

Учреждение образования
**«БЕЛОРУССКИЙ ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ
ИНФОРМАТИКИ И РАДИОЭЛЕКТРОНИКИ»**

УДК 681.3.01:621.391:517.988

ПЕТРОВСКИЙ
Алексей Александрович

**МЕТОДЫ И АЛГОРИТМЫ АНАЛИЗА И СИНТЕЗА ПРОЦЕССОРОВ С
ДИНАМИЧЕСКИ РЕКОНФИГУРИРУЕМОЙ АРХИТЕКТУРОЙ
ПЕРЦЕПТУАЛЬНОГО АУДИО КОДЕРА**

Специальность 05.13.15 – «Вычислительные машины и системы»

АВТОРЕФЕРАТ
диссертации на соискание ученой степени
кандидата технических наук

Минск 2002

Работа выполнена в Учреждении образования «Белорусский государственный университет информатики и радиоэлектроники»

Научный руководитель –
доктор технических наук, профессор Садыхов Р.Х.,
кафедра Электронных вычислительных машин, БГУИР

Официальные оппоненты:
доктор технических наук, профессор Леусенко А.Е.,
кафедра Программного обеспечения информационных технологий, БГУИР
кандидат технических наук, доцент Белый А.А.,
кафедра Системного анализа, БГУ

Оппонирующая организация –
НИРУП «Институт цифрового телевидения»

Защита состоится 31 октября 2002 года в 16 часов на заседании совета по защите диссертаций Д02.15.04 при Учреждении образования «Белорусский государственный университет информатики и радиоэлектроники» по адресу: 220027, г. Минск, П.Бровки, 6, корп. 1, ауд. 232, тел. 2398989.

С диссертацией можно ознакомиться в библиотеке Учреждения образования «Белорусский государственный университет информатики и радиоэлектроники».

Автореферат разослан « 26 » сентября 2002 г.

ОБЩАЯ ХАРАКТЕРИСТИКА РАБОТЫ

Актуальность темы диссертации. Для современной вычислительной техники характерно развитие проблемно-ориентированных вычислительных структур – специализированных процессоров и систем оптимальных для той или иной специальной области применения. По мере совершенствования элементов микроэлектроники разнообразие этих процессоров и систем растёт и их специализация во многих случаях сужается, а масштабы применения расширяются. Среди ряда проблемно-ориентированных процессоров важное место занимают высокопроизводительные процессоры кодирования аудиовизуальной информации мобильных мультимедийных систем. Значительно возросшее внимание к повышению качества реконструированного сигнала обуславливает актуальность исследований в рамках совершенствования методов и аппаратно-программных средств кодирования аудио сигналов в реальном времени. Развитие этих работ в настоящее время идет по пути построения перцептуальных аудио кодеров на основе подходящего частотно-временного преобразования, в частности, банков цифровых фильтров и наиболее гибкого в смысле управления частотно-временным разрешением, пакета дискретного вэйвлет-преобразования (ПДВП).

Один из основных подходов в организации архитектур процессоров для решения задач аудио кодирования в реальном времени базируется на адаптации алгоритма обработки и архитектуры процессора под текущий сигнал. Такие вычислительные системы называются системами с реконфигурируемыми процессорами, а данные подходы обработки сигналов относятся к классу динамических: согласованно с входным сигналом непрерывно (от фрейма к фрейму сигнала) осуществляется адаптация алгоритма обработки; синхронно с трансформацией алгоритма обработки реконфигурируется архитектура процессора. При этом обеспечивается минимальная избыточность в представлении сигнала адаптивным ПДВП и повышение показателя быстройдействия/стоимость аппаратного обеспечения.

Вышеуказанное подчёркивает актуальность исследований для разработки процессоров с динамически реконфигурируемой архитектурой перцептуального аудио кодера на основе адаптивного пакета дискретного вэйвлет-преобразования. Важность проблемы, совершенствование методов и алгоритмов обработки аудио сигналов в реальном времени и их отображения на архитектуру процессора определили цель работы, решаемые задачи и методы исследований.

Связь работы с крупными научными программами, темами. Диссертационная работа выполнена в соответствии с научно-техническими заданиями и планами работ кафедры «Электронные вычислительные машины» и НИЛ 3.1 «Микропроцессорные системы реального времени» Белорусского государственного университета информатики и радиоэлектроники в рамках фундаментальных исследований Министерства образования Республики Беларусь: ГБЦ 97-3077 «Разработка высокопроизводительных модулей для параллельных систем цифровой обработки сигналов на базе ЦПОС процессоров», ГБЦ 99-3066 «Оптимальный синтез реконфигурируемых процессоров вэйвлетного преобразования в системах реального времени», ГБЦ 01-3059 «Разработать методы оптимального синтеза динамически реконфигурируемых вычислительных структур для задач ЦОС», а также при финансовой поддержке Фонда фундаментальных

исследований Республики Беларусь Т99-308 «Новый подход к низкоскоростному кодированию речевых сигналов в условиях зашумления с учётом особенностей слуха человека», хозяйственного договора «Разработать процедуры сжатия речевой информации, обеспечивающие коммерческое качество восстановленной речи», и по договору о международном сотрудничестве организаций Республики Беларусь и Китайской Народной Республики в области науки и технологий: «Run-time reconfigurable architecture for non-uniform filter bank computing and its applications to 1D and 2D signals coding», Institute of Information and Communication Engineering, Zhejiang University, Hangzhou, P.R. China.

Цель и задачи исследования. Целью работы является разработка методов и алгоритмов анализа и синтеза процессоров с динамически реконфигурируемой архитектурой перцептуального аудио кодера мобильных мультимедийных систем.

Поставленная цель определяет следующие задачи исследования:

1. Анализ подходов построения перцептуальных аудио кодеров с целью выбора метода частотно-временного анализа – пакета дискретного вэйвлет-преобразования, обеспечивающего максимальную эффективность системы кодирования.

2. Разработка метода динамической трансформации алгоритма вычисления дискретного вэйвлет-преобразования на основе психоакустической модели в вэйвлет-области.

3. Разработка метода и алгоритма динамической трансформации структуры дерева ПДВП в процессе работы кодера под заданный вычислительный ресурс – реконфигурируемая вычислительная среда.

4. Разработка правила оптимального распределения бит при квантовании вэйвлет-коэффициентов в кодере с учётом эффектов реконструкции сигнала в декодере.

5. Аппаратная реализация процессора с динамически реконфигурируемой архитектурой перцептуального аудио кодера реального времени мобильной мультимедийной системы и методов вычисления оценок его объективного тестирования.

Объект и предмет исследования. Объектом исследования являются процессоры с реконфигурируемой архитектурой перцептуальных аудио кодеров. Предмет исследования – методы кодирования аудио сигналов на основе ПДВП и динамической трансформации алгоритмов под заданный вычислительный ресурс – реконфигурируемая вычислительная среда.

Методология и методы проведенного исследования. Решение рассматриваемых в диссертации задач базируется на методах цифровой обработки сигналов, психоакустики, акустики, математического моделирования, теории вероятностей и математической статистики, методах проектирования проблемно-ориентированных вычислительных средств.

Научная новизна и значимость полученных результатов.

1. На основании анализа существующих подходов к реализации перцептуальных аудио кодеров мультимедийных систем показано, что наиболее эффективным способом реализации аудио кодеров является реконфигурируемая вычислительная среда, структура которой согласованно перестраивается с динамической трансформацией частотно-временного преобразования сигнала – дерева ПДВП.

2. Разработана карта частотно-временного разрешения дерева ПДВП шкалы критических частот и процедура расчёта в вэйвлет-области порогов маскирования восприятия акустической информации человеком.

3. Разработаны метод и алгоритм динамической трансформации структуры дерева ПДВП под заданный вычислительный ресурс и мультiframeовую обработку аудио сигнала в реальном масштабе времени, результатом работы которого является оптимальная структура дерева ПДВП, согласованная с текущим фреймом сигнала (минимум стоимостной функции – временная энтропия вэйвлет-коэффициентов). С другой стороны, алгоритм представляет собой динамическое расписание реконфигурации архитектуры вычислительного ресурса – параллельно-поточного процессора ПДВП.

4. Предложена новая структура ПДВП-кодера аудио сигналов на основе метода динамической декомпозиции ПДВП, которая в отличие от известных ориентирована на обработку сигнала в вэйвлет-области, что исключает частотно-вэйвлетные преобразования психоакустической информации при квантовании вэйвлет-коэффициентов, обуславливающие большую алгоритмическую задержку, и работу в реальном масштабе времени.

5. Разработано правило оптимального распределения бит при квантовании вэйвлет-коэффициентов в кодере с учётом эффектов реконструкции сигнала в декодере – «просачивание» энергии шума квантования в смежные полосы банка фильтров синтеза, что позволяет ограничиться одинаковыми фильтрами как в анализирующем (кодере), так и в синтезирующем (декодере) банках фильтров при соответствующем порядке вэйвлет-функции.

6. Разработаны методы объективной оценки качества реконструированного сигнала перцептуального ПДВП-кодера аудио сигналов на основе периферийной модели уха, что в отличие от известных методик исключает погрешность, обусловленную взаимными частотно-вэйвлетными преобразованиями.

Практическая значимость полученных результатов. Полученные результаты формируют теоретическую и практическую базу для синтеза и анализа процессоров с динамически реконфигурируемой архитектурой перцептуальных аудио кодеров на основе адаптивного пакета вэйвлет-преобразования, функционирующих в реальном времени для мобильных мультимедийных систем. При этом практически реализованы следующие процессоры, устройства и инструментальные программы:

- Экспериментальный модуль системы аудио кодера мобильной мультимедийной системы на основе двух реконфигурируемых ПДВП-процессоров (кодер/декодер), выполненных на двух ПЛИС XC4085XL фирмы Xilinx Inc., а также на ЦПОС TMS320VC33, обеспечивающий скорость передачи от 36 кбит/с до 45 кбит/с. Восстановленный сигнал не содержит артефактов: оценка отношения шума к порогу маскирования около -9 дБ. Данный модуль может найти применение в качестве мониторинговой системы для проектирования новых аудио кодеров, изучения принципов психоакустики и структур реконфигурируемых процессоров.

- Подсистема ввода аудио сигнала – супернаправленный микрофон. Экспериментальные исследования диаграммы направленности для конфигурации из пяти микрофонов показывают перспективность такого построения дешевого

супернаправленного микрофона, который может использоваться также в мобильных системах распознавания речи с малым словарём и в автомобильных телефонах.

- Методы и программа объективной оценки качества восстановленного сигнала перцептуальных аудио кодеров на основе периферийной модели уха, дающие высокую степень коррелированности результатов объективных тестов с субъективными.

- Базовые результаты диссертационной работы внедрены в серийно выпускаемые мобильные абонентские радиостанции «Алтай» и «Вилия» (Молодеченский радиозавод «Спутник»), в мобильную мультимедийную систему автомобиля – система супернаправленного микрофона (Mannesmann-Philips: Car Communication Germany, Wetzlar Germany), переданы ОАО «Минский приборостроительный завод», используются в учебном процессе специальности «Электронные вычислительные средства» (БГУИР).

Основные положения диссертации, выносимые на защиту.

1. Карта частотно-временного разрешения дерева ПДВП шкалы критических частот и процедура расчёта в вэйвлет-области порогов маскирования восприятия акустической информации человеком.

2. Метод и алгоритм динамической трансформации структуры дерева ПДВП в процессе работы кодера под заданный вычислительный ресурс – параллельно-поточный процессор ПДВП, основу которого составляет процедура расчёта порогов маскирования, оценка которых уточняется по мере роста дерева ПДВП.

3. Структура ПДВП-кодера аудио сигналов на основе метода динамической декомпозиции структуры дерева ПДВП, которая в отличие от известных ориентирована на обработку сигналов в вэйвлет-области и работу в реальном масштабе времени.

4. Способ оптимального распределения бит при квантовании вэйвлет-коэффициентов в аудио кодере с учётом эффектов восстановления сигнала в декодере.

5. Параллельно-поточный процессор с динамически реконфигурируемой архитектурой перцептуального аудио кодера реального времени мобильной мультимедийной системы и методы вычисления оценок объективного тестирования аудио кодера.

Апробация результатов диссертации. Результаты, полученные в ходе выполнения исследований, докладывались и обсуждались на: III, IV и VI Международной конференции «Современные средства связи», Нарочь, Беларусь, 1998, 1999, 2001; 1-ой Международной конференции «Техника и технология связи», Минск, Беларусь, 1999; 1-ой Республиканской научно-технической конференции «Цифровое телевидение и радиовещание», Минск-Паубичи, Беларусь, 1999; the 2nd International conference of “Digital Signal Processing in Communication”, Belarussian Polytechnics Academy, Minsk, Belarus, 1999; the 6th International Workshop on Systems, Signals and Image Processing (IWSSIP’99), Bratislava, Slovakia, 1999; the 3rd European Conference “DSP Education and Research”, Paris, France, 2000; 3-й и 4-й Международной конференции «Цифровая обработка сигналов и ее применение» (DSPA’2000, DSPA’2002), Москва, Россия, 2000, 2002; the 110th Audio Engineering Society (AES) Convention, Amsterdam, Netherlands, 2001; the 6th International Conference “Pattern Recognition and Information Processing” (PRIP’2001), Minsk, Belarus, 2001; the 7th European conference on Speech Communication and Technology (EUROSPEECH 2001), Aalborg, Denmark, 2001; the 17th International

Congress of Acoustics (ICA'2001), Rome, Italy, 2001; the 2nd International conference on Neural Networks and Artificial Intelligence (ICNNAI'2001), Minsk, Belarus, 2001; the 9th International conference "Mixed Design of Integrated Circuits and Systems" (MIXDES'2002), Wroclaw, Poland, 2002; the 3rd International conference on Parallel Computing in Electrical Engineering (PARELEC'2002), Warsaw, Poland, 2002; регулярных научно-технических конференциях БГУИР, 1998-2002; научном семинаре Института обработки и передачи данных Аахенского университета, Германия, 1999, 2000.

Опубликованность результатов. По тематике представленной диссертационной работы было опубликовано 26 печатных работ, включая 3 тезиса на научных конференциях и 23 статьи в журналах и материалах международных научных конференций и симпозиумов. Суммарный объем публикаций составляет около 150 страниц.

Структура и объем диссертации. Диссертационная работа состоит из введения, пяти глав, заключения и шести приложений. Она содержит 100 страниц основного текста, 86 рисунков на 58 страницах, 19 таблиц на 10 страницах, шесть приложений на 48 страницах, в списке использованных источников на 11 страницах представлено 132 наименования.

ОСНОВНОЕ СОДЕРЖАНИЕ РАБОТЫ

Во введении показана актуальность темы диссертационной работы, формулируются цель и задачи исследования.

В главе 1 дан детальный анализ подходов построения перцептуальных аудио кодеров, технологий проектирования банков цифровых фильтров и аппаратно-программных реализаций. Общая структура перцептуального аудио кодера показана на рис. 1.



Рис. 1 Общая структура перцептуального аудио кодера

Входной аудио сигнал разбивается на квазистационарные фреймы длительностью от 2 до 50 миллисекунд в зависимости от алгоритмов обработки и методов кодирования. Блок частотно-временного анализа аппроксимирует временные и спектральные параметры аудио сигнала для каждого анализируемого фрейма с учетом шкалы критических частотных полос. В блоке психоакустического анализа оценивается энергия маскирующего сигнала (уровни маскирующих порогов) на базе психоакустической модели. При этом определяются максимальные искажения, возникающие в каждой точке частотно-временной плоскости в процессе квантования и кодирования частотно-временных оценок без введения искусственного артефакта слышимости при восстановлении сигнала. Следовательно, психоакустический анализатор вычисляет частотно-временной параметр невосприятия акустической информации слушателем, который затем передается в блок квантования и кодирования.

Минимальное число бит, необходимое для кодирования воспринимаемого слушателем аудио сигнала, оценивается «перцептуальной энтропией» (PE):

$$PE = \frac{1}{N} \sum_{f=f_l}^{f_h} \max \left(0, \log_2 \frac{|x(f)|}{T(f)} \right),$$

где N – число частотных компонент (каналов) в частотном диапазоне f_l и f_h ; f_l – нижняя частота (например, $f_l = 0$ Гц) диапазона; f_h – верхняя частота (например, $f_h = 22050$ Гц) диапазона; $|x(f)|$ – амплитуда частотной компоненты f ; $T(f)$ – оценка порога маскирования на частоте f .

Показано, что общая концепция построения системы кодирования аудио сигналов неразрывно связана с выбором метода частотно-временного анализа, согласованного с входным сигналом, а максимальная эффективность кодирования определяется разделением спектральных компонент сигнала. Оптимальное распределение бит по каналам банка фильтров может быть оценено по формуле:

$$R_k = R + \frac{1}{2} \log_2 \left(\frac{x_k^2}{T_k^2} \right) - \frac{1}{2} \log_2 \left(\prod_{k=0}^{N-1} \frac{x_k^2}{T_k^2} \right)^{\frac{1}{N}},$$

откуда следует, что мера перцептуальной избыточности определяется отношением:

$$PSFM = \left(\prod_{k=0}^{N-1} \frac{x_k^2}{T_k^2} \right)^{\frac{1}{N}} \left/ \left(\frac{1}{N} \sum_{k=0}^{N-1} \frac{x_k^2}{T_k^2} \right) \right.$$

Как видно, $PSFM$ зависит от распределения по частотному диапазону спектральной энергии, взвешенной энергией порога маскирования. Таким образом, предполагается, что увеличение разрешающей способности частотно-временного преобразования (числа каналов в банке фильтров), должно привести к росту степени компрессии сигнала. Однако, это справедливо не для всех источников сигнала. Дж. Джонстон показал, что для акустического сигнала клавишина эффективность кодирования повышается с ростом числа каналов, но для сигналов кастаньет наблюдается обратный эффект. В данном случае необходимо построить частотно-временное преобразование, характеристики которого зависят от временных изменений сигнала, т.е. обеспечивается требуемое разрешение как по частоте, так и по времени, а не только по частоте.

В отличие от традиционно применяемого для анализа сигналов в аудио кодерах полифазного ДПФ банка фильтров, ПДВП дает двумерную развертку кодируемого аудио сигнала, что обеспечивает возможность анализировать свойства сигнала одновременно во временном и частотном пространствах. Из анализа известных подходов построения перцептуальных аудио кодеров на основе ПДВП следует, что результирующие пороги маскирования не могут быть применены непосредственно в эйвлет-области, т.к. они рассчитываются в частотной области, поэтому необходим перерасчет психоакустической информации для квантования эйвлет-коэффициентов, а также оптимизация структуры дерева ПДВП на основе вычисления полного дерева. Данные обстоятельства обуславливают большую алгоритмическую задержку и невозможность вести обработку в поточном режиме и, как следствие, кодер работает не в реальном времени. Далее в главе показано, что для алгоритмов кодирования аудио сигналов характерны глубокая конвейеризация и высокая параллельность, следовательно, реконфигурируемые вычислительные системы превосходят альтернативные решения на порядок по таким характеристикам как стоимость и скорость вычислений.

Во 2-ой главе решается задача декомпозиции ПДВП путём изменения и минимизации определенной меры качества (стоимостной функции), зависящей от обрабатываемых данных. Целевые функции, основанные на энтропии, наиболее адекватно описывают информационные свойства сигнала. В перцептуальном кодировании аудио сигналов выбирается такая декомпозиция ПДВП, при которой минимизируется скорость передачи с сохранением высокого качества восприятия человеком декодированного сигнала. Декомпозиция «лучшего» дерева преобразования выбирается как можно ближе к шкале барков, т.е. к шкале критических частотных полос, а банк вэйвлет-фильтров, согласованный с критической шкалой частот восприятия акустической информации человеком, является предельной структурой для метода перцептуального кодирования аудио сигнала. Пусть $\{\varphi_n(t): n \in Z\}$ определяет множество структур деревьев ПДВП и пусть $E \subset \{(l,n): 0 \leq l \leq L - \text{масштабный уровень}, 0 \leq n \leq N - \text{номер узла на масштабном уровне}\}$ представляет собой узлы дерева ПДВП, тогда отрезок $[0,1)$ разделяется на диадические интервалы $I_{l,n} = [n2^{-l}, (n+1)2^{-l}]$, которые соответствуют специфическому множеству узлов E . В частности, $\{\varphi_{l,n,k}(t): (l,n) \in E, k \in Z\}$, где $\varphi_{l,n,k}(t) \stackrel{\Delta}{=} 2^{-l/2} \varphi_n(2^{-l}t - k)$ является базовой формой в пространстве сигнала $\overline{\text{span}}\{\varphi_0(t-k): k \in Z\}$. Узел $(l,n) \in E$ дерева ПДВП ассоциируется с частотной полосой, в которой центральная частота и полоса пропускания приблизительно задаются следующими соотношениями: $f_{l,n} = 2^{-l}(GC^{-1}(n)+0.5) \cdot f_s/2$, $\Delta f_{l,n} = 2^{-l} \cdot f_s/2$, где GC^{-1} – обратный код перестановок Грея, f_s – частота дискретизации сигнала.

Эмпирически определена структура дерева ПДВП (Critical Band Wavelet Packet Decomposition (CB-WPD)) $(l,n) \in E_{CB}$, $l=0,8$, где E_{CB} – обозначает множество узлов дерева ПДВП соответствующего CB-WPD. Дерево CB-WPD осуществляет разделение частотного диапазона $[0-22,05 \text{ кГц}]$ аудио сигнала на полосы согласно критической шкале частот (25 барков). Корневой узел $(l,n) = (0,0)$ данного дерева соответствует всему частотному диапазону сигнала. Каждый внутренний узел дерева $(l,n) \in E$, названный узлом предка, делится на два потомка: 1-й потомок и 2-й потомок, ассоциируемые соответственно с высокочастотной и низкочастотной фильтрацией, выходные сигналы (вэйвлет-коэффициенты) которых децимируются в соотношении 2:1: $X_{l,n,k}(t) = \langle x(t), \varphi_{l,n,k}(t) \rangle$, $(l,n) \in E_{CB}$, $k \in Z$. Вычислена его амплитудно-частотная характеристика, определены погрешности аппроксимации центральных частот и ширины каждой критической полосы. Структура дерева ПДВП достаточно точно отражает центральные частоты критических полос, в то же время аппроксимация ширины полос несколько неточная, что обусловлено применением диадического вэйвлет-преобразования – наиболее скоростного преобразования. Далее построена карта частотно-временного разрешения дерева ПДВП шкалы критических частот, для которой разработана процедура расчёта порогов маскирования. Данная карта и процедура являются основой модели восприятия акустической информации человеком, используемой в алгоритмах динамической декомпозиции дерева ПДВП и объективных оценках качества реконструированного сигнала.

Предлагается новый подход динамической трансформации алгоритма (ДТА) на основе двухэтапной обработки фрейма аудио сигнала: на *первом этапе* на основании перцептуальной энтропии, вычисленной в вэйвлет-области, оценивается требуемое количество бит кодирования вэйвлет-коэффициентов в узлах дерева преобразования

$(l, n) \in E: PE_{l,n} = \sum_{k=1}^{K_{l,n}-1} \log_2(2[\text{int}(SRM_{l,n,k})+1])$, [бит/(l, n)], $(l, n) \in E_j$, $k \in Z$, где $SMR_{l,n,k}$ –

отношение среднеквадратического значения вэйвлет-коэффициентов $X_{l,n,k}$ в полосе узла (l, n) дерева E_j к соответствующему маскирующему порогу $T_{l,n}$, равномерно распределенному между $K_{l,n}$ коэффициентами $X_{l,n,k}$, $k=1, \dots, K_{l,n}$ узла (l, n) , определяется на

основе соотношения: $SMR_{l,n,k} = |X_{l,n,k}| / \sqrt{12 \cdot T_{l,n} / K_{l,n}}$. Знаменатель $\sqrt{12 \cdot T_{l,n} / K_{l,n}}$

представляет собой максимальный шаг квантователя $\Delta_{l,n}$ вэйвлет-коэффициентов в узле $(l, n) \in E_j$, а величина $SMR_{l,n,k}$ задает минимальное количество уровней квантования.

Далее осуществляется рост дерева преобразования; на *втором этапе* оценивается информативность структуры дерева ПДВП $J_{l,n} = \sum_{\forall (l,n) \in E} |X_{l,n,k}|^2 \log(1/|X_{l,n,k}|^2)$, $(l, n) \in E_i$, $k \in Z$ –

а именно, как точно частотно-временное разрешение банка фильтров анализа, соответствующего структуре дерева ПДВП, локализовало нестационарности сигналов. Основу алгоритма роста дерева ПДВП составляет модифицированная процедура расчёта порогов маскирования: грубая оценка порогов уточняется по мере роста дерева ПДВП, т.е. с увеличением разрешающей способности по частоте.

Алгоритм. Рост дерева ПДВП.

Пусть решение о декомпозиции узла (l, n) дерева E_j ПДВП будет обозначаться как $split(l, n)$, где l – уровень декомпозиции, т.е. масштабный уровень преобразования, а n есть n -й узел на уровне l . Пусть текущий узел (предок) будет (l, n) , а его потомки определяются как $(l+1, 2n)$ и $(l+1, 2n+1)$, где $l=0, 1, 2, 3, \dots$, $n=0, 1, 2, 3, \dots$.

Шаг 1. Пусть $l=0$, $split(l, n) = \text{YES}$, т.е. задан корневой узел $(0, 0)$ дерева преобразования E_0 – входной фрейм аудио сигнала, перцептуальная энтропия которого равна $PE_{0,0}$.

Шаг 2. Осуществляется декомпозиция входного сигнала на основе ячейки – банка из двух ортонормальных вэйвлет-фильтров.

Шаг 3. Вычисляется перцептуальная энтропия в узлах декомпозиции.

Шаг 4. $l = l+1$. ЕСЛИ $l-1 >$ максимального масштабного уровня дерева *CB-WPD*, ТОГДА STOP – конец роста дерева ПДВП.

Шаг 5. Для каждого узла n уровня l рост дерева E_l ПДВП осуществляется следующим образом:

- выполняется декомпозиция узла предка (l, n) ;
 - вычисляется перцептуальная энтропия в узлах потомков: $PE_{l+1, 2n}$ и $PE_{l+1, 2n+1}$
- ЕСЛИ $PE_{l,n} \geq PE_{l+1, 2n} + PE_{l+1, 2n+1}$, ТОГДА $split(l, n) = \text{YES}$
ИНАЧЕ $split(l, n) = \text{NO}$.

Шаг 6. Переход к шагу 4.

Таким образом, для каждого входного фрейма аудио сигнала каждый узел-предок (l, n) дерева E_j разделяется на два узла-потомка $(l+1, 2n)$ и $(l+1, 2n+1)$, тогда и только тогда,

когда сумма перцептуальной энтропии в узлах-потомках $(l+1,2n)$ и $(l+1,2n+1)$ меньше, чем значение перцептуальной энтропии в узле-предке (l,n) . При этом определяется субоптимальная структура декомпозиции ПДВП. В отличие от известных алгоритмов трансформации ПДВП, рост дерева согласно данному алгоритму осуществляется сверху вниз, без возвратов на меньшие уровни преобразования и необходимости построения полного дерева ПДВП. Экспериментальная проверка эффективности данного алгоритма роста дерева ПДВП и оценки точности порогов маскирования показала, что порог едва различимых искажений приближается к такому же уровню, как и для предельного дерева преобразования E_{CB} , а шум квантования не превышает данный порог.

На основе метода динамической декомпозиции ПДВП разработана новая структура перцептуального ПДВП-кодера аудио сигналов, которая, в отличие от известных ориентирована на обработку сигналов в вэйвлет-области и работу в реальном масштабе времени.

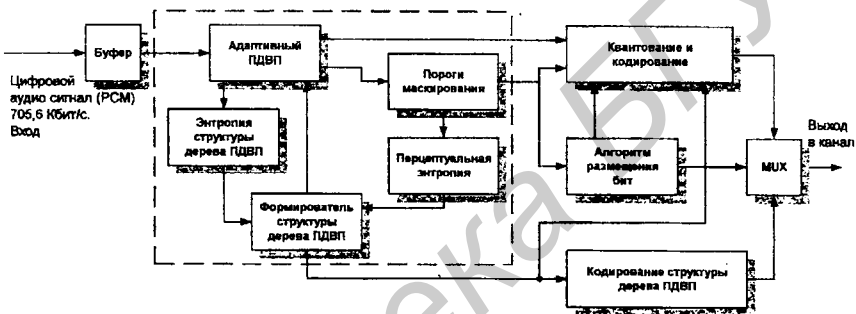


Рис. 2 Структура кодера аудио сигналов на базе адаптивного ПДВП

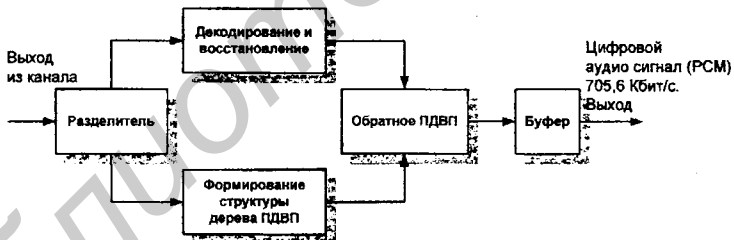


Рис. 3 Структура декодера аудио сигналов

В темпе обработки сигналов в блоке «адаптивный ПДВП» для каждого текущего дерева E_i вычисляются пороги маскирования $T_{l,n}$, значение перцептуальной энтропии $PE_{l,n}$, а также энтропия структуры дерева E_i ПДВП $J_{l,n}$. На основании данной информации в блоке «Формирователь структуры дерева ПДВП» рассчитываются параметры реконфигурации дерева ПДВП. Данный процесс осуществляется поступательно без возвратов на меньшие масштабные уровни преобразования. Следовательно, весь вычислительный процесс идеально отображается на архитектуру параллельно-поточных процессоров. Обработка аудио сигнала выполняется фреймами. Следующим этапом работы кодера является квантование и кодирование коэффициентов оптимального дерева ПДВП $X_{l,n,k}$, $(l,n) \in E$, $k \in Z$. Управление данным процессом осуществляется алгоритмом размещения бит. Наконец, кодируется структура дерева

ПДВП E . Структура ПДВП-декодера аудио сигналов значительно проще структуры кодера и показана на рис. 3, где входная информация разделяется на два потока данных: кодированные вэйвлет-коэффициенты и код структуры дерева ПДВП; формирование новой структуры дерева ПДВП; реконструкция аудио сигнала синтезирующим банком цифровых фильтров, реализованным как обратное адаптивное ПДВП.

В 3-й главе предложена структурная организация реконфигурируемой системы ЦОС для аудио кодера на основе ДТА-подхода, представляющая собой неоднородную вычислительную систему: спецпроцессор, ориентированный на обработку сигнала в частотных полосах, - система мониторинга и ПДВП-процессор с реконфигурируемой архитектурой – аудио кодер. Система мониторинга осуществляет: обработку вэйвлет-коэффициентов $X_{l,n,k}$ в частотных полосах, соответствующих текущей структуре дерева $(l,n) \in E_i$; вычисление психоакустической модели дерева $(l,n) \in E_i$; $T_{l,n}$ и $PE_{l,n}$; получение вектора реконфигурации ПДВП-процессора $r_{l,n}$, $(l,n) \in E_i$; формирование и разборка данных передачи $[code(X_{l,n,k}, (l,n) \in E_i, k \in Z)]$ и приема $\hat{X}_{l,n,k}(code)$, $(l,n) \in E_i, k \in Z$, если используется дуплексный режим работы аудио кодера. Следовательно, основная цель системы мониторинга – получить вектор трансформации $r_{l,n}$ ПДВП-процессора. На процессоре ПДВП с реконфигурируемой архитектурой $r_{l,n}$, $(l,n) \in E_i$ решаются задачи: анализа (кодирование) $X_{l,n,k} = WPD_{l,n}(x_j(t))$, $(l,n) \in E_i$, $x_j(t)$ – j -й фрейм аудио сигнала; синтеза (декодирование) $\hat{x}_j(t) = WPD_{l,n}^{-1}(\hat{X}_{l,n,k})$, $(l,n) \in E_i$, $\hat{x}_j(t)$ – j -й фрейм реконструированного аудио сигнала. В ПДВП-процессоре, в зависимости от масштабного уровня l текущего фрейма, могут находиться в обработке $l-1$ последующих фрейма входного аудио сигнала $x_j(t)$, что обусловлено конвейерной организацией вычислений. Мету информативности структуры дерева ПДВП текущего фрейма сигнала предлагается конструировать на основе временной энтропии вэйвлет-коэффициентов

$$WTE_{E_i} = \sum_{\substack{\min \\ \forall (l,n) \in E_i}} \sum_k \frac{|X_{l,n,k}|}{\sum_{\substack{\min \\ \forall (l,n) \in E_i}} |X_{l,n,k}|} \ln \left(\frac{|X_{l,n,k}|}{\sum_{\substack{\min \\ \forall (l,n) \in E_i}} |X_{l,n,k}|} \right), (l,n) \in E_i, k \in Z, i = \overline{1,8},$$

где $X_{l,n,k} \in (l,n)$ – коэффициенты узла (l,n) дерева E_i . Данная стоимостная функция характеризует энтропию вэйвлет-коэффициентов $X_{l,n,k}$ в узлах (l,n) дерева E_i и отражает изменение во времени информативности ПДВП, отсюда и название временная энтропия вэйвлет-коэффициентов (WTE – wavelet time entropy). Доказано, что её минимизация ведет к сокращению неопределенности и соответственно к увеличению информативности структуры дерева ПДВП. Разработан метод и алгоритм динамической трансформации дерева ПДВП в реальном времени под заданный вычислительный ресурс: параллельно-поточный реконфигурируемый процессор $(m, r_{l,n})$, где m – номер ступени процессора, $r_{l,n} = [(\alpha_1, \beta_0, \beta_1), (\alpha_2, \beta_0, \beta_1, \beta_2, \beta_3), \dots, (\alpha_n, \beta_0, \dots, \beta_n)]$ – вектор реконфигурируемых параметров структуры процессора соответствующей декомпозиции дерева ПДВП $(l,n) \in E_i$, α_i и β_n принимают значения 0 или 1: параметры α_i задают переход на новый масштабный уровень l дерева ПДВП, т.е. включают в обработку сигнала очередную ступень m процессора; группа параметров β_n активизируют узлы n на уровне l дерева E_i . Результатом работы алгоритма является оптимальная структура дерева

ПДВП, согласованная с текущим фреймом сигнала (минимум стоимостной функции – временная энтропия вэйвлет-коэффициентов).

Алгоритм. Динамическая трансформация дерева ПДВП под заданный вычислительный ресурс

Пусть реконфигурируемый вычислительный ресурс $(m, r_{l,n})$ ограничивается величиной C . Предельное дерево ПДВП $CB-WPD$: $(l,n) \in E_{CB}$. Пусть требуемый вычислительный ресурс i -го дерева ПДВП $(l,n) \in E_i$, определяется величиной c_i , а решение о декомпозиции узла предка (l,n) на узлы потомков $(l+1,2n)$ и $(l+1,2n+1)$ будет обозначаться как $split(l,n)$, где l – масштабный уровень преобразования, n – n -й узел на уровне l . Положим, что j является номером фрейма входного аудио сигнала.

Шаг 1. Пусть $j = 1, m = 1, l = 0$ и $split(l,n) = YES, r_{l,n} = YES$, т.е. задан корневой узел $(0,0)$ дерева преобразования – 1-й фрейм входного аудио сигнала, у которого перцептуальная энтропия $PE_{0,0}$ и информативность дерева $WTE_{E_{0,0}}$, а также реконфигурация процессора разрешена.

Шаг 2. $i = j$, т.е. 1-й фрейм входного сигнала определяет рост дерева ПДВП. Осуществляется декомпозиция сигнала на основе ячейки (1-й ступени процессора) – банка из двух ортонормальных вэйвлет-фильтров.

Шаг 3. Вычисляются оценки перцептуальной энтропии в узлах декомпозиции и информативности дерева $WTE_{E_{i,1}}$.

Шаг 4. Проверяется информативность дерева $E_{i,1}$ по сравнению с деревом $E_{0,0}$
ЕСЛИ $WTE_{E_{0,0}} < WTE_{E_{i,1}}$,

ТОГДА это не аудио сигнал, параметры не кодировать, переход к шагу 1.

Шаг 5. $l = l + 1$. ЕСЛИ $l - 1 >$ максимального масштабного уровня дерева $CB-WPD$,
ТОГДА STOP – конец роста дерева $E_{i,m}$ ПДВП.

Шаг 6. Проверить на принадлежность узлов дерева $E_{i,m}$ дереву $CB-WPD$: E_{CB} :
ЕСЛИ $(l,n) \in E_{i,m} = (l,n) \in E_{CB}$, ТОГДА $r_{l,n} = NO$.

Шаг 7. $m = m + 1$. Рост дерева $E_{i,m}$ осуществляется следующим образом.
Для каждого узла n уровня l :

- оценивается и проверяется достаточность вычислительного ресурса $c_{i,m}$:
ЕСЛИ $c_{i,m} > C$, ТОГДА $r_{l,n} = NO$, STOP – конец роста дерева $E_{i,m}$ ПДВП;
- выполняется декомпозиция узла предка (l,n) ;
- вычисляется перцептуальная энтропия в узлах потомках $PE_{l+1,2n}$ и $PE_{l+1,2n+1}$:
ЕСЛИ $PE_{l,n} \geq PE_{l+1,2n} + PE_{l+1,2n+1}$, ТОГДА $split(l,n) = YES, r_{l,n} = YES$
ИНАЧЕ $split(l,n) = NO, r_{l,n} = NO$.

Шаг 8. $j = j + 1$. Читать следующий фрейм аудио сигнала, который обрабатывается в соответствии со структурой дерева $E_{i,m}$ ПДВП.

Шаг 9. Оценивается информативность $WTE_{E_{i,m}}$ дерева $E_{i,m}$:

ЕСЛИ $WTE_{E_{i,m}} > WTE_{E_{i,m-1}}$,

ТОГДА $r_{l,n} = NO$, STOP – конец роста дерева $E_{i,m}$ ПДВП.

Шаг 10. Переход к шагу 5.

STOP. Для входного фрейма i оптимальное дерево преобразования $E_{i,m-1}$, $m = m - 2$,
 $l = l - 2, i = i + 1$. Переход к шагу 5.

С другой стороны, данный алгоритм представляет собой динамическое расписание реконфигурации архитектуры параллельно-поточного процессора ПДВП. Трансформация алгоритма обработки под текущий фрейм сигнала обуславливает минимальную избыточность в представлении сигнала адаптивным ПДВП, а согласованность реконфигурации архитектуры процессора с трансформацией алгоритма обработки приводит к увеличению показателя производительность/стоимость.

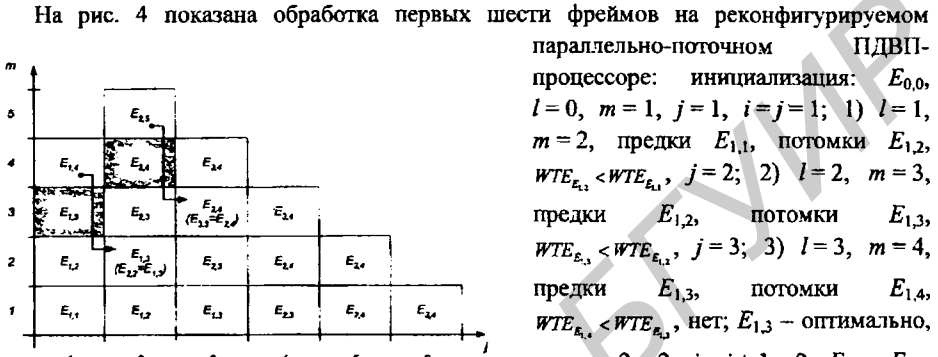


Рис. 4 Схема динамической реконфигурации дерева ПДВП

5) $l=3, m=4$, предки $E_{2,3}$, потомки $E_{2,4}$, $WTE_{E_{2,4}} < WTE_{E_{2,3}}$, $j=5$; 6) $l=4, m=5$, предки $E_{2,4}$, потомки $E_{2,5}$, $WTE_{E_{2,5}} < WTE_{E_{2,4}}$, нет; $E_{2,4}$ – оптимально, $m=m-2=3$, $i=i+1=3$, $E_{3,3}=E_{2,4}$, $l=l-2=2$; 7) $l=3, m=4$, предки $E_{3,3}$, потомки $E_{3,4}$, $WTE_{E_{3,4}} < WTE_{E_{3,3}}$, $j=6$. Здесь по оси ординат отложен номер ступени m поточного процессора, а по оси абсцисс номер фрейма j аудио сигнала «ставшего» на обработку в ПДВП-процессоре в соответствии со структурой дерева $E_{m,i}$ текущего фрейма i . Заштрихованные клетки обозначают оптимальное ПДВП дерево для соответствующего фрейма аудио сигнала, а стрелки показывают на новую структуру дерева нового текущего фрейма – фрейма, по которому определяется информативность $WTE_{E_{m,j}}$ структуры дерева ПДВП. Так, например, пока 3-й фрейм будет текущим, фреймы 4, 5, 6, 7 будут обрабатываться согласно структуре дерева ПДВП, которая формируется данным фреймом аудио сигнала.

Реконфигурируемая параллельно-поточная архитектура процессора ПДВП, состоит из последовательно включённых восьми процессорных модулей и модулей буферной памяти, реконфигурация которых адаптирует архитектуру процессора к структуре дерева ПДВП текущего фрейма сигнала (рис. 5).

На данной реконфигурируемой архитектуре ПДВП-процессора возможна структурная минимизация потребляемой энергии процессором за счёт включения/выключения ступеней процессора. Показано, что реализация базовой операции ПДВП на последовательно включённых умножителе и накапливающем сумматоре является наиболее эффективной для взъелет-фильтров высокого порядка (20 и выше) семейства Добеши. При этом требуется не более 1,5 КЛБ (конфигурационных логических блоков ПЛИС) на порядок фильтра при скорости потока поступления входных данных от 100 до 1000 кГц и в зависимости от типа ПЛИС.

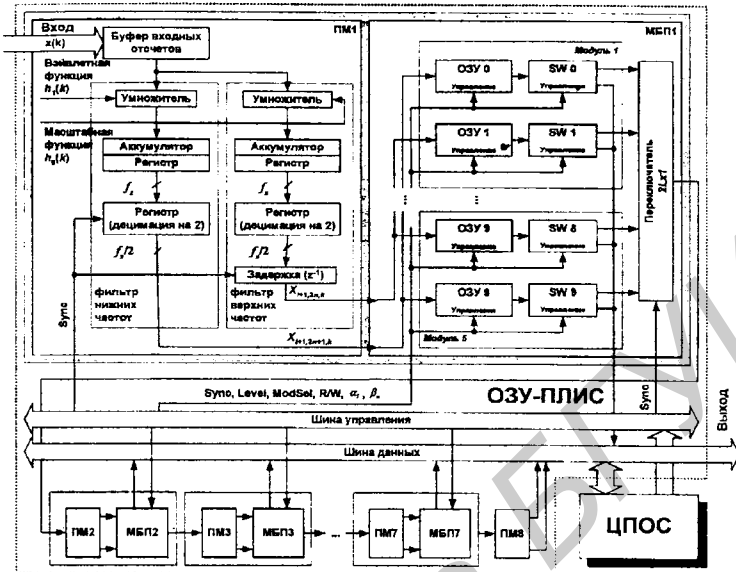


Рис. 5 Архитектура параллельно поточного ЦДВП-процессора

4-я глава посвящена разработке и исследованию системы аудио кодера реального времени мобильной мультимедийной системы, состоящей из подсистемы ввода сигнала – супернаправленного микрофона и подсистемы аудио кодера: ЦОС-процессора системы мониторинга и двух ЦДВП-процессоров (кодер/декодер). Синхронизация работы подсистемы осуществляется host-процессором. Оценка вычислительной сложности алгоритмов и процедур системы мониторинга на фрейм 23 мс показывает, что производительности серийно выпускаемых ЦОС-процессоров класса TMS320VC33 достаточно для реализации системы мониторинга аудио кодера реального времени. Из анализа структурных решений подсистемы ввода сигнала в мобильных мультимедийных системах на основе массива микрофонов (супернаправленного микрофона) следует, что планарная позиция микрофонов в плоскости более предпочтительнее, чем при линейной позиции микрофонов: обеспечивается управление чувствительностью в двух направлениях, а также усиление сигнала значительно больше. В главе описывается практическая реализация встраиваемой системы аудио кодера мобильной мультимедийной системы на основе двух реконфигурируемых ЦДВП-процессоров (кодер/декодер), выполненных на двух ПЛИС XC4085XL фирмы Xilinx Inc., а также на ЦПОС TMS320VC33, который осуществляет мониторинг аудио кодера. Анализ временных затрат выполнения процедур в данном аудио кодере показывает, что вносимая задержка равна размеру фрейма входного аудио сигнала, т.е. 23 мс.

Показано, что «эффект просачивания» энергии шума квантования в смежные полосы банка фильтров синтеза не может пренебрегаться без внесения ущерба в качество восстановленного аудио сигнала. Разработано правило оптимального распределения бит при квантовании вэйвлет-коэффициентов в кодере с учётом эффектов реконструкции сигнала в декодере при коэффициенте децимации $1/\mu_{l,n}$:

$$R_i = R + \frac{1}{5} \log_2 \left(\frac{\sigma_{x_i}^2 \sum_{\substack{k \\ \forall (l,n) \in E_m}} \frac{X_{l,k}}{T_k}}{12 \cdot \mu_i} \Bigg/ \prod_{\substack{k \\ \forall (l,n) \in E_m}} \left[\frac{\sigma_{x_j}^2 \sum_{\substack{k \\ \forall (l,n) \in E_m}} \frac{X_{j,k}}{T_k}}{2 \cdot \mu_j} \right]^{\mu_j} \right),$$

где $i = 1, 2, \dots$, максимальное значение, которое определяется числом частотных полос соответствующего дерева $(l, n) \in E_m$, $\sigma_{x_i}^2$ - дисперсия вэйвлет-коэффициентов i -го узла.

Следует заметить, что данное выражение представляет собой обобщение формулы распределения бит по каналам банка фильтров, которая была получена для идеальных фильтров (банк с равномерными полосами пропускания) в первой главе работы. Выбирая достаточно большой порядок (20-го и более) вэйвлет-функции можно ограничиться одинаковыми фильтрами как в анализирующем, так и в синтезирующем банках фильтров.

Кодирование вэйвлет-коэффициентов осуществляется с помощью кодовых книг Хаффмана. В связи с динамическим характером изменения структуры дерева ПДВП от фрейма к фрейму с целью минимизации числа и объема кодовых книг Хаффмана предлагается формировать их только для каждого масштабного уровня на основе предельного дерева ПДВП, структура которого постоянна. Структуру же дерева ПДВП кодировать только для первого фрейма, а для всех последующих предлагается передавать код ее модификации.

Глава 5 является экспериментальной частью диссертационной работы. Исследуются технические характеристики экспериментального модуля подсистемы аудио кодера и исследовательского модуля подсистемы ввода аудио сигнала (супернаправленный микрофон), анализ которых показывает принципиальную возможность реализации встраиваемых аудио кодеров реального времени для мобильных мультимедийных систем, соответствующих высоким требованиям международных стандартов. Разработаны методы и программа объективной оценки качества реконструированного сигнала перцептуального ПДВП-кодера аудио сигналов в вэйвлет-области на основе периферийной модели уха, базирующейся на модели дерева *CB-WPD* и процедурах расчета порогов маскирования $T_{l,n}(i)$, $(l, n) \in E_{CB}$ в вэйвлет-области: (для i -го фрейма) отношение энергии шума (энергии вносимых искажений) к порогу маскирования - $NMR_{l,n}(i)$ и среднее арифметическое по всем критическим частотным полосам - $NMR_{loc}(i)$, $(l, n) \in E_{CB}$; (для полного сигнала) среднее арифметическое $NMR_{loc}(i)$ - NMR_{total} или как среднее геометрическое $NMR_{loc}(i)$ - NMR_{SEG} ; относительное число искаженных фреймов к общему числу фреймов (MF) кодируемого аудио сигнала (masking flag); частотная полоса сигнала (SB – signal bandwidth), характеризующая искажение тембра, притупление или заглушение звука сигнала. Ниже в таблице приведены объективные оценки перцептуального ПДВП-аудио кодера в сравнении с популярным аудио кодером MPEG-1 уровень III, минимальная скорость передачи которого 64 кбит/с.

Анализ таблицы показывает, что реконфигурированный сигнал имеет достаточно хорошее качество, соответствующее требованиям стандарта ITU-R PEAQ при высокой степени компрессии в 15 и более раз, или, что соответственно, при минимальной скорости передачи от 36 кбит/с до 45 кбит/с. Восстановленный сигнал не содержит

никаких артефактов при оценке $NMR_{total} \approx -9$ дБ; субъективная оценка качества – средняя вероятность предпочтения оригинального аудио сигнала над восстановленным равна около 0.5, что соответствует «прозрачному» кодированию, а процедура расчёта маскирующих порогов по мере роста дерева, согласно алгоритму динамической декомпозиции ПДВП, даёт объективные уровни восприятия акустической информации человеком. Экспериментальные исследования показывают высокую коррелированность результатов объективных и субъективных тестов.

Тестовый материал $f_s = 44,1$ кГц, 16 бит РСМ код	Тип кодера	Минимальная скорость передачи, [кбит/с]	Максимальная степень компрессии	Среднее количество, [бит/отсчет]	Ширина частотной полосы SB, [кГц]	NMR_{total} [дБ]	NMR_{seg} [дБ]	МТ, % искаженных фреймов
Boney-M 15 сек.	ПДВП-кодер	45	15.68	1.02	18.49	-8.297	-0.8699	27.04
	MPEG-1 уровень III	64	11.02	1.45	15.13	4.430	0.4076	99.85
Classic 10 сек.	ПДВП-кодер	42	16.8	0.95	17.62	-6.495	-0.7987	43.45
	MPEG-1 уровень III	64	11.02	1.45	12.58	3.284	0.3145	99.78
Queen 10 сек	ПДВП-кодер	41	17.2	0.92	16.93	-7.597	-0.7765	40.39
	MPEG-1 уровень III	64	11.02	1.45	11.89	4.236	0.3994	97.80
ABBA "Take a chance on me" 2 sec	ПДВП-кодер	36	19.7	0.81	18.58	-4.037	-0.5844	66.25
	MPEG-1 уровень III	64	11.02	1.45	15.39	3.720	0.3312	98.70
ABBA S.O.S. 6 sec	ПДВП-кодер	36	19.5	0.81	18.72	-8.857	-0.9026	20.65
	MPEG-1 уровень III	64	11.02	1.45	15.20	3.740	0.3585	99.63

Приложения содержат краткие сведения из базовых принципов психоакустики; характеристики вэйвлет-функций; VHDL – программу базовой операции на MAC структуре; листинг имитационной программы перцептуального аудио кодера (программа разработана для среды MATLAB 6.1); кодовые книги Хаффмана; акты внедрения результатов диссертационной работы.

ЗАКЛЮЧЕНИЕ

Основные научные и практические результаты диссертационной работы следующие:

1. На основании анализа существующих подходов реализации перцептуальных аудио кодеров мультимедийных систем показано, что наиболее эффективным способом реализации аудио кодеров является реконфигурируемая вычислительная среда, структура которой согласованно перестраивается с динамической трансформацией частотно-временного преобразования сигнала – дерева ПДВП [2,4,6,10,11,16,25].
2. Разработана карта частотно-временного разрешения дерева ПДВП шкалы критических частот и процедура расчёта в вэйвлет-области порогов маскирования восприятия акустической информации человеком [8,17,18].

3. Разработаны метод и алгоритм динамической трансформации структуры дерева ЦДВП под заданный вычислительный ресурс и мультифреймовую обработку аудио сигнала в реальном масштабе времени, результатом работы которого является оптимальная структура дерева ЦДВП, согласованная с текущим фреймом сигнала (минимум стоимостной функции – временная энтропия вэйвлет-коэффициентов) [9,12,18,20,26]. С другой стороны, алгоритм представляет собой динамическое расписание реконфигурации архитектуры вычислительного ресурса – параллельно-поточного процессора ЦДВП [1,10,13,14,15,17].

4. Предложена новая структура ЦДВП-кодера аудио сигналов на основе метода динамической декомпозиции ЦДВП, которая в отличие от известных ориентирована на обработку сигнала в вэйвлет-области, что исключает частотно-вэйвлетные преобразования психоакустической информации при квантовании вэйвлет-коэффициентов, обуславливающие большую алгоритмическую задержку и работу в реальном масштабе времени [17,20].

5. Разработано правило оптимального распределения бит при квантовании вэйвлет-коэффициентов в кодере с учётом эффектов реконструкции сигнала в декодере – «просачивание» в смежные полосы энергии шума квантования, что позволяет ограничиться одинаковыми фильтрами как в анализирующем (кодере), так и в синтезирующем (декодере) банках фильтров при соответствующем порядке вэйвлет-функции [17].

6. Разработаны методы объективной оценки качества реконструированного сигнала перцептуального ЦДВП-кодера аудио сигналов на основе периферийной модели уха [7], которые в отличие от известных методик, исключают погрешность, обусловленную взаимными частотно-вэйвлетными преобразованиями.

7. Осуществлена аппаратная реализация экспериментального модуля аудио кодера с подсистемой ввода аудио сигнала (супернаправленный микрофон) мобильной мультимедийной системы на основе двух реконфигурируемых ЦДВП-процессоров (кодер/декодер), выполненных на двух ПЛИС XC4085XL фирмы Xilinx Inc., а также на ЦПОС TMS320VC33, обеспечивающий скорость передачи от 36 кбит/с до 45 кбит/с [3,5,11,17,21,22,23,24]. Восстановленный сигнал не содержит артефактов: оценка отношения шума к порогу маскирования около -9 дБ. Экспериментальные исследования диаграммы направленности для конфигурации из пяти микрофонов показывают перспективность такого построения дешевого супернаправленного микрофона [19], который может использоваться также в мобильных системах распознавания речи с малым словарём и в автомобильных телефонах.

СПИСОК ОПУБЛИКОВАННЫХ РАБОТ ПО ТЕМЕ ДИССЕРТАЦИИ

Статьи в журналах и научных сборниках

1. Петровский Ал.А. Встраиваемый процессор вэйвлетного преобразования реального времени для систем телекоммуникаций // Специальный выпуск журнала «Известия Белорусской инженерной академии». – Минск, 1998. – №2(6)/1. – С. 7-10.
2. Петровский Ал.А. Графический кроссовый интерпретатор процесса вэйвлетного преобразования в среде MATLAB // Специальный выпуск журнала «Известия Белорусской инженерной академии». – Минск, 1998. – №2(6)/1. – С. 11-13.

3. Petrovsky A.I., Anoshenko A.E. Implementation of decoupled partitioned block frequency domain adaptive filter based on the pipeline multi-DSP systems // Научно-производственный журнал «Вестник связи». – Минск, 1999. – №1. – С. 39-42.
4. Петровский Ал.А., Садыхов Р.Х. Эксперименты с вейвлетным преобразованием в среде MATLAB // Научно-производственный журнал «Вестник связи». – Минск, 1999. – №1. – С. 35-38.
5. Латушко Д.В., Петровский Ал.А. Универсальная система разработки и отладки приложения на базе DSP процессоров с среде Windows // Специальный выпуск журнала «Известия Белорусской инженерной академии». – Минск, 1999. – №1(7)/2. – С. 104-107.
6. Петровский Ал.А. Компрессия аудио сигналов на базе психоакустики: подходы и структуры // Республиканский межведомственный сборник научных трудов «Радиотехника и электроника». – Мн.: БГУИР, 1999. – вып. 24. – С. 140-149.
7. Petrovsky A.A., Bielawski K., Petrovsky A.I. Combined system for acoustic echo and noise reduction based on the psychoacoustically motivated multirate filter bank // Journal of the University of Applied Sciences Mittweida. – Mittweida, Germany, 2000. – № 9. – P. 33-41.
8. Петровский Ал.А. Расчет маскирующих порогов для аудио кодеров на базе пакетного дискретного вейвлетного преобразования // Республиканский межведомственный сборник научных трудов «Радиотехника и электроника». – Мн.: БГУИР, 2000. – вып. 25. – С. 44-57.
9. Петровский Ал.А. Динамическая реконфигурация пакетного вейвлетного преобразования на основе вычисления перцептуальной энтропии. – Сборник научных трудов «Идентификация образов» / ИТК НАН Беларуси. – Минск, 2001. – С. 45-52.
10. Petrovsky A.I. Run-time reconfigurable processor using DAT-based approach // Специальный выпуск журнала «Известия белорусской инженерной академии». – Минск, 2001. – №1(11)/2. – С. 29-31.
11. Петровский Ал.А. Процессор на базе ЦПОС/ПЛИС для кодирования аудиосигналов в реальном времени // Научно-технический журнал «Chip News» – Инженерная микроэлектроника. – Москва, 2002. – №5(68). – С. 32-34.
12. Petrovsky A.I. A., Petrovsky A. Dynamic algorithm transforms for reconfigurable real-time audio coding processor // Proc. “Parallel computing in electrical engineering” – IEEE Computer Soc. Press, NJ, 2002. – P. 231-234.

Материалы научно-технических конференций

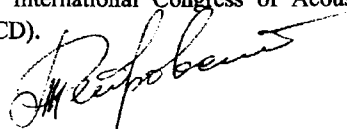
13. Petrovsky A.I.A., Schkredov S.L. The role of Distributed Arithmetic in FPGA-based Signal Processing Embedded System // The Proc. of 2nd International conference of “Digital Signal Processing in Communication”. – Belarussian Polytechnics Academy, Minsk, Belarus, 14-16 June, 1999. – P. 68-71.
14. Petrovsky A.I.A. Multi-DSP based hardware platform for wavelets and filter banks // The Proc. of the International Workshops on System, Signal and Image Processing (IWSSIP'99). – Bratislava, Slovakia 2-4 June 1999. P. 111-114
15. Bielawski K., Petrovsky A.I.A. Dynamic non-uniform filter bank constructing algorithms for reconfigurable speech processing system based on the FPGA-device and TMS320C31 //

Proc. of the 3rd European Conference "DSP Education and Research". – Paris, France, 20-21 Sept. 2000. – 7p. – (CD)

16. Петровский Ал.А. Динамически реконфигурируемый процессор адаптивного пакетного вэйвлетного преобразования для систем кодирования аудио сигналов // Труды 3-й Международной конференции «Цифровая обработка сигналов и ее применение» (DSPA'2000). – Москва, 2000. – С. 231-236.
17. Audio coding with a masking threshold adapted wavelet packet based on run-time reconfigurable processor architecture / Al. Petrovsky, A.A. Petrovsky. – Amsterdam, Netherlands, May 2001. – 8 p. – (preprint / AES-110th Convention; № 5298).
18. Petrovsky Al. Masking threshold adapted wavelet packet (WP) tree with a time-invariant analysis wavelet // The Proc. of 6th International Conference "Pattern Recognition and Information Processing" (PRIP'2001). – Minsk, Belarus, 15-17 May, 2001. – P. 10-14.
19. Martin R., Petrovsky Al., Lotter T. Planar Superdirective Microphone Arrays for Speech Acquisition in the Car // Proc. of 7th European conference on speech communication and technology (EUROSPEECH 2001), vol. 4, 3-7 Sept., Aalborg, Denmark, 2001. – P. 2623-2626.
20. Petrovsky Al., Sadykhov R. and et.al. Dynamic algorithm transform of wavelet packet decomposition structure in audio coders // The Proc. of the 2nd International Conference on Neural Networks and Artificial Intelligence (ICNNAI'2001). – Minsk, Belarus, 2-5 Oct. 2001. – P. 124-129.
21. Petrovsky Al., Sadykhov R. and et.al. DISC (Dynamic Instructor Set Computer) architecture of wavelet packet processor for real-time audio signal compression systems // The Proc. of the 2nd International Conference on Neural Networks and Artificial Intelligence (ICNNAI'2001). – Minsk, Belarus, 2-5 Oct. 2001. – P. 130-135.
22. Петровский Ал.А. Процессор на базе ЦПОС/ПЛИС для кодирования аудио сигналов в реальном времени // Труды 4-й Международной конференции "Цифровая обработка сигналов и её применение (DSPA'2002)". – Москва, Россия, 2002. – т.2. – С. 453-456.
23. Petrovsky Al., Petrovsky A. Reconfigurable processing: the solution to audio coding in real-time // The Proc. of 9th International Conference "Mixed design of Integrated circuits and systems (MIXDES)". – Wroclaw, Poland, June, 2002. – P. 629-634.

Тезисы докладов научно-технических конференций

24. Петровский Ал.А. Программируемый процессор для обработки сигналов в реальном времени // Сборник докладов III Республиканской научной конференции студентов Белоруссии – БГУ, Минск, 14-16 Мая, 1997. – С. 190-191.
25. Петровский Ал.А. Компрессия аудио сигналов на базе адаптивного пакета дискретного вэйвлетного преобразования и психоакустической информации // Труды 1-й республиканской конференции «Цифровое телевидение и радиовещание». – Минск-Раубичи, 14-16 Апреля, 1999. – С. 84-87.
26. Petrovsky Al. Perceptually optimized time-varying wavelet packet decomposition and its applications in acoustic signal processing // 17th International Congress of Acoustics (ICA'2001). – Rome, Italy, 2-7 Sept. 2001 – 2p. – (CD).



дысертациі Пятроўскага Аляксея Аляксандравіча

**«Метады і алгарытмы аналізу і сінтэзу працэсараў з дынамічна
рэканфігуруемая архітэктурай перцэптуальнага аудыа кодэру»**

Ключавыя словы: вэйлет-пераўтварэнне, псіхаакустыка, аудыа кодар/декодар, паралельна-паточны працэсар, лічбовая апрацоўка сігналаў (ЛАС), масіў мікрафонаў.

Аб'ектам даследвання з'яўляюцца працэсары з рэканфігуруемай архітэктурай перцэптуальнага аудыа кодэру. Прадмет даследвання – метады кадзіравання аудыа сігналаў на падставе пакеці дыскрэтнага вэйлет-переутварэння (ПДВП) і дынамічнай трансфармацыі алгарытмаў пад дадзены вылічальны рэсурс – рэканфігуруемае вылічальнае асяроддзе.

Мэтай працы з'яўляецца распрацоўка метадаў і алгарытмаў аналізу і сінтэзу працэсараў з дынамічна рэканфігуруемай архітэктурай перцэптуальнага аудыа кодэру мабільных мультымедычных сістэм.

На падставе псіхаакустычнай мадэлі распрацаваны метады і алгарытмы аналізу і сінтэзу працэсараў з дынамічна рэканфігуруемай архітэктурай для рэалізацыі перцэптуальнага аудыа кодэру мультымедычных сістэм на падставе ПДВП, структура якога перабудоўваецца з дынамічнай трансфармацыяй частотна-часовага плану ПДВП, узгодненага з шкалай крытычных частот утрымання акустычнай інфармацыі чалавекам, пад дадзены вылічальны рэсурс і мультыфрэймавую распрацоўку аудыа сігналаў у рэальным маштабу часа (рэканфігуруемае вылічальнае асяроддзе). Вынікам працы алгарытма дынамічнай трансфармацыі структуры дрэва ПДВП з'яўляецца аптымальная структура дрэва ПДВП, узгодненага з бягучым фрэймам сігналу (мінімум каштоўнай функцыі – часовая энтрапія вэйлет-каэфіцыентаў), якая ў сваю чаргу з'яўляецца дынамічным раскладам рэканфігурацыі архітэктурны паралельна-паточнага працэсара ПДВП. Разпрацавана навейшая структура ПДВП-кодэру/декодэру аудыа сігналаў на падставе метаду дынамічнай дэкампазіцыі ПДВП, якая, у адрозненне ад вядомых, арыентавана на распрацоўку сігналу ў вэйлет-вобласці і працу ў рэальным часе. Правіла аптымальнага размеркавання бітаў пры квантаванні вэйлет-каэфіцыентаў у структуры кодэру, якая ўлічвае эфекты рэканструкцыі аудыа сігналу ў декодэру – “прасачыванне” энергіі шуму квантавання ў суседнія палосы сінтэзіруючага банку фільтраў. Метады аб'ектыўнай ацэнкі якасці рэканструіраванага сігналу перцэптуальнага ПДВП-кодэру аудыа сігналаў на падставе перыферычнай мадэлі вуша даюць высокую ступень карэляванасці вынікаў аб'ектыўных тэстаў з суб'ектыўнымі, ацэнка адносін шуму да парога маскавання ў сігнале, які не змяшчае артэфактаў, каля -9 дБ.

Выканана апаратная рэалізацыя эксперыментальнага модуля аудыа кодеру з падсістэмай увада аудыа сігналу (супернапраўлены мікрафон) мабільнай мультымедычнай сістэмы на падставе дзвух рэканфігуруемых ПДВП-працэсараў (кодар/декодар) і ЛАС-працэсару сістэмы маніторынгу.

РЕЗЮМЕ

диссертации Петровского Алексея Александровича

«Методы и алгоритмы анализа и синтеза процессоров с динамически реконфигурируемой архитектурой перцептуального аудио кодера»

Ключевые слова: вэйвлет-преобразование, психоакустика, аудио кодер/декодер, параллельно-поточный процессор, цифровая обработка сигналов, массив микрофонов.

Объектом исследования являются процессоры с реконфигурируемой архитектурой перцептуальных аудио кодеров. Предмет исследования – методы кодирования аудио сигналов на основе пакета дискретного вэйвлет-преобразования (ПДВП) и динамической трансформации алгоритмов под заданный вычислительный ресурс – реконфигурируемая вычислительная среда

Целью работы является разработка методов и алгоритмов анализа и синтеза процессоров с динамически реконфигурируемой архитектурой перцептуального аудио кодера мобильных мультимедийных систем.

На основе психоакустической модели разработаны методы и алгоритмы анализа и синтеза процессоров с динамически реконфигурируемой архитектурой для реализации перцептуальных аудио кодеров мультимедийных систем на базе ПДВП, структура которого перестраивается с динамической трансформацией карты частотно-временного разрешения ПДВП, согласованной со шкалой критических частот восприятия акустической информации человеком, под заданный вычислительный ресурс и мультифреймовую обработку аудио сигналов в реальном масштабе времени (реконфигурируемая вычислительная среда). Результатом работы алгоритма динамической трансформации структуры дерева ПДВП является оптимальная структура дерева ПДВП, согласованная с текущим фреймом сигнала (минимум стоимостной функции – временная энтропия вэйвлет-коэффициентов), которая в свою очередь является динамическим расписанием реконфигурации архитектуры параллельно-поточного процессора ПДВП. Разработана новая структура ПДВП-кодера/декодера аудио сигналов на основе метода динамической декомпозиции ПДВП, которая, в отличие от известных, ориентированна на обработку сигнала в вэйвлет-области и работу в реальном масштабе времени. Правило оптимального распределения бит при квантовании вэйвлет-коэффициентов в структуре кодера учитывает эффекты реконструкции аудио сигнала в декодере – «просачивание» энергии шума квантования в смежные полосы синтезирующего банка фильтров. Методы объективной оценки качества реконструированного сигнала перцептуального ПДВП-кодера аудио сигналов на основе периферийной модели уха, дают высокую степень коррелированности результатов объективных тестов с субъективными, оценка отношения шума к порогу маскирования в сигнале, не содержащем артефактов, около -9 дБ.

Выполнена аппаратная реализация экспериментального модуля аудио кодера с подсистемой ввода аудио сигнала (супернаправленный микрофон) мобильной мультимедийной системы на основе двух реконфигурируемых ПДВП-процессоров (кодер/декодер) и ЦОС-процессоре системы мониторинга.

SUMMARY

of the dissertation thesis **“Methods and algorithms of analysis and synthesis of perceptual audio coder processors with dynamic reconfigurable architecture”** by Alexey Petrovsky

Keywords: wavelets transform, psychoacoustics, audio coder/decoder, parallel-pipeline processor, digital signal processing, and microphone array.

Processor with dynamic reconfigurable architecture of perceptual audio coder is the object under investigation. The subject for investigation is methods of audio signal coding based on the wavelet packet (WP) and dynamic algorithms transformation (DAT) under the given computation resource – reconfigurable computation environment.

The objective of this work is to design the methods and algorithms of analysis and synthesis of perceptual audio coder processors with dynamic reconfigurable architecture for mobile multimedia systems.

On the base of psychoacoustical model the methods and algorithms of analysis and synthesis of processors with dynamic reconfigurable architecture for realization the perceptual audio coder in mobile multimedia systems based on WP, which tree structure dynamically transformed the time-frequency resolution, under the given computation resource and multi-frame audio signals processing in real time (reconfigurable computation environment) have been developed. The algorithm result of WP tree structure dynamic decomposition is optimal WP tree structure coherented with the current signal frame (minimal cost function – wavelet time entropy) which is also represented as a dynamic reconfiguration schedule of the parallel-pipeline WP processor architecture. A novel WP audio coder/decoder structure based on methods of WP dynamic decomposition, as distinct from known, has been oriented under the signal processing in wavelets domain and work in real time. The optimal bit allocation rule of wavelet coefficients allows the reconstruction effects of audio signal in decoder – “leakage” the quantizer noise energy into the neighboring subbands of the synthesis filter bank. The methods of the objective audio signal quality measurement based on the model peripheral ear have ensure the high order correlation between objective test results with subjective tests and the noise-to-masking estimation for the signal without hearable artifacts is equal -9 dB.

The hardware realization of the experimental module of audio coder together with input audio signal subsystem (superdirective microphone array) for mobile multimedia system based on two reconfigurable WP processors (coder/decoder) and DSP-processor of monitoring systems has been performed.

ПЕТРОВСКИЙ Алексей Александрович

**МЕТОДЫ И АЛГОРИТМЫ АНАЛИЗА И СИНТЕЗА ПРОЦЕССОРОВ С
ДИНАМИЧЕСКИ РЕКОНФИГУРИРУЕМОЙ АРХИТЕКТУРОЙ
ПЕРЦЕПТУАЛЬНОГО АУДИО КОДЕРА**

Специальность 05.13.15 – Вычислительные машины и системы

АВТОРЕФЕРАТ

диссертации на соискание ученой степени
кандидата технических наук

Подписано в печать 18.09.2002.

Формат 60×84 1/16.

Бумага офсетная.

Печать ризографическая.

Усл.печ.л. 1,63.

Уч.-изд.л. 1,7.

Тираж 90 экз.

Зак. 573.

Издатель и полиграфическое исполнение:

Учреждение образования «Белорусский государственный университет
информатики и радиоэлектроники».

Лицензия ЛП № 156 от 05.02.2001.

Лицензия ЛВ № 509 от 03.08.2001.

220013, г.Минск, ул. П.Бровки, 6.