

1. Слайд «Первый слайд»

Здравствуйте уважаемая комиссия, меня зовут Лукьянчикова Александра. Я хочу представить вам свою выпускную работу на тему «Система распознавания изолированных речевых команд»

2. Слайд «Цель работы»

Разработка и программная реализация системы распознавания изолированных речевых команд

3. Слайд «Задачи»

Задача данной работы:

- Изучение методов распознавания изолированных речевых команд
- Разработка программной системы распознавания изолированных речевых команд на основе теории активного восприятия
- Тестирование программной системы

4. Слайд «Структура программной систем»

Данная система состоит из следующих уровней:

- 1) Пользовательский интерфейс – оболочка, которая взаимодействует с пользователем.
- 2) Программная логика – функции, обрабатывающие пользовательские запросы
- 3) Данные – база данных слов, записанная мной для обучения классификаторов работающих в программе.

Рассмотрим подробнее программную логику.

5. Слайд «Программная логика»

Она состоит из следующих подсистем:

- 1) Предварительной обработки сигнала;
- 2) Подсистемы формирования системы признаков;
- 3) Подсистема классификации;

6. Слайд «Блок-схема работы алгоритма программы»

Здесь показан алгоритм работы все программы. Пользователь записывает голосовую команду. Далее идет выделение речевой активности из всего сигнала, предварительная обработка. После формирование признакового описателя, классификация и вывод результата на экран.

7. Слайд «U-преобразование»

8. Слайд «Q-представление»

К сигналу применяется Q-преобразование применяется на этапе предварительной обработки сигнала. Оно суммирует несколько сегментов и в результате получается всего 16 сегментов по числу фильтров.

9. Слайд «Формирования признакового описателя»

На этапе формирования признакового описателя после применения фильтра к каждому сегменту получаем спектральное представление каждого сегмента. Далее по спектру вычисляются замкнутые группы. Затем по списку вычисленных замкнутых групп формируется гистограмма замкнутых групп.

U-преобразование является основным в теории активного восприятия. Оно состоит из интегрального Q-преобразования и дифференциального преобразования. К каждому сегменту сначала применяется интегральное преобразование, а затем к результирующему сигналу применяются фильтры.

10. Слайд «Фильтры»

Так выглядят фильтры для U-преобразования

11. Слайд «Подсистема классификации»

В этой подсистеме происходит сравнение признакового описателя произнесенной команды с эталоном из базы и определения класса произнесенной команды. Для этого используются следующие классификаторы на основе:

- 1) к- Ближайших соседей;
- 2) метода опорных векторов;
- 3) нейронных сетей;

12. Слайд «Пользовательский интерфейс»

Так выглядит интерфейс программы. Три кнопки: Запись – запись команды, Распознавание – распознавание произнесенной команды, Сохранение – сохранение команды. В листе отображается область голосовой активности для распознавания. График для наглядного отображения голосовой активности. Результат распознавания выводится под соответствующей строчкой.

13. Слайд «База данных»

14. Слайд «Тестирование системы»

Тестирование производилось на двух наборах слов:

Первый набор – слова различные по звучанию (например, почта телефон, стол и др.);

Второй набор – слова схожие по звучанию (например, мама, лама, дама и др.)

Для тестирования будет использоваться перекрестная проверка.

15. Слайд «Первый набор»

Здесь указаны все слова, участвующие в первом тесте. Каждое слово записано в базе 20 раз. Обучающая выборка составляет 16 слов, контрольная – 4 слова.

16. Слайд «Классификатор к-ближайших соседей»

Это результаты тестирования на классификаторе к-ближайших соседей. Ошибка составляет 2,5%.

17. Слайд «Классификатор на основе метода опорных векторов»

Результаты тестирования на классификаторе, основанном на методе опорных векторов. Ошибка – 0,5%

18. Слайд «Классификатор на основе нейронных сетей»

Результаты тестирования на классификаторе на основе нейронных сетей. Ошибка – 1,5%

19. Слайд «Точность распознавания первого набора слов»

Самая высокая точность распознавания у классификатора, основанного на методе опорных векторов -99,5%. Далее НН-98,5% и КБЛ.С – 97,5%

20. Слайд «Второй набор слов»

Второй набор слов – это слова схожие по звучанию. Выборка слов такая же как и в первом наборе.

21. Слайд «Классификатор к-ближайших соседей»

Это результаты тестирования на классификаторе к-ближайших соседей. Ошибка составляет 14,5%. Значительная разница. (в первом 2,5%)

22. Слайд «Классификатор на основе метода опорных векторов»

Результаты тестирования на классификаторе, основанном на методе опорных векторов. Ошибка – 8,5% (в первом 0,5%)

23. Слайд «Классификатор на основе нейронных сетей»

Результаты тестирования на классификаторе на основе нейронных сетей. Ошибка – 3,5% (в первом 1,5%)

24. Слайд «Точность распознавания первого набора слов»

Самая высокая точность распознавания у классификатора, основанного на нейросетях - 96,5%. Далее М.Оп.В-91,5% и КБЛ.С – 85,5%

25. Слайд «Демонстрация видео»

А теперь я хочу продемонстрировать видео работы программы.

26. Слайд «Публикации»

По данной теме опубликована одна статья в журнале (Системы управления и информационные технологии). Она докладывалась на двух конференциях.

27. Слайд «Спасибо за внимание»

28. Слайд «Пусто»

29. Слайд «Пусто»

30. Слайд «Метод к-ближайших соседей»

31. Слайд «Метод опорных векторов»

32. Слайд «Нейронные сети»

33. Слайд «Замкнутые группы»

34. Слайд «Гистограмма замкнутых групп»