

Малахова Елена Анатольевна – кандидат технических наук, Украинский государственный университет железнодорожного транспорта, доцент кафедры «Управление эксплуатационной работой», площадь Фейербаха, 7, г. Харьков, 61050; тел.: 066-341-84-81; e-mail: alena_mal@mail.ru

Malakhova Olena – candidate of technical sciences, Associate professor, Ukrainian State University of Railway Transport, Kharkiv, Feuerbach sqr, 7 61050; tel.: 066-341-84-81; e-mail: alena_mal@mail.ru

Фат Карина Эдуардовна – студентка, Український державний університет залізничного транспорту, кафедра «Управління експлуатаційною роботою»; майдан Фейербаха, 7, м. Харків, Україна, 61050; тел.: 099-612-78-85; e-mail: www.karinka.ru08@mail.ru

Fat Karina Eduardovna – студентка, Украинский государственный университет железнодорожного транспорта, кафедра «Управление эксплуатационной работой», площадь Фейербаха, 7, г. Харьков, 61050;

Fat Karyna – student, Ukrainian State University of Railway Transport, Kharkiv, кафедра Department of "Management of operational work"; Feuerbach sqr, 7 61050; tel.: 099-612-78-85; e-mail: ww.karinka.ru08@mail.ru

УДК 004.522

А. А. ШТЕПА, А. Ю. КОЛЛАРОВ

МЕТОД ПОВЫШЕНИЯ ДОСТОВЕРНОСТИ РАСПОЗНАВАНИЯ РЕЧЕВЫХ КОМАНД В СЛОЖНЫХ АКУСТИЧЕСКИХ УСЛОВИЯХ

В данной статье обсуждается применение методов обработки зашумленных сигналов в системах распознавания речевых команд и приводятся некоторые из результатов исследований в этой области. Основной целью исследования является повышения достоверности распознавания речевых команд в системах голосового управления техническими объектами в сложной акустической обстановке путем повышения соотношения сигнал/шум за счет применения пространственного разделения сигналов с помощью нескольких направленных микрофонов и цифровой обработки сигналов на основе адаптивной компенсации помех. Использование современных методов лингвистического распознавания речевых команд совместно с методом адаптивной компенсации для обработки зашумленного сигнала позволяет повысить достоверность распознавания речевых команд.

Ключевые слова: голосовое управление, распознавание речевых команд, адаптивная компенсация, аддитивный шум.

Введение. Управление техническими устройствами с помощью голосовых команд является одной из наиболее востребованных и развивающихся технологий на протяжении последних лет. Основным достоинством голосового интерфейса является естественность этого метода передачи информации для человека. Свободные руки и отсутствие необходимости визуального контроля над используемыми элементами управления снижают утомляемость оператора и повышают эффективность управления в целом.

Управление многими техническими средствами влечет за собой повышенный уровень ответственности, что значительно повышает требование к достоверности распознавания речевых команд управления. Если управление технически сложным объектом происходит в условиях с повышенным зашумлением (технологическое помещение, транспортное средство), его интенсивность может исказить результаты распознавания речевых команд. Таким образом, задача повышения достоверности распознавания речевых команд в сложных акустических условиях является актуальной.

Анализ литературных данных и постановка проблемы. Обзор методов повышения качества и разборчивости зашумленных речевых сигналов показывает, что существует много различных подходов к обработке зашумленной речи [1].

Универсальных методов обработки, которые одинаково хорошо боролись бы с преимущественно нестационарными и стационарными, аддитивными и мультипликативными шумами, существенно повышали бы качество и разборчивость речи, сейчас нет и перспективы их появления не определены [1]. Поэто-

му пути решения обозначенной задачи лежат в плоскости совместного использования взаимодополняющих методов.

Одной из наиболее эффективных групп методов обработки речевых сигналов с аддитивным шумом являются методы, основанные на адаптивной компенсации помех, которые позволяют значительно улучшить качество зашумленных сигналов – на несколько десятков децибел [2]. При наличии же нескольких каналов поступления информации эффективность метода может повыситься за счет применения оптимального весового суммирования дополнительных сигналов.

Для разделения смесей сигналов в работах [3, 4] успешно используется метод анализа независимых компонент (independent component analysis, ICA). Однако, ограничение, накладываемое методом на количество смесей, которое должно быть не большим количеством источников шумов и требование наличия полезного сигнала во всех смесях делают применение этого метода малоэффективным в случае динамического изменения акустической картины в реальных производственных условиях применения многоканальной системы распознавания. В рамках классического весового подхода адаптивной компенсации [3, 4] выходной сигнал получают путем вычитания из сигнала основного канала взвешенной суммы сигналов дополнительных каналов. Ограничением для применения весового подхода является предположение об отсутствии полезного сигнала в дополнительных каналах. Эффективным решением для реализации этого подхода может стать многоканальная система на базе узконаправленных микрофонов, реализующая

принцип пространственного разделения каналов [5, 6]. Весьма интересен и эффективен подход использованный разработчиками слуховых аппаратов Phonak, применяющих многомикрофонные решения и сложные алгоритмы изменения параметров такой системы с целью достижения наибольшей разборчивости речи для разных направлений [7, 8].

Обработанный с целью устранения влияния шумов сигнал подвергают одной из существующих процедур распознавания речевых команд. Весьма эффективными считаются методы распознавания, основанные на процедурах нелинейного временного выравнивания (динамического программирования) [9].

Объект, цель и задачи исследования. *Объектом исследования* является процесс распознавание речевых команд в системах управления техническими устройствами. *Целью работы* является повышения достоверности распознавания речевых команд в системах голосового управления техническими объектами.

Для достижения поставленной цели были поставлены следующие *задачи*: -проанализировать специфику акустических условий характерных для эксплуатации систем голосового управления сложными техническими устройствами в производственной сфере; обосновать применение многомикрофонной системы распознавания речевых команд управления с целью использования преимуществ пространственного разделения сигналов; предложить метод повышения достоверности распознавания речевых команд управления, на основе взаимно дополняющих друг друга существующих методов и разработать его базовый алгоритм.

Анализ методов обработки зашумленных речевых сигналов. Рассмотрим акустические условия, характерные для технологических, складских помещений и кабин транспортных средств. Если техногенные шумы часто имеют стационарный характер и вполне определенную полосу частот, то шумы производимые говорящим человеком, находящимся в одном помещении с оператором, будут представлять наибольшую угрозу для задачи достоверного распознавания команд управления. Кроме того, сам оператор в процессе выполнения своих непосредственных обязанностей может и должен перемещаться в помещении или, по крайней мере, изменять положение головы. В этом случае источник полезного сигнала будет менять свое расположение по отношению к стационарно закрепленным микрофонам системы. Например, в салоне автомобиля водителю следует внимательно следить за обстановкой на проезжей части поворачивая для этого голову, а требовать использования наушников или закрепленного на голове микрофона, отвлекающего, создающего дискомфорт и повышающего утомляемость от водителя нельзя. В этом случае рациональным решением будет применение принципа пространственного разделения сигналов с помощью нескольких узконаправленных микрофонов, диаграммы направленности которых, минимально пересекаясь друг с другом (рис. 1).

Рассмотрим особенности обработки зашумленных речевых сигналов. Адаптивная компенсация помех основана на использовании, помимо основного зашумленного сигнала, который подлежит очистке,

также одного или нескольких дополнительных сигналов – сигналов, которые коррелированы с шумовыми сигналами и одновременно некоррелированы (или слабо коррелированы) с полезным сигналом. С помощью дополнительных сигналов формируется сигнал, выступающий оценкой помехи. Этот сигнал затем вычитается из зашумленного сигнала и результат этой операции рассматривается как оценка полезного сигнала [1].

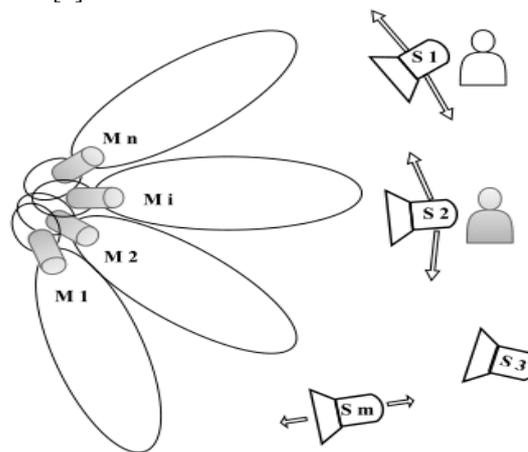


Рис. 1 – Расположение микрофонов (M1, M2, Mi, Mn) и источников звуковых сигналов и шумов (S1, S2, S3, Sm) в многоканальной системе распознавания речевых команд управления

Таким образом, эффективность методов адаптивной компенсации непосредственно зависит от доступности качественных опорных сигналов, удовлетворяющих требованиям коррелированности с шумом и одновременной некоррелированности (или слабой степени коррелированности) с полезным сигналом. На практике получение таких сигналов достигается использованием двух микрофонов, размещаемых таким образом, чтобы один из них был направлен на предполагаемый источник полезного сигнала, а диаграмма направленности второго не захватывала полезный сигнала.

В случае стационарного размещения группы направленных микрофонов важной задачей становится определение основного и вспомогательных каналов. Критерием такого определения может служить соотношение сигнал/шум. Однако, если в помещении разговаривают несколько человек, выделить голос оператора, опираясь на частотную полосу становится трудно. В качестве эффективного критерия может выступать параметр, получаемый на основе применения существующих методов распознавания речевых команд на лингвистическом этапе распознавания. Таким параметром может служить мера близости между полученным системой сигналом и одним из эталонов библиотеки команд. Оценка меры близости между входным речевыми сигналами и эталоном может производиться с помощью метода нелинейного временного выравнивания (динамического программирования) [9]. Метод позволяет сравнивать разные по длительности образцы независимо от темпа речи. Пусть сравниваются два образца сигналов, представленных в виде массива векторов (для речевых сигналов это наборы линейных спектральных корней):

$$X = \{ \bar{x}_0, \bar{x}_1, \dots, \bar{x}_i, \dots, \bar{x}_N \}, \quad (1)$$

$$Y = \{ \bar{y}_0, \bar{y}_1, \dots, \bar{y}_i, \dots, \bar{y}_M \}. \quad (2)$$

Различие между векторами двух образов определяется последовательностью состояний S_k и обозначается:

$$F() = \{ C_0, C_1, \dots, C_k, \dots, C_K \},$$

где C_0 и C_K – начальное и конечные состояния, $F()$ – функция временного выравнивания, которая проецирует временную область одного образа на временную область другого. В методе динамического программирования ищется такая функция $F()$, при которой путь из состояния C_0 в состояние C_K , является оптимальным, т. е. будет обеспечено минимальное накопленное расстояние между двумя образами. При построении оптимального пути, на каждом шаге алгоритма используется основная формула:

$$d_{i,j} = \min \left\{ \begin{array}{l} d_{i,j-1} + r(\bar{x}_i, \bar{y}_j) \\ d_{i-1,j-1} + r(\bar{x}_i, \bar{y}_j) \\ d_{i-1,j} + r(\bar{x}_i, \bar{y}_j) \end{array} \right\}, \quad (3)$$

где $i=0\dots N, j=0\dots M$.

В качестве расстояния между векторами используется взвешенная евклидова метрика:

$$r(\bar{x}, \bar{y}) = \sum_{k=0}^{N_SEC-1} (x_k - y_k)^2, \quad (4)$$

где N_SEC – размерность векторов признаков.

На выходе процедуры сравнения получается некоторое число (мера близости), представляющее собой величину, обратную степени близости между сигналами. В работе [9] предложено решение задачи автоматического отсеивания ложных срабатываний, основанное на нормировании оценок меры близости. Предложенный критерий позволяет оценивать достоверность распознавания текущего слова [9] и может быть использован с целью определения базового канала из числа задействованных.

Алгоритм определения основного канала многоканальной системы распознавания речевых команд. В качестве оценки достоверности $D(i)$ выбора базового канала может служить величина, характеризующая степень близости между сигналом, полученным системой и ближайшим к нему эталоном, оцененная в предположении, что i -тый канал является основным, а остальные – дополнительными. Такая оценка будет величиной обратной по отношению к мере близости. Таким образом, вычисление критерия для оценки достоверности выбора базового канала не будет требовать дополнительных вычислительных мощностей. Кроме того, в качестве альтернативного критерия определения основного канала $D(i)$ может выступать величина характеризующая достоверность идентификации личности оператора как, например, в работе [10]. Идентификация личности по голосу проводится на сходных с распознаванием речевых команд принципах и так же подразумевает вычисление оценки достоверности. Сформируем общий алгоритм определения базового канала многоканальной системы распознавания речевых команд на основе предложенного критерия (рис. 2).

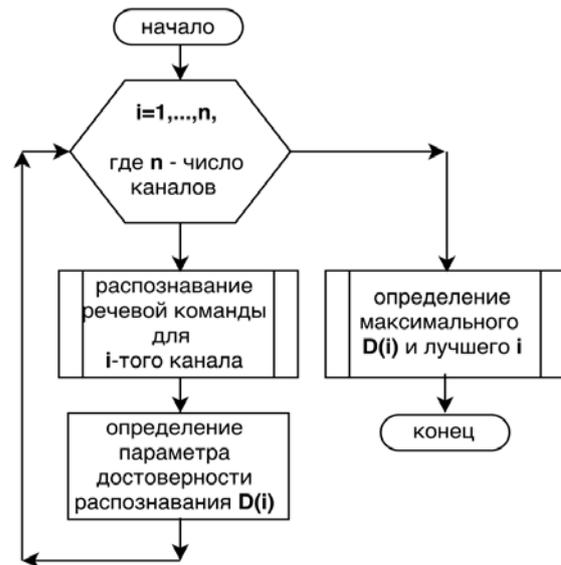


Рис. 2 – Общий алгоритм определения основного канала системы

При наличии в системе n независимых каналов и минимальным взаимным перекрытием диаграмм направленностей микрофонов определение базового канала может быть проведено перебором всех задействованных каналов в предположении, что каждый из них может быть базовым в текущий момент. Недостатком алгоритма является увеличение времени обработки в n раз, что повысит требования к вычислительной мощности системы. Количество каналов системы рационально выбирать более двух. При двух каналах эффективность пространственного разделения будет достаточно низкой, когда источник полезного сигнала займет положение, при котором полезный сигнал попадет в оба канала. В этом случае ни один из каналов не сможет эффективно выступать в качестве дополнительного. Увеличение количества каналов приведет к более выраженному положительному эффекту от пространственного разделения сигнала. Однако, использование большого количества независимых каналов приведет к необоснованному усложнению системы.

Перспективными направлениями для дальнейших исследований в обсуждаемой области могут стать моделирование работы системы с целью определения оптимального числа направленных микрофонов, разработка алгоритмических решений оптимизации вычислительной нагрузки, а также изучение влияния диаграмм направленности микрофонов на эффективность работы системы.

Выводы. На основании проведенного анализа особенностей акустической картины характерной для систем голосового управления в качестве наиболее эффективного в условиях нескольких перемещающихся источников шума предлагается использовать метод адаптивной компенсации совместно с пространственным разделением сигнала, осуществляемым за счет реализации многоканальной системы на базе узконаправленных микрофонов.

В качестве критерия надежного определения основного канала многомикрофонной системы предложено ис-

пользовать величины, характеризующие меру близости полученного сигнала к известному образцу команды, рассчитываемые на лингвистическом этапе распознавания.

Список литературы: 1. Чучупал, В. Я. Цифровая фильтрация зашумленных речевых сигналов [Текст]: Сообщение по программному обеспечению ЭВМ / В. Я. Чучупал, А. С. Чичагов, К. А. Маковкин ; под ред. Ю. И. Журавлева ; Вычислительный центр РАН – М. : Вычислительный центр РАН, 1998. – 52 с. 2. McWhirer, J. S. A Digital Adaptive Noise-Canceller Based on a Stabilizer Version of the Widrow L.M.S. Algorithms [Text] / J. S. McWhirer, K. J. Palmer, J. B. Roberts // A, Proc. 1982, IEEE Int. Conf. ASSP, pp.1384-1387. 3. Ширман, Я. Д. Радиоэлектронные системы: основы построения и теория: [Текст] Справочник / Я. Д. Ширман – М. : Радиотехника, 1998. – 828 с. 4. Ревунова, Е. Г. Разделение сигнальных смесей на основе принципа минимальной длины описания [Текст] / Е. Г. Ревунова // Компьютерні засоби, мережі та системи. – 2005.– № 4 С. 86–93. 5. Ricketts, T. 1999. Comparison of performance across three directional hearing aids [Text] / T. Ricketts, S. Dhar // J Am Acad Audiol 10(4):180-9 – 1983. 6. Gnewikow et al. Real-world benefit from directional microphone hearing aids. // J Rehabil Res Dev 46(5):603-18 – 2005. 7. Nyffeler, M. Auto ZoomControl – Automatic change of focus to speech signals of interest [Electronic resource] / M. Nyffeler // Field Study News. – September 2010. – Available at: http://www.phonakpro.com/content/dam/phonak/b2b/C_M_tools/Librar/y/Field_Study_News/en/. 8. Nyffeler, M. Field Study on User Control of Directional Focus: Benefits of Hearing the Facets of a Full Life [Text] / M. Nyffeler, S. Dechant // Hearing Review. – 2008 – 16(1) – pp. 24-28. 9. Гладышев, К. К. Информативные признаки на основе линейных

спектральных корней в системах распознавания речевых команд : [Текст] : автореф. дис. ... к-та тех. наук : 05.13.01 / К. К. Гладышев [Санкт-Петербургский государственный университет телекоммуникаций им. Проф. Бонч-Бруевича]. – С.-Пб., 2010. – 16 с. 10. Ахмад, Х. М. Математические модели принятия решений в задачах распознавания говорящего [Текст] / Х. М. Ахмад // Вестник ТГТУ – 2008, – Т. 14, №1. – С. 19–31.

Bibliography (transliterated): 1. Chuchupal, V. Ya., Chichagov, A. S., Makovkin, K. A.; In: Zhuravlev, Yu. I. (1998). Tsifrovaia fil'tratsiia zashumlennykh rechevykh signalov. Moscow: Vychislitel'nyi tsentr RAN, 52. 2. McWhirer, J.S., Palmer, K.J., Roberts J.B. (1982). A Digital Adaptive Noise-Canceller Based on a Stabilizer Version of the Widrow L.M.S. Algorithms, Proc. IEEE Int. Conf. ASSP, 1384-1387. 3. Shirman, Ya. D. (1998). Radioelektronnye sistemy: osnovy postroeniya i teoriya: Spravochnik. M.: Radiotekhnika, 828. 4. Revunova, E. G. (2005). Razdelenie signal'nykh smesey na osnove principa minimal'noj dliny opisanija. Komp'juterni zasobi, merezhi ta sistemi, 4, 86–93. 5. Ricketts, T., Dhar, S. (1999). Comparison of performance across three directional hearing aids. J Am Acad Audiol 10(4):180-9. 6. Gnewikow et al. (2005). Real-world benefit from directional microphone hearing aids. J Rehabil Res Dev 46(5):603-18. 7. Nyffeler, M. (2010). auto ZoomControl – Automatic change of focus to speech signals of interest. Field Study News, September: www.phonakpro.com. 8. Nyffeler, M., Dechant, S. (2008). Field Study on User Control of Directional Focus: Benefits of Hearing the Facets of a Full Life. Hearing Review. 16(1):24-28. 9. Gladyshev, K. K. (2010). Informativnye priznaki na osnove lineinykh spektral'nykh kornei v sistemah raspoznavaniia rechevykh komand. S.-Pb., 16. 10. Akhmad, Kh. M. (2008). Decision-Making Mathematical Models for Tasks of Speaker's Recognition. Vestnik TGTU, Vol. 14. № 1, 19–32.

Поступила (received) 20.10.2015.

Відомості про авторів / Сведения об авторах / About the Authors

Штепа Александр Анатольевич – кандидат технических наук, ГВУЗ Донецкий национальный технический университет, доцент кафедры электронной техники; пл. Шибанкова, 2, г. Красноармейск, Украина, 85300;

Штепа Олександр Анатолійович – кандидат технічних наук, ДВНЗ Донецький національний технічний університет, доцент кафедри електронної техніки; пл. Шибанкова, 2, м. Красноармійськ, Україна, 85300;

Shtepa Aleksandr – Candidate of technical science, State Higher Education Establishment “Donetsk National Technical University”, Department of Electronic Engineering; Shybankova Square, 2, Krasnoarmiysk, Donetsk region, 85300, Ukraine; tel.: +38-063-400-65-65; e-mail: A.Shtepa@mail.ru.

Колларов Александр Юрьевич – кандидат технических наук, ГВУЗ Донецкий национальный технический университет, и. о. заведующего кафедрой электрической инженерии; пл. Шибанкова, 2, г. Красноармейск, Украина, 85300; тел.: 050-646-54-38; e-mail: kollarov@ukr.net.

Колларов Олександр Юрійович – кандидат технічних наук, ДВНЗ Донецький національний технічний університет, в. о. завідувача кафедри електричної інженерії; пл. Шибанкова, 2, м. Красноармійськ, Україна, 85300; тел.: 050-646-54-38; e-mail: kollarov@ukr.net.

Kollarov Oleksandr – Candidate of technical science, State Higher Education Establishment “Donetsk National Technical University”, Department of Electrical Engineering; Shybankova Square, 2, Krasnoarmiysk, Donetsk region, 85300, Ukraine; tel.: +38-050-646-54-38; e-mail: kollarov@ukr.net

УДК 004.5:004.78 (045)

Є. Б. АРТАМОНОВ, Г. М. КРЕМЕНЕЦЬКИЙ, А. О. ДЛУЖЕВСЬКИЙ, О. В. ПАНФЬОРОВ

ПІДХОДИ ДО РЕАЛІЗАЦІЇ АПАРАТНО-ПРОГРАМНОГО НАВЧАЛЬНОГО КОМПЛЕКСУ ШРИФТУ БРАЙЛЯ

Розглядається реалізація апаратно-програмного навчального комплексу, який представляє собою тактильний дисплей для виведення шостикрапкового символу в системі Брайля на основі сигналів з комп'ютера та спеціальне програмне забезпечення. Робота з тактильним дисплеєм відбувається завдяки подачі постійного струму через електромеханічні елементи. Особливу увагу приділено компоновці елементів на схемі та можливості використання комплексу у мережі. Результати дослідження можна використовувати для випуску навчальних електронних пристроїв для людей з вадами зору.

Ключові слова: апаратно-програмний комплекс, шрифт Брайля, навчання інвалідів зору, програмне забезпечення.

Вступ. За статистикою, в Україні живе близько 70 тисяч сліпих людей. За відсутності зору сліпим людям доводиться покладатися на інші джерела інформації, такі як звук та тактильні відчуття. Так для чи-

тання одним з найрозповсюдженіших тактильних шрифтів є шрифт Брайля.

В Україні не більше 10 % інвалідів зору володіють шрифтом Брайля, і ця проблема нагадує замкнене

© Є. Б. Артамонов, Г. М. Кременецький, А. О. Длужевський, О. В. Панфьоров. 2015