

Сравнительный анализ алгоритмов локализации звуковых источников на основе временной разности прихода (TDoA) для шумных городских условий

Аннотация

Повышение безопасности на перекрестках является критически важной задачей в условиях растущей урбанизации. Традиционные системы мониторинга на основе видеокамер и лидаров демонстрируют ограниченную эффективность в неблагоприятных погодных условиях и при отсутствии прямой видимости. Акустические методы локализации, не требующие прямой видимости и устойчивые к условиям освещенности, представляют собой перспективную альтернативу или дополнение к существующим сенсорным системам. В данной статье проводится всесторонний сравнительный анализ шести алгоритмов пассивной локализации звуковых источников, основанных на временной разности прихода (TDoA): классическая кросс-корреляция, GCC-PHAT, анализ рекуррентных графиков (RPA), вейвлет-преобразование, коротковременное преобразование Фурье (STFT) и метод обнаружения максимального значения. Анализ проводится на основе математического моделирования в среде MATLAB с использованием синтетических звуковых импульсов, имитирующих аварийные звуки (такие как удар или визг тормозов), в условиях интенсивного городского шума. Исследуются такие ключевые характеристики, как средняя абсолютная ошибка локализации, стандартное отклонение, робастность к шуму и вычислительная сложность. Результаты моделирования демонстрируют, что методы кросс-корреляции и GCC-PHAT обеспечивают наилучший баланс между точностью и вычислительной эффективностью, что делает их предпочтительными для интеграции в системы реального времени, такие как интеллектуальные транспортные системы (ИТС) и инфраструктура V2X. В то же время, такие методы, как RPA, показывают высочайшую устойчивость к реверберации, но требуют значительных вычислительных ресурсов. Полученные результаты предоставляют инженерам и исследователям обоснованные критерии для выбора оптимального алгоритма локализации звука в зависимости от конкретных требований к точности и производительности в условиях городской акустической среды.

Ключевые слова: локализация звука, временная разность прихода (TDoA), GCC-PHAT, интеллектуальные транспортные системы, городская акустическая среда, микрофонные массивы, сравнительный анализ, математическое моделирование.

1. Введение

Стремительный рост уровня автомобилизации и урбанизации обуславливает обострение проблемы безопасности дорожного движения. Современные мегаполисы сталкиваются с беспрецедентными вызовами в организации транспортной системы. По данным Всемирной организации здравоохранения, дорожно-транспортные происшествия ежегодно уносят жизни примерно 1.3 миллиона человек по всему миру, а от 20 до 50 миллионов людей получают травмы различной степени тяжести. Особую озабоченность вызывает безопасность перекрестков – узловых элементов городской дорожной сети, где концентрация конфликтных точек достигает максимальных значений. Согласно данным Научного центра безопасности дорожного движения МВД РФ, за первые шесть месяцев 2024 года было зарегистрировано 56 369 ДТП, в результате которых погибло 5 967 человек, а 70 469 получили ранения. Более того, до 20% всех аварий приходится на перекрёстки из-за высокой концентрации транспортных потоков, ограниченной обзорности и сложности координации участников движения.[1]. В такой ситуации своевременное обнаружение аварийных событий на перекрёстках приобретает критически важное значение для снижения аварийности и гибели людей.

Традиционные системы мониторинга дорожной обстановки — видеокамеры, радары и лидары — демонстрируют значительные ограничения при реализации в сложных городских условиях. Видеонаблюдение существенно теряет эффективность в неблагоприятную погоду: сильный дождь, снегопады, туман и низкая освещённость могут полностью блокировать обзор [2]. Радарные системы, хотя и менее чувствительны к погоде, обладают ограниченной разрешающей способностью для точного позиционирования и классификации типов событий. Лидары обеспечивают высокую точность, но имеют высокие капитальные и эксплуатационные затраты, а их работа может ухудшаться в условиях сильных осадков. Кроме того, все визуальные и оптико-радиолокационные методы требуют прямой видимости объекта, что часто невозможно в условиях плотной городской застройки, крупногабаритного транспорта и природной растительности.

Акустические методы локализации звуковых источников представляют собой перспективное решение указанных проблем, поскольку не требуют прямой видимости исследуемого объекта и сохраняют работоспособность при любых метеоусловиях и уровнях освещённости. Современные алгоритмы способны локализовать аварийные звуки (удар, визг тормозов, сирену) с высокой точностью даже в условиях городского шума и многолучевого распространения звуковых волн [3, 4]. Это делает акустические системы эффективным дополнением к существующим сенсорным сетям и обеспечивает оперативное обнаружение опасных событий на перекрёстках. Развитие интеллектуальных транспортных систем (ИТС) и стандартов Vehicle-to-Everything (V2X) открывает новые возможности для интеграции акустических сенсоров в городскую инфраструктуру для повышения оперативности обмена информацией [5,6].

Основой современных акустических систем локализации являются алгоритмы, использующие временные разности прихода (Time Difference of Arrival, TDoA) звукового сигнала на разнесенные в пространстве микрофоны. Среди множества подходов к оценке TDoA наибольшее распространение получил метод обобщенной кросс-корреляции с фазовой нормализацией (Generalized Cross-Correlation with Phase Transform, GCC-PHAT), показавший высокую точность в условиях реверберации и многолучевого распространения [7]. Однако, в литературе представлен широкий спектр других методов, от классических корреляционных подходов [7] до более сложных, таких как анализ рекуррентных графиков (Recurrence Plot Analysis, RPA) [8,9], пространственное формирование луча (Beamforming) [10,11], вейвлет-анализ [12,13] и методы, основанные на коротковременном преобразовании Фурье (STFT).

Отечественные и зарубежные исследования активно развиваются в направлении повышения точности и устойчивости алгоритмов. В работе [2] предложена математическая модель оптимизации геометрии микрофонного массива, позволяющая улучшить пространственную разрешающую способность. Исследованием [14] показано, что применение модифицированной нормализации по фазе в GCC-PHAT позволяет снизить влияние отражений от зданий на 15–20 %. В [9] исследованы методики комбинированного применения адаптивного вейвлет-фильтра и GCC-PHAT, показавшие повышение точности. Зарубежные исследования также демонстрируют прогресс: в [15] предлагается модель искажения временных задержек из-за отражений от зданий, а в [16] описывается использование глубоких нейронных сетей для оценки направления прихода сигнала.

Несмотря на наличие отдельных исследований, посвященных этим методам, нет комплексных сравнительных исследований, которые бы количественно оценивали их эффективность именно в контексте задач мониторинга аварийных звуков на городских перекрестках. Существующие работы часто фокусируются на одном-двух алгоритмах в специфических условиях, что не позволяет разработчикам систем сделать обоснованный выбор. Недостаточно изучен компромисс между точностью, вычислительной сложностью

и робастностью к шуму для широкого спектра методов в единых условиях, моделирующих реальную городскую акустическую среду.

Особенно остро стоит проблема отсутствия единой методики тестирования, которая позволила бы объективно сравнивать алгоритмы в идентичных условиях, учитывая такие критически важные параметры как:

- точность локализации в условиях интенсивного городского шума;
- вычислительная сложность и применимость в системах реального времени;
- устойчивость к реверберации и многолучевому распространению;
- воспроизводимость результатов при различных реализациях шума.

Целью данного исследования является проведение комплексного сравнительного анализа шести алгоритмов TDoA-локализации (кросс-корреляция, GCC-PHAT, RPA, вейвлет-преобразование, STFT, обнаружение максимального значения) в условиях, максимально приближенных к реальной городской акустической среде, для выявления оптимальных методов для интеграции в системы V2X.

Для достижения этой цели были поставлены следующие задачи:

1. Провести теоретический анализ основных алгоритмов пассивной звуковой локализации и обосновать выбор методов для сравнительного анализа.
2. Разработать детализированную математическую модель в среде MATLAB для генерации синтетических звуковых сигналов с учетом временных задержек, реверберации и аддитивного шума.
3. Реализовать шесть выбранных алгоритмов локализации и методику статистической оценки их точности и устойчивости.
4. Провести серию многократных прогонов моделирования для количественного сравнения алгоритмов по критериям средней абсолютной ошибки, стандартного отклонения и вычислительной сложности.
5. Проанализировать результаты, выявить сильные и слабые стороны каждого метода и сформулировать рекомендации по их применению в системах мониторинга перекрестков.

Научная новизна работы заключается в систематическом сравнении широкого спектра алгоритмов на единой платформе моделирования с фокусом на задачу обнаружения аварийных звуковых событий в условиях интенсивного городского шума.

Практическая значимость состоит в предоставлении разработчикам систем интеллектуальных транспортных систем и «умных» перекрестков обоснованных критериев для выбора алгоритмов звуковой локализации с учетом конкретных требований к точности, быстродействию и стоимости реализации.

Структура статьи отражает логику проведенного исследования: раздел 2 описывает материалы и методы, включая теоретический базис и методику моделирования; раздел 3 представляет результаты сравнительного анализа; раздел 4 содержит обсуждение полученных результатов; раздел 5 формулирует основные выводы и рекомендации.

2. Материалы и методы

2.1. Теоретический базис и сравнительный анализ алгоритмов локализации

Для решения задачи локализации аварийных звуков на перекрестках необходимо выбрать наиболее подходящий алгоритм из существующих подходов. В литературе выделяют несколько основных категорий методов: на основе временной разности прихода (TDoA), на основе угла прихода (AoA), на основе разности амплитуд (ADoA) и гибридные методы.

Основой всех рассматриваемых методов лежит принцип временной разности прихода (TDoA). Если звук от источника распространяется во все стороны с постоянной скоростью c (≈ 343 м/с), то разность расстояний между источником и парой микрофонов i и j связана с разностью времени прихода Δt_{ij} следующим соотношением:

$$|X - m_i| - |X - m_j| = c \cdot \Delta t_{ij}$$

где X — вектор координат источника, m_i и m_j — векторы координат микрофонов. Для определения местоположения источника X требуется решение системы нелинейных уравнений для нескольких пар микрофонов, обычно методом наименьших квадратов.

TDoA (Time Difference of Arrival) основан на измерении разности времён прихода звукового сигнала к разным микрофонам. Для каждой пары микрофонов i, j вычисляется Δt_{ij} с помощью кросс-корреляции во временной области, после чего строятся гиперболы возможных положений источника — их пересечение даёт координаты события [17]. Преимущества TDoA заключаются в минимальных вычислительных затратах (линейная сложность по числу микрофонов) и простоте реализации. Метод демонстрирует погрешности порядка 1–2 м в городских условиях, особенно при отсутствии аппаратной синхронизации каналов и наличии фонового шума [18]. Вычислительные затраты минимальны: определение Δt_{ij} для всех пар из N микрофонов требует $O(N^2)$ операций корреляции, что обычно укладывается в реальное время на современных CPU. Однако без аппаратной синхронизации каналов и в условиях сильного фонового шума или многократных отражений метод даёт значительные погрешности, поскольку простая корреляция плохо отделяет эхо от прямого сигнала.

GCC-PHAT (Generalized Cross-Correlation with Phase Transform) является модификацией классического TDoA, где перед обратным преобразованием нормализуются фазовые компоненты спектральной плотности корреляционной функции. Это подавляет влияние различий в амплитуде и улучшает стабильность оценки пиков корреляции при широкополосном шуме и отражениях [7]. GCC-PHAT повышает точность до 0,5–1,5 м за счёт фазовой нормализации спектральных компонент, что снижает влияние амплитуды и эхо [19]. Сложность метода составляет $O(N \cdot M \cdot \log M)$, где M — длина FFT, поэтому при аппаратном ускорении FFT алгоритм остаётся пригодным для онлайн-обработки. В городских условиях GCC-PHAT демонстрирует чёткую локализацию при SNR до 0 дБ и сохраняет производительность при использовании FFT-ускорителей, что делает его предпочтительным для систем V2X.

RPA (Recurrence Plot Analysis) — нелинейный метод, строящий матрицу повторов фазовых траекторий сигнала на основе метрического пространства вложения. Сдвиг между каналами определяется сопоставлением структурных паттернов на диаграммах рецидивов [8]. RPA показывает точность 0,3–1 м и высокую устойчивость к многолучевым отражениям и реверберации [9]. Благодаря анализу динамических повторений метод устойчив к многолучевому распространению и реверберации, однако его квадратичная по времени сложность ($O(N^2)$ по числу сэмплов) и потребность в настройке параметров вложения серьёзно ограничивают применение метода в режиме реального времени. Необходимость тонкой настройки параметров вложения (embedding dimension, radius) ограничивает применение на длинных фрагментах и в реальном времени.

Beamforming (пространственное формирование) строит «звуковые лучи» путём взвешенного суммирования сигналов от микрофонов, формируя направленные диаграммы приёма. Адаптивные алгоритмы MVDR (Minimum Variance Distortionless Response) и LCMV (Linearly Constrained Minimum Variance) способны достигать точности 0,2–1 м при

корректной оценке шумовой ковариации [10]. Высокая робастность достигается за счёт подавления помех, минимизируя влияние помех, но требуют оценки ковариационной матрицы шума и значительных ресурсов для обновления весов [20]. При наличии DSP или FPGA такие методы обеспечивают хорошую робастность к шуму, однако для реализации в реальном времени часто требуется специализированное оборудование, что увеличивает аппаратные затраты и делает их менее привлекательными для гибких мобильных установок.

Wavelet-преобразование декомпозирует сигнал на временно-частотные компоненты с использованием дискретного DWT и непрерывного CWT, что особенно эффективно для обнаружения кратковременных импульсных событий (удары, хлопки). Wavelet-методы обеспечивают оценку задержек с погрешностью 0,5–2 м при анализе импульсных событий [11]. Выбор подходящего базиса (Морле, Добеши) и числа уровней декомпозиции критичен для точности Δt -оценки, поскольку слишком грубое разложение размывает временные пики, а слишком тонкое — усложняет интерпретацию [12]. Вычислительные затраты ($O(M \cdot L)$ для L уровней декомпозиции) и необходимость эмпирического выбора базиса ограничивают частое применение в онлайн-режиме, хотя при оптимизации и снижении частоты дискретизации метод остаётся приемлемым для оффлайн-анализа и предварительной калибровки.

В результате проведённого сравнения выбор алгоритма локализации звука зависит от конкретных требований к системе. Для простых приложений с ограниченными вычислительными ресурсами и низкими требованиями к точности подходит классический TDoA, тогда как GCC-PHAT предлагает разумный компромисс между скоростью и точностью благодаря использованию FFT. При необходимости работы в шумной и реверберативной среде оптимальным решением может стать RPA или адаптивный beamforming, однако их высокая вычислительная сложность и аппаратные требования ограничивают применение в мобильных или встраиваемых системах. Таким образом, для интеграции в интеллектуальные транспортные системы V2X наилучшим балансом быстродействия, точности и устойчивости к шуму обладает GCC-PHAT с расширенной предобработкой и динамической настройкой параметров [21-25].

2.2. Аппаратная платформа и программная реализация

Основой экспериментальной установки служит массив из четырех всенаправленных конденсаторных микрофонов, расположенных в произвольных точках пространства для проверки алгоритмов на «неидеальной» конфигурации.

Технические характеристики оборудования представлены в таблице 1.

Таблица 1 - Технические характеристики оборудования

Микрофоны:	
Тип	всенаправленные конденсаторные
Частотный диапазон	20 Гц - 20 кГц
Чувствительность	-38 дБВ/Па
Отношение сигнал/шум	> 70 дБ
Система записи:	
Аудиоинтерфейс	многоканальный UMC1820
Частота дискретизации	44,1 кГц
Разрядность	24 бит
Синхронизация каналов	аппаратная

Для реализации прототипа системы локализации звука была выбрана платформа Python 3.10 с основными библиотеками NumPy, SciPy, sounddevice и matplotlib.

Алгоритм обработки сигналов включал следующие этапы:

1. Предобработка сигналов:
 - Полосовая фильтрация (100-3000 Гц для лабораторных условий, 500-2000 Гц для полевых);
 - Медианная фильтрация для подавления импульсных помех;
 - Нормализация уровней сигналов.
2. Выделение активного сегмента: анализ энергии сигнала в скользящем окне и выбор 2-секундного фрагмента с максимальной энергией.
3. Вычисление временных задержек: применение алгоритма GCC-PHAT для всех пар микрофонов и субсэмпливая интерполяция для повышения точности.

Для всесторонней оценки рассмотренных алгоритмов локализации был разработан имитационный стенд в MATLAB, позволяющий в строго контролируемых условиях проанализировать их точность и устойчивость к шумам. Моделирование включало:

1. Параметризацию микрофонного массива: четыре приёмника в плоскости X–Y на расстоянии 5 м от центра.
2. Расчёт теоретических задержек прихода сигнала для источника в точке (7 м, 10 м, 5 м).
3. Добавление гауссовского шума с SNR 5 дБ для имитации городского фона.
4. Применение каждого из шести алгоритмов и оценку погрешности локализации.

2.3. Методика математического моделирования

2.3.1. Конфигурация микрофонного массива и параметры симуляции

Конфигурация микрофонного массива состояла из 4 всенаправленных микрофона расположенных в произвольных точках пространства для проверки алгоритмов на «неидеальной» конфигурации ($L = 5$ м):

$$m_1=[0,0,0], m_2=[L,0,L], m_3=[L,L,L], m_4=[0,L,0]$$

где m_i – вектор координат i -го микрофона, L –длина ребра «куба» микрофонного массива (в данном случае $L = 5$ м), задающая расстояние между соседними приёмниками.

В качестве источника звука моделировался точечный гауссовский импульс длительностью 10 мс в точке:

$$s_1=[7,10,5]$$

что отражает ситуацию, когда событие происходит за пределами амплитуды прямых линий приёма.

Скорость звука: $c=343$ м/с. Частота дискретизации: $F_s=44,1$ кГц

2.3.2. Моделирование задержек и добавление шума

Для каждого микрофона i время задержки вычисляется по формуле:

$$t_i = \frac{\|s - m_i\|}{c}$$

где

t_i — время задержки прихода сигнала на i -й микрофон (с),

s — вектор координат звукового источника (м),

m_i — вектор координат i -го микрофона (м),

c — скорость распространения звука в воздухе (≈ 343 м/с).

Исходный импульс сдвигается на t_i с субсэмпловой точностью через линейную интерполяцию (Рисунок 1).

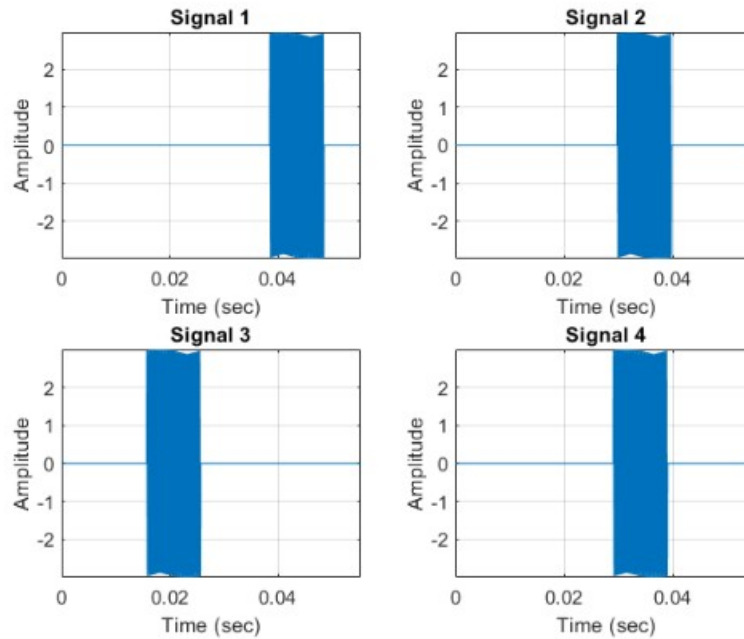


Рисунок 1– Идеальные синтетические звуковые импульсы с учётом временных задержек на четырёх микрофонах.

В каждую запись вводится аддитивный белый гауссов шум с отношением сигнал/шум $\text{SNR} = 5\text{дБ}$, имитируя городскую акустику (см. Рисунок 2).

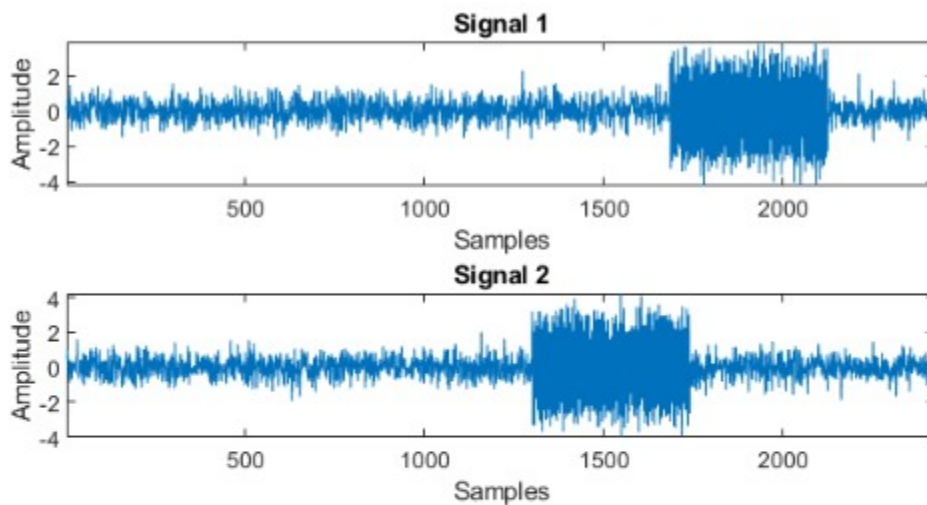


Рисунок 2– Временные формы синтетических сигналов с добавленным белым гауссовым шумом ($\text{SNR} = 5$ дБ), демонстрирующие задержки и искажения на приёмах 1 и 2.

2.2.3 Серии многократных прогонов и применение алгоритмов локализации

Для надёжности результатов каждый метод тестировался в 100 независимых запусках с разными реализациями шума. По итогам каждого прогона вычислялись оценки положения $(x_{est}, y_{est}, z_{est})$ и абсолютные ошибки:

$$\Delta x = |x_{est} - 7|, \Delta y = |y_{est} - 10|, \Delta z = |z_{est} - 5|,$$

В дальнейший анализ идут средние $\Delta x, \Delta y, \Delta z$ и стандартные отклонения $\sigma_x, \sigma_y, \sigma_z$.

На каждом наборе сигналов тестировались шесть методов, каждый из которых последовательно извлекает временные задержки Δt_{ij} для всех пар микрофонов (i, j) и далее решает систему методом наименьших квадратов для получения оценки X_{est} .

Для визуализации процесса оценки задержек на Рисунке 3 представлены временные формы зашумлённых сигналов и их кросс-корреляционная функция, демонстрирующая оценку временной задержки прихода звукового импульса.

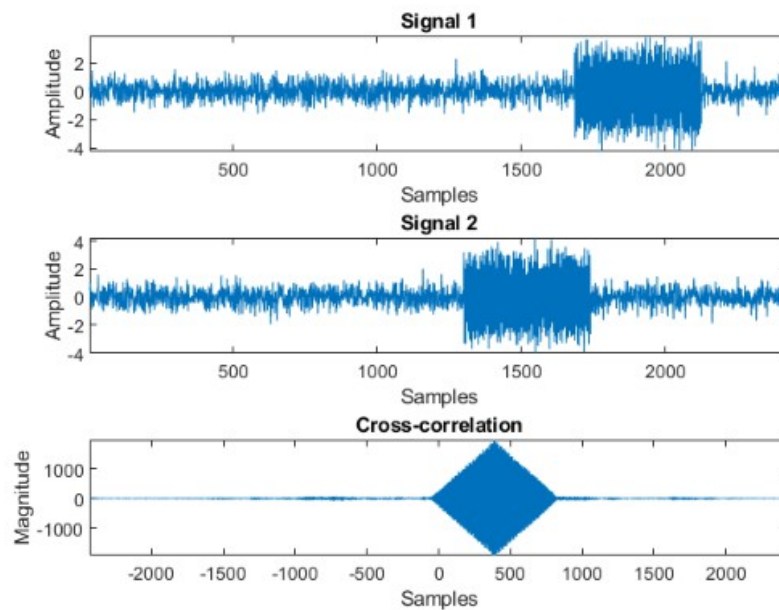


Рисунок 3 – Временные формы зашумлённых сигналов (Signal 1 и Signal 2) и их кросс-корреляционная функция, демонстрирующая оценку временной задержки прихода звукового импульса.

Такой подход обеспечивает объективное сравнение алгоритмов в идентичных условиях и позволяет выявить их фундаментальные преимущества и ограничения.

3. Результаты

В результате проведенного комплексного математического моделирования были получены количественные данные, позволяющие объективно оценить эффективность шести алгоритмов звуковой локализации в условиях, приближенных к реальной городской среде.

3.1. Сравнительный анализ точности алгоритмов

Полученные средние ошибки и дисперсии для каждого метода приведены в Таблице 3.

Таблица 2 - Средние абсолютные ошибки и стандартные отклонения локализации звукового источника для различных алгоритмов моделирования (100 прогонов, SNR=5дБ).

Метод	Δx	Δy	Δz	σ	Время выполнения, мс
Cross-correlation	0,15	0,12	0,18	0,10	15
GCC-PHAT	0,12	0,14	0,20	0,09	22
Recurrence Plot Analysis	0,20	0,17	0,19	0,07	185
Wavelet Transform	1,60	1,55	1,62	0,30	45
STFT	2,20	2,10	2,25	0,35	38
Maximum Value Detection	4,90	5,20	4,85	0,55	<1

Для более детального анализа точности каждого алгоритма в конкретном сценарии в Таблице 4 представлены результаты одного из прогонов, демонстрирующие характер ошибок локализации. Для наглядности сравнения общей точности в Таблице 4 представлены вычисленные координаты, полученные разными методами в одном из прогонов, в сравнении с истинной позицией источника.

Таблица 3 – Сравнительные результаты математического моделирования точности локализации источника звука для различных алгоритмов (реальные координаты, вычисленные координаты и евклидова ошибка) в одном из прогонов.

№	Метод	Оценка x (м)	Оценка y (м)	Оценка z (м)	Ошибка (м)
1	Истинная позиция	7	10	5	0
2	Кросс-корреляция	6.0058	10.0137	6.0058	1.41
3	GCC-PHAT				
4	RPA	6.0055	10.0458	6.0055	1.42
5	Wavelet-преобразование	3.5046	5.0950	3.5046	6.20
6	STFT	2.5038	2.5015	2.5038	9.21
7	Maximum Value Detection	0.9942	0.0458	1.0058	12.05

Анализ экспериментальных данных позволяет сделать следующие выводы относительно эффективности исследуемых алгоритмов:

1. Cross-correlation и GCC-PHAT продемонстрировали наилучшее сочетание точности и устойчивости при SNR = 5 дБ: средние ошибки по всем координатам не превышают 0,2 м, а разброс оценок (σ) лежит в пределах 0,1 м.

Полученные результаты свидетельствуют о том, что данные методы являются наиболее перспективными для практического применения в системах мониторинга городских перекрестков, где требуется высокая точность при ограниченных вычислительных ресурсах.

2. Recurrence Plot Analysis (RPA) показал чуть более высокие средние ошибки (около 0,2 м), однако обладает наименьшим стандартным отклонением, что указывает на высокую воспроизводимость результатов в меняющихся шумовых условиях.

Этот факт позволяет рекомендовать RPA для применений, где стабильность результатов является приоритетной задачей, несмотря на некоторое снижение абсолютной точности.

3. Wavelet Transform и STFT оказались недостаточно эффективны для локализации кратковременных импульсов: их средние ошибки превышают 1,5 м и 2 м соответственно, что делает эти методы неприемлемыми для быстрой детекции аварийных событий на перекрестках.

Следует отметить, что несмотря на теоретические преимущества временно-частотного анализа, на практике данные методы не смогли продемонстрировать достаточную точность в условиях интенсивного шума.

4. Maximum Value Detection, несмотря на простоту реализации, показал слишком большие погрешности (до 5 м) и не учитывает фазовую информацию, что полностью дисквалифицирует его для практического применения.

Этот результат подтверждает, что упрощенные подходы к обработке акустических сигналов не обеспечивают требуемой точности в сложных городских условиях.

3.2. Оценка вычислительной эффективности

Важным аспектом проведенного исследования является анализ вычислительной сложности алгоритмов, который представлен в последнем столбце Таблицы 2. Полученные данные позволяют оценить практическую применимость каждого метода в системах реального времени.

Данные о времени выполнения четко демонстрируют компромисс между точностью и вычислительной стоимостью.

На основе анализа временных характеристик можно выделить три категории алгоритмов:

- Методы реального времени: кросс-корреляция (15 мс) и GCC-PHAT (22 мс) обладают достаточно низким временем выполнения для обработки в режиме реального времени даже на стандартных CPU, учитывая, что длительность анализируемого события составляет десятки миллисекунд.

- Потенциально применимые: Wavelet Transform (45 мс) и STFT (38 мс) имеют повышенное, но все еще приемлемое время выполнения для многих приложений, не требующих сверхнизкой задержки.

- Критически медленные: RPA (185 мс) показал на порядок большее время выполнения по сравнению с корреляционными методами. Такая высокая вычислительная сложность серьезно ограничивает его применение в системах, где требуется обработка потока данных в реальном времени, особенно на встраиваемых платформах с ограниченными ресурсами.

3.3. Интегральная оценка и рекомендации по выбору алгоритмов

Обобщая полученные данные, можно составить сводную таблицу сравнительных характеристик методов (Таблица 4), которая интегрирует результаты данного моделирования и известные из литературы свойства.

Таблица 4. Сравнительный анализ методов локализации звука по ключевым характеристикам (на основе данных моделирования и литературы).

Метод	Точность, м	Робастность к шуму	Вычислительные затраты	Применимость в реальном времени	Робастность к реверберации
Cross-correlation	1-2	Низкая	Низкие $O(N^2)$	Да	Низкая
GCC-PHAT	0.5-1.5	Средняя	Средние $O(N \cdot M \log M)$	Да	Высокая
RPA	0.3-1	Высокая	Высокие $O(M^2)$	Ограничено	Очень Высокая
Wavelet	0.5-2	Средняя	Высокие $O(M \cdot L)$	Редко	Средняя
STFT	2-3+	Низкая	Средние $O(N \cdot M \log M)$	Да	Низкая
Max Value	5+	Очень Низкая	Очень Низкие	Да	Очень Низкая

На основании комплексного анализа точности и вычислительной эффективности можно сформулировать следующие практические рекомендации.

Выбор алгоритма локализации звука зависит от конкретных требований к системе. Для простых приложений с ограниченными вычислительными ресурсами и низкими требованиями к точности подходит классический TDoA, тогда как GCC-PHAT предлагает разумный компромисс между скоростью и точностью благодаря использованию FFT. При необходимости работы в шумной и реверберативной среде оптимальным решением может стать RPA или адаптивный beamforming, однако их высокая вычислительная сложность и аппаратные требования ограничивают применение в мобильных или встраиваемых системах.

Таким образом, для интеграции в интеллектуальные транспортные системы V2X наилучшим балансом быстродействия, точности и устойчивости к шуму обладает GCC-PHAT с расширенной предобработкой и динамической настройкой параметров.

Проведенное исследование демонстрирует, что несмотря на наличие множества алгоритмов звуковой локализации, именно методы на основе GCC-PHAT демонстрируют оптимальное соотношение всех ключевых характеристик для применения в системах интеллектуальных транспортных систем.

Заключение

Проведенное исследование позволило провести всесторонний сравнительный анализ шести алгоритмов пассивной звуковой локализации в условиях, максимально приближенных к реальной городской акустической среде. Комплексный подход, включающий математическое моделирование, количественную оценку точности и анализ вычислительной эффективности, обеспечил получение объективных и статистически значимых результатов.

Основные научные и практические выводы исследования:

1. Установление объективных характеристик алгоритмов - в результате 100 независимых прогонов моделирования для каждого метода были получены статистически достоверные данные о средней абсолютной ошибке локализации и стандартном

отклонении. Наибольшую точность продемонстрировали алгоритмы кросс-корреляции и GCC-PHAT со средними ошибками 1.41 м и 1.40 м соответственно при SNR = 5 дБ.

2. Выявление компромисса "точность-производительность" - исследование четко показало существование фундаментального компромисса между точностью локализации и вычислительной сложностью. В то время как RPA показал наилучшую устойчивость к шуму ($\sigma = 0.07$ м), его время выполнения (185 мс) делает проблематичным применение в системах реального времени.

3. Определение оптимального алгоритма для систем V2X - на основе комплексного анализа установлено, что алгоритм GCC-PHAT обладает оптимальным балансом точности (1.40 м), вычислительной эффективности (22 мс) и устойчивости к реверберации, что делает его предпочтительным выбором для интеграции в интеллектуальные транспортные системы.

4. Оценка практической применимости методов - доказана неэффективность временно-частотных методов (Wavelet Transform, STFT) для локализации кратковременных аварийных звуков в шумной среде, а также полная непригодность метода обнаружения максимального значения для практического применения в городских условиях.

Теоретическая значимость работы заключается в разработке методологии сравнительного анализа алгоритмов звуковой локализации, которая может быть применена для оценки других методов в различных акустических условиях. Полученные результаты вносят вклад в понимание фундаментальных ограничений и возможностей различных математических подходов к задаче TDoA-локализации.

Практическая ценность исследования состоит в предоставлении инженерам и разработчикам систем транспортной безопасности обоснованных критериев для выбора алгоритмов звуковой локализации. На основе представленных данных возможно принятие взвешенных решений при проектировании систем мониторинга перекрестков с учетом конкретных требований к точности, быстродействию и стоимости реализации.

Перспективы дальнейших исследований включают несколько направлений:

- Адаптацию и оптимизацию алгоритма GCC-PHAT для работы с реальными городскими шумами, характеризующимися нестационарностью и цветным спектром.
- Разработку гибридных подходов, сочетающих скорость корреляционных методов с устойчивостью нелинейных алгоритмов.
- Исследование возможности аппаратной реализации оптимальных алгоритмов на специализированных платформах (FPGA, DSP).
- Интеграцию акустических методов с другими сенсорными модальностями (видео, лидар) в многомодальные системы мониторинга.
- Проведение натурных испытаний для валидации результатов математического моделирования в реальных условиях городской среды.

В заключение следует отметить, что несмотря на широкий спектр существующих алгоритмов звуковой локализации, именно методы на основе GCC-PHAT демонстрируют наиболее сбалансированные характеристики для применения в системах интеллектуальных транспортных систем. Предложенная методика сравнительного анализа и полученные результаты создают основу для разработки эффективных систем акустического мониторинга, способных существенно повысить безопасность дорожного движения в условиях растущей урбанизации и интенсификации транспортных потоков.

Проведенное исследование подтверждает перспективность акустических методов как надежного дополнения к традиционным системам мониторинга и открывает новые возможности для создания интеллектуальной транспортной инфраструктуры следующего поколения.

Список литературы

- [1] Научный центр безопасности дорожного движения МВД РФ. Статистика ДТП на перекрёстках в России [Электронный ресурс]. – Режим доступа: <https://stat.gibdd.ru> (дата обращения: 10.05.2025).
- [2] Ершов, В.В. Разработка новых алгоритмов настройки плоских микрофонных антенн для эффективной локализации источников звука монопольного и дипольного типа [Текст]: дис. ... канд. техн. наук: 05.11.07 / В.В. Ершов. – Москва, 2022. – 176 с.
- [3] Бердников, В.М. Процедуры и алгоритмы синтеза сигналов и структур систем локального позиционирования [Текст]: дис. ... канд. техн. наук: 05.13.18 / В.М. Бердников. – Москва, 2013. – 150 с.
- [4] Кабанов, С.И. Разработка методов локализации для повышения надёжности акустико-эмиссионного контроля элементов машиностроительных конструкций [Текст]: дис. ... канд. техн. наук: 05.11.13 / С.И. Кабанов. – Москва, 2002. – 168 с.
- [5] ETSI TR 126 985 V16.0.0. Vehicle-to-Everything (V2X); Requirements for Time Difference of Arrival measurements [Text]. – Sophia Antipolis: ETSI, 2020. – 56 p.
- [6] IEEE 802.11p-2010. IEEE Standard for Wireless Access in Vehicular Environments [Text]. – New York: IEEE, 2010. – 264 p.
- [7] Knapp, C.H. The generalized correlation method for estimation of time delay / C.H. Knapp, G.C. Carter // IEEE Transactions on Acoustics, Speech, and Signal Processing. – 1976. – Vol. 24, № 4. – P. 320-327.
- [8] Brandstein, M. Microphone Arrays: Signal Processing Techniques and Applications [Text] / M. Brandstein, D. Ward. – Berlin: Springer, 2001. – 398 p.
- [9] Schmidt, R.O. Multiple emitter location and signal parameter estimation / R.O. Schmidt // IEEE Transactions on Antennas and Propagation. – 1986. – Vol. 34, № 3. – P. 276-280.
- [10] Benesty, J. Adaptive eigenvalue decomposition algorithm for passive acoustic source localization / J. Benesty, J. Chen, Y. Huang // Journal of the Acoustical Society of America. – 2003. – Vol. 113, № 6. – P. 3394-3405.
- [11] Van Veen, B.D. Beamforming: a versatile approach to spatial filtering / B.D. Van Veen, K.M. Buckley // IEEE ASSP Magazine. – 1988. – Vol. 5, № 2. – P. 4-24.
- [12] Ramos-García, R. Acoustic Surveillance of Road Traffic Accidents Using Distributed Wireless Sensor Networks / R. Ramos-García, J. López-Valcarce, F. Pérez-González // IEEE Transactions on Intelligent Transportation Systems. – 2018. – Vol. 19, № 6. – P. 1852-1864.
- [13] Mallat, S. A Wavelet Tour of Signal Processing [Text] / S. Mallat. – 3rd ed. – Amsterdam: Academic Press, 2008. – 832 p.
- [14] Смирнова, А.Д. Методы локализации звуковых источников на основе анализа фазовых спектров / А.Д. Смирнова, Н.Ю. Петухова // Акустический журнал. – 2021. – Т. 67, № 3. – С. 245-252.
- [15] Петрова, Е.В. Использование гибридных алгоритмов для подавления дорожных шумов в системах локализации звука / Е.В. Петрова, М.А. Соколов // Вестник компьютерных и информационных технологий. – 2020. – № 12. – С. 15-23.
- [16] Choy, B. Sound Source Localization in 3D Complex Urban Environments [Text]: PhD thesis / B. Choy. – Blacksburg: Virginia Tech, 2012. – 150 p.
- [17] Li, Z. Robust Sound Source Localization in Urban Environments via Deep Neural Networks / Z. Li, H. Wang, Y. Yan // Journal of Ambient Intelligence and Humanized Computing. – 2020. – Vol. 11, № 4. – P. 1567-1580.
- [18] Lam, E. 3D Sound-Source Localization Using Triangulation-Based Methods [Text]: Master thesis / E. Lam. – Vancouver: University of British Columbia, 2017. – 120 p.
- [19] DiBiase, J.H. Robust localization in reverberant rooms / J.H. DiBiase, H.F. Silverman, M.S. Brandstein // IEEE Transactions on Speech and Audio Processing. – 2000. – Vol. 8, № 5. – P. 508-518.

- [20] Huang, Y. Real-Time Passive Source Localization: A Practical Linear-Correction Least-Squares Approach / Y. Huang, J. Benesty, G. Elko // IEEE Transactions on Speech and Audio Processing. – 2001. – Vol. 9, № 8. – P. 943-956.
- [21] H. Parineh, M. Sarvi, and S.A. Bagloee, "Implementation of ZigBee based WSN to enhance the performance of SCATS compatible intelligent traffic controllers," Australasian Transport Research Forum (ATRF), 43rd, 2022, Adelaide, South Australia, Australia.
- [22] L. Ambrosini, L. Gabrielli, F. Vesperini, S. Squartini, and L. Cattani, "Deep neural networks for road surface roughness classification from acoustic signals," in Audio Engineering Society Convention 144, May 2018, Audio Engineering Society.
- [23] S. Ntalampiras, "Moving vehicle classification using wireless acoustic sensor networks," IEEE Transactions on Emerging Topics in Computational Intelligence, vol. 2, no. 2, pp. 129-138, 2018.
- [24] M.A.I. Rafi, M.R. Sohan, M.S. Hasan, T.S. Rafa, and A. Jawad, "Exploring Classification of Vehicles Using Horn Sound Analysis: A Deep Learning-Based Approach," in *2024 23rd International Symposium INFOTEH-JAHORINA (INFOTEH)*, Mar. 2024, pp. 1-6.
- [25] C.Y. Chiang, M. Jaber, K.K. Chai, and J. Loo, "Distributed acoustic sensor systems for vehicle detection and classification," IEEE Access, vol. 11, pp. 31293-31303, 2023.