реализациЯ слухового аппарата на мобильной
вычислительной платформе

доц. Вашкевич М.И., доц. Азаров И.С., проф. Петровский А.А.

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

Аннотация. В статье предлагается алгоритм обработки речевого сигнала для повышения разборчивости у людей страдающих тугоухостью. В качестве целевой аппаратной платформы выступает современный смартфон iPhone. Задачей алгоритма является корректировка спектральной огибающей речевого сигнала для улучшения звуковосприятия, а также компенсация нарушения ощущения громкости. Для разработанного алгоритма приводится блок-схема, а также примеры обработки тональных и речевых сигналов.

Введение. Тугоухость является одной из важнейших проблем в современном обществе и продолжает оставаться областью активных исследования многих научных центров и лабораторий. По данным всемирной организации здравоохранения на 2005 год 278 миллионов человек в мире имели умеренные или серьезные нарушения слуха. 80% из них живет в странах с низким и среднем уровнем дохода. В силу данных обстоятельств обеспечение всех нуждающихся средствами улучшения слуха является весьма затруднительным. Возможным выходом из данной ситуации является перенесение функции слухового аппарата на современные мобильные платформы (смартфоны). Такое решение является эффективным поскольку смартфоны лидируют среди средств коммуникации и имеют очень широкое распространение. Ключевым требованием к вычислительной платформе для реализации функции слухового аппарата является возможность обработки речевого сигнала в режиме реального времени с задержкой не более 15 мс [1]. Анализ вычислительных платформ современных смартфонов показывает, что указанному требованию в настоящее время удовлетворяет только смартфоны фирмы iPhone.

**Особенности вычислительной платформы iPhone**. В результате экспериментов было установлено, что смартфоны позволяют выполнять обработку звуковых сигналов внося при этом задержку порядка 10-15 мс, что вполне удовлетворяет требованием рассматриваемой задачи. Операционная система iOS позволяет обрабатывать входной звуковой сигнал отдельными кадрами по $R$ отсчетов. Размер кадра, как правило, колеблется в пределах от 128 до 1024 отсчетов и выбирается равным степени числа 2. Для реализации функции слухового аппарата с целью уменьшения задержки лучше выбирать размер кадра равным 128 или 256.

**Обработки речевого сигнала с целью повышения разборчивости**. Обработка сигнала заключается в выполнении двух основных процедур: 1) коррекция огибающей спектра (целевые значения усиления вычисляются на основе аудиограммы пользователя) и 2) динамическая компрессия сигнала с целью компенсации нарушения ощущения громкости.

Коррекция огибающей спектра может быть выполнена при помощи КИХ фильтра с фиксированной частотной характеристикой. Фильтр синтезируется с использованием правил расчета целевого усиления POGO, NAL-R или Berger [2]. Для уменьшения вычислительной сложности фильтрацию необходимо выполнять в частотной области. Для этой цели предлагается использовать метод перекрытия с суммированием [3]. Схема алгоритма обработки сигнала для реализации функции слухового аппарата показан на рисунке 1.

Рисунок 1 – Схема обработки сигнала для реализации слухового аппарата

На вход поступает дискретизированный сигнал $x[n]$, который разбивается на последовательные, неперекрывающиеся кадры $x^{(m)}[r]$:

|  |  |
| --- | --- |
| $$x^{(m)}\left[r\right]=x\left[mR+r\right], $$ |  |

где $m$ – номер кадра, $r=1,…R$ – индекс отсчета внутри кадра.

Для выполнения линейной фильтрации методом перекрытия с суммированием входной кадр расширяется последовательностью из $R$ нулевых отсчетов. К полученной последовательности применяется алгоритм быстрого преобразования Фурье (БПФ) для перевода сигнала из временной области в частотную. Результат преобразования умножается на заранее рассчитанную частотную характеристику КИХ фильтра-корректора. Для перевода сигнала во временную область используется алгоритм обратного быстрого преобразования Фурье (ОБПФ). Результатом преобразования является частичная свертка кадра входного сигнала $x^{(m)}[r]$ с импульсной характеристикой фильтра-корректора. Частичная свертка имеет два участка $y\_{1}^{(m)}[r]$ и $y\_{2}^{(m)}\left[r\right].$ Текущий результат фильтрации, который соответствует кадру $m$, формируется путем суммирования

|  |  |
| --- | --- |
| $$y^{(m)}\left[r\right]=y\_{1}^{(m)}\left[r\right]+y\_{2}^{(m-1)}\left[r\right]. $$ |  |

Следующим этапом обработки сигнала является компрессия динамического диапазона (КДД) с целью компенсации нарушения ощущения громкости. Главной задачей КДД является автоматическое регулирование коэффициента усиления сигнала. Характер КДД определяется функцией вход/выход компрессора, которая показывает как должен изменятся уровень выходного сигнала в зависимости от уровня входного. В работе используется компрессор широкого динамического диапазона (англ. WDRC – *wide dynamic range compression*) характеристика вход/выход которого показана на рисунке 2.



Рисунок 2 ‑ Характеристика вход/выход КДД

Динамический диапазон входного сигнала условно разбит на три неперекрывающиеся области: "шум", "тихие звуки" и "речь". Под "шумом" понимается собственный шум внутренней схемы платформы iPhone, которая отвечает за прием и дискретизацию входного сигнала. Это сигнал малой интенсивности, который при прохождении компрессора не усиливается. Выше уровня "шума" находятся "тихие звуки". Сигналы, относящиеся к этой категории, плохо различимы людьми, страдающими тугоухостью, поэтому при их обнаружении компрессор КДД начинает работать в режиме расширения (англ. – *expansion*), который характеризуется коэффициентом расширения $XR$. Если через $Δx$ обозначить изменение уровня входного сигнала, а через $Δy$ – выходного, то

|  |  |
| --- | --- |
| $$XR=Δx/Δy. $$ |  |

Главной особенностью режима расширения КДД является то, что коэффициент $XR$ всегда меньше единицы. Например, для коэффициента расширения $XR$=0,5 изменение входного уровня сигнала на 10 дБ приведет к изменению выходного уровня на 20 дБ.

Если уровень входного сигнала попадает в категорию "речь" (рисунок 2), то активизируется режим компрессии динамического диапазона, который характеризуется коэффициентом компрессии

|  |  |
| --- | --- |
| $$CR=Δx/Δy, $$ |  |

который всегда больше единицы. Например, при коэффициенте $CR$=2, изменение входного уровня сигнала на 10 дБ приведет к изменению выходного уровня на 5 дБ.

Таким образом, характеристика вход/выход компрессора имеет вид кусочно-линейной кривой, имеющей характерные точки: $TK\_{exp}$ – переход к режиму расширения и $TK\_{comp}$ – переход в режим компрессии. Кроме характеристики вход/выход компрессор имеет параметры времени атаки и восстановления [4].

В соответствие с рисунком 1 на вход компрессора поступает сигнал $y^{(m)}[r]$ от фильтра-корректора. Задача компрессора состоит в расчете линейной функции усиления $g[r]$ для формирования выходного сигнала:

|  |  |
| --- | --- |
| $$s^{(m)}\left[r\right]=y^{(m)}\left[r\right]⋅g\left[r\right], r=1,…R.$$ |  |

Поскольку на функцию $g[r]$ наложено ограничение линейности, то достаточно определить значение $g[1]$ и $g[R]$, а остальные значения можно найти путем линейной интерполяции. Ниже приведен псевдокод для определения $g[1]$ и $g[R]$:

|  |
| --- |
| $P\_{start}=10log\_{10}(p)$;for $r=1,2,…,R$ $p\_{inst}=\left(y^{\left(m\right)}\left[r\right]\right)^{2}$; if ($p\_{inst}>p$) then $p=αp+\left(1-α\right)p\_{inst}$;  else $p=βp+\left(1-β\right)p\_{inst}$;  end if;end for;$P\_{end}=10log\_{10}(p)$; $P\_{start}^{out}$=IO\_Func($P\_{start}$);$P\_{end}^{out}$ =IO\_Func($P\_{end}$);$g\left[1\right]=10^{(P\_{start}^{out}-P\_{start})/20}$;$g\left[R\right]=10^{(P\_{end}^{out}-P\_{end})/20}$; |

где $α$ – параметр экспоненциального усреднения, зависящий от времени атаки,

 $β$ – параметр экспоненциального усреднения, зависящий от времени восстановления.



Рисунок 3 ‑ Обработка тестовых тональных сигналов компрессором динамического диапазона

В приведенной программе переменная $p$ – хранит текущее (среднее) значение уровня мощности сигнала. Для первого кадра сигнала ($m=0$) $p$ присваивается значение $2^{-10}$, в последующем значение $p$ сохраняется от кадра к кадру. Таким образом, $P\_{start}$ для $m$-го кадра равно значению $P\_{end}$ ($m$-1)-го кадра. Текущее значение мощности сигнала вычисляется путем экспоненциального усреднения. Параметр экспоненциального усреднения выбирается в зависимости от того, что происходит атака (нарастание) или отпускание (спад) сигнала. Через IO\_Func() в программе обозначена функция вход/выход компрессора.

На рисунке 3 показан пример обработки тестовых тональных сигналов в компрессоре динамического диапазона. Пример иллюстрирует характер изменения коэффициента усиления $g[r]$ в зависимости от уровня входного сигнала, а также поясняет значения понятий времени атаки и времени восстановления.

Пример обработки речевого сигнала предложенным алгоритмом показан на рисунке 4. Видно, что совместное использование фильтра-корректора и компрессора динамического диапазона позволяют выполнить уровне- и частотно-зависимую обработку сигнала. Следует обратить особое внимание на результат обработки шипящего звука "с". Значительное усиление этого звука обусловлено тем, что основная часть его энергии лежит в верхней части частотного диапазона, который плохо воспринимается людьми, страдающими тугоухостью. По тем же причинам значительное усиление получил взрывной звук "п".



Рисунок 4 ‑ Обработка речевого сигнала при помощи разработанного алгоритма

Вывод. В работе рассмотрена возможность применения вычислительной платформы iPhone для реализации слухового аппарата и предложен соответствующий алгоритм обработки сигнала. Основной целью алгоритма является выполнение уровне- и частотно-зависимого усиления речевого сигнала для коррекции огибающей спектра и компенсации нарушения ощущения громкости.

Литература

1. Bäuml R.W., Sörgel W. Uniform polyphase filter banks for use in hearing aids: design and constraints / Proc. of 16th European Signal Processing Conference (EUSIPCO’2008), Lausanne, Switzerland, – 2008.
2. Фонлантен А., Хорст А. Слуховые аппараты – Ростов н/Д.: Феникс, 2009. – 304 с.
3. Рабинер Р., Гоулд Б. Теория и применение цифровой обработки сигналов: Пер. с англ. М.: Мир, 1978. – 848 с.
4. Вашкевич М.И., Петровский А.А. Компрессор речевых сигналов для слуховых аппаратов на основе кохлеарного банка фильтров // 14-я межд. конф. Цифровая обработка сигналов и её применения, Россия, Москва, 28-30 марта 2012 г. – Т.1, – С. 87-91.

Speech processing algorithm for implementing a hearing aid on mobile platform

Maxim Vashkevich, Elias Azarov, Alexander Petrovsky

Belarusian State University of Informatics and Radioelectronics

Abstract. The paper presents a speech processing algorithm design for improving speech intelligibility for hearing-impaired persons. An iPhone have been chosen as a target platform for algorithm implementation. The goals of proposed algorithm are 1)  correction of spectral envelope for enhancing speech perception and 2) compensation of volume sensitivity. A block diagram and examples of synthetic and natural signals processing are presented.