УДК 621.865: 681.518

DOI 10.35211/1990-5297-2020-9-244-76-80

В. Н. Скакунов, Л. В. Жога, С. Е. Терехов, В. Ю. Бархатов

**СИСТЕМА АКУСТИЧЕСКОЙ ЛОКАЦИИ
МОБИЛЬНОГО АВТОНОМНОГО РОБОТА**

Волгоградский государственный технический университет

svn@vstu.ru

Рассматриваются способы звуковой локации и алгоритмы поиска источников звука акустической системой автономного мобильного робота. Предлагается схема реализации встраиваемой акустической системы, приводятся результаты экспериментальных исследований.

Ключевые слова: автономный мобильный робот, массив микрофонов, способы и алгоритмы звуковой локализации, система технического слуха.

© Скакунов В. Н., Жога Л. В., Терехов С. Е., Бархатов В. Ю., 2020

V. N. Skakunov, L. V. Zhoga, S. E. Terekhov, V. U. Barhatov
**SYSTEM OF ACOUSTIC LOCATION OF A WALKING ROBOT
 HUMANITARIAN APPOINTMENT**

Volgograd State Technical University

The methods of sound location and search algorithms for sound sources by the acoustic system of an autonomous mobile robot are considered. The implementation scheme of the embedded speaker system is proposed, the results of experimental studies are presented.

Keywords: autonomous mobile robot, array of microphones, methods and algorithms for sound localization, technical hearing system.

Введение

В настоящее время для повышения эффективности локационных систем, наряду с радарами, системами локального позиционирования и технического зрения [1, 2, 3], развивается пространственно-временная обработка акустических сигналов. Методы обработки акустических сигналов на основе массива микрофонов, разнесенных в пространстве, помимо прямых задач управления мобильным роботом, могут служить для идентификации звуковых сигналов, обнаружения активности источников звука и определения их координат [4, 5, 6, 10]. В работе рассматривается алгоритмическая и аппаратная реализуемость алгоритмов локализации источника звука в условиях ограниченных вычислительных ресурсов, встраиваемой системы управления роботом.

В основе многих известных решений по локализации лежит метод оценки временной задержки сигнала (DTE – time delay estimated), принимаемого парой сдвоенных микрофонов с последующим определением направления на источник сигнала и расстояния до него [4, 5].

Способы локализации источника звука

В идеальной схеме, иллюстрирующей метод DTE, измерения проводятся двумя разнесенными микрофонами с круговыми диаграммами направленности. Источник излучения и положение робота считаются статичными. При измерении принимаются следующие упрощения: однородная изотропная окружающая среда, источник и приемники сигналов располагаются в одной плоскости, предполагается отсутствие шумов и реверберации. В реальной среде сигналы шума должны учитываться, поэтому для сигналов, излучаемых источником звука, обычно используются следующая модель [4, 5]:

$$\begin{aligned}x_1(t) &= ks_1(t) + b_1(t) \\x_2(t) &= ks_1(t + \tau) + b_2(t),\end{aligned}$$

где $x_1(t)$ и $x_2(t)$ – сигналы, записанные микрофонами M1, M2 соответственно; $s_1(t)$ – звуковые сигналы; $b_1(t)$ и $b_2(t)$ – сигналы аддитивного шума микрофонов M1 и M2, добавленные к звуковым сигналам; k – коэффициент ослабления сигнала при распространении в воздухе; τ – разница во времени прихода сигналов. Предполагается, что $s(t)$ и $b_i(t)$ – независимые случайные гауссовские процессы. Для вычисления угла направления звуковой волны измеряется разница во времени прихода сигналов τ . Из геометрической интерпретации схемы определения направления (рис. 1) при плоской падающей звуковой волне с помощью двух микрофонов, разнесенных на расстояние d , угол пеленга Theta равен:

$$\text{Theta} = \arccos \frac{r}{d} = \arccos \frac{c\tau}{d}.$$

Рис. 1. Геометрическая интерпретация метода DTE

Более точно пространственное положение источника звука может быть определено в системе с тремя микрофонами [6], расположенными на концах равностороннего треугольника (рис. 2).

Рис. 2. Схема определения угла пеленга в системе с тремя микрофонами

Для каждой пары микрофонов используется та же схема оценки времени задержки. Три соответствующих угла вычисляются следующим образом:

$$\begin{aligned}\alpha &= \arccos \frac{c\tau_{12}}{d} \\ \beta &= \arccos \frac{c\tau_{23}}{d} \\ \theta &= \arccos \frac{c\tau_{21}}{d},\end{aligned}$$

где α, β, θ – углы, показанные на рис. 2; d – расстояние между микрофонами; $\tau_{12}, \tau_{23}, \tau_{21}$ – времена задержки между соответствующими парами микрофонов.

Чтобы оценить угол Theta, который указывает направление на источник звука, используются следующие уравнения [6]:

$$\begin{aligned}\text{Theta} &= \alpha + 60 \\ \text{Theta} &= -\alpha + 60 \\ \text{Theta} &= \beta \\ \text{Theta} &= -\beta \\ \text{Theta} &= \theta + 120 \\ \text{Theta} &= -\theta + 120.\end{aligned}$$

После вычисления шести значений углов Theta выбираются два ближайших друг к другу угла, вычисляется их среднее значение, которое и определяет направление на источник звука.

Простое увеличение числа микрофонов в массиве повышает точность вычисления угла пеленга. Например, в работе [7] система локализации источника звука построена на крестообразной микрофонной решетке из четырех приемников сигнала. Известны также системы с пятью, шестью и большим числом микрофонов в массиве. В работе [8] показана возмож-

ность использования произвольного числа микрофонов для определения координат источника звука. Однако и в этом случае измеряемой величиной является время задержки прихода сигнала от источника до пары микрофонов. Вычисление координат источника осуществляется методами численного решения систем дифференциальных уравнений, что в виду алгоритмической сложности затрудняет применение этого подхода во встраиваемых системах.

При сравнении схем звуковой локализации по методу DTE с массивами микрофонов более сложной топологии, как правило, используется оценка времени задержки для пары микрофонов, как основы в последующих геометрических преобразованиях с целью вычисления угла пеленга. В тоже время увеличение числа микрофонов усложняет вычислительные алгоритмы и требует существенного увеличения времени на обработку сигналов, снимаемых с микрофонов.

Алгоритмы звуковой локализации

Многие известные алгоритмы определения времени задержки сигналов от источника звука основаны на кросскорреляционном анализе и относятся к одному из двух основных подходов обработки измерительных данных – разностно-временной анализ временных рядов, оцифрованных звуковых сигналов и обобщенная кросскорреляционная обработка, включающая дискретное преобразование Фурье и операции со спектрами цифровых сигналов. Эти методы более известные, как TDOA (time difference of arrival) объединяют множество разных реализаций.

Существуют и более сложные алгоритмы обработки данных, но при ограниченных ресурсах встроенной системы управления роботом и простых приемниках звука преимущественно с круговыми диаграммами направленности рассматривались алгоритмы с наименьшей вычислительной сложностью.

Разностно-временной метод. Наиболее простой способ локализации, основан на определении взаимной корреляции двух временных рядов. В этом методе для оцифрованных в N точках сигналов от двух микрофонов (рис. 1) $x_1(n)$ и $x_2(n)$ вычисляется коэффициент взаимной корреляции по выражению [5, 6]:

$$r_{12}(k) = \frac{1}{N} \sum_{n=0}^{N-1} x_1(n)x_2(n),$$

где k – один из M возможных наборов данных.

Положение источника в двумерном пространстве находится по максимальному значению взаимно-корреляционной функции, когда разность углов прихода плоской звуковой волны на микрофоны М1 и М2 равна нулю:

$$\tau_{12} = \arg(\max(r_{12}(j))).$$

Метод прост, допускает реализацию в реальном масштабе времени, но влияние шумов настолько велико, что использовать его практически невозможно. Существенно лучшую точность обеспечивают алгоритмы обобщенной взаимной корреляции (GCC – Generalized Cross Correlation).

Метод обобщенной взаимной корреляции. В методе GCC преобразования выполняются над спектрами звуковых сигналов. После оцифровки входные сигналы фильтруются в некотором конечном окне, что позволяет избежать смешивания сигналов от различных источников и частично устранить влияние реверберации [4, 5, 6, 10]. Для вычисления временной задержки необходимы переходы из временной области в частотную и обратно, осуществляемые с помощью прямого и обратного дискретного преобразования Фурье (ДПФ). Для двух микрофонов М1 и М2, принимающих сигналы $x_1(n)$ и $x_2(n)$ (n номер выборки во временной области), через ДПФ вычисляются спектры $X(k)$ и $X_2(k)$ (k – индекс частотной выборки). Спектральная плотность перекрестной мощности сигналов определяется в виде:

$$\Phi_{x_1x_2}(n) = X_1(n)X_2^*(n),$$

где $X_2^*(n)$ – прямое преобразование Фурье над комплексно-сопряженными числами.

Кросс-корреляция между этими сигналами задается уравнением [5]:

$$R_{x_1x_2}(j) = \frac{1}{M} \sum_{k=0}^{M-1} \Phi_{x_1x_2}(k) e^{\frac{j2\pi nk}{M}},$$

где M – длина выборки, на которой вычисляется спектральная плотность; j – сдвиг между дискретными значениями двух сигналов. Значение $R_{x_1x_2}(j)$ может быть вычислено для диапазона возможных отрицательных и положительных сдвигов. Задержка, при которой функция взаимной корреляции $R_{x_1x_2}(j)$ имеет максимум, может быть вычислена по формуле

$$\tau_{12} = \frac{1}{F_s} \arg\max_n R_{x_1x_2}(j).$$

Функция взаимной корреляции в отдельных случаях имеет довольно размытый максимум. Поэтому для достижения более резкого пика

у функции взаимной корреляции и большей точности в оценке задержки вводится весовая функция Ψ_Φ , с учетом которой задержка равна [4, 5, 6]:

$$\tau_{12} = \frac{1}{F_s} \arg\max_n \sum_{k=0}^{M-1} \Psi_\Phi(k) R_{x_1x_2}(k) e^{\frac{j2\pi nk}{M}}.$$

В работах [4, 6, 10] показано, что в голосовом диапазоне частот лучший результат дает весовая функция преобразование фазы (PHAT) вида:

$$\Psi_\Phi = \frac{1}{|\Phi_{x_1x_2}(n)|}.$$

Результаты экспериментальных исследований

Цель экспериментов заключалась в оценочном определении ошибки измерений и временных затрат в базовом варианте с двумя микрофонами. При построении макета системы звуковой локации использовались малогабаритные микрофоны с круговыми диаграммами и простые вычислительные алгоритмы. В качестве приемников использовались электретные микрофоны ЕМ-9745, МАХ9814, цифровой микрофон МР34DT05ТРА и активный микрофон ШОРОХ-7.

Макет построен на 32-разрядном микроконтроллере STM32F407. Контроллер имеет достаточно высокую производительность и развитый физический интерфейс. Временная задержка вычислялась по алгоритму обобщенной взаимной корреляции с преобразованием фазы (GCC-PHAT). Спектральные преобразования выполнялись по схемам прямого и обратного быстрого преобразования Фурье (БПФ).

Полосовые фильтры проектировались как линейные фильтры с окном Блэкмана. Параметры фильтра настраивались на весь частотный диапазон (50–12 000) Гц и отдельные скользящие окна размером в 500 Гц в том же диапазоне. Источниками звука служили моногармонические и полигармонические сигналы. Как базовый алгоритм, определяющий время вычисления по методу GCC-PHAT, применялись методы расчеты прямого и обратного БПФ. При числе выборок в 1024 на кратковременный анализ по БПФ потребовалось не более 5,20 мс. Время цифровой фильтрации и корреляционного анализа при выполнении простых операций попарного сложения и суммирования чисел с фиксированной точкой оказалось пренебрежимо мало по сравнению с операциями БПФ.

Невысокая скорость передвижения мобильных роботов позволяет рассчитывать на работу в реальном масштабе времени. Ошибка вычисления угла направления на источник для электретных микрофонов не больше $7,5^\circ$ на расстоянии до 1,5 м. Более стабильные и точные измерительные данные получаются с цифрового выхода микрофона MP34DT05TRA, в котором аналоговый сигнал с МЭМС-датчиков встроенным контроллером преобразуется в PDM-поток данных, т. е. импульсно-модулированный сигнал. В микроконтроллере PDM-сигналы от двух микрофонов в таймер/счетчиках, установленных в режим захвата, переводятся в двоичный код. Помехозащищенность PDM-сигнала значительно выше аналогового напряжения, но и в этом случае дальность обнаружения источника звука не превышает 2,5 м. Существенно лучшие характеристики получены при использовании активного микрофона ШОРОХ-7 со встроенным АРУ. Расстояние, различимое системой, увеличилось до 5–6 м, а усредненная оценка угла (ошибка) оказалась в диапазоне $3,5$ – $4,2^\circ$.

При проведении экспериментов с целью проверки реализуемости и точности алгоритма GCC-RHAT выполнения функции локализации в реальном масштабе времени варьировались сочетания параметров: расстояние между микрофонами, частота дискретизации, время наблюдения, частота источника звука, порядок оконных фильтров. Результаты эксперимента подтверждают возможность построения встроенной системы звуковой локации на выбранных компонентах.

БИБЛИОГРАФИЧЕСКИЙ СПИСОК

1. Processing of Data from the Camera of Structured Light for Algorithms of Image Analysis in Control Systems of Mobile Robots / В. Н. Скакунов, В. А. Беликов, В. В. Жога, И. А. Несмиянов // *Advances in Mechanical Engineering : Selected Contributions from the Conference «Modern Engineering: Science and Education»* (Saint Petersburg, Russia, June, 2014) / ed. by A. Evgrafov. – [Switzerland] : Springer International Publishing, 2016. – P. 149–158. – (Book Series: Lecture Notes in Mechanical Engineering).
2. Мобильные технологические роботы и тренажеры: интеграционное программное обеспечение группового взаимодействия / В. П. Андреев [и др.] // *Информационно-измерительные и управляющие системы*. – № 4. – 2013. – С. 74–79.
3. Жога, В. В. Алгоритм адаптивного управления приводом шагающего робота на основе нечеткой логики / В. В. Жога, С. Е. Терехов, В. Н. Скакунов // *Известия ВолгГТУ : научный журнал* № 3 / ВолгГТУ. – Волгоград, 2019. – (Серия «Роботы, мехатроника и робототехнические системы»). – С. 12–18.
4. Ронжин, А. Л. Сравнение методов локализации пользователя многомодальной системы по его речи / А. Л. Ронжин, А. А. Карпов // *Изв. вузов. Приборостроение*. – 2008. – Т. 51. – № 11. – С. 41–47.
5. Knapp, C. H. The generalized correlation method for estimation of time delay / C. H. Knapp, G. C. Carter // *IEEE Trans. Acoustics Speech Signal Proc.* 1979. Vol. 24, p. 320–327.
6. Hasan, Khaddour. A Comparison of Algorithms of Sound Source Localization Based on Time Delay Estimation // *Electrorevue*. – 2011, Vol 2, № 1, p. 31–37.
7. Real-time Sound Source Localization on a Mobile Robot Based on the Guided Spectral-Temporal Position Method / Xiaofei Li et al // *International Journal of Advanced Robotic Systems*. – 2012. – Vol. 9.
8. Львов, А. В. Распределенная микроконтроллерная система акустической локации / А. В. Львов, М. Н. Агапов, А. И. Тищенко // *Радиоэлектроника*. – 2010. – № 11.
9. Разработка многомодального информационного киоска / А. Л. Ронжин, А. А. Карпов, Ан. Б. Леонтьева, Б. Е. Костюченко // *Тр. СПИИРАН*. – 2007. – Вып. 5. – С. 227–245.
10. Acoustic Source Localization Based on Generalized Crosscorrelation Time-delay Estimation / L. Chen et al // *Procedia Engineering*. – 2011. – № 15.