

УДК 004.934

В. В. Савченко, профессор, зав. каф., e-mail: svv@lunn.ru,
Нижегородский государственный лингвистический университет, г. Нижний Новгород

Распознавание речевых команд методом фонетического декодирования слов с подавлением фонового шума

Предложен новый, помехоустойчивый метод фонетического декодирования слов для коммерческих (недорогих и надежных) систем голосового управления робототехникой. В его основу положена идея "исправления" выборочных оценок автокорреляционных матриц гласных фонем путем вычитания из их элементов на главной диагонали оценки дисперсии шумового фона. Показано, что достигаемый эффект состоит в многократном (11,8 дБ и более) увеличении отношения сигнал/шум на входе решающего устройства.

Ключевые слова: речь, распознавание речи, речевая команда, голосовое управление, шум, помехи, борьба с помехами, защита от помех, помехозащищенность

Введение

Метод фонетического декодирования слов (МФДС) предложен в работах [1–4] для повышения точности и быстродействия автоматического распознавания речи (АРР) в системах голосового управления робототехникой. Со ссылкой на информационную теорию восприятия речи [5] данный метод изначально его авторы позиционировали как альтернативу каноническому подходу к АРР, реализованному в большой группе известных методов [6–9] на основе математического аппарата скрытых марковских моделей речевого сигнала [10]. Их общий принципиальный недостаток — использование многозатратной процедуры выравнивания слов по темпу речи диктора. В основу же МФДС положена идея выделения множества информативных, в частности гласных или вокализованных, фонем $\{x_r\}$ из фонетической базы данных (ФБД) русского языка на роль конечной группы объектов АРР с наиболее ярко выраженными различиями в теоретико-информационном смысле [4]. В данном случае алгоритм АРР сводится к фонетическому послоговому декодированию слов на сильно редуцированном множестве гласных фонем, а процедура динамического выравнивания слов утрачивает всякий смысл. Именно в этом преимуществе авторы МФДС первоначально видели его главное обоснование. Между тем, существует и еще один важный мотив, возможно, не менее актуальный для практики, а именно [3] повышение помехоустойчивости АРР с использованием редуцированного МФДС. Цель статьи — исследование эффективности метода

при действии интенсивного внешнего (фонового) шума, а также поиски путей повышения его помехозащищенности. Указанная цель направлена на преодоление актуальной научно-технической проблемы — защиты АРР от внешних помех [11] на пути к широкому распространению систем и технологий голосового управления робототехникой в быту, на транспорте и на производстве.

1. Постановка задачи

Пусть задана аддитивная модель речевого сигнала

$$x = x_r + \eta, \quad r = \overline{1, R}, \quad (1)$$

на интервале квазистационарности его текущего вокализованного отрезка длиной T_r . При частоте дискретизации сигнала F в результате будем иметь $N_r = FT_r$ отсчетов данных. Рассматривая этот сигнал в режиме неперекрывающегося "скользящего окна" длиной в n отсчетов, получаем $L_r = N_r/n$ независимых в совокупности массивов данных $x_l = x_{r,l} + \eta_l$ размерностью n каждый. Здесь $x_{r,l}$, η_l — n -векторы (столбцы) отсчетов r -й фонемы x_r и фонового шума η соответственно в дискретном времени $l = 1, 2, \dots, L_r$. Так, например, при $F = 8$ кГц, $T_r = 80 \dots 120$ мс и $n = 20$ (рекомендуемые [1–4] значения параметров МФДС) получим объем выборки многомерных наблюдений над речевым сигналом (1) $L_r = 30 \dots 50$, что представляется вполне достаточным для обеспечения высокой помехоустойчивости МФДС.

Предположим далее [3, 12, 13], что и полезный сигнал x_r , и шум η из выражения (1) подчиняются

многомерному гауссовскому закону распределения с нулевыми средними и автокорреляционными $(n \times n)$ -матрицами (АКМ) \mathbf{K}_r и $\mathbf{K}_\eta = \sigma_\eta^2 \mathbf{I}_n$ соответственно, где $n = 10...20 \ll N$, $\sigma_\eta^2 = \text{const}$ — дисперсия или средняя мощность шума, предполагаемого стационарным; \mathbf{I}_n — единичная $(n \times n)$ -матрица. В таком случае гауссовским же будет и закон распределения суммарного речевого сигнала (1) с АКМ, равной $\mathbf{K}_x = \mathbf{K}_r + \sigma_\eta^2 \mathbf{I}_n$. Задача АРР при этом сводится [13] к многоальтернативной проверке статистических гипотез:

$$W_r: \mathbf{K}_x = \mathbf{K}_r, r = \overline{1, R}, \quad (2)$$

в отношении АКМ наблюдаемой фонемы x_r . Это стандартная задача статистической радиотехники [14]. Ее оптимальное (в байесовском смысле) решение приводит [12—14] к R -канальной обработке речевого сигнала (1) с предварительной настройкой каждого канала на заданную ФБД $\{\mathbf{K}_r\}$.

В терминах общесистемного критерия минимума величины информационного рассогласования (ВИР) по Кульбаку—Лейблеру [15] запишем следующее выражение для оптимальной решающей статистики:

$$\rho_r(x) = 0,5[\text{tr}(\widehat{\mathbf{K}}_x \cdot \mathbf{K}_r^{-1}) - \ln|\widehat{\mathbf{K}}_x \cdot \mathbf{K}_r^{-1}| - n], \quad r = \overline{1, R}, \quad (3)$$

где

$$\widehat{\mathbf{K}}_x = L^{-1} \sum_{l=1}^L x_l x_l^T \quad (4)$$

— статистическая оценка АКМ речевого сигнала (1) на интервале действия гласной фонемы x_r ; $\text{tr}(\cdot)$, $|\cdot|$ — соответственно след и определитель квадратной $(n \times n)$ -матрицы, $(\cdot)^T$ — символ транспонирования векторов. При этом учтена естественная центрированность речевого сигнала $\mathbf{M}\{x_r\} = 0$ на выходе речевого тракта диктора [12]. Здесь $\mathbf{M}\{\cdot\}$ — символ математического ожидания. А при учете осуществляемой по МФДС [1—5] нормировки минимальных собственных чисел эталонных АКМ $\{\mathbf{K}_r\}$ к некоторому постоянному уровню $\sigma_0^2 = \text{const}$, перепишем выражение (3) в упрощенном виде:

$$\rho_r(x) = 0,5[\text{tr}(\widehat{\mathbf{K}}_x \cdot \mathbf{K}_r^{-1}) - n]. \quad (5)$$

Или в терминах обеляющего фильтра (ОФ) $(n - 1)$ -го порядка [4]

$$y_{r,l} = a_r^T x_l, l = \overline{1, L}, \quad (6)$$

будем иметь выражение для ВИР [5]

$$\rho_r(x) = 0,5[\sigma_r^2(x)/\sigma_0^2 - 1], r = \overline{1, R}, \quad (7)$$

где

$$\sigma_r^2(x) = L^{-1} \sum_{l=1}^L y_{r,l}^2 \quad (8)$$

— выборочная дисперсия отклика r -го ОФ (6) на речевой сигнал (1);

$$a_r = \mathbf{K}_r^{-1} e / (e^T \mathbf{K}_r^{-1} e) —$$

n -вектор его коэффициентов; $e = (1, 0, \dots, 0)^T$. Здесь $e^T \mathbf{K}_r^{-1} e = \rho_r$ — это средняя мощность речевого сигнала на интервале действия r -й фонемы. Выражения (2)—(8) в совокупности определяют общую формулировку МФДС в задаче АРР. Решение задачи (2) в пользу одной из гипотез W_v , $v \leq R$, принимается здесь по признаку достижения минимизации ВИР из (6) при равенстве $r = v$. Исследуем его помехоустойчивость.

2. Приведенное отношение сигнал/шум

По логике применения МФДС [1—4] точность АРР может быть охарактеризована в общем случае вероятностью перепутывания v -й и r -й фонем:

$$\begin{aligned} \beta_{v \rightarrow r} &\triangleq P\{\rho_r(x) < \rho_v(x) | x \in x_v\} = \\ &= P\{\sigma_r^2(x) < \sigma_v^2(x) | x \in x_v\}, v \neq r \leq R, \end{aligned}$$

которая в первом приближении сводится к следующему выражению [14]:

$$\beta_{v \rightarrow r} = [1 - \Phi_{L,L}(1 + 2\rho_{v/r})]. \quad (9)$$

Здесь обозначено

$$\begin{aligned} \rho_{v/r} &\triangleq \rho_r(x) | x \in x_v = \\ &= 0,5 \left[\frac{\mathbf{M}\{\sigma_r^2(x_v)\}}{\sigma_0^2} - 1 \right] = 0,5 \left[\frac{\sigma_{r/v}^2}{\sigma_0^2} - 1 \right] \quad (10) \end{aligned}$$

— среднее (статистическое) значение ВИР (7) между v -й и r -й фонемами в отсутствие шума; $\Phi_{L,L}(\cdot)$ — интегральная функция Φ -распределения (Фишера) с (L, L) степенями свободы. В качестве примера рассмотрим случай характерных на множестве гласных фонем [3, 4] значений параметров МФДС: $L = 30$ и $2\rho_{v/r} = 0,7...1,0$. Согласно таблицам Φ -распределения [16] в таком случае будем иметь $\beta_{v,r} = 0,07...0,03$, т. е. достаточно мало по меркам фонетического декодирования слов. Ситуация меняется при учете фонового шума из выражения (1).

В этом случае выражение (2) может быть переписано следующим образом:

$$\mathbf{K}_x = \mathbf{K}_r + \sigma_\eta^2 \mathbf{I}_n. \quad (11)$$

Тогда из выражения (10) получим

$$\rho_{v/r} = 0,5 \left[\frac{\sigma_{r/v}^2 + \sigma_{r/\eta}^2}{\sigma_0^2} - 1 \right], \quad v, r \leq R,$$

где $\sigma_{r/\eta}^2 = (a_r^\top \cdot a_r) \sigma_\eta^2 = c_r \sigma_\eta^2$ — приведенная к выходу ОФ (6) дисперсия шума. И как результат, запишем следующее равенство:

$$\beta_{v,r} = \left[1 - \Phi_{L, L} \left(\frac{1 + 2\rho_{v/r} + q_r^{-2}}{1 + q_r^{-2}} \right) \right], \quad (12)$$

где

$$q_r^2 \triangleq (\sigma_0^2 / \sigma_{r/\eta}^2) - \quad (13)$$

приведенное (к выходу ОФ) отношение сигнал/шум (ОСШ). Как видим, величина ОСШ может служить объективной характеристикой точности АРР по МФДС при действии фонового шума определенной дисперсии σ_η^2 .

В связи с введенным понятием отметим, что речевой вокализованный сигнал x_r сильно ослабляется в настроенном на него ОФ (6). Поэтому его приведенная средняя мощность резко уменьшается: $\sigma_0^2 \ll \rho_r$. И наоборот, белый шум на выходе того же ОФ в несколько раз усиливается по своей дисперсии σ_η^2 , так как применительно к гласным фонемам выполняется соотношение $(a_r^\top \cdot a_r) \gg 1$. Нетрудно понять, в таком случае, что при действии интенсивных внешних помех, когда ОСШ сокращается до уровня $q^2 \ll 1$, вероятность перепутывания любых пар фонем согласно (12) резко возрастает и в пределе достигает максимума 0,5. А это означает, что принимаемые согласно МФДС решения в пользу той или иной гипотезы в задаче (2) при действии даже не слишком интенсивного шума по существу принимаются наугад. Понятно, что для всех других методов АРР полученные оценки помехоустойчивости, по крайней мере, не улучшаются. Поэтому далее предпринята попытка всемерно ослабить указанную проблему, хотя бы, применительно к МФДС.

3. Коэффициент подавления шума

Отталкиваясь от выражения (11), преобразуем оценку АКМ из (5) к модифицированному виду

$$\hat{\mathbf{K}}_{x/\eta} = \hat{\mathbf{K}}_x - \sigma_\eta^2 \mathbf{I}_n. \quad (14)$$

По существу, этим взят на вооружение известный [17] метод подавления шума — минимального

собственного числа АКМ. Однако он реализуется в (14) с одной принципиальной оговоркой: дисперсия фонового шума σ_η^2 не является минимальным собственным числом для суммарной матрицы \mathbf{K}_x из выражения (11), как это имеет место в упомянутом здесь методе. В нашем случае ее оценка $\hat{\sigma}_\eta^2$ может быть получена, причем с достаточно высокой степенью точности, по стандартной формуле средней квадратичной величины (8) — в паузах между слогами текущей речевой команды. Напомним при этом [3, 4], что послоговое произнесение команд диктором является принципиальной особенностью МФДС в задачах голосового управления робототехникой.

При учете оценки (14) перепишем выражение (5) следующим образом:

$$\rho_{r/\eta}(x) = 0,5 [\text{tr}(\hat{\mathbf{K}}_x - \hat{\sigma}_\eta^2 \mathbf{I}_n) \mathbf{K}_r^{-1}] - n]. \quad (15)$$

Из этого выражения по аналогии с (7) получим модифицированное выражение для решающей статистики МФДС вида

$$\begin{aligned} \rho_{r/\eta}(x) &= 0,5 [(\sigma_r^2(x) + (\hat{\sigma}_{r/\eta}^2 - \sigma_{r/\eta}^2)) / \sigma_0^2 - 1] = \\ &= 0,5 [(\sigma_r^2(x) + c_r(\hat{\sigma}_\eta^2 - \sigma_\eta^2)) / \sigma_0^2 - 1] \end{aligned} \quad (16)$$

при очевидном равенстве $\sigma_\eta^2 = \mathbf{M}(\hat{\sigma}_\eta^2)$ для несмещенной оценки дисперсии. Решение здесь, как и прежде в МФДС, принимается по принципу минимизации решающей статистики. Нетрудно понять на качественном уровне, что точность такого решения существенно возросла по сравнению с первоначальным вариантом алгоритма (7)—(9) — за счет предпринятого в (14) и (15) подавления шума. Дадим количественную характеристику достигаемого выигрыша по ОСШ на входе решающего устройства. Определим этот выигрыш значением коэффициента подавления фонового шума

$$\gamma \triangleq (\hat{\sigma}_\eta^2 - \sigma_\eta^2) / \sigma_\eta^2 \quad (17)$$

в модифицированном алгоритме (16).

Следуя известной схеме вычислений [14], раскроем выражение (17) в явном виде $\gamma = L^{-1} \chi_L^2 - 1$, где χ_L^2 — случайная величина, распределенная по закону χ^2 -Пирсона с L степенями свободы [16]. Ее математическое ожидание равно объему многомерной выборки наблюдений L , а среднее квадратичное отклонение равно $\sqrt{2L}$. Квадрат их отношения $h^2 = L/2$ и определяет в конечном итоге значение достигаемого выигрыша в ОСШ при применении модифицированного МФДС (15) с подавлением шума η . Например, при $L = 30 \dots 50$ выигрыш составит 11,8...14,0 дБ. Чем больше объем многомерной выборки речевого сигнала на интервалах его вокализации гласными фонемами, тем выше эф-

фективность предложенного метода защиты АРР от шумовых помех.

Заключение

Системы голосового управления робототехникой привлекают к себе внимание исследователей на протяжении нескольких десятков лет в связи, главным образом, с почти неограниченной сферой их возможного применения: от бытовых приборов и проектов типа "умного дома" до систем специального назначения. К сожалению, их распространение на практике до настоящего времени сдерживается недостаточно высокой степенью надежности АРР в условиях действия интенсивных внешних акустических помех, таких как шум улицы, звуки от проезжающего транспорта, речь посторонних лиц и т.п.

МФДС — это один из новейших методов АРР, основанный на принципах слоговой фонетики русского языка и нацеленный на преодоление проблемы защиты АРР от помех за счет использования ограниченного числа $R < \infty$ минимальных речевых единиц — с максимальными отличиями друг от друга в информационной метрике Кульбака—Лейблера. Помехоустойчивость АРР еще более возрастает в предложенной выше модификации МФДС с автоматическим подавлением фонового шума. Показано, что достигаемый при подавлении шума эффект состоит в многократном (11,8 дБ и более) увеличении ОСШ на входе решающего устройства. На основании данного результата можно сделать вывод о значительных перспективах расширения сферы распространения голосового управления робототехникой за счет использования предложенного метода АРР.

Таким образом, благодаря проведенному исследованию предложен на основе МФДС новый, эффективный метод защиты АРР от помех, предназначенный для использования в условиях действия интенсивного внешнего шума, когда известные технологии голосового управления не отвечают требованиям к их точности и надежности. Если говорить об основных ограничениях на его оптимальность и область применения, то они связаны, прежде всего, с предположением о гауссовском распределении речевого сигнала в постановочной части статьи. Однако указанное ограничение только на первый взгляд кажется чрезмерно жестким. В работе [14] строго показано, что гауссовский закон в его адаптивном варианте теоретически строго обусловлен общесистемным принципом максимальной энтропии наблюдений для широкого круга алгоритмов, использующих моменты распределения не выше второго порядка. Отметим, что подавляющее большинство известных методов АРР относится, как известно [10], именно к такому кругу алгоритмов,

поскольку основывается на методах корреляционной и спектральной обработки речевого сигнала.

Список литературы

1. Патент № 111944 (полезная модель). Устройство для фонетического анализа и распознавания речи / В. В. Савченко, А. В. Савченко, Д. Ю. Акатьев / Роспатент: по заявке № 2011125526/08 от 21.06.2011. — Оpubл. 27.12.2011, бюл. № 6.
2. Савченко А. В., Савченко В. В. Разработка быстродействующих алгоритмов автоматического распознавания голосовых команд с регулируемой точностью и надежностью на основе принципов слоговой фонетики русского языка и метода фонетического декодирования слов в информационной метрике Кульбака—Лейблера: Тезисы докладов Всероссийской научно-технической конференции и выставки, посвященной итогам реализации федеральной целевой программы "Исследования и разработки по приоритетным направлениям развития научно-технологического комплекса России на 2007—2013 годы": М.: МИСиС, 2013. С. 184—185.
3. Савченко А. В. Адаптивный алгоритм распознавания речи на основе метода фонетического декодирования слов в задаче голосового управления // Информационные технологии. 2013. № 4. С. 34—39.
4. Савченко В. В., Савченко А. В. Метод фонетического декодирования слов в информационной метрике Кульбака—Лейблера для систем автоматического анализа и распознавания речи с повышенным быстродействием // Информационно-управляющие системы. 2013. № 2. С. 7—12.
5. Савченко В. В., Пономарев Д. А. Оптимизация фонетической базы данных по группе дикторов на основе информационной теории восприятия речи // Информационные технологии. 2009. № 12. С. 7—12.
6. Pat. US 6301560 B1, Int. CI G10L15/22. Discrete speech recognition system with ballooning active grammar / Inventor: Masters S. P. / Assignee: Microsoft Corporation. — Pub. Date 09.10.2001.
7. Schuster M. Speech Recognition for Mobile Devices at Google // Proc. of the 11th Pacific Rim international conference on Trends in Artificial Intelligence, 2010. LNCS Vol. 6230. С. 8—10.
8. Pat. US 8175883 B2, Int. CI G10L21/00 (2006.01). Speech recognition system and method / Inventor: Grant R., Gregor P. / Assignee: Nuance Communications Inc. — Pub. Date 08.05.2012.
9. Pat. US 0016678 A1, Int. CI G10L21/00 (2006.01). Intelligent automated assistant / Inventor: Gruber T., Cheyer A., Kittlaus D., Guzzoni D., Brigham C., Giuliano R., Bastea-Forte M., Sessler H. / Assignee: Apple Inc. — Pub. Date 19.01.2012.
10. Rabiner L. A tutorial on Hidden Markov Models and Selected Applications in Speech Recognition // Proc. of the IEEE. 1989. Vol. 77 (2). P. 257—285.
11. Крашенинников В. Р., Армер А. И., Крашенинникова Н. А., Кузнецов В. В., Хвостов А. В. Некоторые задачи, связанные с распознаванием речевых команд на фоне интенсивных шумов // Инфокоммуникационные технологии. 2008. Т. 6, № 1. С. 72—75.
12. Levinson S. C. Mathematical models for speech technology. Chichester: England: John Wiley&Sons Ltd, 2005. 261 p.
13. Савченко В. В. Автоматическая обработка речи по критерию минимума информационного рассогласования на основе метода обеляющего фильтра // Радиотехника и электроника. 2005. Т. 50, № 3. С. 309—314.
14. Савченко В. В. Различение случайных сигналов в частотной области // Радиотехника и электроника. 1997. Т. 42, № 4. С. 426—431.
15. Kullback S. Information Theory and Statistics. New York: Dover Publications, 1997. 399 p.
16. Мюллер П., Нойман П., Шторм Р. Таблицы по математической статистике: Пер. с нем. М.: Финансы и статистика, 1982. 278 с.
17. Кошелев В. И., Андреев В. Г. Оценка собственных значений в задаче обработки эхо-сигналов // 100-летие начала использования электромагнитных волн для передачи сообщений и зарождения радиотехники: тезисы докладов Международной конференции, г. Москва, 4—6 мая 1995 г.: М.: Изд-во журнала "Радиотехника", 1995. С. 180—181.

The Speech Recognition Method of Phonetic Decoding of Words with Background Noise Cancellation

This paper studies the usage of the known words phonetic decoding method in the task of automatic speech recognition in robotics voice control systems. Our goal is to improve the efficiency of this method for acute problem of modern speech recognition techniques, namely, insufficient accuracy in case of the presence of intensive background noise. The achievement of this goal can cause an extensive use of voice control systems in many practical applications. Our study is based on modern theoretical computer science. In this paper we proposed the novel noise-proof version of phonetic decoding method for its application in commercial (inexpensive and reliable) intelligent systems with voice control of robotics in life, on transport and on production in the presence of intensive external noise. Our modification explores the idea of correcting the estimated covariance matrices of vowels by subtraction from their diagonal elements of the noise variance. Analytical expression of noise suppression coefficient is synthesized. It is shown that the proposed approach allows to multiple increase of the signal-to-noise-ratio (more than 11,8 dB). In conclusion we discuss the possible future expansion of speech recognition and voice control technologies on the basis of proposed method and give practical recommendations of its usage.

Keywords: speech, speech recognition, speech recognition, speech command, voice-activated control, noise, interference, fight against interference, protection against interference, noise immunity

References

1. **Pat. № 111944** (poleznaja model'). *Ustrojstvo dlja foneticheskogo analiza i raspoznavanija rechi*, V. V. Savchenko, A. V. Savchenko, D. Ju. Akat'ev, Rospatent: po zajavke № 2011125526/08 ot 21.06.2011, opubl. 27.12.2011, bjul. no. 6 (in Russian).
2. **Savchenko A. V., Savchenko V. V.** Razrabotka bystrodejstvujushih algoritmov avtomaticheskogo raspoznavanija golosovyh komand s reguliruemoj tochnost'ju i nadezhnost'ju na osnove principov slogovoj fonetiki russkogo jazyka i metoda foneticheskogo dekodirovanija slov v informacionnoj metrike Kul'baka—Lejblera: Tezisy dokladov Vserossijskoj nauchno-tehnicheskoy konferencii i vystavki, posvjashhennoj itogam realizacii federal'noj celevoj programmy "Issledovanija i razrabotki po prioritetyim napravlenijam razvitiija nauchno-tehnologicheskogo kompleksa Rossii na 2007—2013 gody", Moscow, MISiS, 2013, pp. 184—185.
3. **Savchenko A. V.** Adaptivnyj algoritm raspoznavanija rechi na osnove metoda foneticheskogo dekodirovanija slov v zadache golosovogo upravlenija, *Informacionnye tehnologii*, 2013, no. 4, pp. 34—39.
4. **Savchenko V. V., Savchenko A. V.** Metod foneticheskogo dekodirovanija slov v informacionnoj metrike Kul'baka—Lejblera dlja sistem avtomaticheskogo analiza i raspoznavanija rechi s povyshennym bystrodejstviem, *Informacionno-upravljajushhie sistemy*, 2013, no. 2, pp. 7—12.
5. **Savchenko V. V., Ponomarjov D. A.** Optimizacija foneticheskoy bazy dannyh po grupe diktov na osnove informacionnoj teorii vosprijatija rechi, *Informacionnye tehnologii*, 2009, no. 12, pp. 7—12.
6. **Pat. US 6301560 B1**, Int. CI G10L15/22. *Discrete speech recognition system with ballooning active grammar* / Inventor: Masters S. P. / Assignee: Microsoft Corporation. — Pub. Date 09.10.2001.
7. **Schuster M.** *Speech Recognition for Mobile Devices at Google*, Proc. of the 11th Pacific Rim international conference on Trends in Artificial Intelligence, 2010, LNCS, vol. 6230, pp. 8—10.
8. **Pat. US 8175883 B2**, Int. CI G10L21/00 (2006.01), *Speech recognition system and method* / Inventor: Grant R., Gregor P., Assignee: Nuance Communications Inc., pub. Date 08.05.2012.
9. **Pat. US 0016678 A1**, Int. CI G10L21/00 (2006.01), *Intelligent automated assistant* / Inventor: Gruber T., Cheyer A., Kittlaus D., Guzzoni D., Brigham C., Giuli R., Bastea-Forte M., Saddler H. Assignee: Apple Inc., pub. Date 19.01.2012.
10. **Rabiner L.** A tutorial on Hidden Markov Models and Selected Applications in Speech Recognition, *Proc. of the IEEE*, 1989, vol. 77 (2), pp. 257—285.
11. **Krashennnikov V. R., Armer A. I., Krashennnikova N. A., Kuznecov V. V., Hvostov A. V.** Nekotorye zadachi, svjazannye s raspoznavaniem rechevnyh komand na fone intensivnyh shumov, *Infokommunikacionnye tehnologii*, 2008, vol. 6, no. 1, pp. 72—75.
12. **Levinson S. C.** *Mathematical models for speech technology*, Chichester, England: John Wiley & Sons Ltd., 2005, 261 p.
13. **Savchenko V. V.** Avtomaticheskaja obrabotka rechi po kriteriju minimuma informacionnogo rassoglasovanija na osnove nietoda obelajushhego fil'tra, *Radiotekhnika i jelektronika*, 2005, vol. 50, no. 3, pp. 309—314.
14. **Savchenko V. V.** Razliczenie sluchajnyh signalov v chastotnoj oblasti, *Radiotekhnika i jelektronika*, 1997, vol. 42, no. 4, pp. 426—431.
15. **Kullback S.** *Information Theory and statistics*, Dover Publications, New York, 1997, 399 p.
16. **Mjuller P., Nojman P., Shtorm R.** *Tablicy po matematicheskoj statistike*, Per. s nem., Moscow, Finansy i statistika, 1982, 278 p.
17. **Koshelev V. I., Andreev V. G.** Ocenka sobstvennyh znachenij v zadache obrabotki jeh-signalov, *100-letie nachala ispol'zovanija jelektronnym voln dlja peredachi soobshhenij i zarozhdenija radiotekhniki: tezisy dokladov Mezhdunarodnoj konferencii*, g. Moscow, 4—6 maja, 1995, Moscow, Izd-vo zhurnala "Radiotekhnika", 1995, pp. 180—181.

Адрес редакции:

107076, Москва, Стромьинский пер., 4

Телефон редакции журнала (499) 269-5510

E-mail: it@novtex.ru

Технический редактор *Е. В. Конова*.

Корректор *Е. В. Комиссарова*.

Сдано в набор 09.11.2015. Подписано в печать 22.12.2015. Формат 60×88 1/8. Бумага офсетная.

Усл. печ. л. 8,86. Заказ IT116. Цена договорная.

Журнал зарегистрирован в Министерстве Российской Федерации по делам печати, телерадиовещания и средств массовых коммуникаций.

Свидетельство о регистрации ПИ № 77-15565 от 02 июня 2003 г.

Оригинал-макет ООО "Авансед солюшнз". Отпечатано в ООО "Авансед солюшнз".

119071, г. Москва, Ленинский пр-т, д. 19, стр. 1.