

Методика подавления цветных шумов в речевом сигнале

Ермоленко Т.В.,
кандидат технических наук

Фёдоров Е.Е.,
кандидат технических наук



В задачах, связанных с распознаванием речи и идентификацией диктора, важную роль играет предварительная очистка сигнала от шума. В статье предлагается модификация известного метода спектрального вычитания – метода максимального правдоподобия, а также проводится численное сравнение реализованной методики со стандартными пакетами и методами, осуществляющими шумоочистку. Предложенный подход подавления шума в речевом сигнале основан на непрерывном вейвлет-преобразовании, продукционных правилах и учитывает акустические особенности фонетических классов звуков речи.

Noise clipping plays significant role in the problems of speech recognition and speaker identification. A modification of spectral subtraction method based on maximum likelihood estimation is proposed in this article. The numerical comparison of the realized technique, standard packages, and noise suppression methods is carried out. The offered approach of noise suppression in speech signal is based on continuous wavelet transformation, production rules and acoustic features of wide phonetic classes of speech sounds.

Введение

Анализ исследований. Для борьбы с помехами, рассредоточенными по спектру и пересекающимися с областью спектра речи, применяют методы спектрального вычитания [1–3] и пороговую вейвлет-обработку [4, 5].

Методы спектрального вычитания используют преобразование Фурье, их суть состоит в следующем. Вычисляется спектр шума $Y(k)$, после чего зашумлённый сигнал разбивают на фреймы, на каждом p -ом фрейме вычисляется спектр $X_p(k)$, из которого удаляется спектр шума. По полученному спектру $S_p(k)$ сигнал восстанавливается с помощью обратного преобразования:

$$S_p(k) = \begin{cases} H_p(k)X_p(k), & |X_p(k)|^v - \alpha |Y(k)|^v > 0, \\ \beta X_p(k), & \text{иначе} \end{cases}$$

где параметры α , β , v и функция $H_p(k)$ зависят от метода фильтрации.

Различают следующие методы: фильтрация по Берути, Шварцу и Макхоулу, основанная на вычитании энергетического спектра шума; винеровская фильтрация; метод максимального правдоподобия (MLEE); метод, использующий сглаживающий фильтр; метод EVRC (Enhanced Variable Rate Coder), реализованный в детекторе речи для мобильных телефонов; фильтрация по Болу, основанная на вычитании модуля амплитуды шума; метод, созданный на основе правила Эфраима и Малаха (EMSR).

В работе [1] были исследованы различные методы спектрального вычитания на белом и коричневом шуме. Стабильно хорошие результаты в обоих случаях показал метод MLEE.

Методы обработки сигналов, основанные на теории вейвлет-преобразований, лучше адаптированы к локальным свойствам сигнала и, в отличие от оконного преобразования Фурье, обеспечивают подвижное частотно-временное окно, которое сужается при высокой центральной частоте и расширяется при низкой. Теоретически доказано, что на всём множестве допустимых временных оконных функций спектрального оценивания вейвлет обладает лучшей разрешающей способностью по частоте и по времени [6, 7].

Традиционно методы пороговой вейвлет-обработки осуществляют подавление шумовой составляющей в сигнале с помощью быстрого вейвлет-преобразования и порогов [4, 5, 8]. Эти методы реализованы в пакете программ «Wavelet Toolbox» системы Matlab [4, 5]. Процесс понижения уровня шумов состоит из трёх этапов и включает в себя: разложение по вейвлет-базису, преобразование коэффициентов разложения, восстановление сигнала по преобразованным коэффициентам. На этапе преобразования детализирующие коэффициенты, не превышающие порог по абсолютному значению, обнуляются. Остальные детализирующие коэффициенты при использовании мягкого порога уменьшаются по модулю на его значение, при использовании жёсткого порога – остаются неизменными.

Однако быстрое вейвлет-преобразование имеет ряд ограничений [6, 8], которые снимаются при использовании более информативного непрерывного вейвлет-преобразования. Кроме того, все вышеперечисленные методы понижения уровня шума не учитывают фонетическую классификацию звуков речи, хотя на этапе предварительной обработки это могло бы понизить ошибки дальнейшего распознавания и повысить разборчивость речи.

Постановка задачи. На основе методов пороговой вейвлет-обработки разработать методику подавления цветного шума в речевом сигнале, учитывающую акустические особенности широких фонетических классов (ШФК) звуков речи.

Осуществить численное сравнение разработанной методики со стандартными пакетами и методами, осуществляющими шумоочистку.

1. Вейвлет-преобразование сигнала

Непрерывное вейвлет-преобразование одномерного сигнала $x(t)$ состоит в его разложении по базису, сконструированному из базового вейвлета $\psi(t)$ посредством масштабных изменений и смещений [9, 10], и может быть записано в виде:

$$CWT(a, b) = \int_{-\infty}^{\infty} x(t) |a|^{-1/2} \psi\left(\frac{t-b}{a}\right) dt. \quad (1)$$

Обратное вейвлет-преобразование имеет вид [6, 8]:

$$\tilde{x}(t) = \frac{1}{C_{\psi}} \int_{-\infty}^{\infty} \int_0^{\infty} CWT(a, b) |a|^{-1/2} \psi\left(\frac{t-b}{a}\right) \frac{dad b}{a^2}, \quad (2)$$

где C_{ψ} – нормирующий множитель.

Преобразование Фурье $\Psi(\omega)$ базового вейвлета является локализованной функцией с центром $\langle \omega \rangle$ и радиусом Δ_{ψ} , а вейвлет-преобразование для каждого значения масштабирующей переменной a в частотной области представляет собой полосовой фильтр с центральной частотой $\langle \omega \rangle / a$ и шириной полосы пропускания $2\Delta_{\psi} / a$ [10]:

$$win_a = [\langle \omega \rangle / a - \Delta_{\psi} / a; \langle \omega \rangle / a + \Delta_{\psi} / a]. \quad (3)$$

При обработке дискретных сигналов $x(n)$ конечной длины N возникает необходимость в использовании процедур квантования переменных a и b [10]:

$$a = a_0^i, \quad b = nb_0 a_0^i, \quad i, n \in \mathbb{Z}, \quad 1 < a_0 < 2^{1/4}, \quad b \neq 0, \quad (4)$$

а также в переходе к численному интегрированию в (1) и (2). Такой выбор шага изменения масштаба a_0 обусловлен исследованиями, проведёнными в [11], и предотвращает появление значительных колебаний на графике амплитудно-частотной характеристики фильтра.

В результате вейвлет-спектр и обратное вейвлет-преобразование вычисляются следующим образом:

$$d_{ml} = \sum_{n=0}^{N-1} x(n) a_0^{-m/2} \psi_{ml}(n) \Delta t, \quad l \in \overline{0, N-1}, \quad m \in \overline{i_{\min}, i_{\max}}, \quad (5)$$

$$\tilde{x}(n) = C_{\psi}^{-1} \sum_{m=i_{\min}}^{i_{\max}} \sum_{l=0}^{N-1} d_{ml} \psi_{ml}(n) \frac{\Delta a \Delta b}{a_0^{2m}} = C_{\psi}^{-1} \sum_{m=i_{\min}}^{i_{\max}} \sum_{l=0}^{N-1} d_{ml} \psi_{ml}(n) b_0 (a_0 - 1), \quad (6)$$

где Δt – шаг квантования по времени, i_{\min}, i_{\max} – номер минимального и максимального уровня разложения,

$$\psi_{ml}(n) = a_0^{-m/2} \psi(a_0^{-m} n - b_0 l).$$

Выбор начального (i_{\min}) и конечного (i_{\max}) уровней разложения, на которых проводятся обработка и анализ сигнала, осуществляется с учётом границ частотного диапазона сигнала $[f_1; f_2]$ (для речевых сигналов $f_1=100$ Гц; $f_2=8$ кГц), а также частоты дискретизации f_d .

Пользуясь (3)–(4), для банка вейвлет-фильтров можно получить значения границ полос пропускания на каждом уровне разложения m :



$$win_m = [\langle \omega \rangle / a_0^m - \Delta\Psi / a_0^m; \langle \omega \rangle / a_0^m + \Delta\Psi / a_0^m], i_{\min} \leq m \leq i_{\max}.$$

Таким образом, увеличение уровня разложения приводит к понижению центральной частоты соответствующего фильтра; следовательно, для анализа частотного диапазона $[f_1; f_2]$ с помощью банка вейвлет-фильтров необходимо, чтобы:

$$\frac{\langle \omega \rangle}{a_0^{i_{\min}}} = \frac{2\pi f_2}{f_d}, \quad \frac{\langle \omega \rangle}{a_0^{i_{\max}}} = \frac{2\pi f_1}{f_d},$$

откуда получаем границы диапазона уровней разложения:

$$i_{\min} = \text{int} \left(\log_{a_0} \frac{f_d \langle \omega \rangle}{2\pi f_2} \right), \quad i_{\max} = \text{int} \left(\log_{a_0} \frac{f_d \langle \omega \rangle}{2\pi f_1} \right),$$

где int – округление до ближайшего целого.

2. Подавление цветного шума в речевом сигнале

Методика подавления цветного шума в сигнале базируется на методах вейвлет-анализа, шумоочистки, продукционных правил, наряду с адаптацией к шуму учитывает акустические особенности ШФК звуков речи и состоит из четырёх этапов:

- вычисление порогов по образцу шума (обучение шуму);
- маркировка фреймов сигнала;
- удаление шума из вейвлет-образа сигнала;
- восстановление сигнала по обновлённым коэффициентам.

2.1. Обучение шуму

На этапе обучения шуму выполняется:

- 1) Вейвлет-разложение (5) сигнала $\varepsilon(n)$ длиной N_ε , содержащего образец шума, по уровням $i = \overline{i_{\min}; i_{\max}}$.

Затем этот сигнал разбивается на фреймы длиной ΔN отсчётов. В пределах одного фрейма сигнал является однородным. Для речевых сигналов длина фрейма $\Delta N \approx \Delta T f_d$, т.е. зависит от периода основного тона ΔT , который составляет не более 0,02с (50 Гц).

- 2) Для каждого s -го фрейма вычисляется мера контрастности (7):

$$C(m, s) = \lg \left(\frac{E_s(m)}{\sum_{j=j_{\min}}^m E_s(j)} \right), \quad m \in \overline{i_{\min} + 1; i_{\max}}, \quad (7)$$

где $E_s(m)$ – энергия вейвлет-спектра сигнала $\varepsilon(n)$ на уровне разложения m :

$$E_s(m) = \sum_{n=(s-1)\Delta N}^{s\Delta N} d_{mn}^2, \quad s \in \overline{1; \text{int}(N_\varepsilon / \Delta N)}. \quad (8)$$

Величина (7) характеризует распределение энергии спектра по уровням разложения и позволяет проводить анализ временной динамики энергии спектра сигнала. Операция логарифмирования даёт сглаженные характеристики, что предотвращает случайные всплески анализируемых величин в сигнале.

3) На каждом m -ом уровне разложения вычисляются несмещённые оценки математического ожидания $Aver(m)$ и дисперсии $D(m)$ величины $\{C(m, s)\}_{s=1}^{\text{int}(N_\varepsilon / \Delta N)}$, в результате чего получаются пороги

$$\alpha(m) = Aver(m) - \sqrt{D(m)}, \quad \beta(m) = Aver(m) + \sqrt{D(m)}$$

и усреднённые энергетические характеристики шума

$$E_\varepsilon(m) = \frac{\sum_{n=0}^{N_\varepsilon - 1} d_{mn}^2}{N_\varepsilon}. \quad (8)$$

2.2. Классификация фреймов

Классификация фреймов речевого сигнала $x(n)$, содержащего шум, выполняется на основе (7). При этом каждый s -ый фрейм может содержать один из четырёх ШФК звуков: только шум ($s \in Noise$); вокализованный звук ($s \in Voc$); шумный глухой щелевой или смычно-щелевой звук (Sh); шумный глухой смычный звук ($s \in P$).

Спектр вокализованных звуков, имеющих формантную структуру, сосредоточен, в основном, в частотной области от 100 Гц до 2.6 кГц [12].

Шумные глухие щелевые звуки и аффрикаты нестационарны, не имеют определённых спектральных параметров и характеризуются большой интенсивностью в диапазоне высоких частот от 4 кГц до 8 кГц в зависимости от звука и диктора [12–14].

Для проведения классификации выделяют два множества масштабов (уровней разложения):

— $M_{voc} = \{m : m_{voc} \leq m \leq i_{max}\}$ соответствует диапазону частот, где сосредоточена энергия вокализованных звуков (100–300 Гц), поэтому значения функции (7) при $m \in M_{voc}$ для фреймов, содержащих вокализованный звук, достаточно велики и превышают значения (7), полученные для фреймов сигнала $\varepsilon(n)$;

— $M_{sh} = \{m : i_{min} \leq m \leq m_{sh}\}$ соответствует высокочастотной части спектра (более 4 кГц), в которой сосредоточена энергия шумных глухих щелевых или смычно-щелевых звуков.

Для классификации каждого s -го фрейма сигнала $x(n)$ необходимы следующие действия.

1) Выполнить вейвлет-преобразование (5) сигнала $x(n)$.

2) Вычислить меру контрастности $C(m,s)$ для каждого s -го фрейма сигнала $x(n)$ согласно (7).

3) Отнести s -ый фрейм к одному из ШФК звуков по следующим правилам:

$$\forall m : \alpha(m) \leq C(m,s) \leq \beta(m) \rightarrow s \in \text{Noise} \vee P,$$

$$(\forall m \in M_{sh} : \alpha(m) > C(m,s)) \wedge (\exists n \in M_{voc} : C(n,s) > \beta(n)) \rightarrow s \in \text{Voc},$$

$$\exists m \in M_{sh} : \beta(m) < C(m,s) \rightarrow s \in \text{Sh},$$

а также получить массив маркировки фреймов $\{a_s\}_{s=1}^{N/\Delta N}$ на основе их классификации:

$$a_s = \begin{cases} 1, & s \in \text{Noise} \vee P \\ 2, & s \in \text{Sh} \\ 3, & s \in \text{Voc} \end{cases}$$

4) Для $i \in \overline{3, N/\Delta N - 4}$ провести корректировку маркировки фреймов:

4а) если один или два фрейма помечены неверно внутри одного звука:

$$a_{i-2} = a_{i-1} \wedge a_{i-1} \neq a_i \wedge a_{i-1} = a_{i+1} \wedge a_{i+1} = a_{i+2} \rightarrow a_i = a_{i+1}$$

(например, 11211→11111),

$$a_{i-2} = a_{i-1} \wedge a_{i-1} \neq a_i \wedge a_{i-1} \neq a_{i+1} \wedge a_{i-1} = a_{i+2} \wedge a_{i+2} = a_{i+3} \rightarrow$$

$$\rightarrow a_i = a_{i+2} \wedge a_{i+1} = a_{i+2}$$

(например, 112211→11111 или 112311→11111),

$$a_{i-2} = a_{i-1} \wedge a_{i-1} \neq a_i \wedge a_i \neq a_{i+1} \wedge a_{i+1} \neq a_{i+2} \wedge a_{i+2} \neq a_{i+3} \wedge a_{i+3} = a_{i+4} \rightarrow$$

$$\rightarrow a_i = a_{i+3} \wedge a_{i+1} = a_{i+3} \wedge a_{i+2} = a_{i+3}$$

(например, 1121211→111111, или 1121311→111111, или 1123211→111111);

4б) если один или два фрейма помечены неверно на границе звуков:

$$a_{i-3} = a_{i-2} \wedge a_{i-2} = a_{i-1} \wedge a_{i-1} \neq a_i \wedge a_i \neq a_{i+1} \wedge a_{i-1} \neq a_{i+1} \wedge a_{i+1} = a_{i+2} \wedge$$

$$\wedge a_{i+2} = a_{i+3} \rightarrow a_i = a_{i+1}$$

(например, 1112333→1111333),

$$a_{i-3} = a_{i-2} \wedge a_{i-2} = a_{i-1} \wedge a_{i-1} \neq a_i \wedge a_i \neq a_{i+2} \wedge a_{i-1} \neq a_{i+1} \wedge a_{i+1} \neq a_{i+2} \wedge$$

$$\wedge a_{i-1} \neq a_{i+2} \wedge a_{i+2} = a_{i+3} \wedge a_{i+3} = a_{i+4} \rightarrow a_i = a_{i-1} \wedge a_{i+1} = a_{i+2}$$

(например, 11122333→11113333),

$$a_{i-3} = a_{i-2} \wedge a_{i-2} = a_{i-1} \wedge a_{i-1} \neq a_i \wedge a_i = a_{i+2} \wedge a_{i-1} \neq a_{i+1} \wedge a_{i+1} \neq a_{i+2} \wedge$$

$$\wedge a_{i-1} \neq a_{i+2} \wedge a_{i+2} = a_{i+3} \wedge a_{i+3} = a_{i+4} \rightarrow a_{i+1} = a_{i+2}$$

(например, 11132333→11133333),

$$a_{i-3} = a_{i-2} \wedge a_{i-2} = a_{i-1} \wedge a_{i-1} \neq a_i \wedge a_i \neq a_{i+2} \wedge a_{i-1} = a_{i+1} \wedge a_{i+1} \neq a_{i+2} \wedge a_{i-1} \neq a_{i+2} \wedge a_{i+2} = a_{i+3} \wedge a_{i+3} = a_{i+4} \rightarrow a_i = a_{i-1}$$

(например, 11121333→11111333).

2.3. Удаление шума из вейвлет-образа сигнала

На этом этапе с использованием модифицированного метода MLEE и (8) производится изменение вейвлет-коэффициентов в соответствии с типом s -го фрейма. Возможны следующие варианты.

1) $a_s = 1$ (фрейм содержит шум):

$$\tilde{d}_{mn} = 0, \quad n \in \overline{(s-1)\Delta N, s\Delta N}, \quad m \in \overline{i_{\min} + 1; i_{\max}};$$

2) $a_s = 2$ (фрейм содержит шумные глухие щелевые и смычно-щелевые звуки):

$$\tilde{d}_{mn} = \begin{cases} \frac{1}{2} \left(1 + \sqrt{\frac{d_{mn}^2 - E_\varepsilon(m)}{d_{mn}^2}} \right) d_{mn}, & (d_{mn}^2 > E_\varepsilon(m)) \wedge m \in M_{sh}, \\ 0, & \text{иначе} \end{cases}$$

$$n \in \overline{(s-1)\Delta N, s\Delta N}, \quad m \in \overline{i_{\min} + 1; i_{\max}};$$

3) $a_s = 3$ (фрейм содержит вокализованные звуки):

$$\tilde{d}_{mn} = \begin{cases} \frac{1}{2} \left(1 + \sqrt{\frac{d_{mn}^2 - E_\varepsilon(m)}{d_{mn}^2}} \right) d_{mn}, & (d_{mn}^2 > E_\varepsilon(m)) \wedge m \in M_{voc} \\ 0, & \text{иначе} \end{cases},$$

$$n \in \overline{(s-1)\Delta N, s\Delta N}, \quad m \in \overline{i_{\min} + 1; i_{\max}}.$$

2.4. Восстановление сигнала

По обновлённым коэффициентам вейвлет-спектра с помощью обратного преобразования (6) сигнал восстанавливается:

$$\tilde{x}(n) = \sum_{m=i_{\min}}^{i_{\max}} \sum_{l=0}^{N-1} \tilde{d}_{ml} \psi_{ml}, \quad n \in \overline{0, N-1}.$$

Нормирующий множитель C_ψ вычисляется как отношение максимумов:

$$C_\psi = \frac{\max_n x(n)}{\max_n \tilde{x}(n)}.$$

Тогда окончательно имеем:

$$y(n) = C_\psi \tilde{x}(n), \quad n \in \overline{0, N-1}.$$

3. Численное исследование эффективности подавления шума с помощью разработанной методики

Для проведения численного исследования предложенная методика была реализована в программном комплексе.

Для предложенной методики использовался вейвлет Морле. В качестве сигнала было выбрано слово «Саша» с частотой дискретизации 22050 Гц, 8 бит, моно. Этот сигнал зашумлялся коричневым, розовым и белым шумом с помощью программы Adobe Audition.

На рис. 1–10 приведены: а) исходный сигнал (рис. 1), зашумлённый коричневым шумом (рис. 2); б) сигналы, очищенные с помощью аппроксимированного непрерывного вейвлет-преобразования (авторская методика) (рис. 3), стандартных методов спектрального вычитания и пакетов работы со звуком (рис. 4–10).

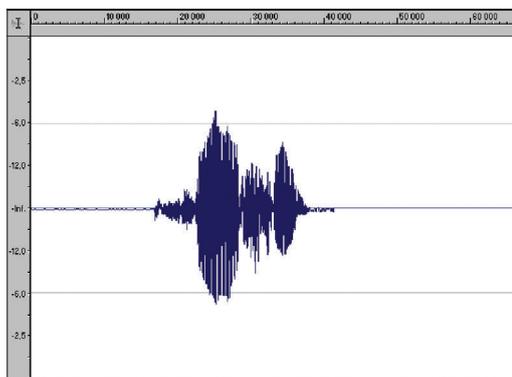


Рис. 1. Исходный сигнал

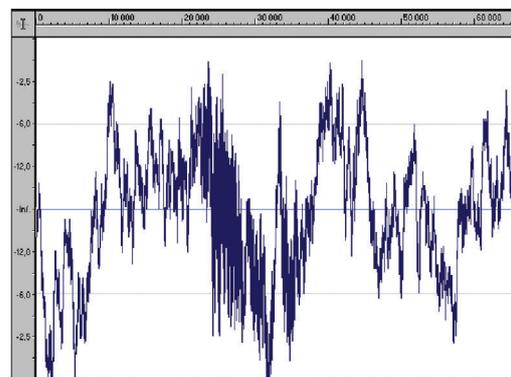


Рис. 2. Сигнал, зашумлённый коричневым шумом

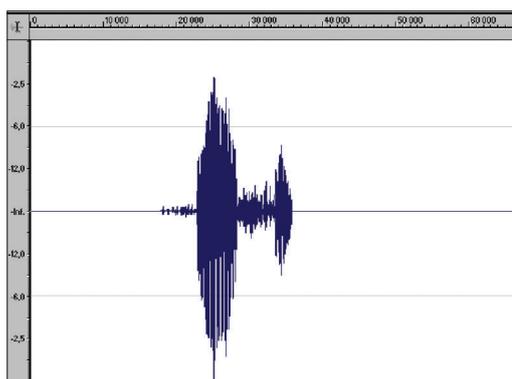


Рис. 3. Сигнал, очищенный с помощью вейвлет-преобразования Морле

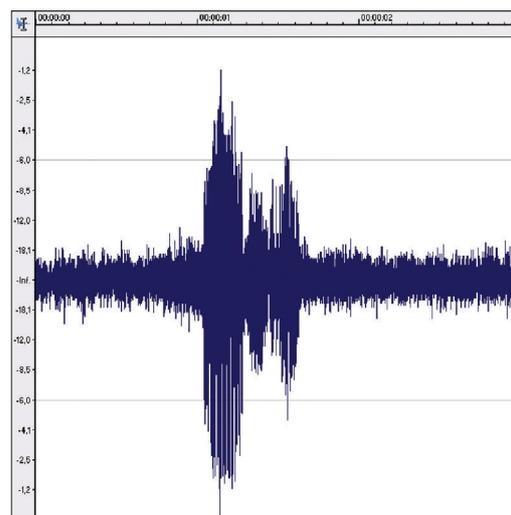


Рис. 4. Сигнал, очищенный с помощью MLEE

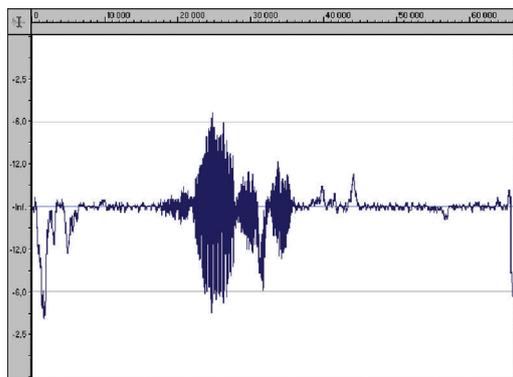


Рис. 5. Сигнал, очищенный с помощью пакета Clear Voice Denoiser

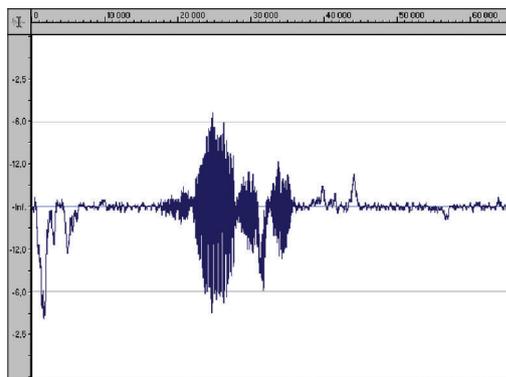


Рис. 6. Сигнал, очищенный с помощью пакета Adobe Audition v1

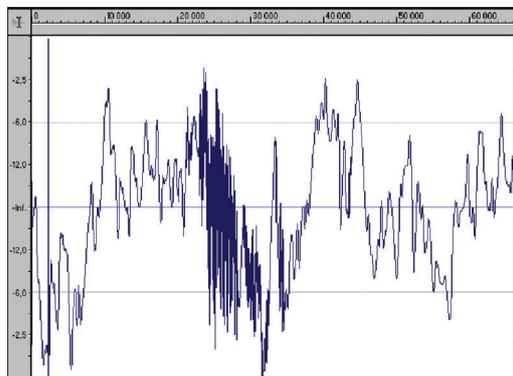


Рис. 7. Сигнал, очищенный с помощью жёсткого порога пакета Wavelets Extension Mathcad v11

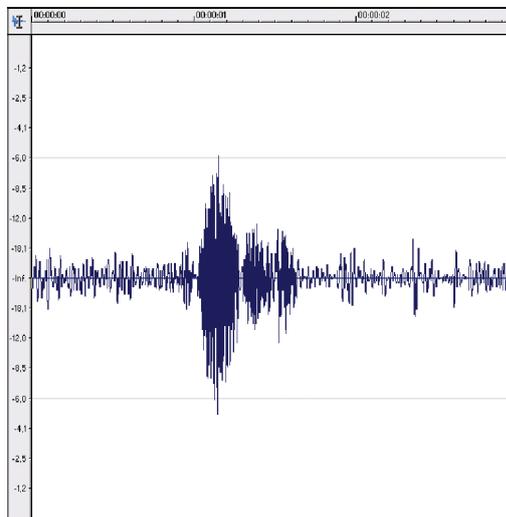


Рис. 8. Сигнал, очищенный с помощью фильтрации по Болу

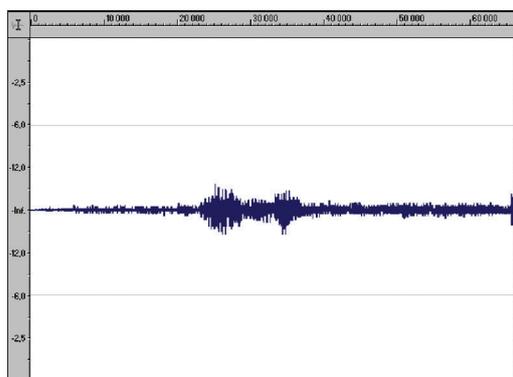


Рис. 9. Сигнал, очищенный с помощью EMSR

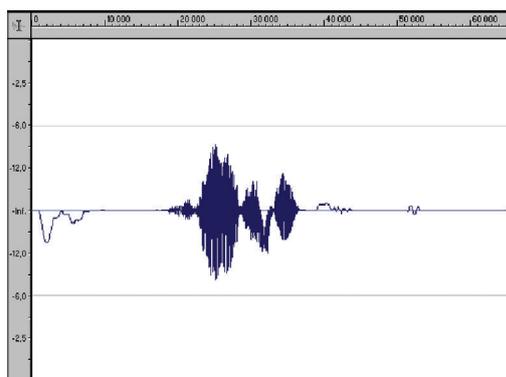


Рис. 10. Сигнал, очищенный с помощью пакета Nero WaveEditor v3

Чтобы провести оценку эффективности и сравнительный анализ результатов метода обесшумливания, реализованного в авторском программном комплексе, и методов обесшум-

ливания, реализованных в стандартных пакетах работы со звуком, авторами введена мера близости очищенного сигнала к исходному:

$$\delta = \sum_{s=1}^{N/\Delta N} \left(\frac{\sum_{n=0}^{\Delta N} (x_s^{source}((s-1)\Delta N + n) - x_s^{clear}((s-1)\Delta N + n))^2}{\sum_{n=0}^{\Delta N} (x_s^{source}((s-1)\Delta N + n))^2} \right),$$

где $x_s^{source}(n)$ – s -й фрейм исходного сигнала, $x_s^{clear}(n)$ – s -й фрейм очищенного сигнала, ΔN – длина фрейма.

Наилучшим будем считать метод, у которого $\delta \rightarrow \min$.

В таблице 1 приведены результаты численного исследования эффективности работы методов обесшумливания сигнала, зашумлённого цветными шумами. Как видно из таблицы 1, наилучшие результаты даёт авторская методика.

Таблица 1

Результаты численного исследования

Методики и пакеты	Значение δ		
	белый, отношение сигнал/шум 18дб	коричневый, отношение сигнал/шум 12дб	розовый, отношение сигнал/шум 25дб
Пакеты:			
Adobe Audition	12377	29419	24622
Clear Voice Denoiser	18122	356624	2840980
Nero WaveEditor	27442	14628	1342810
Wavelets Extension	17402	34514695	5918960
Методики:			
авторская	619796	12392900	6278350
одноканальный адаптивный фильтр	343049	114899000	9123650
Болла	41266	237512	387731
MLEE	59067	87095	180327
EMSR	11204	131935	202621
сглаживающий фильтр	113203	763875	8322940
EVRC	17402	34514700	5918960
жесткий порог	619796	12392900	6278350

Заключение

Новизна. В статье была предложена модификация метода MLEE подавления цветного шума в сигнале, основанная на пороговой вейвлет-обработке и продукционных правилах. Благодаря учёту акустических особенностей ШФК звуков речи, подобный подход на этапе предварительной обработки позволяет понизить ошибки дальнейшего распознавания и повысить разборчивость речи.

Ермоленко Т.В., Фёдоров Е.Е.

Методика подавления цветных шумов в речевом сигнале

Практическое значение. Основные результаты данной работы предназначены для реализации в системах распознавания речи и идентификации диктора, а также для создания интеллектуальных систем управления, в которых команды поступают на естественном языке.

Литература

1. *Фёдоров Е.Е.* Модели и методы преобразования речевых сигналов. Донецк: Норд-Пресс, 2006. 260 с.
2. *Reccione M.C.* The enhanced variable rate coder: Toll quality speech for CDMA // International Journal of Speech Technology. 1999. № 2. P. 305–315.
3. *Thiemann J.* Acoustic Noise Suppression for Speech Signals using Auditory Masking Effects. Montreal: McGill University, 2001. 83 p.
4. *Дьяконов В.П.* Вейвлеты. От теории к практике. М.: СОЛОН-Р, 2002. 448 с.
5. *Дьяконов В., Абраменкова И.* MATLAB. Обработка сигналов и изображений: Специальный справочник. СПб.: Питер, 2002. 608 с.
6. *Чуи К.* Введение в вейвлеты: Пер. с англ. М.: Мир, 2001. 412 с.
7. *Zhenilo V.R., Zhenilo M.V., Kalyuzhny D.N.* Fourier-Gauss Transform: Speech Signal Decomposition into Sonels. // Proc. International Conf. on Speech and Computer (SPECOM'2007). Moscow (Russia). 2007. P. 259–29.
8. *Малла С.* Вейвлеты в обработке сигналов: Пер. с англ. М.: Мир, 2005. 671 с.
9. *Дремин И.М., Иванов О.В., Нечитайло В.А.* Вейвлеты и их использование // Успехи физических наук. 2001. Т. 171, №5. С. 465–501.
10. *Астафьева Н.М.* Вейвлет-анализ: основы теории и некоторые приложения // Успехи физических наук. 1998. №11. С. 1145–1170.
11. *Шитов А.Б.* Разработка численных методов и программ, связанных с применением вейвлет-анализа для моделирования и обработки экспериментальных данных: Дис. канд. техн. наук: 05.13.18. М., 2001. 125 с.
12. *Златоустова Л.В.* Фонетические единицы русской речи. М.: МГУ, 1981. 108 с.
13. *Фёдоров Е.Е., Шевцова И.А.* Численное исследование шипящих согласных звуков // Искусственный интеллект. 2004. №4. С. 661–665.
14. *Фёдоров Е.Е., Шевцова И.А.* Количественный анализ шумных глухих щелевых и смычно-щелевых звуков // Искусственный интеллект. 2005. №3. С. 308–313.

Ермоленко Т.В.

Кандидат технических наук, научный сотрудник отдела распознавания речевых образов Института проблем искусственного интеллекта МОН и НАН Украины. Распознаванием и обработкой речевых сигналов занимается с 2002 года.
etv@iai.donetsk.ua

Фёдоров Е.Е.

Доцент кафедры специализированных компьютерных систем Донецкой академии автомобильного транспорта.

В 2003 году защитил кандидатскую диссертацию в Институте проблем искусственного интеллекта МОН и НАН Украины. Автор более 75 научных публикаций, в том числе трёх монографий, посвящённых моделям и методам преобразования речевых сигналов. Основная область интересов: идентификация и верификация диктора, распознавание и синтез речи, анализ и синтез языковых конструкций, вибродиагностика и шумодиагностика, медицинская диагностика (анализ биосигнала).