

УДК 621.395.92

## АЛГОРИТМ ПОДАВЛЕНИЯ ШУМА И АКУСТИЧЕСКОЙ ОБРАТНОЙ СВЯЗИ НА ОСНОВЕ СПЕКТРАЛЬНОГО ВЫЧИТАНИЯ В СЛУХОВОМ ПРОТЕЗЕ НА БАЗЕ СМАРТФОНА

М.И. ВАШКЕВИЧ, И.С. АЗАРОВ, А.А. ПЕТРОВСКИЙ

Белорусский государственный университет информатики и радиоэлектроники  
П. Бровки, 6, Минск, 220013, Беларусь

Поступила в редакцию 21 ноября 2016

В работе предлагается совмещенный алгоритм подавления шума и акустической обратной связи. Алгоритм основан на спектральном вычитании и является устойчивым к резким изменениям параметров пути распространения акустической обратной связи, что делает его подходящим для использования в слуховых протезах на основе смартфонов.

*Ключевые слова:* шумоподавление, подавление акустической обратной связи, слуховой аппарат.

### Введение

Широкое распространение смартфонов создало основание для развития карманных слуховых аппаратов в новом качестве. В последнее время было выпущено большое количество приложений для портативных мультимедиа устройств, позволяющих выполнять диагностику и коррекцию слуха [1]. Использование смартфона в качестве слухового аппарата имеет множество преимуществ, среди которых гибкость используемых алгоритмов, большой энергетический и вычислительный ресурс, возможность персональной настройки при помощи встроенной аудиометрии без помощи сурдолога, возможность использования произвольной внешней гарнитуры и психологический комфорт.

Известно, что 52 % людей, пользующихся слуховыми аппаратами, имеют необходимость их применения в шумовой обстановке [2]. Исследования показали, что шумоподавление значительно ослабляет вредное воздействие на слух [3]. Кроме окружающего шума крайне нежелательным является эффект акустической обратной связи (АОС), который возникает тогда, когда обработанный сигнал попадает из динамиков в микрофон.

Адаптивная фильтрация на основе метода наименьших квадратов (МНК) на сегодняшний день является наиболее используемой техникой подавления акустической обратной связи [4–9]. Однако, этот подход является неэффективным для реализации на смартфоне, поскольку при использовании внешней гарнитуры путь распространения АОС быстро изменяется. Показано, что акустика помещения вносит значительный вклад в путь распространения [10], однако выполнить моделирование акустики помещения путем адаптивной фильтрации в реальных условиях не представляется возможным. В настоящей работе предлагается оригинальный алгоритм подавления акустической обратной связи, основанный на спектральном вычитании. Алгоритм использует взвешивающее правило, специально выведенное для сигналов обратной связи, и позволяет дополнительно выполнять шумоподавление, которое ослабляет остаток АОС и фоновый шум.

### Алгоритм подавления шума и акустической обратной связи

Схема обработки сигнала показана на рис. 1. Для декомпозиции на частотные составляющие используется ДПФ-модулированный банк фильтров, выполненный в виде структуры со сложением с перекрытием [11, 12], имеющий  $K$  каналов и коэффициент

децимации  $M = K/2$ . Входной речевой сигнал  $x(n)$  разбивается на комплексные субполосные сигналы банком фильтра анализа (БФА).

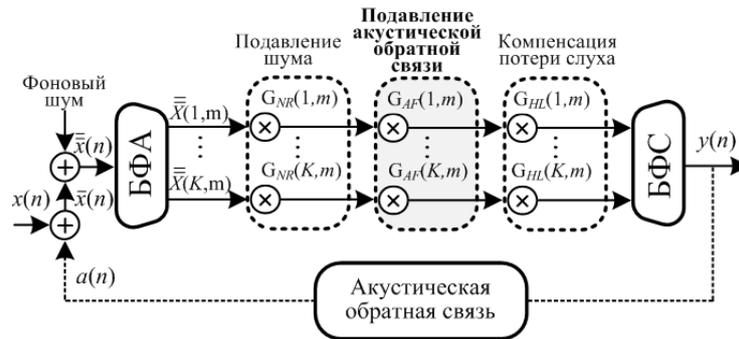


Рис. 1. Обработка сигнала в слуховом протезе на базе смартфона

Субполосные сигналы прореживаются по времени и формируют последовательность частотно-временных отсчетов  $X(k, m)$ , где  $k$  – частотный индекс, а  $m$  – временной индекс. Отсчеты умножаются на коэффициенты усиления  $G_{HL}(k, m)$ , компенсирующие потерю слуха, рассчитываемые в соответствии с выбранным стандартом [3] либо установками пользователя. Обработанный выходной сигнал  $y(n)$  восстанавливается из обработанных отсчетов при помощи банка фильтра синтеза (БФС).

На вход попадает окружающий шум и акустическая обратная связь и предполагается, что субполосный отсчет  $\bar{X}(k, m)$  является суммой чистой речи  $X(k, m)$ , акустической обратной связи  $A(k, m)$  и фонового шума  $N(k, m)$ :  $\bar{X}(k, m) = X(k, m) + A(k, m) + N(k, m)$ . Минимальную амплитуду чистого речевого сигнала можно оценить вычитая амплитуду шума и обратной акустической связи из амплитуды входного сигнала. Вычитание эквивалентно умножению на соответствующие коэффициенты  $G_{NR}(k, m)$  и  $G_{AF}(k, m)$ . Обозначим сумму речевого сигнала и акустической обратной связи  $X(k, m) + A(k, m)$  как  $\bar{X}(k, m)$ , тогда

$$G_{NR}(k, m) = \frac{|X(k, m)| + |A(k, m)|}{|X(k, m)| + |A(k, m)| + |N(k, m)|} = \frac{|X(k, m)|}{|\bar{X}(k, m)|}, \quad G_{AF}(k, m) = \frac{|X(k, m)|}{|X(k, m)| + |A(k, m)|} = \frac{|X(k, m)|}{|\bar{X}(k, m)|}. \quad (1)$$

Амплитуда чистого речевого сигнала может быть восстановлена из зашумленного следующим образом:  $|X(k, m)| = |\bar{X}(k, m)| G_{NR}(k, m) G_{AF}(k, m)$ . Вычисление искомых коэффициентов основано на статистической оценке амплитудного спектра шума и акустической обратной связи.

Алгоритм оценки амплитудного спектра шума основан на рекурсивном усреднении, параметры которого управляются минимальными значениями амплитуды [13]. Оценка амплитудного спектра шума  $|\hat{N}(k, m)|$  вычисляется путем усреднения амплитуд предыдущих отсчетов сигнала при помощи следующего выражения:

$$|\hat{N}(k, m)| = \tilde{\alpha}_d(k, m-1) |\hat{N}(k, m-1)| + (1 - \tilde{\alpha}_d(k, m-1)) |\bar{X}(k, m)|, \quad (2)$$

где  $\tilde{\alpha}_d(k, m) = \alpha_d + (1 - \alpha_d) \hat{p}(k, m)$  изменяющийся параметр сглаживания, который зависит от условной вероятности присутствия речи  $\hat{p}(k, m)$ ,  $\alpha_d$  – параметр сглаживания, который определяет время усреднения при отсутствии речи ( $0 < \alpha_d < 1$ ). Вероятность присутствия речи оценивается путем рекурсивного усреднения бинарного признака  $I(k, m)$ :

$$\hat{p}(k, m) = \alpha_p \hat{p}(k, m-1) + (1 - \alpha_p) I(k, m), \quad I(k, m) = \begin{cases} 1, & S_r(k, m) > \delta \\ 0, & S_r(k, m) \leq \delta \end{cases} \quad (3)$$

где  $S_r(k, m)$  – соотношение между текущей амплитудой кратковременного спектра и минимальной амплитудой  $S_r(k, m) = |X(k, m)| / |\overline{X}_{\min}(k, m)|$ . Значение минимальной амплитуды  $|\overline{X}_{\min}(k, m)|$  оценивается следующим образом. Используются две переменные минимума, для которых устанавливаются начальные значения:  $|\overline{X}_{\min}(k, 0)| = |\overline{X}(k, 0)|$  и  $|\overline{X}_{\min}(k, 0)| = |\overline{X}(k, 0)|$ . Для каждого последующего отсчета текущее значение переменных получается путем сравнения текущей амплитуды и существующего минимума:

$$|\overline{X}_{\min}(k, m)| = \min\{|\overline{X}_{\min}(k, m-1)|, |\overline{X}(k, m)|\}, \quad |\overline{X}_{\min}(k, m)| = \min\{|\overline{X}_{\min}(k, m-1)|, |\overline{X}(k, m)|\}. \quad (4)$$

Каждые  $L$  кадров (т.е. когда  $m$  кратно  $L$ ) переменные минимума обновляются следующим образом:  $|\overline{X}_{\min}(k, m)| = \min\{|\overline{X}_{\min}(k, m-L)|, |\overline{X}(k, m)|\}, \quad |\overline{X}_{\min}(k, m)| = |\overline{X}(k, m)|$ .

Параметр  $L$  определяет разрешение поиска локального минимума. Локальный минимум определяется на окне длиной не менее чем  $L$  отсчетов, но не более чем  $2L$  отсчетов. Хороший практический результат достигается на окне длиной 0,5–1,5 с. Коэффициент  $G_{nr}(k, m)$  вычисляется следующим образом:

$$G_{NR}(k, m) = \max\left\{\sqrt{\frac{|\overline{X}(k, m)|^2 - \nu |\hat{N}(k, m)|^2}{|\overline{X}(k, m)|^2}}, 10^{-RL/20}\right\}, \quad (5)$$

где  $\nu$  – коэффициент вычитания ( $1 < \nu < 6$ ),  $RL$  – настраиваемый параметр, определяющий желаемый уровень остаточного шума в дБ.

Амплитуда сигнала обратной связи нарастает стремительно и переводит систему в нелинейный режим прежде чем алгоритм шумоподавления успеет отреагировать. Для того, чтобы адекватно оценить амплитудный спектр обратной акустической связи необходимо использовать ее отличительные свойства. В обоих режимах (линейном и нелинейном) акустическая связь проявляется в виде квазипериодического сигнала, который является результатом рекурсивного суммирования выхода с периодом  $t_0$ . Периодическая природа акустической обратной связи гарантирует, что ее спектральные компоненты равноудалены друг от друга в частотной области с равным шагом  $f_0 = 1/t_0$  и, следовательно, акустическая обратная связь оказывает влияние только на некоторое подмножество отсчетов полезного сигнала  $X(k, m)$ :

$$\overline{X}(k, m) = X(k, m) + A(k, m) \text{ если } \min\left|k - \frac{f_0 \nu}{f_s / K}\right| \leq d, \text{ иначе } \overline{X}(k, m) = X(k, m), \quad (6)$$

где  $\nu$  – номер гармоники сигнала обратной связи,  $f_s$  – частота дискретизации и  $d$  – смещение частоты, которое зависит от оконной функции, использованной в банке фильтров анализа. Для чистой речи математическое ожидание амплитудного спектра близких частотных компонент примерно равно  $E[X(k, m)] \approx E[X(k \pm d, m)]$  для любой достаточно малой величины  $d$ . Используя (6), и предполагая, что акустическая обратная связь всегда увеличивает среднее значение амплитудного спектра, т.е.  $E[\overline{X}(k, m)] \geq E[X(k, m)]$ , получим

$$E[X(k, m)] \approx \min_{-d < i < d} [E[\overline{X}(k+i, m)]] - E[A(k, m)] \approx E[\overline{X}(k, m)] - \min_{-d < i < d} [E[\overline{X}(k+i, m)]] \quad (7)$$

В соответствии с (6) и (7) ожидаемый коэффициент увеличения амплитуды сигнала за счет обратной связи  $E[\overline{X}(k, m)] / E[X(k, m)]$  можно оценить из кратковременного амплитудного спектра в некоторой окрестности отсчета  $\overline{X}(k, m)$ . Предлагается использовать следующую меру оценки коэффициента увеличения амплитуды  $\chi(k, m)$ , использующую  $l$  предыдущих отсчетов по времени и  $2d$  соседних отсчетов по частоте:

$$\chi(k, m) = \min_{-l+1 \leq j \leq 0} |\bar{X}(k, m+j)| / \min_{-d \leq i \leq d} \left[ \max_{-l+1 \leq j \leq 0} |\bar{X}(k+i, m+j)| \right]. \quad (8)$$

Чтобы избежать завышения оценки уровня обратной связи в числителе используется минимальный отсчет, а в знаменателе максимальный. В соответствии с полученными экспериментальными данными  $\chi > 1$  означает присутствие обратной связи с вероятностью 95 %, однако, появление обратной связи не определяется в случае, когда она намного тише речи.

Подавление обратной связи выполняется плавно с использованием сглаживающего параметра  $\alpha_{AF}$  ( $0 < \alpha_{AF} < 1$ ), который определяет время усреднения и изменяемого параметра  $\tilde{\alpha}_{AF} = \alpha_{AF} + (1 - \alpha_{AF})(\chi(k, m)/\chi_{th})^\beta$ , где  $\beta$  – регулировка баланса между реакцией на тихую и громкую обратную связь и  $\chi_{th}$  – пороговое значение для включения режима сильного подавления. Коэффициент подавления  $G_{AF}(k, m)$  вычисляется следующим образом:

$$G_{AF}(k, m) = \begin{cases} \tilde{\alpha}_{AF} G_{AF}(k, m-1) + \frac{(1 - \tilde{\alpha}_{AF})}{\max(\chi(k, m), 1)}, & \chi(k, m) < \chi_{th} \\ 1 / \chi(k, m), & \chi(k, m) \geq \chi_{th} \end{cases}. \quad (9)$$

Когда  $\chi(k, m)$  превышает  $\chi_{th}$ , алгоритм определяет наличие сильной обратной связи и в этом случае коэффициент подавления обновляется мгновенно, а затем медленно отпускается.

### Результаты экспериментов

Предложенные взвешивающие правила реализованы в виде комбинированного алгоритма подавления шума и акустической обратной связи. Для оценки эффективности алгоритма использовалась модель распространения обратной связи, аналогичная приведенной в работе [7]. Путь распространения моделировался в виде фильтра с конечной импульсной характеристикой из 279 коэффициентов. Использовалась частота дискретизации сигнала 44100 Гц, длина окна анализа 512 отсчетов, что соответствует алгоритмической задержке в 5,8 мс. Для всех частотных полос использовался один коэффициент компенсации потери слуха  $G_{HL}$ , для подавления обратной акустической связи использовались значения параметров  $\alpha_{AF} = 0,997$ ,  $\beta = 0,15$ ,  $\chi_{th} = 10$ , для шумоподавления  $\alpha_d = 0,95$ ,  $\alpha_p = 0,2$ ,  $L = 172$  (окно поиска минимума примерно 1 с),  $\delta = \sqrt{5}$ ,  $v = 2$ ,  $RL = 9$ .

В качестве первого эксперимента оценивалась устойчивость системы подавления АОС и качество выходного сигнала. Для этого определялся добавочный к предельному коэффициент усиления  $\Delta G$  [8] с применением предложенного алгоритма подавления обратной связи и алгоритма адаптивной фильтрации (279 коэффициентов фильтра) на основе метода наименьших квадратов [4]. Результат работы всей системы оценивался в терминах сегментного соотношения сигнал/шум (СССШ). Шум оценивался путем вычисления разности между выходным сигналом (с обратной акустической связью и включенным подавлением) и выходным сигналом в идеальных условиях (без обратной связи и подавления). Полученные значения СССРШ приведены в таблице 1. 'US' означает неустойчивость системы.

СССШ для различных добавочных усиления  $\Delta G$

$\Delta G$ , дБ	Без подавления, дБ	LMS, дБ	Предложенный алгоритм подавления АОС, дБ
0	8,12	17,66	12,27
8	US	5,35	11,10
16	US	1,72	9,25
24	US	US	4,56

Предложенный алгоритм подавления обратной связи обеспечивает значительно более высокие значения СССРШ по сравнению с алгоритмом на основе МНК для всех случаев, где присутствует обратная связь и сохраняет устойчивость системы даже для максимального добавочного усиления 24 дБ.

В качестве второго эксперимента выполнялась оценка эффективности алгоритма в шумовой обстановке, для чего использовался речевой сигнал с добавлением розового шума. Реализованный алгоритм шумоподавления заметно улучшает показатели СССШ (для  $\Delta G = 12$  дБ разница составляет примерно 3–5 дБ), подавляя одновременно фоновый шум и остаток АОС, что значительно повышает субъективное качество обработанной речи.

Для того, чтобы оценить работу алгоритма в реальных условиях эксплуатации создан макет на основе персонального компьютера со стандартными наушниками и микрофоном, работающий в реальном масштабе времени. Макет помещен в большую комнату с выраженным эффектом реверберации. Предложенный алгоритм подавления АОС показал на макете хороший результат, сохраняя устойчивость системы при  $\Delta G > 12$  дБ, в то время как алгоритм на основе МНК обеспечил устойчивость при  $\Delta G < 3$  дБ.

### Заключение

Предложен алгоритм подавления фонового шума и акустической обратной связи, основанный на спектральном вычитании. В соответствии с результатами экспериментов предложенный алгоритм является устойчивым к быстрому изменению пути распространения акустической обратной связи, обеспечивает значительное добавочное усиление и высокое качество обработанной речи. Благодаря своим свойствам алгоритм может быть использован в слуховых аппаратах на базе смартфонов.

## COMBINED NOISE AND ACOUSTIC FEEDBACK REDUCTION ALGORITHM BASED ON SPECTRAL WEIGHTING RULES FOR SMARTPHONE HEARING AID

M.I. VASHKEVICH, E.S. AZAROV, A.A. PETROVSKY

### Abstract

The paper presents a combined noise and acoustic feedback reduction algorithm. The algorithm is based on spectral subtraction and is robust to rapid changes in acoustic feedback path which makes it suitable for using in a smartphone-based hearing aid.

*Keywords:* noise reduction, acoustic feedback reduction, hearing aid.

### Список литературы

1. *Ismaili J.* // Education and Information Technologies. 2016. Vol. 21, iss. 82. P. 1–17.
2. *Bertoli S., Staehelin K., Zemp E. et. al.* // International journal of audiology. 2009. Vol. 48, № 4. P. 183–195.
3. *Vonlanthen A., Arndt H.* Hearing Instrument Technology for the Hearing Health Care Professional. Delmar Cengage Learning, 2006.
4. *Maxwell J.A., Zurek P.M.* // IEEE Transactions on Speech and Audio Processing. 1995. Vol. 3, № 4, P. 304–313.
5. *Schepker H., Tran L., Nordholm S. et. al.* // IEEE International Conference on Acoustic, Speech, and Signal Processing. Shanghai, China, March 20–25, 2016. P. 231–235.
6. *Strasser F., Puder H.* // IEEE Signal Processing Letters. 2016. Vol. 23, № 7. P. 979–983.
7. *Schepker H., Doclo S.* // IEEE/ACM Transactions on Audio, Speech and Language Processing. 2016. Vol. 24, iss. 2, P. 366–377.
8. *Vicen-Bueno R., Martinez-Leira A., Gil-Pita R. et. al.* // IEEE Transactions on Instrumentation and Measurement. 2009. Vol. 58, № 9. P. 3177–3190.
9. *Vashkevich M., Azarov E., Petrovsky A.* // European Signal Processing Conference. Bucharest, Romania, August 27–31, 2012. P. 2787–2791.
10. *Kates J.M.* // Journal Acoust. Soc. Am. 2001. Vol. 109, № 1. P. 367–378.
11. *Crochiere R.E., Rabiner L.R.* Multirate Digital Signal Processing. Prentice-Hall Inc., 1983.
12. *Rosa-Zurera M., Gil-Pita R., Alexandre Cortizo E. et. al.* // Materials of 10<sup>th</sup> international conference on Pattern Recognition and Information Processing «WOLA Filter Bank Design Requirements in Hearing Aids». Minsk, Belarus, May 19–21, 2009. P. 215–218.
13. *Cohen I., Berdugo B.* // IEEE Signal Processing Letters. 2002. Vol. 9, № 1. P. 12–15.