

ОПТИМИЗАЦИЯ ФУНКЦИИ ОКНА БАНКА ФИЛЬТРОВ ДЛЯ КОДИРОВАНИЯ АУДИОСИГНАЛОВ

Осинов К.С.

Введение

В настоящее время наиболее перспективными являются такие направления повышения эффективности сжатия аудио информации как построение банков фильтров для преобразования временного сигнала в частотную область, выбор хорошей психоакустической модели для определения оптимального распределения битовых ресурсов, а также поиск эффективных методов квантования и сжимающего кодирования передаваемых данных. В данной работе рассматривается проблема оптимизации банка фильтров.

Банк фильтров служит для преобразования сигнала из временной формы в частотную. Он может быть интерпретирован как совокупность полосовых фильтров, полосы пропускания которых практически не перекрываются и покрывают весь доступный частотный диапазон сигнала. Банк фильтров, по сути, делит спектральную область на частотные поддиапазоны и формирует последовательные блоки коэффициентов разложения, описывающих последовательные фрагменты (кадры) сигнала во временной области. Чтобы избежать искажений на стыке кадров, кадры выбирают перекрывающимися во временной области. Тем не менее, известны структуры банков фильтров [1], позволяющие обеспечивать практически точное восстановление сигнала при числе спектральных компонент равном числу отсчетов во временной области. К числу таких преобразования относятся, например, модифицированное косинусное преобразование (МДКП), ортогональные преобразования с перекрытиями (ОПП) [2] и преобразования на основе квадратичных зеркальных фильтров (КЗФ) [1]. Банки на основе ОПП применяются, в частности, в современном стандарте сжатия аудио сигналов AAC [3], а банки на основе КЗФ в сочетании с МДКП использованы в стандарте MPEG, Layer 3, или MP3 [4].

В настоящей работе мы рассматриваем каскадный банк фильтров, построенный как последовательное применение ОПП и МДКП аналогично тому, как в аудиокодере MP3, где на первом этапе используется КЗФ, а на втором – МДКП. Рассмотрим подробно первую ступень преобразования.

Формула для получения коэффициентов $S_k^{(t)}$ преобразования ОПП выглядит как [2]

$$S_k^{(t)} = \sqrt{\frac{2}{K}} \sum_{n=0}^{N-1} w_n x_n^{(t)} e^{\frac{\pi}{K} (k+\frac{1}{2})(n+\frac{K+1}{2})}, \quad k = 0..K-1 \quad (1)$$

где t – номер кадра, K – размер спектра на выходе, (w_0, \dots, w_{N-1}) – функция окна, $(x_0^{(t)}, \dots, x_{N-1}^{(t)})$ – отсчеты

Рассматривается задача оптимизации банка фильтров на основе ортогонального преобразования с многократными перекрытиями для аудиокодера. Предлагается использовать комплексный критерий качества и итеративный алгоритм для нелинейной оптимизации параметров фильтра-прототипа. Полученный банк фильтров выигрывает, по сравнению с известными банками фильтров, как по точности восстановления, так и по качеству разделения спектральных компонент.

входного сигнала, N – длина входного кадра. Кратность перекрытия L равна $L = N / K$.

Основными свойствами фильтра, характеризующими эффективность его применения к задаче аудио кодирования, являются:

– точность восстановления, которая вычисляется как

$$E_{pr} = \sum_{s=0}^{K-1} \sum_{n=0}^{M/2-1} \left[\sum_{i=0}^{2K-2s-1} w(n+iM) w(n+iM+2sM) - \delta(s) \right]^2 \quad (2)$$

– энергия сигнала в полосе за пределами полосы пропускания

$$E_s = \frac{1}{\pi} \int_{\omega_s}^{\pi} |H(e^{j\omega})|^2 d\omega. \quad (3)$$

Первый параметр, очевидно, важен для того, чтобы в отсутствие квантования спектральных коэффициентов точность прямого и обратного преобразования была достаточно высокой для исключения слышимых искажений.

Второй параметр говорит о качестве разделения соседних спектральных компонент и существенным образом влияет на битовую скорость, необходимую для кодирования сигнала с заданным уровнем искажений. Задача оптимизации банка фильтров состоит в уменьшении ширины переходной полосы и энергии полосы не пропускания при ограничении на точность восстановления сигнала. Рассматриваемый банк фильтров представляет собой набор идентичных фильтров, представляющих собой частотные сдвиги так называемого фильтра-прототипа. Характеристики фильтра-прототипа, в свою очередь, полностью определяются видом функции окна. Поэтому задача оптимизации банка фильтров сводится к оптимизации функции окна. В общем случае, в силу симметрии, число параметров, по которым производится оптимизация, очень велико (оно равно половине длины кадра). В связи с этим непосредственное применение стандартных методов оптимизации невозможно.

В следующем параграфе мы рассмотрим известные методы построения окон преобразования и опишем новый подход.

Подходы к построению фильтра-прототипа

Метод Кайзера

В работе [5] предлагается выбирать окно преобразования с большим перекрытием как покомпонентное произведение функции Кайзера на функцию $\sin x / x$, что даёт хорошую аппроксимацию идеального НЧ-фильтра. Получаемая функция окна имеет вид

$$h(x) = \frac{f}{\sqrt{M}} S\left(\frac{x}{2KM}, \alpha\right) \frac{\sin\left[\frac{f\pi(x-KM)}{M}\right]}{\left[\frac{f\pi(x-KM)}{M}\right]}, \quad (4)$$

где

$$S(x, \alpha) = \frac{I_0(\alpha\sqrt{1-x^2})}{I_0(\alpha)}, |x| \leq 1 \quad (5)$$

– функция окна Кайзера, а $I_0(x)$ – модифицированная функция Бесселя нулевого порядка.

Как следует из формулы, функция окна Кайзера определяется двумя параметрами: α и f . Задача оптимизации сводится к поиску этих параметров, минимизирующих функцию точности восстановления (2). Пример окна длины 512, построенного данным методом, приведен на рис. 1.

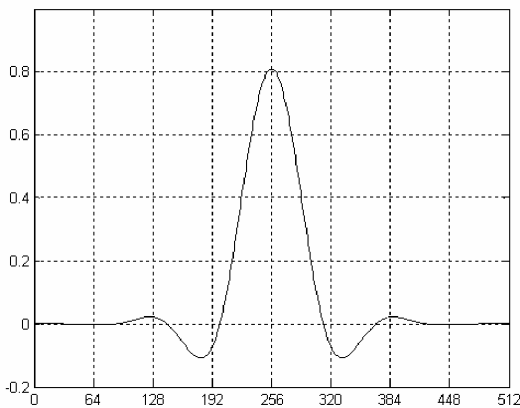


Рис. 1. Пример окна Кайзера ($K=512, L=16$)

На сегодняшний день этот метод представляется наиболее практичным, как с точки зрения сложности выполнения оптимизации, так и по характеристикам получаемых фильтров.

Оптимизация с использованием метода углов поворота (butterfly angles)

Данный метод описан Г. Малваром в книге [1], [6].

На первом шаге строится окно, представляющее собой сумму косинусов, задаваемую формулой

$$w(n) = \sum_{i=0}^{K-1} a_i \cos\left[f_i\left(n - KM + \frac{1}{2}\right)\right], n = 0, 1, \dots, 2KM - 1 \quad (6)$$

где a_i – амплитуды а f_i – частоты соответствующих косинусов.

В качестве начальных значений a_i и f_i выступают произвольные, например, случайные числа. Затем при помощи алгоритма безусловной нелинейной оптимизации минимизируется параметр точности восстановления. По построенному таким образом окну вычисляется набор так называемых углов поворота.

Под углами поворота понимаются значения на диагоналях матриц вида

$$\begin{aligned} C_k &= \text{diag}\{\cos \theta_{0k}, \cos \theta_{1k}, \dots, \cos \theta_{M/2-1,k}\}; \\ S_k &= \text{diag}\{\sin \theta_{0k}, \sin \theta_{1k}, \dots, \sin \theta_{M/2-1,k}\}, \\ D_k &= \begin{bmatrix} -C_k & S_k J \\ JS_k & JC_k J \end{bmatrix}, \end{aligned} \quad (7)$$

где матрица D_k используется для вычисления соответствующей степени преобразования ОПП (см. рис. 2).

Наконец, углы поворота θ_{ik} подбираются так, чтобы минимизировать E_s . Отметим, что после перехода от представления окна в виде (6) к представлению в виде набора углов поворота (7) окно приобретает свойство идеального восстановления, и в ходе дальнейшей оптимизации это свойство сохраняется.

Таким образом, построение фильтра при помощи данного метода всегда даёт фильтр с идеальным восстановлением. Поэтому как критерий для качества фильтра при построении используется только энергия сигнала в полосе непропускания. Начальные условия для построения случайны, а найденные оптимумы – локальны, поэтому для получения «хорошего» фильтра требуются многократные итерации алгоритма с различными начальными условиями. К сожалению, для фильтров с большой степенью перекрытия и большим количеством каналов получить хорошие фильтрующие свойства не удаётся, поэтому применительно к задаче кодирования аудиосигналов этот метод проигрывает методу Кайзера.

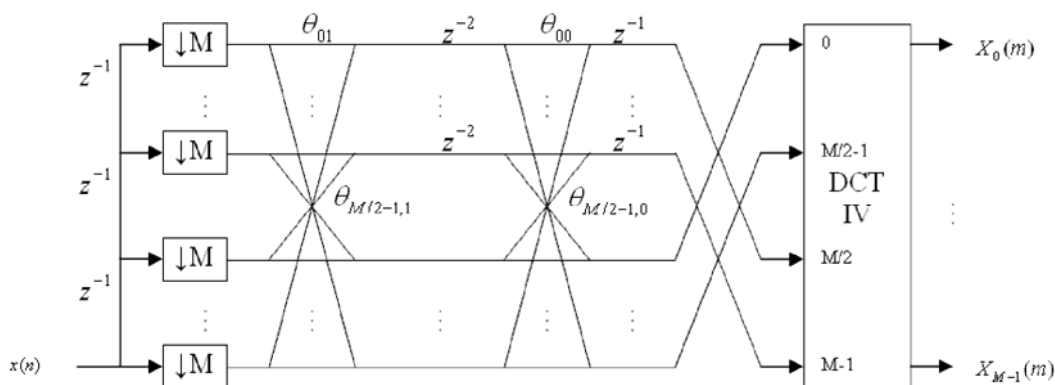


Рис. 2. Прямое преобразование ОПП, $K=2$

Предлагаемый алгоритм построения и оптимизации фильтра-прототипа

Предлагается отказаться от свойства идеального восстановления, а для построения фильтра использовать комбинацию двух описанных выше методов. Метод Кайзера (4) используется для получения «хорошего» начального фильтра, а при помощи суммы косинусов выбирается аддитивная добавка аналогичная (6), улучшающая качества построенного фильтра. В итоге формула окна принимает вид

$$h(x) = \frac{f}{\sqrt{M}} S\left(\frac{x}{2KM}, \alpha\right) \frac{\sin\left[\frac{f\pi(x-KM)}{M}\right]}{\left[\frac{f\pi(x-KM)}{M}\right]} + \sum_{i=0}^{K-1} a_i \cos\left[f_i\left(n-KM + \frac{1}{2}\right)\right], n = 0, 1, \dots, 2KM - 1. \quad (8)$$

Параметры α_i и f_i подбираются при помощи процедуры безусловной нелинейной оптимизации (алгоритм Нелдера-Мида), минимизирующей параметр точности E_{pr} с условием сохранения параметра фильтрации E_s . При поиске оптимальных значений параметров использовалась стандартная библиотека пакета Matlab 7.0.

Пример оптимизации

В качестве примера рассмотрим преобразование ОПП с параметрами $K=512, L=16$. Частотная характеристика фильтра показана на рис. 3 вместе с частотными характеристиками фильтра Кайзера и МПЗ-фильтра. Как видно из рисунка, энергия сигнала за пределами полосы пропускания заметно меньше, чем для других фильтров. В результате, применение такого фильтра может существенно сказаться на эффективности кодирования.

На рис. 4 приведена полученная в результате оптимизации аддитивная добавка к окну Кайзера.

Численные характеристики рассмотренных в работе фильтров сведены в таблицу 1. Из представленных ре-

зультатов видно, что по сравнению с фильтром Кайзера имеется выигрыш по обоим параметрам – и по точности восстановления и по подавлению энергии за пределами полосы пропускания.

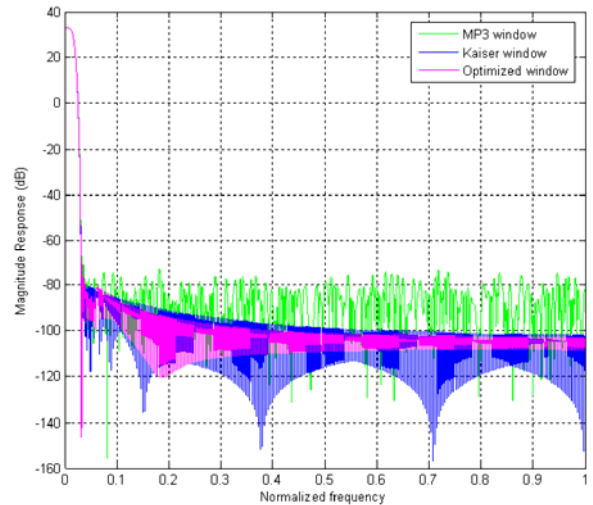


Рис. 3. Частотный отклик фильтров

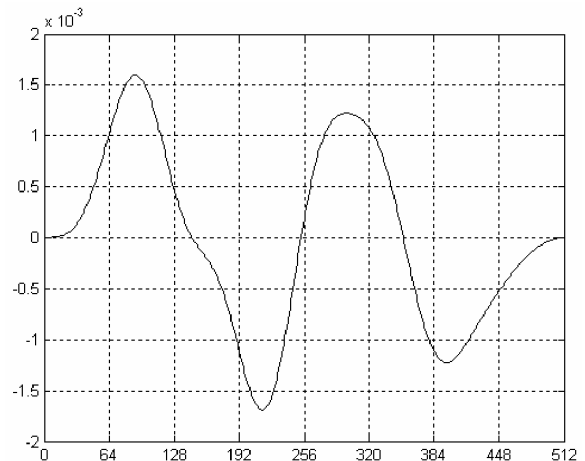


Рис. 4. Аддитивная оптимизирующая добавка

Таблица 1.

Точность восстановления и энергия полосы непропускания

	E_{pr}	E_s , дБ
Окно МПЗ	1.83 е-04	84.26
Метод Кайзера	1.272 е-05	96.84
Метод Малвара	0	41.1
Предложенный метод	1.872 е-06	98.8

Заключение

Предложен алгоритм оптимизации окна, полученного при помощи метода Кайзера. В примере достигается улучшение параметра восстановления почти на порядок, без ухудшения фильтрующих свойств. Достоинством метода является невысокая сложность алгоритма.

Литература

1. Henrique S. Malvar «Signal Processing with Lapped Transforms», Artech house 1991.
2. Henrique S. Malvar, David H. Staelin «The LOT: transform coding without blocking effects» IEEE Transactions on acoustics,

- speech, and signal processing, vol. 37, №4, April 1989.
3. R. Koenen «Overview of the MPEG-4 Standard» ISO/IEC JTC1/SC29/WG11 N2323, July 1998.
4. Стандарт IS11172-3 ISO/IEC JTC1/SC29/WG11 MPEG. Information Technology – Coding of Moving Pictures and Associated Audio for Digital Storage Media at up to About 1.5 Mbit/s, Part 3: Audio 1992. (MPEG-1)
5. P. P. Vaidyanathan «Multirate Systems and Filter Banks» Englewood Cliffs, NJ: Prentice-Hall Signal Processing Series, 1993.
6. Henrique S. Malvar «Extended Lapped Transforms: Properties, Applications, and Fast Algorithms» IEEE Transactions on signal processing, vol. 40, №11, November 1992.