области DSP – фирмы BDTI

(<http://www.bdti.com/>) – оценку в 13170 баллов, позволяя ему уверенно лидировать по скорости вычислений в классе процессоров с фиксированной точкой. От ранее выпущенных моделей с той же маркировкой C6455 новое устройство будет отличаться только более высокой частотой. Стоимость кристалла составляет \$245*. Подготовлен и стартовый набор разработчика (DSK) на процессоре TMS320C6455 с частотой 1.2 ГГц стоимостью \$495*.

Нарастивая мощь процессоров C64x, компания Texas Instruments одновременно стремится и к расширению номенклатуры изделий. Процессоры TMS320C6454 и C6452, относясь к линейке наиболее быстрых моделей C645x, по сравнению с базовым устройством C6455 обладают несколько меньшей функциональностью и доступны по меньшей цене. В частности, процессор TMS320C6452 являясь решением, направленным на оптимизацию соотношения цена/производительность. При стоимости, существенно меньшей цены процессора C6455, процессор C6452 с тактовой частотой 900 МГц дает возможность быстро и без дополнительных затрат перейти от использования популярной серии C641x к наиболее производительной линейке C645x. Семейства C641x и C645x хоть и относятся к одному классу процессоров C64x, но все же имеют существенные различия. Прежде всего, семейство C645x строится на базе ядра C64x+, что уже на уровне архитектуры дает выигрыш по производительности и сокращение объемов программных кодов на 20-30%. Объем памяти уровней L1 и L2 увеличен вдвое. На кристалле размещается дополнительная коммуникационная периферия. Все это делает семейство C645x более привлекательным и сделать переход к более перспективному классу устройств наименее затратным позволяет процессор TMS320C6452. Стоимость кристалла при частоте 720 МГц составляет \$94*. Выпускается также оценочный модуль (EVM) на базе устройства C6452 по цене \$1295*.

Компания Texas Instruments не оставляет без внимания и менее мощные линейки процессоров класса C64x. Ведь во многих случаях именно в этих продуктах может быть найдено изделие с оптимальным соотношением экономичности и производительности. Этой концепции отвечают новые модели процессоров TMS320C6424 и TMS320C6421. Приведем лишь кратко их характеристики. Процессоры C6424/C6421 работают на частотах от 400 до 600 МГц в зависимости от конкретной модели. Производительность оценивается в 3200 – 4800 MIPS (миллионов инструкций в секунду). Внутренняя память уровня L1 имеет объем 32 Кбайт кэш программ/80 Кбайт кэш данных для процессора C6424 и 16 Кбайт кэш программ/16 Кбайт кэш данных для процессора C6421. Память уровня L2 составляет 128 Кбайт/64 Кбайта для C6424/C6421, соответственно. Снижено напряжение питания ядра и периферии, которое может достигать 1.05 В и 1.8 В вместо прежних 1.2 В и 3.3 В. Стоимость устройств составляет от \$18* до \$28* для процессора C6424 и от \$10* до \$15* для C6421.

Сигнальные контроллеры платформы C2000 с плавающей точкой

Платформа C2000 сигнальных процессоров компании Texas Instruments ориентирована на использование в системах промышленного контроля, управления двигателями, цифровых источниках питания, преобразователях солнеч-

ной энергии и других приложениях. Этот класс устройств совмещает богатый состав периферийных узлов и простоту и удобство программирования, характерные для микроконтроллеров, с вычислительной эффективностью сигнальных процессоров. Поэтому фирма-производитель использует для них термин сигнальные контроллеры.

Традиционно, указанный класс задач, возлагаемых на сигнальные контроллеры, решался в классе обработки с фиксированной точкой. В 2007 году компания Texas Instruments заявила о появлении новой линейки процессоров в семействе C2000 – TMS320C2833x, работающих в формате с плавающей точкой. Плавающая точка – это, прежде всего, существенное упрощение процесса разработки программного обеспечения и повышение его эффективности за счет расширения динамического диапазона и увеличения точности представления чисел. Стоимость кристаллов лежит на уровне цен на процессоры с фиксированной точкой. Тактовая частота процессора TMS320F2833x составляет 150 МГц, обеспечивая производительность 300 MFLOPS (миллионов операций с плавающей точкой в секунду). На кристалле размещается богатый набор периферийных устройств: 12-разрядный 16-канальный АЦП с быстродействием 12.5 миллионов выборок в секунду; 18 каналов ШИМ; интерфейсы CAN, I2C, UART, SPI и McBSP. Обеспечивается совместимость по кодам с базовыми моделями TMS320C28x.

Для новой линейки процессоров разработан стартовый набор F28335 eZdsp. Он включает плату на базе процессора TMS320F28335, расширенную синхронную память SRAM 128x16, интерфейсами CAN 2.0 и RS-232 и другими модулями. Программное обеспечение, поставляемое вместе с набором, включает среду Code Composer Studio, примеры готовых программ и другое. Стоимость набора составляет \$495.02*.

Лидер по производительности среди процессоров с плавающей точкой

Основной класс процессоров компании Texas Instruments, работающих в формате с плавающей точкой, – это семейство TMS320C67x. Эти процессоры ориентированы на аудиоприложения профессионального уровня, системы биометрии, медицинское оборудование, обработку изображений, работу с трехмерной графикой, распознавание речи и другие приложения.

Компания Texas Instruments повышает мощность наиболее высокопроизводительного процессора этого семейства TMS320C7627, увеличивая его тактовую частоту до 350 МГц. Оставаясь полностью совместимым по кодам с другими моделями семейства, процессор TMS320C6727 с новой частотой обеспечивает скорость вычислений 2100 MFLOPS (миллионов операций с плавающей точкой в секунду). Совместимость с другими моделями позволит легко встроить новый процессор в уже разработанные изделия и повысить их эффективность или расширить функциональность. К другим характеристикам процессора относятся: память данных/программ уровней L1/L2 – 32/256/384 Кбайт; последовательные аудиопорты McASP – 3 шт.; хост-порт; порты I2C/SPI – 2/2 шт.; напряжение питания ядра/периферии – 1.2/3.3 В, корпус BGA, 256 выводов, 17x17 мм. Стоимость устройства составляет \$29.35*.

Начать работу с процессором TMS320C6727 можно,

приобретя набор разработчика профессиональных аудиосистем (PADK). Это набор программно-аппаратных инструментальных средств, позволяющих создавать собственные решения в кратчайшие сроки. Стоимость набора - \$1995*

Платформа DaVinci – основа систем обработки видео и изображений следующего поколения

Пожалуй, основное внимание в своих текущих разработках в области DSP компания Texas Instruments в настоящий момент уделяет платформе разработки систем обработки видео и изображений – DaVinci, делая ставку на возрастающую востребованность приложений, связанных с обработкой мультимедиа-данных в реальном масштабе времени. В 2007 году существенно расширена номенклатура процессоров DaVinci. Компания стремится предоставить разработчику возможность выбора такого устройства, которое бы оптимально подходило для его конкретной задачи. Это необходимо, чтобы DaVinci давала максимальный эффект по снижению стоимости разработки, заменяя одним кристаллом целый набор электронных компонентов и обеспечивала оптимальное соотношение производительности, функциональности и цены.

Напомним, что базовым семейством платформы DaVinci стали процессоры TMS320DM644x, за которыми были также выпущены процессоры TMS320DM643x. Об этих моделях мы рассказывали в выпуске журнала «Цифровая обработка сигналов» №4 за 2006 год (стр. 48-53). Процессоры TMS320DM644x строятся на базе ядра DSP, ядра микроконтроллера, сопроцессора обработки видео и изображений, подсистемы обработки видеоданных и блока периферийных устройств. Они представляли наиболее мощный класс процессоров DaVinci и ориентировались на применение в видеотелефонах, автомобильных бортовых компьютерах, цифровых фотокамерах и других системах. Процессоры TMS320DM643x не включали в свой состав ядро микроконтроллера и обеспечивали снижение стоимости технологии DaVinci до уровня \$10*. Эта линейка процессоров была ориентирована на чувствительные к стоимости элементной базы приложения, например, системы машинного зрения, робототехнику и другие, одновременно, включая необходимые для такого использования компоненты.

В прошедшем году к указанным линейкам добавились устройства TMS320DM646x, TMS320DM647/648 и TMS320DM355.

Процессоры DM646x представляют направление развития платформы DaVinci по пути наращивания мощности. Первый процессор этого класса TMS320DM6447, выпускаемый в настоящий момент в инженерных образцах, заявлен, как первое однокристалльное решение, обеспечивающее в реальном времени передовые стандарты кодирования/декодирования для приложений видео высокой четкости (HD Transcoding) при поддержке богатого набора разнообразных форматов данных. По сравнению с имеющимися альтернативными решениями система на кристалле TMS320DM6447 дает 10-кратный выигрыш по производительности и 10-кратное снижение цены в указанных приложениях. Современные тенденции развития систем видео высокой четкости требуют от вычислительной элементной базы высочайшей мощности и поддержки различ-

ных форматов входных и выходных данных. Процессор TMS320DM6447 полностью отвечает этим требованиям. Он строится на базе ядра DSP 64x+ и ядра микроконтроллера ARM926. В отличие от семейства DM644x процессор DM6467 включает сопроцессор обработки видео и изображений высокой четкости (HD Video Imaging Coprocessor), что и служит основой его эффективности в указанных приложениях. Стоимость процессора составляет \$35.95*. Оценочный модуль DM6447 DVEVM должен быть выпущен в первом квартале 2008 года.

Процессоры TMS320DM647/648, аналогично процессорам семейства TMS320DM643x, не включают в свой состав ядро микроконтроллера. В их основе лежит подсистема DSP на базе ядра C64x+ с тактовой частотой 720/900 МГц в зависимости от модификации. Данные модели процессоров ориентированы на применение в многоканальных видеосистемах безопасности, а также других приложениях обработки видеоданных с высокой вычислительной эффективностью, в том числе, цифровых системах видеозаписи, системах технического зрения и так далее.

Интересным и активно рекламируемым решением платформы DaVinci является еще один процессор – TMS320DM355. Оставаясь мощным и функциональным представителем своего семейства, этот процессор отличается экономичностью и малыми энергозатратами. В его состав не включено ядро сигнального процессора. Вместо этого он строится на базе подсистемы обработки видео, сопроцессора, реализующего стандарты MPEG-4/JPEG, и ядра микроконтроллера. Тактовая частота составляет 216 или 270 МГц в зависимости от модификации. Благодаря своей архитектуре, процессор становится оптимизированным для портативных и других приложений, требующих малого потребления энергии, выполняющих обработку видеоданных в формате HD (высокой четкости). Производителем заявлено, что использование процессора TMS320DM355 способно вдвое увеличить срок службы батарей питания в приложениях видео высокой четкости по сравнению с альтернативными решениями, существующими на данный момент. Процессор рекомендуется применять в следующих системах: цифровых фотоаппаратах и видеокамерах, портативных мультимедийных плеерах, системах видеонаблюдения, цифровых фоторамках, видеотелефонах и других. Стоимость кристалла – \$12.60*.

Данная статья написана на основе информации, предоставленной компаниями Texas Instruments и Analog Devices на их сайтах: <http://www.ti.com> и <http://www.analog.com>. Дополнительная информация по продукции может быть получена у их официальных представителей. Для TI это: «Сканти-Рус» (<http://www.scanti.ru>), «Компел» (<http://www.compel.ru>), «EBV Elektronik» (<http://www.ebv.com>), «Spoerle» (<http://www.spoerle.com>), «Silica» (<http://www.silica.com>) и «ITC Electronics» (<http://www.itc-electronics.com>). Компанию ADI представляют «АВТЭК» (<http://www.autex.ru>), «АРГУССОФТ Компани» (<http://www.components.argussoft.ru>) и «ЭЛТЕХ» (<http://www.eltech.spb.ru>).

* **Примечание** Все приводимые цены следует считать только ориентировочными. Цены на продукцию компаний, действующие на территории Российской Федерации, можно узнать у их официальных представителей, перечисленных выше.

ЦОС В УНИВЕРСИТЕТАХ: УНИВЕРСИТЕТСКАЯ ПРОГРАММА КОМПАНИИ TEXAS INSTRUMENTS, часть 2

О появлении новой рубрики журнала «ЦОС в университетах» мы сообщали в прошлом выпуске журнала «Цифровая обработка сигналов» №4 за 2007 год. В том же номере была открыта рубрика, которая посвящена обзору Университетской программы компании Texas Instruments (TI), проводимой на территории Европы и России и обеспечивающей поддержку образовательных и научно-исследовательских работ учебных заведений в области передовых методов цифровой обработки сигналов и информационных технологий реального времени. Компания TI является мировым лидером в области технологий цифровой обработки и участие в организуемой ею Университетской программе способно послужить мощным импульсом к развитию данного направления в российских ВУЗах. Данная статья является второй частью обзора Университетской программы TI, посвященной краткому описанию ряда средств, доступных по этой программе среди аналоговых устройств и микроконтроллеров; рассказу об обучающих дисках, помогающих преподавателям и научным сотрудникам повысить свой уровень знаний и подготовить учебные материалы; сообщению о вознаграждении, на которое могут рассчитывать наиболее активные студенты; описанию механизма подключения к Университетской программе TI.

Основным механизмом реализации Университетской программы компании TI является предоставление оборудования ВУЗам на льготных условиях. Кроме модулей цифровой обработки сигнала на базе DSP различных платформ, о некоторых из которых мы рассказывали в первой части статьи, предлагается и ряд средств, использующих аналоговые устройства, микроконтроллеры и преобразователи данных. В частности, внимания заслуживают аппаратно-программные средства для оценки эффективности аналого-цифровых и цифро-аналоговых преобразователей и инструментов по оптимизации энергетических затрат проектируемой системы. Тестовые платы АЦП/ЦАП позволяют вести оценку эффективности преобразователей данных различного типа, выбирая наилучший для конкретного применения. Тестовая плата подключается к персональному компьютеру и специальное программное обеспечение демонстрирует основные характеристики функционирования АЦП/ЦАП. Для целей оптимизации энергопотребления также предлагается визуализация потребляемой мощности. Данные инструменты (оценочные платы + программное обеспечение) могут быть чрезвычайно полезны не только при реализации конкретных проектов, но и в целях обучения.

Другим классом оборудования, предлагаемого для поддержки научной и образовательной работы в ВУЗах, являются продукты компании Chipcon, ориентированные на построение беспроводных телекоммуникационных сетей с использованием протокола ZigBee. Непосредственно необходимый состав оборудования в этом классе следует обсуждать особо, однако, в любом случае, продукты Chipcon способны оказать существенную поддержку процессу обучения или научных исследований по программам, связанным с системами беспроводной связи и передачи данных.

Большой популярностью в инженерных и научных разработках пользуются микроконтроллеры семейства MSP430 компании TI. В рамках Университетской программы предлагается полный набор инструментальных средств для работы с микроконтроллерами этого семейства, в том числе среда программирования MSP430 – Code Composer Essentials версии 2.0. В качестве набора разработчика программных кодов для MSP430 можно использовать миниатюрную плату, совмещенную в одном корпусе с USB-эмулятором, подключаемую непосредственно к USB-разъему персонального

компьютера и не требующую для функционирования и отладки дополнительных проводов питания или интерфейса с ПК. Кроме того, этот продукт позволяет после проведения отладки легко отсоединить модуль MSP430 от USB-эмулятора и использовать его непосредственно в конечном оборудовании.

Возвращаясь к средствам программирования цифровых сигнальных процессоров, отметим наличие специальных обучающих дисков (Teaching ROMs), предназначенных для ознакомления с тремя основными платформами DSP компании TI: C6000, C5000 и C2000. На дисках можно найти лекции по изучению основ архитектур и принципов программирования DSP, руководства по начальному освоению процессоров, материалы для самостоятельной работы с подсказками и комментариями, примеры программ с доступными исходными кодами. В данный момент в наличии имеются только англоязычные версии дисков, однако, ведется работа по переводу их материалов на русский язык. Диски предоставляются бесплатно для участников Университетской программы.

Поговорим теперь о поощрениях, предоставляемых компанией TI студентам, демонстрирующим активную и успешную работу с продуктами компании. Стимул в образовательном процессе играет действительно важную роль для обучаемых студентов. В качестве основной награды компания TI предлагает персональную версию интегрированной среды Code Composer Studio с бесплатной поддержкой обновлений в течение 5 лет. Такие награды предоставляются только наиболее активным и плодотворно работающим студентам, обучаемым в лаборатории, организованной в рамках Университетской программы TI, по мнению руководителя.

Подводя итог обзору Университетской программы, ответим на вопрос: что же нужно, чтобы подключиться к программе и какие результаты можно от этого ожидать.

В первую очередь, вам необходимо принять обоснованное решение о целях вашего участия в программе TI и о месте ваших работ в ВУЗе. Далее постарайтесь тщательно разобраться с предлагаемыми TI продуктами и выбрать то, что вам действительно нужно. При необходимости вы можете обратиться за консультацией в центр технической поддержки (см. первую часть статьи) или к координатору Университетской программы TI в России и странах СНГ.

Запрос на предоставление оборудования от компании TI должен очень кратко рассказывать о целях вашего проекта, о конкретных ожидаемых вами результатах, о том, планируете ли вы сделать результаты «открытыми» и других аспектах проекта, а также перечень запрашиваемого вами оборудования. Каждый Ваш запрос компания TI рассматривает индивидуально.

Запросы отправляются в электронном виде менеджеру и координатору Университетской программы TI.

Информационное сообщение подготовлено на основе материалов доклада менеджера Университетской программы TI в странах Восточной и Западной Европы и Северной Африки Роберта Оуэна (rcwo@ti.com) и координатора Университетской программы TI в России и странах СНГ Алексея Петровского (тел. +375(29)6522470, E-mail: petrovsky@bsuir.by). Еще раз обращаем ваше внимание на открытие нового раздела сайта журнала «Цифровая обработка сигналов» под названием «ЦОС в университетах». На его страницах мы намерены давать информацию об университетских программах фирм-производителей оборудования ЦОС и рассказывать о научных и учебных лабораториях, имеющих и появляющихся в российских ВУЗах. Все заинтересованные организации приглашаем к сотрудничеству.

МИКРОКОНТРОЛЛЕР USIP® ОТ INNOVA CARD: НОВАЯ ПЛАТФОРМА ДЛЯ СИСТЕМ С ВЫСОКИМИ ТРЕБОВАНИЯМИ К БЕЗОПАСНОСТИ

Кирпичников А.П., Киришин Е.А.

Один из авторов прочитал приведённое выше объявление, напечатанное скромно мелким шрифтом, (в "Volt&Nuts", кажется) и грустно подумал о всех разработчиках с наивной верой закрывающих программу во внутренней флэш любезно предоставленным изготовителем «битом секретности» – «на века». А ведь алгоритмы в современном мире часто много ценнее и долговечнее, чем сами изделия. Недаром с тех пор произошли некоторые сдвиги в области их закрытия – многие солидные фирмы предлагают более серьёзные технологии «секретности» (например, компания Analog Devices ввела механизм Security LockBox® в DSP-процессорах Blackfin, что позволяет использовать двухуровневое шифрование для аутентификации и закрытия кода; Atmel объявил о создании специализированной CryptoMemory® EEPROM и т.д.).

Так существуют ли вообще микропроцессоры, гарантирующие стойкость от всех современных методов взлома (включающих анализ работы внутренних шин по записи электромагнитного фона; принуждение аппаратной части к ошибкам и раскрытию секретов воздействием «рентгена» – как в зубоврачебном кабинете, бесчисленные зонды и щупы для подключения к «интимным» местам кристалла сквозь тонкую «обёртку» корпуса и пр.), и, если да, то насколько специфика сохранения тайны ограничивает при этом реальные возможности такого вычислителя – об этом наша статья.

Микроконтроллер USIP®

USIP был разработан во Франции в 2004 году для нужд сохранения информации наиболее ценящейся в Европе – банковской тайны. Причём расчётная стоимость «взлома» самыми современными средствами, экономящими силы и время, и без учёта стоимости оных, по проекту должна составлять не менее € 50 тыс. (а в идеале, достигать €1 млн.), а время, при достаточном количестве образцов, – тысячи человеко-часов. Основным применением микросхемы должны являться различные платёжные терминалы и иные электронные средства идентификации и доступа (ключи, радио ID и пр.). На борту контроллера может быть размещена любая из существующих криптографических систем (наиболее примитивная – AES, даже аппаратно встроена, чтобы не занимать ресурсы процессора). При этом микросхема продолжает являться достаточно мощным вычислительным средством с разветвлёнными интерфейсами, позволяющим автономно решать большинство задач, возникающих при создании компактного интеллектуального прибора обработки информации.

Ядро. В качестве платформы для построения системы было выбрано защищённое ядро MIPS 4Ksd, как наиболее отвечающее требованиям защиты информации, как с учётом топологии так и производительности.

Ядро этого типа разрабатывалось для поддержки продуктов с множеством приложений, в частности, базирующихся на высокоэффективных алгоритмах шифрования и виртуальных машинах. 4Ksd реализует архитектуру MIPS32 с поддержкой SmartMIPS ASE (для ускорения алгоритмов шифрования, таких как RSA, ECC, DES и AES) и механизмом оптимизации

«Если вы забыли программу своего микроконтроллера и первоисточники утеряны, обратитесь к нам, и мы прочитаем её для вас...»

Объявление в американском журнале конца 90-х

плотности кода MIPS16E. При этом ядро имеет дополнительные блоки, благодаря которым достигается увеличение быстродействия на вышеуказанных алгоритмах. Так, имеется автономный блок умножения/деления 32x16 и 32x32, выполняющий операции умножения и деления с накоплением за один и два такта частоты ядра соответственно.

MIPS 4Ksd имеет пятиступенчатый конвейер с обратной связью и работает на тактовой частоте до 96 МГц (в новых моделях – до 175 МГц) с возможностью выполнения инструкций в каждом такте. За счет наличия дополнительных блоков вычисления пиковая производительность достигает 1,30 MIPS/МГц, что выгодно отличает кристалл от других ядер, имеющих обычно 0.5-0.8 MIPS/МГц. В состав ядра также входят 2 блока кэш-памяти инструкций и данных, по 8 Кбайт каждый.

Память. Контроллер содержит в себе три типа памяти: ROM (128 Кб), RAM (128 Кб) и Flash (256 Кб). Память ROM содержит загрузчик Secure Boot Loader (предназначен для обновления прошивки и записи ключей для алгоритмов шифрования) и библиотеку низкоуровневого доступа к периферии Hardware Abstraction Layer (HAL).

Доступ к памяти ROM и RAM осуществляется в режимах 8-, 16- и 32- бит за один такт частоты ядра.

В USIP® имеется контроллер внешней памяти, поддерживающий память SDRAM (32 Мб, максимальная частота 96 МГц) и до четырех слотов памяти SRAM (4x32Мб).

Адресное пространство памяти линейное и составляет 4Гб. Имеется механизм MMU, с использованием которого возможно разграничение доступа к памяти по страницам (размером от 1 Кб до 16 Мб) с режимом доступа чтение/запись/выполнение. Это достигается с использованием конфигурируемых блоков TLB.

Достаточно быстрая встроенная Flash осуществляет запись страницы (2048 бит) за 1 мс, есть дополнительные возможности приостановки записи/стирания, проверки содержимого по 128-бит и 16-бит сигнатурам. Чтение Flash возможно в двух режимах (синхронный и асинхронный), отличающихся скоростью и энергопотреблением.

Прерывания и DMA. Ядро MIPS 4Ksd имеет шесть уровней приоритетов внутренних прерываний, что позволяет создавать системы с гибким распределением и иерархией процессов. Для каждого уровня в USIP® жестко закреплен определённый набор периферии. Для обработки внешних прерываний имеется отдельный вход для быстрого обслуживания. Кроме того, есть возможность настроить каждый из 32-х имеющихся выводов GPIO в качестве источника внешнего прерывания. Отдельное прерывание – NMI, которое используется в системе безопасности для генерации сброса микросхемы в случае атаки.

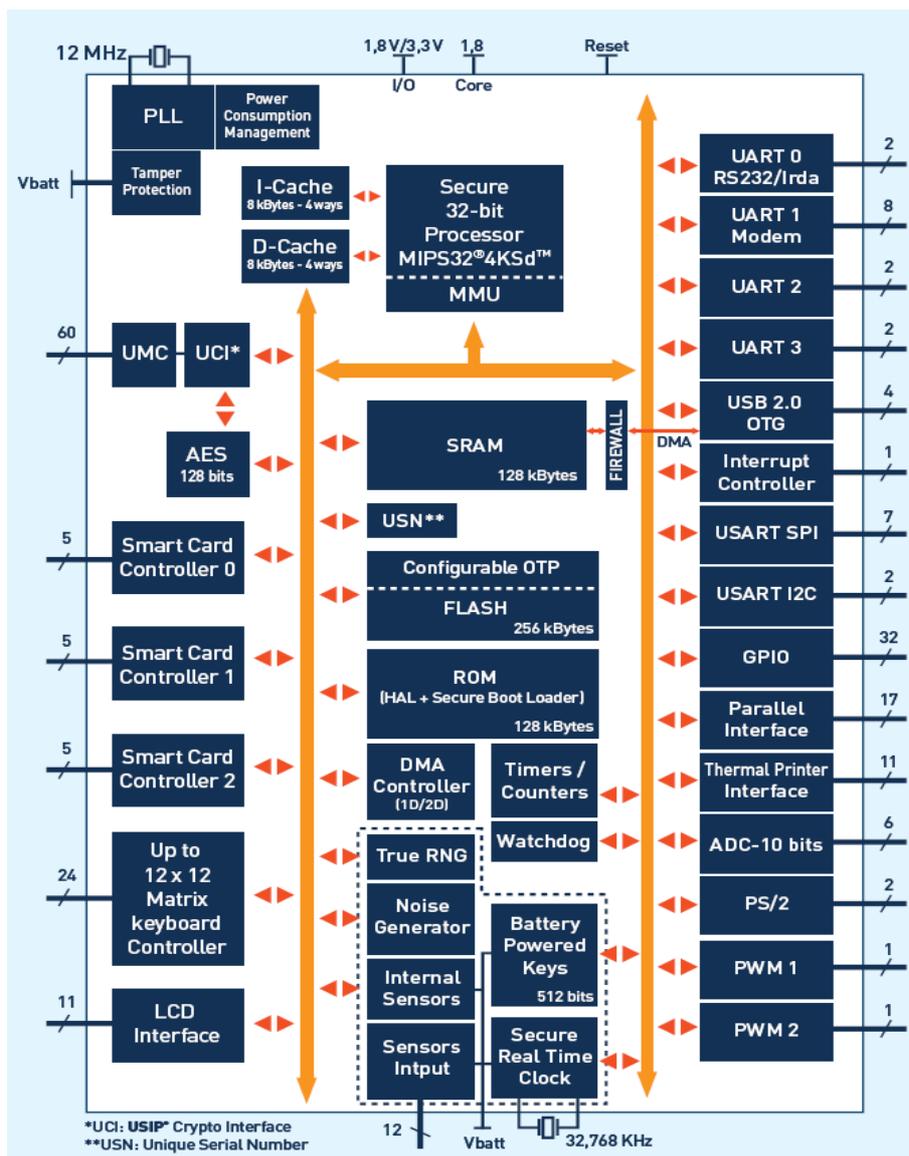


Рис. 1. Структурная схема микроконтроллера USIP®.

Благодаря достаточно развитому механизму DMA существует возможность передачи данных память - память, память - периферийные устройства (кроме тех устройств, в которых в качестве базовой используется частота, отличная от частоты PLL). Режимы 1D и 2D с положительным и отрицательным шагом позволяют организовать «разворот» буферов в памяти, а также упаковку/распаковку без использования ресурсов процессора. Эти возможности могут значительно повысить быстродействие при различных преобразованиях форматов данных. При обмене данными с периферией существует возможность упаковки/распаковки данных на лету. Еще одна возможность контроллера DMA – приостановка передачи с последующим возобновлением.

Периферия. USIP представляет собой систему на кристалле, поэтому большинство интерфейсов передачи данных – внешние, такие как UART и USB. Интерфейсы обмена в пределах устройства ограничены минимумом: SPI и I²C.

Из коммуникационных интерфейсов в USIP® присутствуют UART (2 двухпроводных, один с поддержкой IrDA и один с полным набором сигналов для подключения модема, до 2 Мбит/с, имеются буфера FIFO), USB 2.0 OTG (16 конечных точек), I²C (Philips 2.1, 100/400 КГц), SPI (1.5 МГц,

настраиваемый режим Motorola и изменяемая длина слов), SPP (IEEE1284, совместимый с IBM PC Centronix), PS/2 (для подключения клавиатуры и мыши IBM PC).

Интерфейсы, обеспечивающие ввод-вывод: интерфейс смарт-карт (три интерфейса, работающих в синхронном и асинхронном (ISO-7816 UART) режимах, есть режим эмуляции смарт-карты), интерфейс термо-принтера, контроллер клавиатуры (для подключения матриц с размером до 12x12 линий), LCD-интерфейс.

В USIP® также присутствует набор таймеров: Watchdog с системой защиты, два 32-битных PWM и четыре 16-битных таймера общего назначения.

Для ввода аналоговых сигналов имеется 10-бит АЦП (разрядность может изменяться от 2 до 10 бит) с максимальной частотой дискретизации 400 КГц, использующий в качестве опоры одно из питаний: 1.8 В или 3.3 В.

Питание, тактовая частота и энергопотребление. Кристалл имеет несколько вводов питания 3.3 В и 1.8 В: линии питания ядра, периферии, памяти, аналоговых устройств и устройств защиты. USIP® довольно экономичен: на частоте 96 МГц, при всех включенных периферийных устройствах, микросхема потребляет около 160 мА (90% по линии ядра (1.8 В) и 10% по линии IO (3.3В)).

Для управления энергопотреблением в USIP® реализовано четыре режима. Два настраиваемых рабочих режима задают набор включенной периферии, режим переключения частоты ядра и режим Ultra-low power (например, для приостановки устройства в режим ожидания по USB).

Механизмы обеспечения безопасности

Механизмы защиты в USIP® присутствуют на различных уровнях и практически в каждом функциональном блоке. Можно выделить следующие уровни защиты: физический, логический и уровень организации процесса разработки и производства.

В качестве пассивной физической защиты используется специальный металлический экран на поверхности кристалла, к нему же подключён дополнительный активный датчик «на вскрытие» кристалла.

Активную защиту обеспечивает набор внешних и внутренних датчиков системы. Задача внутренних датчиков – постоянный мониторинг параметров системы и выполнение необходимых мер (сброс микросхемы, генерация прерывания NMI, стирание защищённой области памяти т.п.) при обнаружении выхода параметров за допустимые пределы. Среди параметров слежения – входная частота PLL, напряжение питания ядра, температура корпуса, доза облучения и др. Внешние датчики позволяют разработчику подключать их по своему усмотрению для отслеживания угрозы внешнего проникновения.

Микроконтроллер USIP® имеет вход резервного питания, используемого для питания механизмов защиты в случае отключения основного источника.

Ядро MIPS 4Ksd также обладает функциями защиты данных. Например, имеются механизмы рандомизации заполнения кэша и его скремблирование. Передача данных по внутренним шинам также осуществляется в режиме скремблирования, что защищает от утечки данных путем анализа электромагнитного излучения (ЭМИ) схемы.

Внешняя память защищена механизмом USIP® Crypto Interface. При его включении все данные, передаваемые по шине аппаратно, шифруются алгоритмом AES-128. Шифрование происходит «прозрачно» для программы микроконтроллера, не занимая ресурсов процессора и не добавляя задержек в передачу данных. К интерфейсу UCI возможно также подключение контроллера Ethernet, что позволяет пересылать зашифрованные данные с большой скоростью на большие расстояния.

Для организации защиты данных на логическом уровне в ядре 4Ksd предусмотрен механизм MMU, который, с использованием конфигурируемых TLB (Translation Lookaside Buffers) обеспечивает разграничение доступа к областям памяти процессора.

Для реализации алгоритмов эффективного шифрования и скремблирования в USIP® имеется блок генерации случайных чисел, соответствующий стандартам NIST 800-22 и DIEHARD.

Разработчики позаботились также о защите периферии. Многие периферийные устройства и интерфейсы имеют встроенные механизмы защиты, устраняющие уязвимые места схемы. Так, например, интерфейс USB имеет встроенный файрвол, а контроллер клавиатуры защищён от сканирования ЭМИ.

Одним из видов вторжения является злонамеренный доступ к разрабатываемому устройству со стороны самих разработчиков одной из стадий разработки. К примеру, у Innova Card, как разработчика аппаратной части имелась бы потенциальная возможность проникнуть в микросхему после того, как устройство на базе USIP уже создано и получить доступ к «защищён-

ной» информации. Однако разработчики предусмотрели защиту «от себя» и других участников разработки. Механизм защиты на уровне процесса разработки и производства состоит в следующем: в USIP® дополнительно предусмотрен механизм смены фаз жизненного цикла микроконтроллера. При переходе от предыдущей стадии к последующей производится смена набора ключей шифрования, что предотвращает возможность доступа к микросхеме со стороны участников предыдущих стадий жизненного цикла.

Для обеспечения безопасного обновления программного обеспечения и загрузки ключей защиты в USIP® реализован загрузчик Secure bootloader с динамической аутентификацией на основе алгоритма AES.

Средства разработки

Средства разработки традиционно представлены программным обеспечением для PC, отладочной платой и эмулятором. Для системы Windows имеется набор программного обеспечения USIP® HDE – среда разработки Eclipse с набором средств компиляции MIPS.

Особый интерес для разработчиков может представлять платформа Linux. Разработчики компании Innova Card портировали ядро Linux на USIP® и создали комплект «Linux Development Environment», включающий в себя гибкую систему конфигурирования компиляции ядра, файловой системы и библиотек, а также среду отладки. Установка и настройка платформы для Linux на USIP происходит довольно быстро, что позволяет выпустить продукцию на рынок за достаточно краткие сроки, обеспечив при этом большую функциональность устройства.

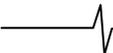
Innova Card предоставляет отладочный комплект UEK – USIP® Essential Kit. Это плата с установленной микросхемой USIP®, содержащая память SDRAM, SRAM, 2 FLASH, 2 интерфейса RS-232, Ethernet, USB OTG, JTAG, дополнительные разъемы для подключения внешнего оборудования (клавиатура, мышь и т.п.) и плат расширения. Кроме того, контакты ввода-вывода USIP® выведены на специальные площадки, к которым имеются две платы-защелки для внешнего доступа. Также компанией разработан комплект «EZ-Cert for EMV L1 Type Approval». Это комплект, состоящий из платы, подключаемой к UEK и позволяющей получить доступ к двум смарт-картам и четырём модулям безопасного доступа (Security Access Modules (SAMs)) одновременно.

Для отладки программного обеспечения по JTAG используется эмулятор FS2 компании First Silicon. Имеется две версии этого эмулятора: с интерфейсами LPT и USB. Скорость доступа составляет 10 МГц.

Программное обеспечение

Имеется множество решений для USIP® от Innova Card и третьих фирм. Рассмотрим некоторые из них:

- UCL Crypto-library от Innova Card. Библиотека включает следующие алгоритмы шифрования: RSA (до 2048 бит), DES, 3DES, AES-128, SHA-1, SHA-256, RIPEMD-160, MD5, HMAC, TRNG/PRNG, ECC, генерация ключей и Diffie-Hellmann.
- SmartSoft EMV Level 2 approved kernel – библиотека, которая предоставляет интерфейс для работы со смарт-картами. Библиотека написана с учетом требований стандарта EMV L2, поэтому её использование упрощает процесс сертификации. Эту библиотеку рекомендуют использовать совместно с пакетом Innova Card EZ-Cert L1. EMV L2 Kernel представляет собой один из уровней иерархии программного обеспечения для USIP®,



устанавливающийся поверх операционной системы (если она имеется) и библиотек Innova Card EMV L1 и HAL. В свою очередь, библиотека состоит из трех уровней: обмена данными между терминалом и смарт-картой; функций парсинга и конвертации данных с карты; функций поддержки требований стандарта EMV2000 4.1. Пакет EZ-Cert2 – это программная часть EMV L2 Kernel, поддержка при интеграции EMV L2 Kernel на систему разработчика и помощь при прохождении сертификации EMV L2.

- Trango – “гипервизор” – одна из перспективных разработок фирмы Trango Systems. Так, если вам требуется защита системы с большой функциональностью и при этом выпустить продукцию необходимо в кратчайшие сроки, то можно за основу вашей платформы взять Linux, «закрыв» его системой Trango.

Trango – система, подобная виртуальной машине, которая устанавливается на USIP®. Эта система создает необходимое количество «виртуальных окружений» для запуска в них операционных систем либо отдельных программ. Главная задача гипервизора – разграничение доступа к отдельным ресурсам микроконтроллера. Например, возможно разрешить использование USB для окружения, в котором выполняется Linux, а доступ к какому-либо UART открыть только для Windows CE. По заявлению разработчиков, накладные расходы на гипервизор Trango составляют не более 3%. Применение этой системы – интеграция открытых (заведомо незащищенных) операционных систем наподобие Линукс в системы со специальными требованиями к безопасности. Ранее задача разграничения доступа к защищенным данным решалась за счет установки двух микроконтроллеров. Использование Trango позволяет это сделать на одной микросхеме.

- jTOP® for Terminals фирмы Trusted Logic – межплатформенное ПО (middleware), совместимое со стандартами GPD и STIP, с набором программных средств и сервисов. Состав библиотек включает в себя: функции стандарта EMV L2, алгоритмы шифрования, стеки коммуникации USB и TCP/IP, поддержка виртуальной машины java.

USIP® в ряду других микроконтроллеров

Глубокая специализация кристалла в области обеспечения безопасности не настолько драматически повлияла на его характеристики как обычного контроллера или вычислителя системы, что позволяет сравнивать его с наиболее успешными образцами из обычного ряда контроллеров. Например, USIP® имеет достаточно схожую структуру с 32-битными микроконтроллерами на базе ядра ARM9. В качестве примера для сравнения был рассмотрен достаточно новый микроконтроллер от STMicroelectronics – STR912FAW44.

Механизмы организации памяти, прерываний и DMA выполнены аналогично. Отличия имеются лишь в деталях реализации. Выборка из Flash организована также, как и в других микроконтроллерах: имеется 128-битный регистр предвыборки, заполняющийся за 4 такта частоты процессора, что позволяет в каждом такте получать одно 32-разрядное слово при условии последовательного расположения во Flash.

В USIP® имеется достаточно развитый набор функциональных устройств и периферии. При этом отдельно взятые периферийные устройства лишь незначительно уступают в функциональной насыщенности микроконтроллерам на базе ARM9. Так, фиксированный множитель PLL не позволяет плавно настраивать частоту, что затрудняет управление энергопотреблением, однако в какой-то степени это компенсируется сравнительно низким общим потреблением микросхемы и возможностью ис-

пользовать четыре режима управления питанием. Несмотря на то, что контроллер в общении с периферией ориентирован на защиту информации, присутствуют также и некоторые достоинства периферийных устройств. Например, имеется возможность достаточно гибко изменять приоритеты устройств на внутренней шине микроконтроллера; устройство DMA предоставляет дополнительные возможности (режим 2D, положительный и отрицательный шаг, приостановка передачи и пр.).

Поскольку специфика микросхемы – обеспечение безопасности, в USIP® сделан упор на увеличение производительности именно алгоритмов шифрования. Так, например, благодаря дополнительным вычислительным блокам, производительность алгоритмов AES, 3DES и SHA-1 на 50-100% выше, чем на микроконтроллере STR912FAW44, имеющем аналогичную тактовую частоту 96 МГц. Для сравнения: генерация 1024-битной сигнатуры RSA занимает менее 35мс.

Следующая версия кристалла USIP2 будет работать на частотах до 175 МГц, что позволит сравнивать этот микроконтроллер с процессорами ARM и MIPS, работающими на частотах до 300 МГц.

Главное достоинство USIP® в сравнении с другими микроконтроллерами – это, несомненно, его защищенность, заложенная в концепцию микроконтроллера изначально. Недостатки же, связанные с некоторыми ограничениями в периферийных устройствах, – лишь временное явление (добавить функциональность в периферию гораздо проще, чем изменять сами принципы архитектуры!). Подобные недостатки могут быть легко исправлены в следующих ревизиях микросхемы, и можно говорить о достаточно высоком потенциале этого кристалла.

Заключение

USIP®, представляя собой готовую платформу для создания платёжных терминалов (POS) и других устройств обеспечения банковской тайны, является одновременно современным высокопроизводительным кристаллом с разветвлённой периферией и может являться базой для построения самых разных устройств, защищающих ноу-хау разработчика на самом высоком уровне.

Обладая низкой стоимостью (€12-15), микросхема, при достаточно высокой производительности, предоставляет широкий набор периферии, позволяющий дополнительно сократить число внешних микросхем и уменьшить стоимость изделия в целом. Основное же преимущество данного кристалла – это система защиты, заложенная в архитектуру системы изначально, что позволяет совместить производительность, гибкость и защищенность в одном корпусе и обеспечить соответствие всем современным требованиям безопасности.

К настоящему времени на основе USIP уже выпущено несколько продуктов, успешно прошедших сертификацию на соответствие современным стандартам безопасности (EMV, PCI и др.)

Ссылки

1. Официальный сайт компании Innova Card: <http://www.innova-card.com>.
2. Официальный сайт компании Analog Devices: <http://www.analog.com>
3. Официальный сайт компании Trusted Logic: <http://www.trusted-logic.com>.
4. Официальный сайт компании Trango Systems: <http://www.trango-systems.com>.
5. Официальный сайт компании SmartSoft: <http://www.smartsoft-it.com>.