

ИССЛЕДОВАНИЕ НИЗКОСКОРОСТНЫХ АУДИО КОДЕРОВ НА ОСНОВЕ ГИБРИДНОЙ МОДЕЛИ ОПИСАНИЯ СИГНАЛА

Белорусский государственный университет информатики и радиоэлектроники

Герасимович В.Ю.

Петровский А.А. – д.т.н., профессор

Аннотация – В рамках доклада рассматриваются алгоритмы аудио кодирования на основе гибридной (синусоиды – шум – транзиенты) модели описания сигнала. Проводиться анализ существующих подходов к решению данного вопроса, и предлагается разрабатываемая модель аудио кодера.

Гибридная модель описания сигнала, представленная в [1], позволяет построить на ее основе масштабируемый аудио кодер. Разделение сигнала на три составляющие (синусоидальную, переходную и шумовую) позволяет найти и применить индивидуальную схему сжатия к каждой из них. Общая схема аудио кодера на базе гибридной модели описания сигнала представлена на рисунке 1.

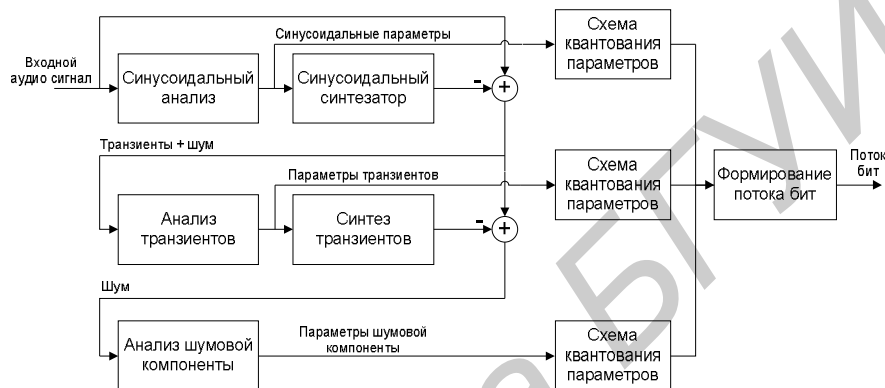


Рисунок 1 – Общая схема аудио кодера на базе гибридной модели описания сигнала

Как видно на приведенной схеме, первым этапом работы кодера является анализ входного аудио сигнала и выделение синусоидальных параметров из него. В качестве анализатора была взята модель получения мгновенных параметров сигнала описанная в [2]. Следующий шаг – сепарация сигнала на синусоидальную составляющую и остаток. На этом этапе проводился поиск энергетических пиков и выделение соответствующих параметров в качестве синусоидальных. На этап анализа переходных компонентов поступает сигнал, являющийся результатом вычитания входного сигнала и синтезированной синусоидальной составляющей. Анализ транзитных компонентов описан в [1]. Последним звеном параметризации сигнала является анализ шумовой компоненты (финального остатка входного сигнала). Шумовая составляющая моделировалась с помощью линейного предсказания, то есть, происходит поиск спектральной огибающей.

Параметрами синусоидальной составляющей сигнала служат амплитуды, частоты и фазы. В низкоскоростных аудио кодерах информация о фазе сигнала не передается, а вычисляется декодером. Амплитуды, в предлагаемой модели, квантуются с помощью векторного квантования с расщеплением. В экспериментах было использовано пять кодовых книг (КК) по 1024 уровня в каждой, однако можно учесть особенности восприятия аудио сигнала и варьировать количество уровней КК. В силу специфики используемой для анализа модели, для передачи сигнала на скорости менее 10 кбит/с есть возможность зафиксировать информацию о частоте на стороне декодера.

Параметрами, описывающими переходную составляющую, служат атомы [1]. Они представляют собой набор из позиции каждого атома и его вес. В силу того, что значение позиции атома должно передаваться в декодер без потерь, эти значения не сжимаются и напрямую передаются. Значение веса, в свою очередь, квантуется скалярно и, затем, передается на сторону декодера.

Шумовая составляющая параметризуется с помощью алгоритма линейного предсказания. Выходные коэффициенты фильтра предсказателя преобразуются в LSF , которые затем квантуются с помощью векторного квантования. Синтезируется данная компонента фильтрацией белого шума фильтром предсказателем с вычисленными ранее коэффициентами.

Для улучшения качества восстановленного сигнала в разрабатываемой модели кодера проводятся эксперименты с использованием в схеме кодирования психоакустической модели.

Список источников:

1. Petrovsky A.I., Azarov E. and Petrovsky A., Hybrid signal decomposition based on instantaneous harmonic parameters and perceptually motivated wavelet packets for scalable audio coding, Elsevier, Signal Processing, Vol. 91, Issue 6, Fourier Related Transforms for Non-Stationary Signals, pp. 1489-1504, June 2011.
2. Azarov E., Vashkevich M. and Petrovsky A., Instantaneous pitch estimation based on RAPT framework, Proc. the 20th European Signal Processing Conference (EUSIPCO'2012), pp. 2787-2791, August 2012.
3. Verma T.S., Meng T.H.Y., A 6 kbps to 85 kbps scalable audio coder, Proc. ICASSP, vol. 2, pp. 877-880, 2000.