

Belt_seq[3]	336	500	190,47	ПЛИС Spartan xc3s200
Belt_par[3]	211	500	289,59	ПЛИС Spartan xc3s200
Программная[5]	682	1000	187,68	Процессор Pentium III
Последовательная[4]	66	112,4	217,98	ПЛИС Virtex xc6vlx130t
Конвейерная[4]	114(1)	217,4	243,65(27 827)	ПЛИС Virtex xc6vlx130t
Уменьшенная конвейерная[4]	114(14)	217,4	243,65(1 987,6)	ПЛИС Virtex xc6vlx130t
С параллелизмом на уровне такта	225	454.54	258,58	ПЛИС Virtex xc6vlx130t

Примечание – Для конвейерных реализаций в скобках указана производительность после заполнения конвейера, а также количество тактов необходимых для формирования каждого следующего результата.

Таблица 2. Аппаратные затраты при размещении на кристалле FPGA

Варианты реализации	Slices	Триггеров	LUTs	RAMs	Устройство
Belt_seq[3]	750	649	1392	9	ПЛИС Spartan xc3s200
Belt_par[3]	1070	302	2050	28	ПЛИС Spartan xc3s200
Последовательная[4]	423	847	1173	-	ПЛИС Virtex xc6vlx130t
Конвейерная[4]	9157	10209	14816	112	ПЛИС Virtex xc6vlx130t
Уменьшенная конвейерная[4]	4829	5598	7948	28	ПЛИС Virtex xc6vlx130t
С параллелизмом на уровне такта	399	1362	937	2	ПЛИС Virtex xc6vlx130t

Список использованных источников:

1. СТБ 34.101.31-2007. Государственный стандарт Республики Беларусь. Криптографические алгоритмы шифрования и контроля целостности.
2. Бибило, П. Н. VHDL. Эффективное использование при проектировании цифровых систем / П. Н. Бибило, Н. А. Авдеев. – М.: СОЛОН-Пресс, 2006. – с. 344
3. Поляков, А. С. Характеристики аппаратной реализации некоторых симметричных алгоритмов шифрования / А. С. Поляков, В. Е. Самсонов // Информатика : ежеквартальный научный журнал, 2011. – №1.
4. Ланкевич, Ю. Ю. Процессор алгоритма шифрования «Belt» на базе ПЛИС / Ю. Ю. Ланкевич // Информационные технологии и системы 2013 (ИТС 2013) : материалы международной научной конференции, БГУИР, Минск, Беларусь, 23 октября 2013 г. / редкол.: Л. Ю. Шилин [и др.]. – Минск : БГУИР, 2013. – с. 190-191.
5. Агиевич, С. В. Алгоритм блочного шифрования BelT / С. В. Агиевич, В. А. Галинский, Ю. С. Харин, Н. Д. Микулич. Управление защитой информации, т.6, №4, 2002. – с.407–412

СПОСОБ ФОРМИРОВАНИЯ ВЫСОКОЧАСТОТНОЙ СОСТАВЛЯЮЩЕЙ СПЕКТРА СИГНАЛА В НИЗКОСКОРОСТНОМ АУДИОКОДЕРЕ

*Белорусский государственный университет информатики и радиоэлектроники
г. Минск, Республика Беларусь*

Герасимович В. Ю.

Петровский А. А. – д-р. техн. наук, профессор

Работа посвящена описанию варианта повышения качества реконструированного сигнала аудиокодера [1] при низких скоростях битового потока. Рассматривается способ формирования высокочастотной составляющей спектра выходного аудиосигнала путем копирования и масштабирования информации среднечастотной части спектра сигнала.

Принцип сжатия аудиоданных с потерями состоит в выделении из входной последовательности данных, которые позволяют компактно представить и синтезировать на стороне декодера выходной сигнал. Основная задача кодера заключается в поиске наиболее важных для восприятия человеком параметров. Так как количество выбираемых данных ограничено скоростью битового потока (битрейта) аудиокодера, часть информации о входном сигнале теряется. В силу особенностей восприятия человеком аудиосигнала, высокочастотные компоненты спектра менее perceptually важны, нежели низко- и среднечастотные [2]. Поэтому, при низких скоростях битового потока (т.е. малом количестве бит, выделяемых для передачи сигнала), параметры, соответствующие данной полосе могут быть проигнорированы и опущены, что повлечет за собой уменьшение качества реконструированного сигнала. Данный материал посвящен вопросу формирования высокочастотной составляющей спектра в аудиокодеке на основе согласованной подгонки

(СП)[1] при низких скоростях битового потока.

Алгоритм СП представляет любой сигнал $x(t)$ в виде линейной комбинации частотно-временных функций (называемых атомами) $g_{\gamma_n}(t)$, выбираемых из избыточного словаря D . В разрабатываемом кодере словарь атомов строится для каждого фрейма входного сигнала из самого сигнала с помощью пакета дискретного вейвлет преобразования (ПДВП). Индексы вектора $g \in D$, $l, n \in E$ – узлы дерева ПДВП (где l – номер уровня дерева, n – номер узла). Каждый узел дерева декомпозиции соответствует определенной частотной полосе. Алгоритм СП, в данном случае, оптимизирован таким образом, что отбираются только перцептуально важные атомы из словаря (на основе информации о порогах маскирования). Остановка выбора параметров осуществляется по достижению заданного количества, которое определяется целевым битрейтом.

Суть проблемы заключается в том, что при низких скоростях битового потока, ресурса атомов не хватает на выбор параметров из полос, эквивалентных высоким частотам. На рисунке 1 показан пример для 200 атомов отбираемых алгоритмом СП.

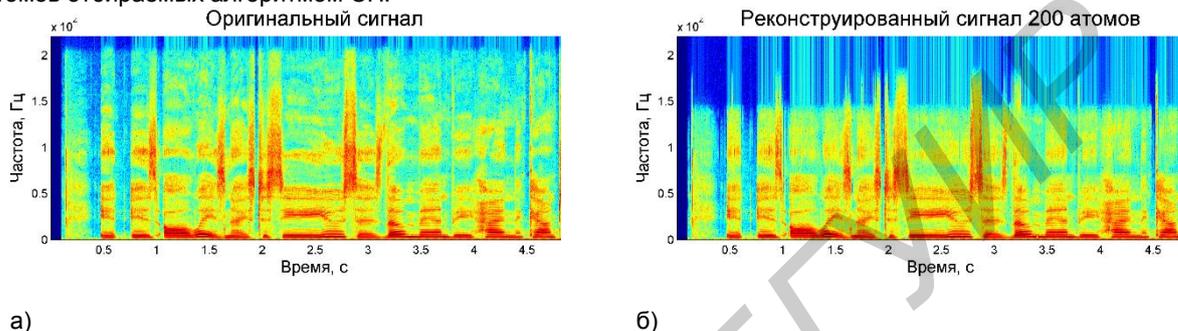


Рис. 1. Спектрограммы: а) – исходного сигнала, б) – реконструированного 200-ми атомами

Вариантом решения является перенос параметров из предыдущих каналов, в которых есть отобранные атомы, и их масштабирование к целевой энергии канала. При этом, вместо того, что бы передавать набор параметров для канала, декодеру для синтеза сигнала необходима только информация об энергии соответствующей частотной полосы. Схожий в своей идее метод применяется в [3], однако в рассматриваемом случае он применен в вейвлет области и масштабирование производится только на энергию канала, а не по спектральной огибающей. Первоисточником идеи является [4], однако применяемый в данной работе метод менее эффективен, нежели *SpectralBandReplication*, но, вместе с тем, менее ресурсоемкий, что критично при работе в реальном масштабе времени.

Для того, что бы оценить эффективность описываемого метода использовалась объективная оценка качества реконструированного сигнала *PEMO-Q*[5]. Данная метрика оценивает перцептуальное подобие сигналов (*PerceptualSimilarityMeasure* – *PSM*). Для того, что бы соотнести данную объективную оценку качества сигнала с субъективной оценкой *SDG* (*Subjective Difference Grade*), *PSM* можно отобразить в шкалу объективного различия (*Objective Difference Grade* – *ODG*), которая коррелирует с *SDG* и определяется следующим образом: 0.0 – не воспринимаемые искажения, -1.0 – воспринимаемые, но не раздражающие, -2.0 – немного раздражающие искажения, -3.0 – раздражающие, -4.0 – очень раздражающие искажения. Для эксперимента было взято три звуковых образца: вокал (*Suzan Vega, es01*), речь на немецком языке (*es02*), речь на английском языке (*es03*). Образцы были закодированы с использованием 200 – 400 атомов, с шагом в 50 атомов, что приблизительно соответствует скорости битового потока от 36,4 до 70,8 кбит/с с шагом в 8,6 кбит/с. Результаты приведены на рисунке 2.

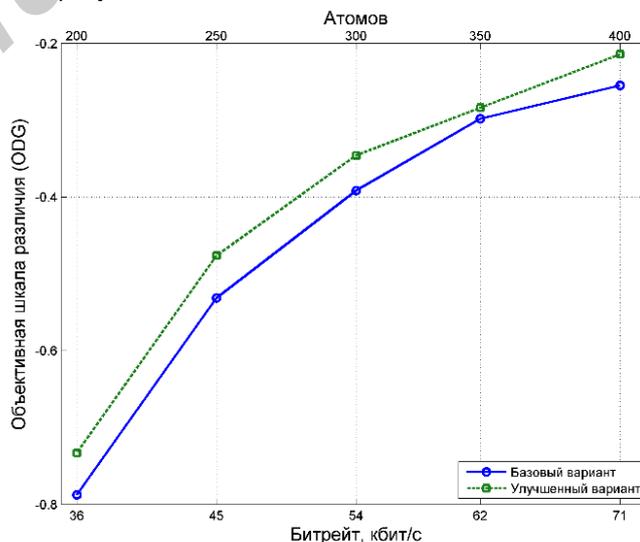


Рис. 2. Экспериментальные результаты

Как видно из экспериментальных результатов, приведенный способ формирования высокочастотной составляющей спектра сигнала даёт прирост качества восстановленного сигнала. При этом, в качестве дополнительных данных, которые необходимо передать декодеру для реконструкции выступает энергия в целевой частотной полосе, т.е. один параметр на полосу. В проведенных экспериментах перенос осуществлялся в три полосы, следовательно на каждый фрейм требовалось только три дополнительных параметра, что незначительно увеличит битрейт, при этом позволит получить более высокое качество выходного сигнала.

Список использованных источников:

1. Petrovsky, Al. Scalable parametric audio coder using sparse approximation with frame-to-frame perceptually optimized wavelet packet based dictionary / Al. Petrovsky, V. Herasimovich, A. Petrovsky // AES 138th Convention, paper 9264 – Warsaw, Poland, 2015, May 7 - 10.
2. Zwicker, E. Psychoacoustics: facts and models / E. Zwicker, H. Fastl. – Berlin: Springer-Verlag, 1990. – 354 p.
3. Valin, J.-M. High-Quality, Low-Delay Music Coding in the Opus Codec / J.-M. Valin, G. Maxwell, T. Terriberry, K. Vos // AES 135th Convention, paper 8942 – New York, USA, 2013, October 17 - 20.
4. Dietz, M. Spectral Band Replication, a novel approach in audio coding / M. Dietz, L. Liljeryd, K. Kjörling, O. Kunz // AES 112th Convention, paper 5553 – Munich, Germany, 2002, May 10 - 13.
5. Huber, R. PEMO-Q – A New Method for Objective Audio Quality Assessment Using a Model of Auditory Perception / R. Huber, B. Kollmeier // IEEE Transactions on audio, speech, and language processing, vol. 14. – November, 2006. – pp. 1902-1911.

СИСТЕМА ИДЕНТИФИКАЦИИ ДИКТОРА В АКУСТИЧЕСКИХ ШУМАХ НА ОСНОВЕ АНТРОМОРФИЧЕСКОЙ ОБРАБОТКИ РЕЧЕВОГО СИГНАЛА

Белорусский государственный университет информатики и радиоэлектроники
г. Минск, Республика Беларусь

Кручок Д. Н.

Петровский А. А. – д-р. техн. наук, профессор

Рассматривается система идентификации диктора в акустических шумах с использованием антропоморфической обработки речевого сигнала. Полученный характеристический вектор на базе данного преобразования используется в качестве признаков для системы идентификации диктора. В качестве решающих правил применяются нейронные сети прямого распространения. Приводятся результаты распознавания разработанной системы идентификации диктора.

В работе [1] был предложен алгоритм получения характеристического вектора, который пытается смоделировать процессы обработки речевого сигнала человеческим ухом. Рассматриваемый алгоритм состоит из этапов, представленных на рис. 1, а и частично моделирует процесс обработки звукового сигнала слуховой системой человека.



Рис. 1. – Сравнение этапов обработки сигнала: а – схема рассматриваемого алгоритма экстракции; б – процесс обработки звукового сигнала в слуховой системе человека

Полученные кохлеарные кепстральные коэффициенты (ККК) используются в системе идентификации в качестве признаков. В качестве решающих правил системы идентификации диктора наиболее часто используются следующие методы: векторное квантование, гауссовские смеси, нейронные сети и метод опорных векторов. В данном исследовании были выбраны нейронные сети прямого распространения для получения сравнительных результатов распознавания диктора с результатами, представленными в работе [1], в которой были использованы модели гауссовых смесей. Для экспериментальной оценки системы идентификации диктора в акустических шумах, был спроектирован идентификатор дикторов на основе нейронных сетей прямого распространения. Модель системы распознавания представлена на рис. 2 и имеет два режима работы: обучение и идентификация. Архитектура используемой сети выбиралась постепенно от