Компрессор речевых сигналов для слуховых аппаратов на основе Банка кохлеарных фильтров

асп. Вашкевич М.И., проф. Петровский А.А.

Белорусский государственный университет информатики и радиоэлектроники
ул. П.Бровки, 6, БГУИР, каф. ЭВС, 220013, Минск Беларусь, e-mail: {vashkevich, palex}@bsuir.by.

Аннотация. Предлагается многополосный компрессор динамического диапазона речевых сигналов, ориентированный на применение в слуховых аппаратах. Для разделения сигнала на частотные составляющие используется банк кохлеарных фильтров. Основная идея использования кохлеарного банка фильтров заключается в том, что он вносит малую групповую задержку и позволяет учитывать особенности слуховой системы человека. Обе характеристики очень важны для применения компрессора в слуховом аппарате. Для уменьшения вычислительной сложности компрессора колеарный банк фильтров предлагается реализовывать в виде гребенки БИХ-фильтров, с регулируемой полосой пропускания и центральной частотой.

Введение. Компрессия динамического диапазона (КДД) используется для уменьшения разницы в уровнях компонент с большой и малой интенсивностью в аудио-сигнале. В цифровых слуховых аппаратах КДД применяется для отображения широкого динамического диапазона речевого сигнала в суженный динамический диапазон остаточного слуха [1]. Люди с потерей слуха, как правило, страдают от понижения порога слышимости, потери разборчивости и уменьшения динамического диапазона. Поэтому для компенсации потери слуха необходимо решить задачу частотно-зависимого усиления сигнала и компрессию динамического диапазона. Специфическим требованием для слуховых аппаратов является малая алгоритмическая задержка на обработку сигнала. Для решения указанных задач в работе предлагается компрессор речевых сигналов на основе кохлеарного банка фильтров. Подробно описывается только задача компрессии динамического диапазона, частотно-зависимое усиление сигнала в субполосох может быть легко интегрировано в предлагаемый метод обработки.

Компрессия динамического диапазона. В данном разделе рассматриваются основные концепции, используемые при КДД. Общая схема предлагаемого компрессора представлена на рисунке 1.



Рисунок 1 – Структура компрессора

Главной задачей компрессора является автоматический выбор коэффициента усиления субполосных сигналов в зависимости от уровня мощности сигнала в данной полосе. Субполосы с высокой мощностью ослабляются, а с малой мощностью усиливаются. Благодаря такой обработке тихие звуки становятся слышимыми, а громкие звуки не вызывают дискомфортных ощущений. КДД как правило имеет следующие параметры [1]:

* порог компрессии (CT);
* коэффициент сжатия (CR);
* время атаки и отпускания;
* коэффициент усиления слухового аппарата ($G\_{dB}$).

Порог компрессии (CT), измеряемый в децибелах, определяет точку изгиба характеристики вход/выход компрессора, после которой алгоритм КДД становится активным. Если уровень входного сигнала ниже CT, то выходной сигнал усиливается линейно. В случае, когда уровень выходного сигнала выше порога компрессии (CT) коэффициент усиления компрессора уменьшается. Параметр CR определяет степень сжатия динамического диапазона. Например, значение CR=5 (или 5:1) означает, что на каждые 5 дБ увеличения уровня входного сигнала, уровень выходной сигнал возрастет на 1 дБ. На рис. 2,а приводится характеристика вход/выход компрессора для параметров CR=2, CT=70 дБ и $G\_{dB}$=10 дБ. Данный график определяет соотношение между входным и выходным уровнями звукового давления (sound pressure level – SPL) в компрессоре.

SPL рассчитывается по следующей формуле [2]:

|  |  |
| --- | --- |
| $$SPL\left(x[n]\right)=20log\_{10}\sqrt{\frac{1}{m}\sum\_{i=0}^{m-1}x^{2}\left[n-i\right]}=10log\_{10}\left(\frac{1}{m}\sum\_{i=0}^{m-1}x^{2}\left[n-i\right]\right),$$ | (1) |

где $x[n]$ – входной сигнал, а $m$ – длина временного окна (в отсчетах), используемого для вычисления мощности сигнала. Временное окно должно быть короче, чем период стационарности речевого сигнала (5-20 мс), но длиннее, чем период одного речевого колебания (0,3-3 мс для 0,3-3 кГц). Поэтому в работе предлагается выбрать длину окна равную 4 мс, что составляет 64 отсчета при частоте дискретизации 16 кГц.



|  |  |
| --- | --- |
| *а)* | *б)* |

Рисунок 2 – а) характеристика вход/выход компрессора, б) АЧХ кохлеарного банка фильтров

Банк кохлеарных фильтров. Потеря разборчивости речи, как правило, появляется в результате изменения частотной характеристики уха человека. Компенсация потери разборчивости требует применения различных коэффициентов усиления в различных частотных полосах. В данной работе для разбиения сигнала на субполосные составляющие предлагается использовать банк кохлеарных фильтров, основанный на SDCM-модели – Second order Difference Cochlear Model (разностная кохлеарная модель второго порядка) [3]:

|  |  |
| --- | --- |
| $$y\_{k}\left(n\right)+b\_{1,k}y\_{k}\left(n-1\right)+b\_{2,k}y\_{k}\left(n-2\right)=A\_{k}α\_{0,k}[u\_{s}\left(n\right)-u\_{s}(n-2)],$$ | (2) |

где $b\_{1,k},b\_{2,k}, a\_{0,k}$ и $A\_{k}$ – параметры, определяемые исходя из физических характеристик базилярной мембраны, таких как жесткость, затухание, длина, масса и др. в позиции $x\_{k}$, $k$ – номер сегмента базилярной мембраны после дискретизации; $y\_{k}(n)$ – перемещение, или пучность, базилярной мембраны в позиции $x\_{k}$; $u\_{s}(n)$ – входной синусоидальный сигнал, характеризующий скорость перемещения стремечка.

Применяемый банк обладает свойством неравнополосности и согласован с работой улитки уха человека. Банк фильтров состоит из цифровых рекурсивных фильтров второго порядка с высокой степенью перекрытия полос пропускания:

|  |  |
| --- | --- |
| $$H\_{k}\left(z\right)=\frac{a\_{0,k}(1-z^{-2})}{1+b\_{1,k}z^{-1}+b\_{2,k}z^{-2}}, k=0…K-1,$$ | (3) |

$b\_{1,k},b\_{2,k}, a\_{0,k}$ – параметры из выражения (2). Для реализации полосовых фильтров определяемых выражением (3) предлагается использовать перестраиваемые полосовые фильтры 2-го порядка [4, 5]:

|  |  |
| --- | --- |
| $$H\left(z\right)=a\_{0}\frac{1-z^{-2}}{1+\left(a\_{0}-1\right)gz^{-1}+(1-2a\_{0})z^{-2}}.$$ |  |

Коэффициенты $a\_{0}$ и $g$ зависят только от ширины полосы пропускания фильтра $Δω$ и центральной частотой $ω\_{0}$, соответственно:

|  |  |
| --- | --- |
| $$a\_{0}=\frac{ΔωΔt}{2+ΔωΔt}, g=2\cos(ω\_{0}Δt).$$ |  |

В данных выражениях $Δt$ – период дискретизации.

В работе использовался 22-канальный кохлеарный банка фильтров, АЧХ которого показана на рис. 2,б. Групповая задержка в банке фильтров колеблется в пределах 0,8–1,5 мс.

Алгоритм многополосной КДД. По полученным в банке фильтров субполосным сигналам производится оценка мощности сигнала на частотах соответствующих центральным частотам фильтров. Для уменьшения вычислительной сложности расчета мощности предлагается использовать экспоненциальное усреднение:

|  |  |
| --- | --- |
| $$P\_{k}^{in}\left[n\right]=\frac{1}{m}\sum\_{i=0}^{m-1}x\_{k}^{2}\left[n-i\right]≅βP\_{k}\left[n-1\right]+\left(1-β\right)x\_{k}^{2}\left[n\right], k=0…K-1,$$ | (4) |

где $P\_{k}^{in}[n]$ – мощность сигнала на выходе *k*-го канала банка фильтров, $x\_{k}[n]$ – сигнал на выходе *k*-го канала банка фильтров, *K* – число каналов банка фильтров, *m* – длина временного окно усреднения (в рассматриваемом случае *m*=64, что соответствует 4 мс при $f\_{s}$=16 кГц). Коэффициент $β$ зависит от *m* и может быть оценен как $(m-1)/m$.

Расчет субполосных коэффициентов усиления $g\_{0},g\_{1},…,g\_{K-1}$ производится в два этапа. На первом этапе полученные оценки мощности в субполосах переводятся в децибелы $P\_{k,dB}$ и выполняется рассчет *ожидаемых* уровней мощности с использованием характеристики вход/выход компрессора [1]

|  |  |
| --- | --- |
| $$P\_{k, dB}^{out}=\left\{\begin{array}{c}P\_{k,dB}^{in}+G\_{dB}, если P\_{k,dB}^{in}<CT \\CT+\frac{1}{CR}\left(P\_{k,dB}^{in}-CT\right)+G\_{dB}, если P\_{k,dB}^{in}\geq CT.\end{array}\right.$$ | (5) |

Далее рассчитываются *целевые* коэффициенты усиления в субполосах, как разница между входным и ожидаемым уровнем мощности в субполосах.

|  |  |
| --- | --- |
| $$G\_{k,dB}=P\_{k,dB}^{out}-P\_{k,dB}^{in},$$ | (6) |

На втором этапе происходит сглаживание коэффициентов усиления. Это необходимо во избежание резких перепадов в выходном сигнале. Блок, выполняющий сглаживание и учитывающий время атаки и время отпускания, реализован согласно алгоритму, предложенному в [2]. Иллюстрация работы алгоритма во время атаки сигнала приведена на рис. 3.



Рисунок 3 – Время атаки

После умножения субполосных сигналов на найденные коэффициенты усиления производится синтез сигнала путем суммирования всех компонент. На последнем этапе алгоритма КДД выполняется нормализация мощности выходного сигнала. Для этого выходе схемы (рис. 1) предусмотрен буфер на 4 мс, в котором накапливается обработанный в банке фильтров сигнал. Блок расчета общего коэффициента усиления производит оценку мощности входного сигнала и по характеристике вход/выход компрессора (рис. 2,а) определяет *целевую* выходную мощность. В случае, если выходная мощность не соответствует целевой рассчитывается коэффициент $g\_{total}$, который корректирует выходную мощность сигнала.

Результаты экспериментов. Проверка предлагаемого алгоритма многополосной компрессии динамического диапазона производилась на речевом сигнале (рис. 4).

Исходный сигнал состоял из двух частей – громкого и тихо участка. После обработки в компрессоре обе части приблизительно выровнялись по громкости. Видимые в спектре на рис, 4,б искажения, появившиеся в результате нелинейной обработки в компрессоре, не оказывают ощутимого влияние на разборчивость речи и узнаваемость диктора (оценка PESQ=3.168). Приведенные результаты подтверждают работоспособность предлагаемого алгоритма.

Вывод. В работе предложен метод многополосной компрессии динамического диапазона речевых сигналов для слуховых аппаратов. Особенностью метода является применение в нем банка кохлеарных фильтров. Использование кохлеарного банка фильтров позволяет выполнить обработку речевого сигнала, согласованную со слуховой системой человека. Кроме того кохлеарный банк фильтров эффективен с точки зрения вычислительных затрат, а также обладает малой групповой задержкой. Указанные свойства чрезвычайно важны для практического применения компрессора в слуховом аппарате. Работоспособность компрессора динамического диапазона подтверждена результатами MATLAB-моделирования.



|  |  |
| --- | --- |
| *а)* | *б)* |

Рисунок 4 – а) Сигнал на входе компрессора, б) Сигнал после обработки в компрессоре

Литература

1. K. Ngo, S. Doclo, A. Spriet, M. Moonen, J. Wouters, and S. Holdt Jensen An integrated approach for noise reduction and dynamic range compression in hearing aids / Proc. of 16th European Signal Processing Conference (EUSIPCO’2008), Lausanne, Switzerland, – 2008.
2. Y.-T. Kuo, T.-J. Lin, W.H. Chang, Y.-T. Li, and C.-W. Liu Complexity-effective auditory compensation for hearing-aids / IEEE International Symposium Circuits and Systems, Seattle, USA, – 2008. – P. 1472–1475.
3. W. G. Wan, A. A. Petrovsky, C. X. Fan A two-dimentional non-linear cochlear model for speech processing: response to pure tones / 6th International Fase-Congress, Zurich, Switzerland, – 1992. – P. 233–236.
4. А.А. Петровский, Ю.А. Ганушкин Метод синтеза каскадной структуры цифровых полосовых фильтров из однотипных звеньев / Радиотехника и электроника. – 1986. - № 15. - С.26-31.
5. A. Petrovsky, Synthesis of high order digital bandpass filters with tunable center frequency and bandwidth / Proc. of 8th European Signal Processing Conference (EUSIPCO’98), Trieste, Italy, – 1996. – P. 1527–1530.

Compressor Of speech signals for hearing aid based on Coсhlear filter bank

Maxim Vashkevich, Alexander Petrovsky

Belarusian State University of Informatics and Radioelectronics

Abstract. The paper describes a dynamic range compressor of speech signals for hearing aids. The compressor uses cochlear filter bank to separate signal into subbands. The main idea of using cochlear filter bank is that it introduces low signal delay and allows taking into account features of human auditory system. Both characteristics a very important for hearing aid applications. In order to reduce computational complexity of compressor cochlear filter bank implements as set of IIR filters with tunable passband and central frequency.