

ПОВЫШЕНИЕ КАЧЕСТВА РЕЧЕВОГО СИГНАЛА НА ОСНОВЕ ПСИХОАКУСТИЧЕСКОГО ПОДХОДА

Аношенко А.Е., Петровский А.А.

Белорусский государственный университет информатики и радиоэлектроники
220027, Минск, ул. П.Бровки, 6 (Беларусь)
E-mail: anoshenko@mail.ru, palex@it.org.by

Реферат. В данной работе описывается комбинированная система подавления эхо сигнала и снижения уровня шумов окружающей среды на основе метода спектрального вычитания, учитывающего психоакустическое восприятие человека. Особенностью предлагаемой системы является то, что обработка речевого сигнала осуществляется только в частотной области. При этом значительно снижается вычислительная сложность системы в целом по сравнению с другими аналогичными разработками. В работе описываются результаты моделирования работы системы, а также проводится сравнение с традиционным подходом к решению проблемы эхо и шума подавления.

Введение

Большинство разработанных систем подавления эхо сигнала и редактирования шума состоят из двух подсистем: предсказателя эхо сигнала на основе адаптивного фильтра нормализованного метода наименьших квадратов (НМНК) во временной области и подсистемы редактирования шума на основе метода спектрального вычитания или его модификации. Анализ различных вариантов построения данных систем проведен в [1].

В данной работе предлагается вариант комбинированной системы подавления эхо сигнала и шумов окружающей среды, в которой отсутствует отдельная подсистема подавления эха и подавление всех искажающих составляющих входного сигнала осуществляется на основе метода спектрального вычитания с учетом психоакустических особенностей слуха человека.

Система подавления эха и шумов окружающей среды на основе метода спектрального вычитания с психоакустической мотивацией

Использование систем подавления эхо сигнала и шумов окружающей среды построенной на основе адаптивного фильтра во временной области и метода спектрального вычитания [1] имеет следующие недостатки:

- 1) высокая вычислительная сложность адаптивного фильтра;
- 2) при высоком уровне шумов окружающей среды снижается скорость сходимости адаптивного фильтра;
- 3) из-за использования в адаптивном фильтре градиентного метода поиска минимума функции при увеличении порядка фильтра ухудшается качество подавления эхо сигнала. Таким образом, данный алгоритм хорошо справляется с эхо сигналами невысокой длительности и плохо – с эхо сигналами большой длительности.

Избавиться от данных недостатков позволяет использование для подавления эхо сигнала метода спектрального вычитания. Основная идея алгоритма заключается в том, что эхо сигнал рассматривается как дополнительный вид шумов и снижается одновременно с шумом по единой методике. Главный недостаток метода спектрального вычитания заключается в появлении в выходном сигнале “музыкального” тона. Этот эффект воспринимается слушателем как присутствие в сигнале неестественных “металлических” ноток. От эффекта “музыкального” тона позволяют избавиться модификации алгоритмов спектрального вычитания, учитывающие психоакустические особенности восприятия человека [2, 3].

Предлагаемая система подавления акустического эха и шумов окружающей среды показана на рис. 1.

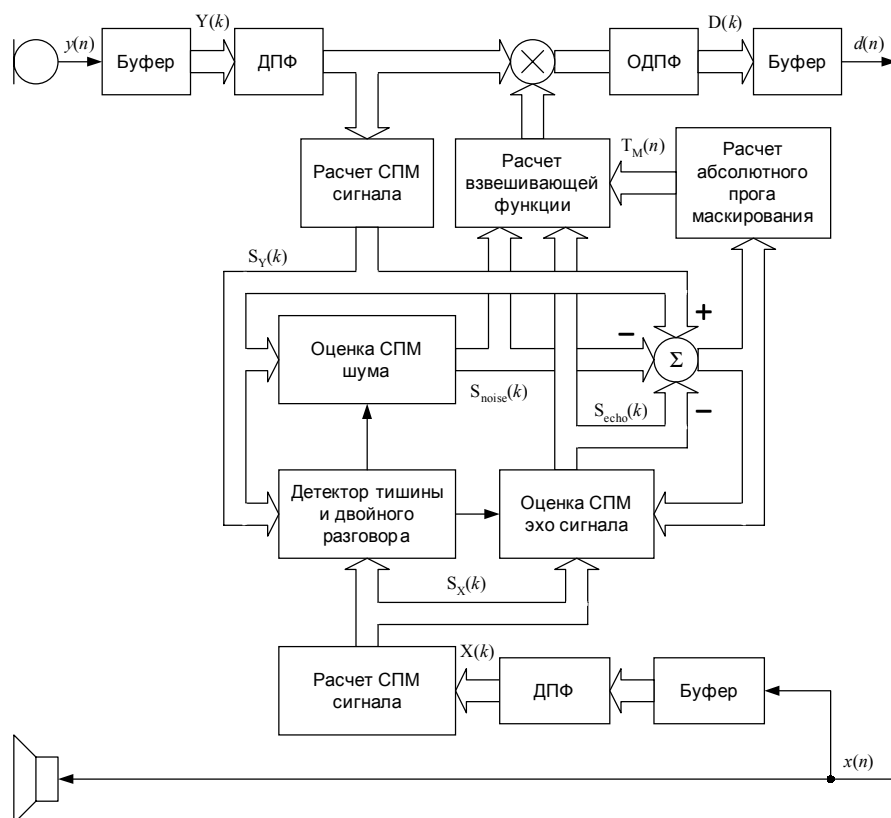


Рис. 1. Комбинированная система подавления акустического эха и шумов окружающей среды.

Отличительной особенностью предлагаемой системы является то, что в ней используется только предсказание спектральной плотности мощности эхо сигнала. По сравнению с подходом используемым в большинстве аналогичных разработок (например [3]), построенных на базе предсказания эхо сигнала во временной области, такой вариант позволяет упростить систему и значительно снизить вычислительную сложность.

Оценки СПМ эхо сигнала и СПМ шумового сигнала осуществляется согласно методике [4]. Абсолютный порог маскирования речевого сигнала рассчитывается по методике приведенной в [5]. Взвешивающая функция вычисляется по формуле:

$$G(k, f) = 1 - \frac{S_{noise}(k, f) + S_{echo}(k, f) - T_M(k, f)}{S_Y(k, f)}, \quad (1)$$

где $S_Y(k)$ – СПМ входного сигнала ближнего диктора $y(n)$; $S_{echo}(k)$ – оценка СПМ эхо сигнала удаленного диктора; $S_{noise}(k)$ – оценка СПМ шумов окружающей среды; $T_M(k)$ – абсолютный порог маскирования чистой речи, f – частота.

Выходной сигнал определяется следующим выражением

$$d(k, f) = \text{ОДПФ}(G(k) \cdot Y(k)), \quad (2)$$

где $Y(k)$ – ДПФ входного сигнала $y(n)$.

Заключение

Сравнение предложенной системы с комбинированной системой использующей для подавления эхо сигнала НМНК адаптивный фильтр и для шумоподавления метод спектрального вычитания с психоакустической мотивацией [3] показало следующие результаты:

– для невысокой длительности акустического эха (порядок фильтра равен 256) обе системы дают практически одинаковое качество подавления эхо сигнала и шумов окружающей среды. Ослабление эхо сигнала $ERLE$ составляло порядка 35-38 дБ и ослабление уровня шума порядка 30-35 дБ (см. рис. 1а);

– при возрастании длительности эха наблюдалось значительное снижение уровня ослабления эхо сигнала для системы на базе НМНК алгоритма и незначительное для предлагаемой системы (см. рис. 2а). Предлагаемая система имеет более низкую вычислительную сложность: выигрыш составляет около 10 раз для низких порядков фильтра и растет с его увеличением. Недостатком алгоритма является его более низкая начальная скорость сходимости, что обусловлено использованием блочных алгоритмов.

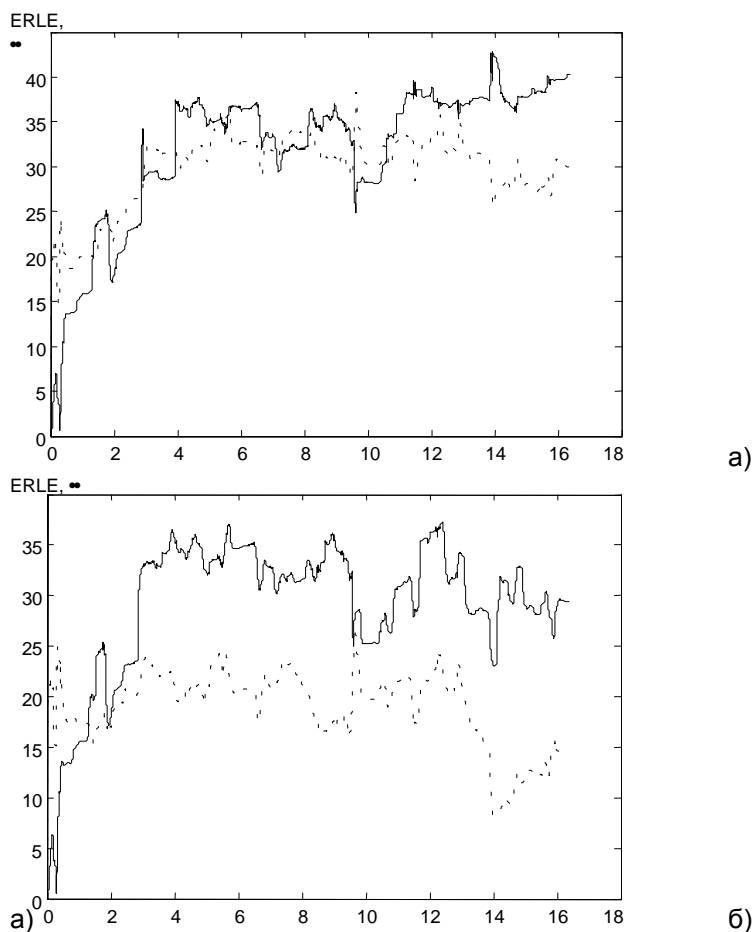


Рис.2. Результаты моделирования систем на базе НМНК адаптивного фильтра (пунктирная линия) и предлагаемой системы (сплошная линия).

Библиография

1. A.Petrovsky, K.Bielawski, A.Anoshenko, Hands-free radiotelephony communication devices with combine front end processing systems: global approaches in the time and frequency domain. Journal of the University of Applied Sciences Mittweida Wissenschaftliche Zeitschrift der Hochschule Mittweida (FH). // Proc. IWKM'98, Band C, Nr.3, 1998, pp.135-142.
2. N. Virag, Speech enhancement based on masking properties of the auditory system. // Proc. IEEE ICASSP, Detroit, MI, May, 1995, pp. 796-799.
3. S.Gustafsson, R. Martin, P.Jax, P. Vary, A Psychoacoustic Approach to Combined Acoustic Echo Cancellation and Noise Reduction // IEEE Transactions on Speech and Audio Processing, Vol. 10, No. 5, July 2002, pp. 245–256.
4. Петровский А.А., Аношенко А.Е. Метод подавления эхо сигнала и шумов окружающей среды на основе спектрального вычитания с психоакустической мотивацией // 4-я Международная конференция “Цифровая обработка сигналов и ее применение”, Москва, 27 февраля – 1 марта 2002, том. 2 – с.450-452.
5. Петровский А.А., Аношенко А.Е. Комбинированная система подавления эха и шумового сигналов на основе обработки в частотной области с использованием психоакустического подхода // 2-я Международная конференция “Цифровая обработка сигналов и ее применение”, Москва, 21-24 сентября 1999, том. 1 – с. 166-169.



SPEECH ENHANCEMENT BASED ON THE PSYCHOACOUSTIC APPROACH

Anoshenko A., Petrovsky A.

The Byelorussian State University of Informatics and Radioelectronics
220027, Minsk, P.Brovki st., 6 (Belarus)
E-mail: anoshenko@mail.ru, palex@it.org.by

The system, presented in this work, is intended for use in hands free communication devices. In such system the signal from a microphone, will be strong distorted. The basic sources of distortions in an input signal are the noise of an environment and acoustic echo from loudspeakers. The system, submitted in this work, of applicability for clearing an input signal of the given deforming components. In the offered combined system the echo signal and noise of an environment cancellation is absent a separate subsystem of echo suppression and the suppression of all deforming components of an input signal is carried out on the basis of a method of spectral subtraction with the account of psychoacoustics features of the man hearing. Use algorithm updating of spectral subtraction which is taking into account psychoacoustics features of the man perception allows to get rid of effect of "musical" tone which is characteristic for traditional algorithm of spectral subtraction.

The offered system carries out cancellation of noise and editing an echo of a signal as follows:

1. The blocks of the input data with overlapping collect, then their Fourier discrete transformation and spectral density of capacity pays off.
2. On a technique described in [1] there are estimations of an echo signal power spectral density (PSD) and of a noise signal PSD.
3. On a technique given in [2] the absolute masking threshold of pure speech pays off.
4. The weighing function pays off:

$$\mathbf{G}(k, f) = 1 - \frac{\mathbf{S}_{noise}(k, f) + \mathbf{S}_{echo}(k, f) - \mathbf{T}_M(k, f)}{\mathbf{S}_Y(k, f)}, \quad (1)$$

where $\mathbf{S}_Y(k)$ is PSD of near end input signal $y(n)$; $\mathbf{S}_{echo}(k)$ is the estimation of echo signal PSD; $\mathbf{S}_{noise}(k)$ is the estimation of environment noise PSD; $\mathbf{T}_M(k)$ is the absolute masking threshold of pure speech, f is frequency.

5. The spectral weighing of an input signal is carried out.

The given system allows to reach the reducing of an echo signal of the order ERLE = 35-38 dB at low echo signal duration and 30-35 дБ at high duration, and also the reducing of noise level of the order 30-35 дБ. Thus the offered system allows receiving 10 multiple prizes in computing complexity in comparison with the traditional decisions on base of NLMS algorithm.

Библиография

1. Anoshenko A.E., Petrovsky A.A Method of Echo Cancellation and Noise Reduction on the Basis of Spectral Subtraction with Psychoacoustics Motivation // The 4 International Conference "Digital signal processing and its applications", Moscow, 2002 February 27 – March 1, Vol.2 – pp. 452-453.
2. Petrovsky A.A., Anoshenko A.E. Combined echo and noise cancellation system based on the frequency domain processing with using the psycoacoustically approach. // The 2nd International Conference "Digital signal processing and its applications", Moscow, 1999 September 21-24, Vol. 1 – pp. 166-169.