

1. Здравствуйте, меня зовут Поляков Игорь.Позвольте представить вашему вниманию выпускную квалификационную работу: Модели и методы монауральной локализации направления на источник звука.

Актуальность работы. В настоящее время актуальна проблема разработки систем оценки направления на источник звука. Существующие системы, основанные на бинауральном методе, требуют установок крупного размера, некоторое количество записывающих модулей и, как следствие, повышенное энергопотребление и низкую мобильность. Системы, основанные на монауральном методе, сравнительно мало изучены и нуждаются в более тщательном исследовании. Такие системы могут использоваться как в масштабах определения местонахождения крупных объектов – например, самолетов, подводных лодок, так и в качестве сенсоров для различных устройств, роботов, охранных систем.

2. Цель и Задачи работы.

Целью данной работы является разработка моделей и методов монауральной локализации направления на источник звука.

Для достижения поставленной цели необходимо решить следующие задачи:

- Проанализировать известные подходы к решению задачи локализации направления на источник звука. Общих подходов два: бинауральный и монауральный. При бинауральном подходе используется несколько звукоприемников, при монауральном подходе используется только один звукоприемник.

- Разработать информационную модель системы монауральной локализации направления на источник звука.

- Разработать программную реализацию системы и исследовать её возможности.

3. Научная новизна.

В ходе выполнения работы получены следующие новые результаты

- Разработана информационная модели системы монауральной локализации направления на источник звука, отличающаяся использованием в качестве признакового описания сигнала полных групп и позволяющая

сократить длительность сигнала для определения направления на источник звука.

- разработан алгоритм вычисления полных и замкнутых групп по спектру звукового сигнала.

4. Модель работы системы. Основой для информационной модели служит теория активного восприятия изображений, предложенная профессором Владимиром Александровичем Утробиним.

Первый этап - обучение. На вход системы подается звуковой сигнал, в соответствии с этапами преобразований в системах распознавания образов, он проходит предварительную обработку, а затем происходит этап формирования системы признакового описания этого сигнала. Описанию присваивается наименование направления, с которого пришел данный сигнал и все это сохраняется в базе данных. Так происходит со всеми сигналами, пришедшими с различных возможных направлений.

Второй этап - тестирование. Этап предварительной обработки сигнала и формирование системы признаков происходит аналогично обучению. Далее происходит этап классификации. Признаковое описание тестируемого сигнала сравнивается по очереди с описаниями сигналов, хранящихся в базе данных и на основе того, к какому значению сохраненных сигналов ближе тестируемый сигнал, выводится результат.

5. Предварительная обработка сигнала.

На этапе предварительной обработки входящий сигнал разбивается на сегменты, по каждому из которых вычисляется Q-преобразование. Предварительно весь сигнал нормируется по амплитуде к отрезку $[0;1]$. Затем суммируются амплитуды, относящиеся к каждому сегменту сигнала.

6. Формирование системы признаков.

На этапе формирования системы признаков сигнал (g), полученный на предыдущем этапе, разбивается на множество сегментов длиной 16 отсчетов. К каждому сегменту применяется система фильтров. Таким образом, формируется спектральное представление каждого сегмента. Далее по спектральному представлению сегмента формируется описание с помощью одной из структур, входящих в алгебру групп: с помощью операторов, полных групп, либо замкнутых групп. Для объединения данных,

полученных от разных сегментов анализируемого сигнала, вычисляется гистограмма элементов структур, использованных при создании описания сегмента. Это позволяет получить представление о сигнале в целом.

7. Ускорение формирования системы признаков. Вместо того, чтобы каждый раз на основе значений операторов высчитывать наличие полных и замкнутых групп, составляется база соответствий между всеми возможными сочетаниями значений операторов и соответствующими ими полными и замкнутыми группами. Это сравнимо с табулированием функции и хранением её значений в памяти без повторных вычислений. Хранение такой информации ускорит этап формирования системы признаков, при этом количество значений не превышает 33 тысячи.

8. Звукозаписывающая установка. На этом рисунке представлены два варианта звукозаписывающей установки, использующие различные виды звукоотражателей, придающих характерные особенности звуковым сигналам, приходящим с разных направлений. В первом случае это стенка с изогнутыми краями (похожие звукоотражатели применялись в известных работах), а во втором случае звукоотражатель - это имитация человеческого уха (наружное ухо с каналом, длиной ~ 3 см). Предположительно, используя масштабы человеческого уха, удастся повысить чувствительность системы при работе со звуками человеческой речи.

9. Процесс работы системы. Исследуемая область - диапазон в 180 градусов. Он разбивается на 13 направлений с шагом в 15. При обучении системы с каждого направления производится запись продолжительностью в 1 минуту. Итого создается 13 записей. Каждая такая запись разбивается на части по 2 секунды. Эти части описываются гистограммами, в соответствии с вышеописанными алгоритмами. Гистограмме присваивается номер направления и она сохраняется в базу данных. Итого с одной минуты обучающего сигнала в базе сохраняется 30 записей.

(На этапе тестирования, с какого-либо из направлений совершается запись, длиной в 20 секунд и разбивается на части по 2 секунды, с каждым из которых происходит процесс признакового описания (составление вектора, состоящего из значений фиксированного набора признаков (полных групп, к примеру) на данном сигнале). На этапе классификации, каждое описание двухсекундного сигнала сравнивается с описаниями, хранящимися в базе данных. Определяется, к какому направлению относится каждый

двухсекундный сигнал. Совокупность результатов определяет ответ программы.)

10. Известные подходы к решению задачи монауральной локализации направления на источник звука.

На данном слайде показаны известные результаты монауральной локализации направления на источник звука. Эксперименты производились с различными типами звуков, натуральными и искусственными. Длительность записей составляла 5 минут. Были использованы различные виды звукоотражателей (А, В, С, D), представляющие собой вариации стенки различной фактуры.

Также на слайде показаны полученные результаты исследования разработанной системы. В таблицах приведена средняя ошибка определения угла направления в градусах.

Сравнивая известные результаты с полученными можно отметить, что результаты полученные мной и в известной работе сравнимы. Однако в предлагаемой системе длительность записываемых сигналов сокращена. В данном случае это одна минута с каждого направления при обучении и по 20 секунд с направления при тестах.

11. На этом слайде более подробно показаны результаты эксперимента, показывающие зависимость точности работы системы от различных параметров.

12. На этом слайде более наглядно показана зависимость точности работы системы от типа использованного звукоотражателя и от длительности тестируемого сигнала. Сверху - звукоотражатель - стенка (ошибка больше), снизу - ухо.

13. Заключение.

В ходе выполнения работы

- Создана система монауральной локализации направления на источник звука;

- Разработана программная реализация модели системы;

- Выполнено тестирование программной системы, включающее в себя проведение экспериментов с реализованной моделью, поиск алгоритма ее работы, дающего наилучший результат, а так же сравнение конечных результатов с аналогами;

- Все представленные в работе методы реализованы в виде программного обеспечения, способного работать на большом количестве распространенных персональных вычислительных машин; По результатам исследований опубликованы тезисы.

14. Публикации

По теме работы имеются следующие публикации

15. Спасибо за внимание