

акустические характеристики фильтра. На следующем этапе, значения функции $h(a, b)$ волосковых клеток для каждой полосы усредняются, при этом используется окно разной длины в отличие от дискретного преобразования Фурье, где размер окна фиксирован. Чем выше частота, тем короче окно – это предотвращает от сглаживания информации длинным окном.

$$S(i, j) = 1/d \sum_{b=l}^{l+d-1} h(i, b), l = 1, L, 2L, \dots; \forall i, j$$

где $d = \max\{3, 5\tau_i; 20\text{мс}\}$ – длина окна; τ_i – период центральной полосы i -й полосы и $L = 10\text{мс}$ – размер окна сдвига.

Для учёта нелинейности были использованы веса функции громкости [2]:

$$y(i, j) = S(i, j)^{1/3}.$$

И на последнем этапе применяется дискретное косинусное преобразование (ДКП), для уменьшения размера полученных признаков. Эти признаки получили название кохлеарные кепстральные коэффициенты (ККК) и они являются тем пространством, в котором будет происходить идентификация диктора.

Для экспериментальной оценки признаков ККК для текстонезависимой идентификации в акустических шумах, был спроектирован идентификатор дикторов на основе нейронных сетей. Также было произведено сравнение результатов распознавания ККК признаков с существующими мел-частотными кепстральными коэффициентами (МКК) и характеристиками, полученные с помощью алгоритма RASTA (Relative Spectral Transform – Относительное спектральное преобразование) в задаче идентификации диктора в акустических шумах. В качестве базы была составлена речевая база, основанная на базе SSC[3]. Было подготовлено 3 множества: обучающее, настраиваемое и тестовое. Сравнение результатов идентификации диктора с различными уровнями шумов (представленный тип шума – bubble шум) представлено на рисунке 2.

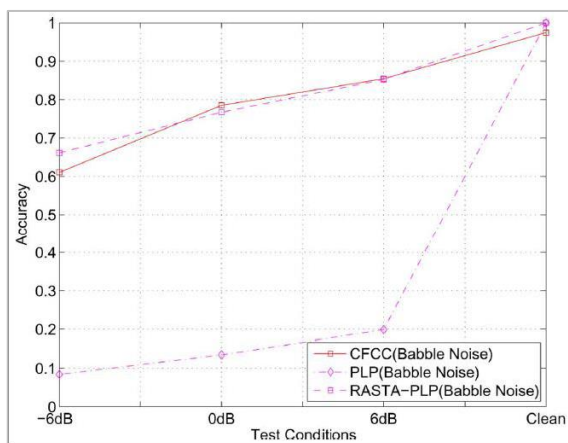


Рис.2 – Сравнение результатов распознавания

Проведённые эксперименты демонстрируют, что в условиях наличия акустических шумов различного типа (белый шум, bubble шум, машинный шум), рассмотренный алгоритм экстракции характеристического вектора справляется с задачей идентификации значительно лучше, чем алгоритмы извлечения МКК и относительного спектрального преобразования при наличии шума любого рассматриваемого типа.

Список использованных источников:

1. The PASCAL CHiME speech separation and recognition challenge [Электронный ресурс] – Электронные данные. – Режим доступа : <http://spandh.dcs.shef.ac.uk/projects/chime/PCC/results.html>.
2. Q. Li, An auditory-based feature extraction algorithm for robust speaker identification under mismatched conditions / Proc. IEEE Transactions Audio, Speech, and Language Processing., – New Paltz, NY, 2011.
3. The PASCAL CHiME speech separation and recognition challenge [Электронный ресурс] – Электронные данные. – Режим доступа : <http://spandh.dcs.shef.ac.uk/projects/chime/PCC/datasets.html>.

КОМБИНИРОВАННАЯ СИСТЕМА УЛУЧШЕНИЯ РАЗБОРЧИВОСТИ АУДИО СИГНАЛА В АГРЕССИВНОЙ ШУМОВОЙ СРЕДЕ

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

Белый В.Л.

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

Современные устройства и технологии позволяют передавать, принимать и воспроизводить аудио сигналы, находясь в различных условиях окружающей среды. Это приводит к возникновению проблемы разборчивости принятого,

воспроизведенного или переданного аудио сигнала в связи с высоким уровнем окружающего шума. Исходя из данной проблемы, очень важно разработать систему, которая будет эффективно подавлять уровень шума, используя различные методы и подходы, такие как усиление уровня интересующего аудио сигнала, не допуская, при этом, увеличения уровня сигнала до критически опасного значения для слуховой системы человека, а также алгоритмы, направленные на обработку, непосредственно, шумовых компонент, без внесения изменений в аудио сигнал.

Основной задачей данной работы, являлось проектирование комбинированной системы улучшения разборчивости аудио сигнала в агрессивно шумовой среде, которая способна осуществлять шумоподавление, как для полученного и воспроизведенного аудио сигнала, так и для передаваемого аудио сигнала. Зашумленный аудио сигнал представляет собой сумму окружающего шума, прошедшего через определенный путь в ухо человека и самого интересующего сигнала.

Для достижения приемлемых результатов было решено проводить два этапа редактирования шумов. Данный подход позволяет получить более гибкий и эффективный алгоритм, а также, не допуская достижения критически опасного уровня аудио сигнала, избавиться от шумов, чей уровень достаточно высок. Первым этапом работы, является алгоритм активного шумоподавления (Active noise control (ANC)), интегрированный в звуковоспроизводящую систему.

На рисунках 1 и 2 приведены структурная схема ANC и конструктивное расположения компонентов системы:

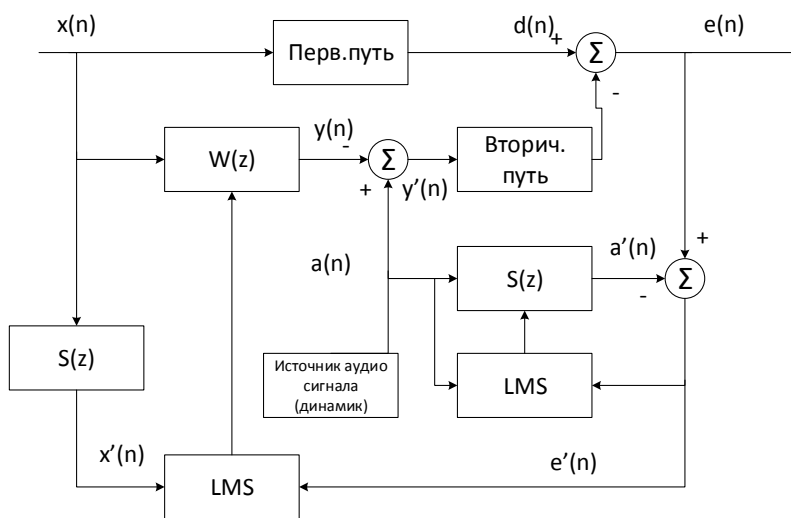


Рис. 1 – Структурная схема ANC



Рис. 2 – Конструктивное расположения компонентов

Алгоритм ANC основан на субполосной адаптивной фильтрации, которая в свою очередь функционирует на базе NLMS-алгоритма:

$$w(n+1) = w(n) + \mu \frac{u(n)}{u(n)^T u(n)} e(n)$$

где $w(n)$ – набор коэффициентов фильтра, $u(n)$ – входящий сигнал, $e(n)$ – ошибка, μ - шаг приращения.

Вторым этапом шумоподавления является маскирование, которое осуществляется в частотно-временной области, с применением психоакустических принципов, за счет использования ANC. На втором этапе необходимо маскировать только остаточный шум, что позволяет достичь достаточно высокого уровня шумоподавления, т.к. наиболее значимые шумовые компоненты были подавлены на этапе активного шумоподавления, а также не допустить достижения критически высокого уровня аудио сигнала при его усилении, в следствие того, что остаточный шум не обладает высокой амплитудой.

На рисунках 3 приведена структурная схема алгоритма маскирования:

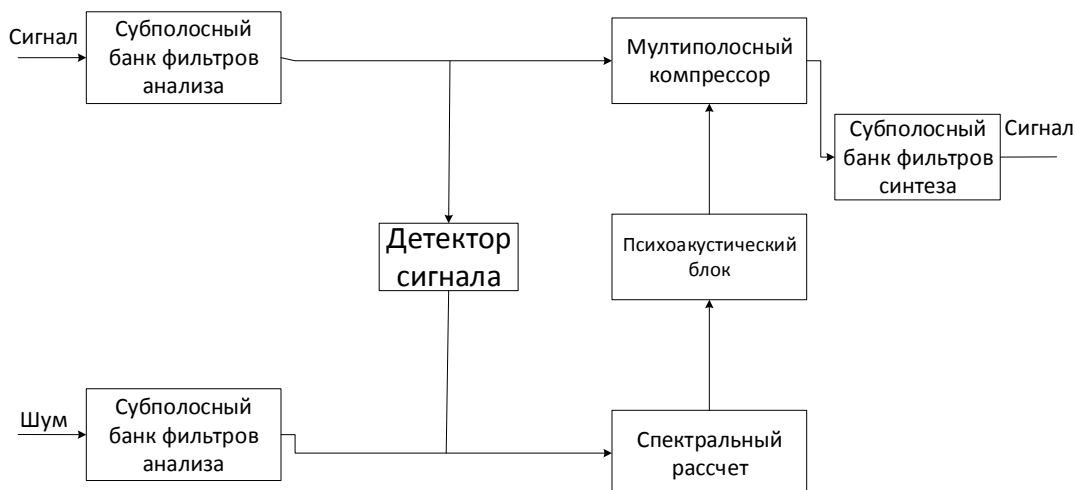


Рис. 3 – Структурная схема алгоритма маскирования

Для создания математических моделей был использован язык технических вычислений MatLab. Для проектирования прототипа устройства была использована процессорная система на базе ARM-ядра.

В результате построена система улучшения разборчивости аудио сигнала в агрессивной шумовой среде. Основными преимуществами является двуэтапность обработки, которая позволяет применять данную систему в различных шумовых средах, для различных аудио сигналов и препятствует достижению критически опасного уровня аудио сигнала для слуховой системы человека, также данная система может работать для передачи аудио сигнала, подавляя шумы в сигнале.

Список использованных источников:

1. Конг-Аик Ли. Субполосная адаптивная фильтрация, теория и применение / Конг-Аик Ли, Вун-Сенг Хан, Сен М. Куо. // – JohnWileyandSons, Ltd, 2009. – 324 с.
2. Петровский, А. А. Цифровые банки фильтров: анализ, синтез и применение в мультимедиа системах / А. А. Карпушкин, М. Парфенюк, А. Борович, М.З. Лившиц // Уч. метод. пособие для студентов учреждений, обеспечивающих получение высшего образования по специальности "Электронные вычислительные средства". – Минск, 2006. – 81 с.
3. Bjarnason, E. Analysis of filtered-X LMS algorithm. / Bjarnason E. // - IEEE Trans. Speech audio Process., 3(6), 1995, 504-514.
4. Davari, P. Designing a new robust on-line secondary path modeling technique for feedforward active noise control systems/ Davari P, Hassanpour H. // - Signal processing, 89(6), 2009, 1195 – 1204.
5. Rees, L. Psychoacoustic modelling for active noise control system/ Rees L., Elliot S.J.// - Proc. Inst. Acoust., 26(pt.2), 2004.

СЛУХОВОЕ ПРЕОБРАЗОВАНИЕ. ЕГО ИСПОЛЬЗОВАНИЕ В ИДЕНТИФИКАЦИИ ДИКТОРА

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

Кручок Д. Н.

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

В современных автоматических системах распознавания (АСР) одной из главных задач является обеспечение распознавания в условиях с различной акустической обстановкой. Для достижения высоких показателей эффективности распознавания в акустических шумах, целесообразно использовать алгоритмы, моделирующие процесс обработки звукового сигнала слуховой системой человека. Основой для таких алгоритмов служат различные преобразования, одним из которых является слуховое преобразование.

Слуховое преобразование – преобразование, которое переводит сигнал из временной области в набор выходов банка фильтров. Частотные характеристики и распределение полос банка фильтров такие же, как и у базилярной мембраны в улитке в среднем ухе человека. Когда звук попадает в человеческое ухо, акустическая энергия с внешнего уха превращается в механическую энергию через среднее ухо, представленное тремя косточками. Когда последняя косточка, стремечко, колеблется, то это приводит в движение жидкость внутри улитки, создавая бегущие волны на базилярной мембране (БМ).