

КОМП'ЮТЕРНІ ЗАСОБИ, МЕРЕЖІ ТА СИСТЕМИ

T. Khashan, J. Pisarenko

SPEECH AND SPECIAL SIGNALS DIRECTION FINDING FOR THE TECHNICAL HEARING SYSTEM IN THE INTELLECTUAL ROBOTICS

The improved method of acoustic signal detection is offered. A distinctive feature of the algorithm is automatic detection of current research object borders.

Пропонується удосконалений метод пеленгування акустичного сигналу. Зсув фаз та обчислення різниці амплітуд знаходять на поточному об'єкті дослідження, який вибирається автоматично, базуючись на висновках про квазіперіодичність та енергетичну завершеність досліджуваного фрагмента акустичного сигналу.

Предлагается усовершенствованный метод пеленгации акустического сигнала. Сдвиг фаз и вычисление разности амплитуд определяют на текущем объекте исследования, который выбирается автоматически, основываясь на выводах о квазипериодичности и энергетической завершенности исследуемого фрагмента акустического сигнала.

© Т.С. Хашан, Ю.В. Писаренко,
2008

УДК 004.896, 004.934

Т.С. ХАШАН, Ю.В. ПИСАРЕНКО

ПЕЛЕНГАЦИЯ РЕЧЕВЫХ И СПЕЦИАЛЬНЫХ СИГНАЛОВ СИСТЕМОЙ ТЕХНИЧЕСКОГО СЛУХА В ИНТЕЛЛЕКТУАЛЬНОЙ РОБОТОТЕХНИКЕ

За последние два десятилетия значительно увеличилось внимание ряда стран к новым принципам построения и создания интеллектуализированных мобильных робототехнических средств. Такие средства действуют как автономные мобильные носители аппаратуры, адаптированные к выполнению сложных и опасных заданий, когда доступ к аварийному объекту людей-экспертов принципиально невозможен или очень опасен из-за высокого уровня риска аварии. Практическая эффективность использования таких средств на сегодня низкая из-за недостатка либо полного отсутствия сенсорных устройств у большинства дистанционно управляемых мобильных роботов. Сенсорные устройства, например, система технического слуха, важны для робота с функциями адаптации к быстро меняющимся условиям внешней среды в процессе выполнения программы обследования. Актуальность исследований в этой области подтверждает всплеск активности лабораторий промышленно развитых стран (Япония, США, Корея, Швеция, Россия) [1–7]. Подобные разработки ведутся также в Украине.

В Институте проблем искусственного интеллекта (Украина, г. Донецк) разработана базовая модель системы технического слуха (СТС) для малых мобильных роботов, ее модификация выполнена в ИПМ им. М.В. Келдыша РАН. В общем случае СТС включает в себя два электретных микрофона и двухкаскадные усилители для двух каналов на мик-

росхеме LM324N. Имеется несколько вариантов исполнения усилителя – в двухканальной или четырехканальной версии. Система слуха в целом может включать два, четыре или матрицу таких сенсоров. В качестве модели для совместного использования информации от датчиков и управления двигателем в ИПМ им. М.В. Келдыша РАН (Россия, Москва) разработан мобильный робот «Аргонавт». Полное техническое описание робота «Аргонавт» и СТС можно найти в [8–10].

На рис. 1 представлена схема функционирования интеллектуальной системы (ИС) управления мобильным роботом с СТС.

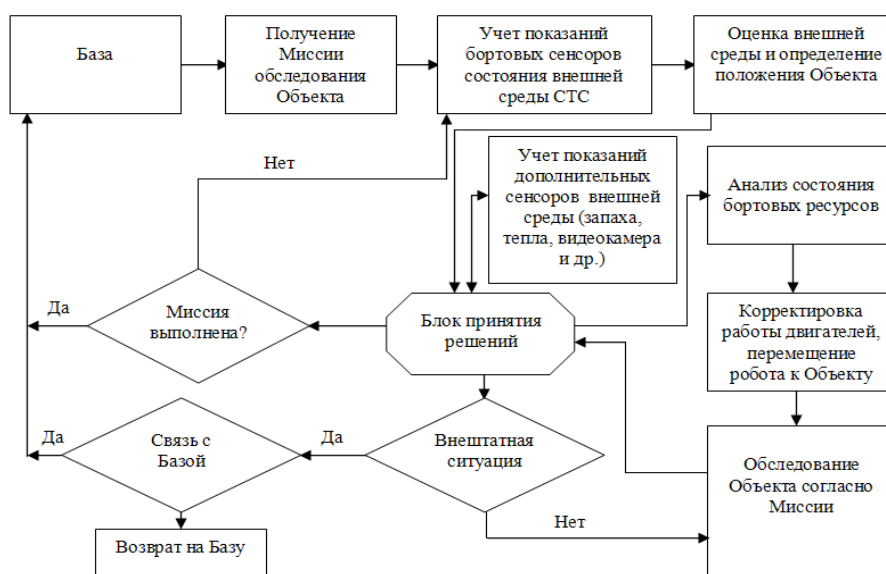


РИС. 1. Схема функционирования ИС управления мобильным роботом с СТС

Как видно из рис. 1, ИС управления мобильным роботом включает в себя блок принятия решения (БПР), в котором в зависимости от поступившей информации от СТС выбирается алгоритм функционирования мобильного робота. СТС не только «собирает» акустическую информацию, но и анализирует ее. В зависимости от возникшей ситуации и поставленной задачи данные передаются БПР. Следует отметить, что первоочередной задачей, которую должна решать СТС, является пеленгация источника звука (ИЗ) (под ИЗ здесь будем понимать речевые и специальные сигналы – свисток, дудочку).

Анализ литературных источников показал, что проблема локализации ИЗ роботом или бортовой системой решается с помощью оценки разницы фаз для разных приемников Interaural Phase Difference (IPD) и разницы в интенсивности для них Interaural Intensity Difference (IID). Значения IPD и IID вычисляются с помощью функции Head-Related Transfer Function (HRTF), которая зависит от изменения внешних условий и формы корпуса робота, в котором установлены

приемные устройства. В работах [1–3] предложен метод локализации источника звука без применения метода HRTF, основанный на «акустической» эпполярной геометрии Auditory Eripolar Geometry (AEG). В основе указанных методов лежит вычисление спектра на окнах априорно заданной длины или текущих объектах исследования (ОИ). В работах [8–10] разность фаз определяется с помощью корреляционной функции или функции рассогласования, текущий ОИ также выбирается статистическим путем. Очевидно, что фиксированное назначение интервала исследования влияет на точность пеленгации (в [1–3] точность пеленгации ИЗ составляет $7\text{--}10^\circ$, в [8–10] – $2\text{--}5^\circ$). Следовательно, разработка метода, позволяющего автоматически определять границы текущих ОИ без априорной информации пеленгуемого сигнала, а также разработка модифицированного метода пеленгации, является актуальной.

1. МОДИФИКАЦИЯ МЕТОДА ПЕЛЕНГАЦИИ

Схема решения задачи пеленгации ИЗ (рис. 2) включает в себя: оцифровку принятого сигнала (получение образа звука) и очистку принятого дискретного сигнала от шума (в данной статье этот этап не рассматривается); пеленгацию ИЗ – определение границ принятого ИЗ; выбор ОИ; анализ сигнала на ОИ; вычисление фазового сдвига, координат и расстояния до ИЗ.



РИС. 2. Структурная схема пеленгации ИЗ

1.1. Метод автоматического определения границ принятого источника звука

Пусть для решения задачи автоматического определения границ дана функция $U_j = \{u_j(n)\}$, $n = \overline{1, N}$ – образ звуковой волны (ОЗВ), или амплитудно-временное представление волнового процесса (АВП ВП), где $j = \overline{2, C}$, C – количество сенсоров (рис. 3, $C = 2$). Известно, что на $[t_0^j; t_1^j] \subseteq [1; N]$ имеет место

образ изолированного речевого сообщения (или специального сигнала) – ИЗ.

Необходимо определить границы образа ИЗ $[t_0^j; t_1^j]$. На рис. 3, 4, 5 изображения а) соответствуют отображению информации для левого канала, изображения б) соответствуют отображению информации для правого канала.

Данная задача будет решаться следующим образом. В качестве объекта исследования целесообразно рассматривать следующие характеристики АВП ВП:

1) последовательность характеристик функции U_j : границы квазипериодов (КП) – $[l_{j1}^i; l_{j6}^i]$, где $i = \overline{1, M}$, M – количество полных колебаний (КП) на АВП ВП; d_{ji} – длина КП; d_{ji}^+, d_{ji}^- – соответственно длина положительных и отрицательных полувольт КП; $r_{ji}, r_{ji}^+, r_{ji}^-$ – соответственно амплитуда КП, амплитуда положительных и отрицательных полувольт КП; $l_{\max}^{ji}, l_{\max}^{ji}$ – соответственно адреса локальных и глобальных экстремумов на КП; K_{ji} – КП;

2) энергетически завершенные фрагменты (ЭЗФ) 1-го уровня $R_{js^1}^1$, где $s = \overline{1, L}$, L – количество ЭЗФ в ОЗВ; $[lr_{j1}^s; lr_{j5}^s]$ – границы ЭЗФ; D_{js} – длины ЭЗФ; $R_{js^2}^2$ – ЭЗФ 2-го уровня.

Характеристики КП получены с учетом двух уровней АЦП, которые отождествляются с уровнем «ноль» вольт. Методы определения границ КП, ЭЗФ и их характеристик подробно описаны в [11].

Как показали исследования, для определения границ образа волнового процесса, на котором $[t_0^j; t_1^j] \subseteq [1; N]$ имеет место образ изолированного ИЗ (рис. 3), необходимо рассматривать последовательности ЭЗФ второго уровня (рис. 5), полученного на основе анализа последовательностей ЭЗФ первого уровня R_{js}^1 (рис. 4). Левая граница R_{j1}^2 и правая граница $R_{js^2}^2$ являются границами ИЗ, т. е. $[t_0^j; t_1^j]$. На рис. 6 показаны найденные границы ИЗ.

1.2. Метод автоматического выбора границ текущих объектов исследования

В рамках одного сигнала, излучаемого одним источником (диктором, свистком, дудочкой), границы ОИ следует выбирать из соображения однородности характеристик исследуемого отрезка сигнала. В работах [11, 12] предложен метод автоматического определения границ текущего ОИ. В основе данного алгоритма лежит гипотеза о том, что ОИ определяются как один или несколько последовательных временных отрезков, которые похожи по своей внутренней структуре, в течение которых происходит одно полное колебание. Последовательность таких ОИ называют квазипериодическими структурами (КС). Последовательность КС с однородными характеристиками формируют однородную КС, согласованную с последовательностями ЭЗФ.

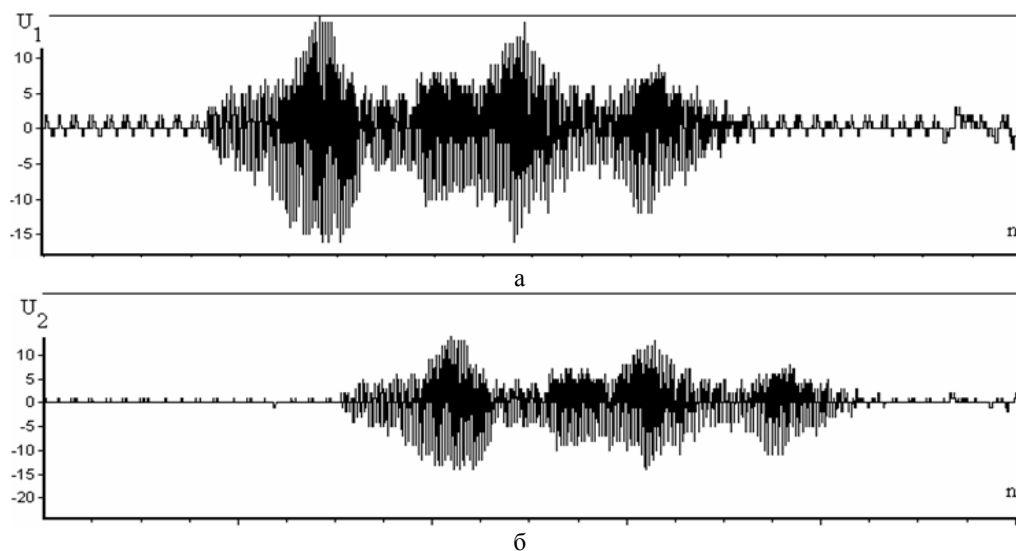


РИС. 3. АВП ВП принятого ИЗ (шум + слово «налево»): а – образ сигнала, полученный левым каналом; б – образ сигнала, полученный правым каналом

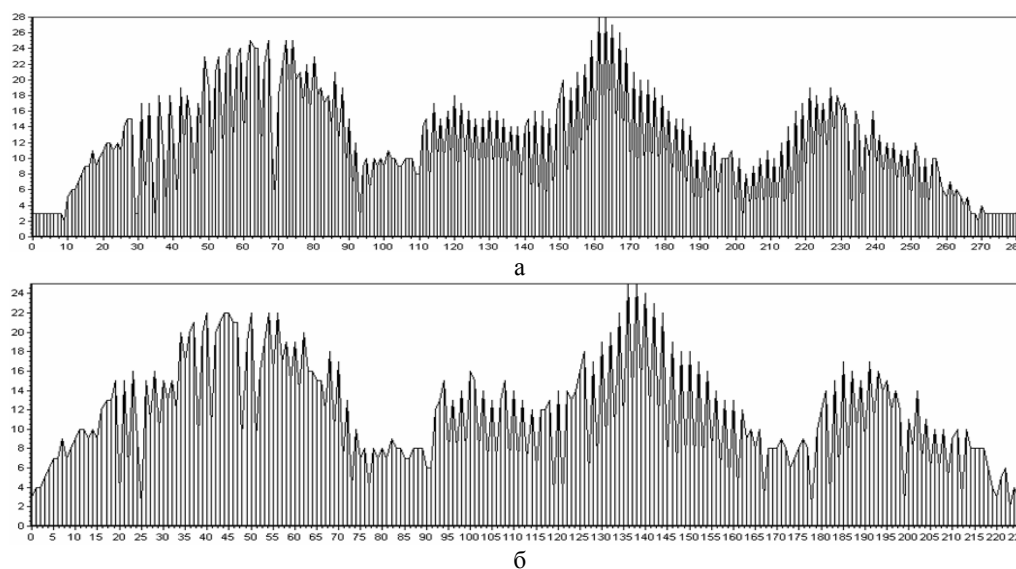


РИС. 4. ЭЗФ 1-го уровня принятого ИЗ (шум + слово «налево»): а – для образа сигнала, полученного левым каналом; б – для образа сигнала, полученного правым каналом

Как показано на рис. 2, алгоритм решения задачи автоматического определения границ текущих ОИ состоит из следующих блоков:

- определение границ слова, а также определение границ КП и их характеристик, определение границ КС, согласованных с ЭЗФ 1-го уровня;

– анализ «близости» КП внутри каждой КС: рассматриваются КП в каждой КС, определяется количество похожих КП. Правило определения «близости» характеристик КП рассмотрено в [11, 12];

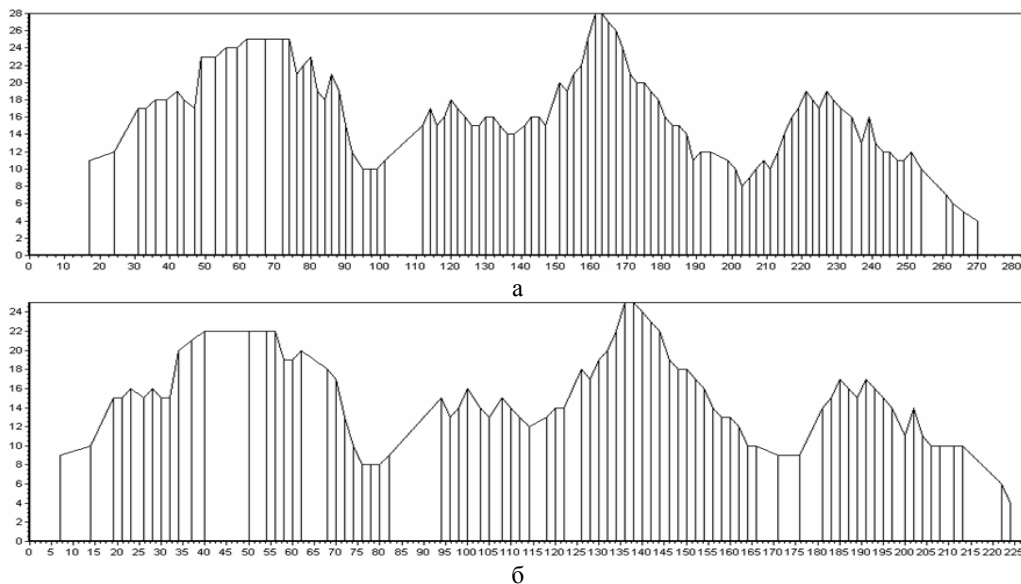


РИС. 5. ЭЗФ 2-го уровня принятого ИЗ (шум + слово «налево»): а – для образа сигнала, полученного левым каналом; б – для образа сигнала, полученного правым каналом

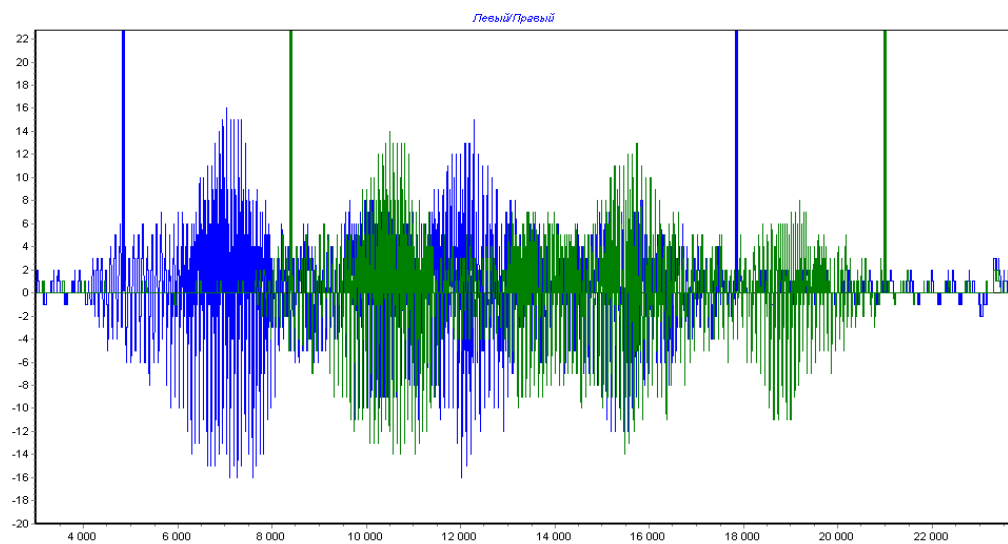


РИС. 6. Определение границ слова «налево» на АВП ВП для левого и правого каналов

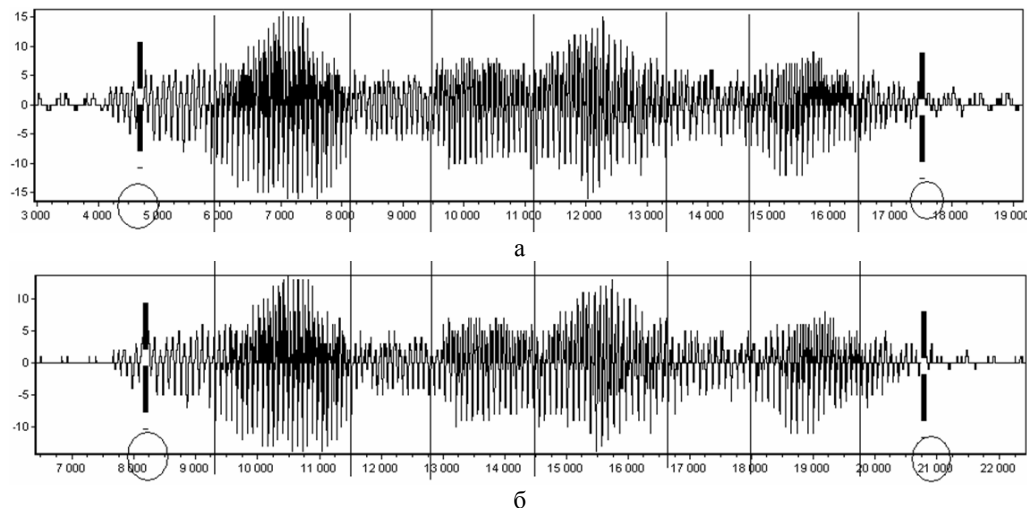


РИС. 7. Границы текущих ОИ: а – для образа сигнала, полученного левым каналом; б – для образа сигнала, полученного правым каналом

– анализ близости КС между собой: после рассмотрения всех КС и определения близких между собой КП внутри каждой структуры, анализируются все найденные КС между собой на близость характеристик. Отметим, что «похожие» КС между собой могут быть определены на некотором расстоянии друг от друга;

– фиксирование границ однородных КС (ОКС): после полного анализа всех КС принимается решение о фиксировании границ ОКС – $[L_j^{1k}; L_j^{2k}]$ (рис. 7). При этом $k = \overline{1, W}$ – количество ОКС на $[t_0^j; t_1^j]$. На рис. 7 вертикальными линиями отмечены границы ОКС.

1.3. Пеленгация источника звука

В основе метода пеленгации лежит определение разности хода акустической волны от источника сигнала (рис. 8) до приемных микрофонов по временному сдвигу (разности фаз) принятого микрофонами сигнала на текущих ОИ, границы которых определены согласно п. 1.2.

На рис. 8 обозначены: A (ИС) – источник сигнала; $M1$ и $M2$ – приемные микрофоны; $s1$ и $s2$ – расстояния от ИЗ до приемных микрофонов (длина хода волны до $M1$ и $M2$); a – разность хода волн, равная отрезку $[S; M2]$; $2b$ – база микрофонов, соответственно, они имеют координаты в принятой на рис. 8 системе координат $M1(-b; 0)$, $M2(b; 0)$; l – пеленгационный луч (и одновременно l обозначает расстояние от начала координат до ИЗ); φ – угол пеленга; $a1$ – отрезок $[S1; M2]$, приближенно представляющий разность хода волн до микрофонов (соотношение 2); $l1$ – вспомогательный луч, на нем строится отрезок $a1$.

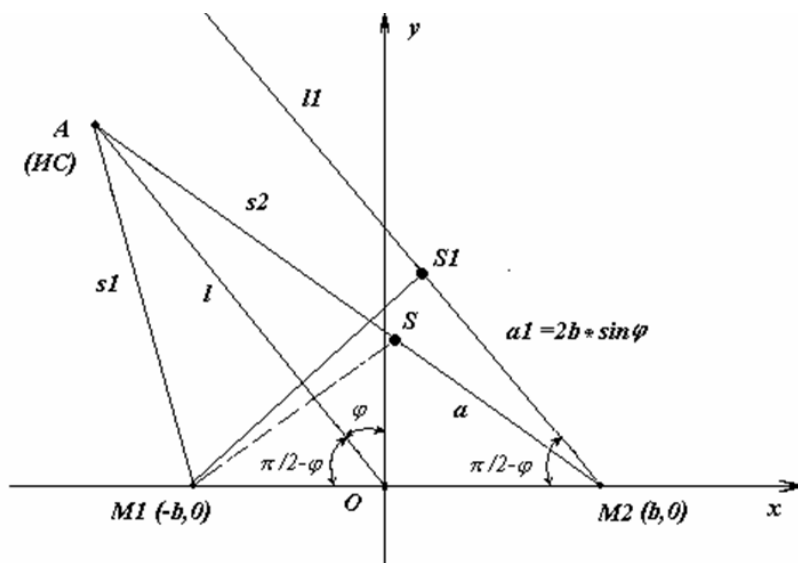


РИС. 8. Схема пеленгации. Линия положения – ветвь гиперболы Н

Отметим, что треугольник $A, M1, M2$ – основной пеленгационный треугольник, определяющий схему пеленгации.

$$a1 = 2b \cdot \sin \varphi, \tag{1}$$

$$\varphi = v\tau, \tag{2}$$

где τ – временной сдвиг, который определяется в свою очередь как

$$\tau = (L_1^k - L_2^k) / f. \tag{3}$$

В данной работе скорость звука принята постоянной (340 м/с), не зависящей от параметров среды распространения (температуры, влажности и т. п.), $b = 0.2$ м, $f = 44100$ – частота дискретизации сигнала.

Временной сдвиг определен на основе сравнительного анализа близости как минимум одной ОКС для двух каналов.

2. ЭКСПЕРИМЕНТЫ

Проверка предложенного метода пеленгации выполнялась на макете СТС, установленной на вращающейся платформе. При проведении экспериментов решалась задача пеленгации одиночного неподвижного ИЗ. По предложенным методикам в разделе 1 были разработаны соответствующие алгоритмы, на основании которых создан программный комплекс для обработки, анализа и пеленгации ИЗ.

Базу программного комплекса составляют: модуль понижения шума (в данной статье не рассматривается); модуль автоматического поиска границ текущего ОИ, который включает в себя определение границ слова и определение границ ОИ; модуль пеленгации и локализации ИЗ (см. рис. 2). Все модули программного комплекса реализованы на языке С и являются кроссплатформенными.

Проведено численное исследование алгоритмов: в экспериментах принимали участие 6 дикторов (2 женских и 4 мужских голоса), каждым были произнесены команды «налево», «прямо», «стоп». Расстояние между диктором и СТС для каждого случая было разным и варьировалось в пределах 3–6 метров. Всего было проведено более 50 экспериментов, результаты части из них приведены в таблице (для удобства результаты экспериментов округлялись).

	Налево		Прямо		Стоп	
	Угол	Расстояние	Угол	Расстояние	Угол	Расстояние
	30°	3,0 м	-30°	3,0 м	45°	3,0 м
Ж1	30°	2,5 м	-29°	4,0 м	45°	4,0 м
Ж2	31°	3,0 м	-33°	4,0 м	44°	3,5 м
М1	31°	3,0 м	-30°	3,5 м	43°	3,5 м
М2	32°	3,0 м	-32°	2,5 м	44°	3,5 м
М3	30°	3,0 м	-29°	3,0 м	44°	3,5 м

Для решения задачи определения расстояния до ИЗ использовались следующие общеизвестные методы: платформа СТС поворачивается так, чтобы определяемый фазовый сдвиг, а вместе с ним и разность хода волны a равнялись нулю: $a = 0$; СТС поворачивается на некоторый угол, и выполняет повторное измерение:

$$l^2 = \frac{a^2 \cdot b^2 - \frac{a^4}{4}}{\sqrt{4b^2 \cdot \sin^2 \varphi - a^2}}, \quad (4)$$

где l – расстояние до ИЗ, м.

Отметим, что во всех экспериментах направление на ИЗ определяется в диапазоне $-90^\circ + 90^\circ$.

ЗАКЛЮЧЕНИЕ

В процессе решения указанной задачи разработан модифицированный алгоритм пеленгации ИЗ. Пеленгация предполагает определение сдвига фаз и вычисление разности амплитуд на текущем ОИ, который выбирается автоматически, основываясь на выводах о квазипериодичности и энергетической завершенности исследуемого фрагмента акустического сигнала. Отличительной особенностью алгоритма определения границ текущего ОИ является отсутствие в нем абсолютных порогов. Выводы относительно различных фрагментов АВП осуществляются исключительно на основе анализа их характеристик. Границы принятого сигнала микрофонами определяются без использования пороговых методов. Результаты данной работы могут быть использованы при решении задач автоматической обработки акустических сигналов, например в распознавании речи.

Среднее абсолютное значение ошибки пеленгации одиночного неподвижного ИЗ составило $1-3^\circ$.

Авторы статьи планируют установить разработанную СТС на борт других роботов, разрабатываемых в Институте кибернетики им. В.М. Глушкова НАН Украины.

1. *Nakadai K., Lourens T., Okuno H. G., Kitano H.* Active Audition for Humanoid // Proc. of the 17-th National Conf. on Artificial Intelligence and 12-th Conf. on Innovative Applications of Artificial Intelligence. (July 30 – August 03), 2000. – P. 832 – 839.
<http://portal.acm.org/citation.cfm?id=1226186>
2. *Murata N., Ikeda Sh.* An on-line algorithm for blind source separation on speech signals // Proc. of 1998 Intern. Symp. on Nonlinear Theory and its Applications (NOLTA'98). (Crans-Montana, Switzerland. September' 98), 1998. – P. 923 – 926.
<http://www.ism.ac.jp/~shiro/publications/index.html>
3. *Nakagawa Y., Okuno H. G., Kitano H.* Using vision to improve sound source separation // Proc. of the 16-th national conf. on Artificial intelligence and the 18-th Innovative applications of artificial intelligence conf. innovative applications of artificial intelligence American Association for Artificial Intelligence. – Menlo Park, CA, USA, 1999. – P. 768 – 775.
<http://portal.acm.org/citation.cfm?id=315149.315450&coll=GUIDE&dl=GUIDE&CFID=69330058&CFTOKEN=92390024>
4. *Brooks, A. Rodney, C. Breazeal, M. Marjanovic, B. Scassellati, M. Williamson.* The Cog Project: Building a Humanoid Robot, in C. Nehaniv, ed., Computation for Metaphors, Analogy and Agents, Vol. 1562 of Springer Lecture Notes in Artificial Intelligence, Springer-Verlag, 1998.
<http://www.ai.mit.edu/projects/humanoid-robotics-group/publications.html>
5. *The Honda humanoid robot ASIMO.*
<http://world.honda.com/ASIMO/>
6. *Человекоподобный робот из Швеции.*
<http://www.prorobot.ru/01/robot-chelovekopodobniy2.php>
7. *ЗАО «Андроидные роботы»*
<http://www.rusandroid.ru/>
8. *Павловский В.Е., Мягков А.С., Хашан Т.С., Павловский В.В.* Концепция, моделирование и разработка слуховых сенсоров для роботов // Информационно-измерительные и управляющие системы. – М.: Радиотехника, 2006. – № 1–3.
<http://www.radiotec.ru/catalog.php?cat=jr9&itm=2006-1-3>
9. *Pavlovsky V.E., Polivtsev S.A., Khashan T.S.* Intelligent Technical Audition and Vision Sensors for Walking Robot Telepresence Functions // Proc. Of 7-th Int. conf. On Climbing and Walking Robots CLAWAR'2004. Madrid, Spain, 22–24 September 2004. Springer Berlin Heidelberg. P. 387–397.
<http://www.springerlink.com/content/u065567257r7171v/>
10. *Павловский В.Е., Поливцев С.А., Хашан Т.С.* Интеллектуальные сенсоры технического слуха для мобильных роботов // Тр. шк.-конф. "Адаптивные роботы-2004" (8-11 июня 2004 г.). – Санкт–Петербург: Изд-во ЦНИИ РТК, 2004. – С. 65–71.
11. *Хашан Т.С.* Сегментация речевого сигнала // Искусственный интеллект. – 2007. – № 4 – С. 386–397.
12. *Мышко С.В., Блохин С.В., Доценко А.В.* Определение на амплитудно-временном представлении звуковых волн элементарных и однородных квазипериодических структур // Биофизика сложных систем. –1993. – № 1. – С. 92–100.

Получено 01.07.2008