

Система голосового управления технологическими комплексами

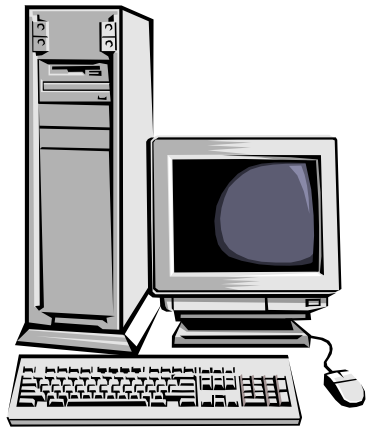
Рассказова С.И.

Цель работы: Разработка системы голосового управления технологическими комплексами.

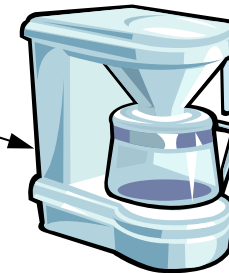
Задачи:

- Провести анализ речевых сигналов и их параметров
- Провести исследование и классификацию методов распознавания речи и голосового управления
- Разработать модели распознавания голосовых команд управления технологическим комплексом
- Разработать аппаратно-программный комплекс голосового управления технологическим комплексом
- Провести отработку системы на экспериментальном комплексе на базе робота МП-9С

Области применения систем голосового управления

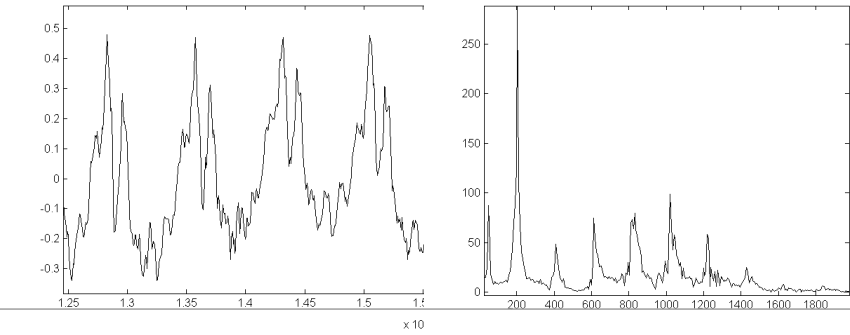
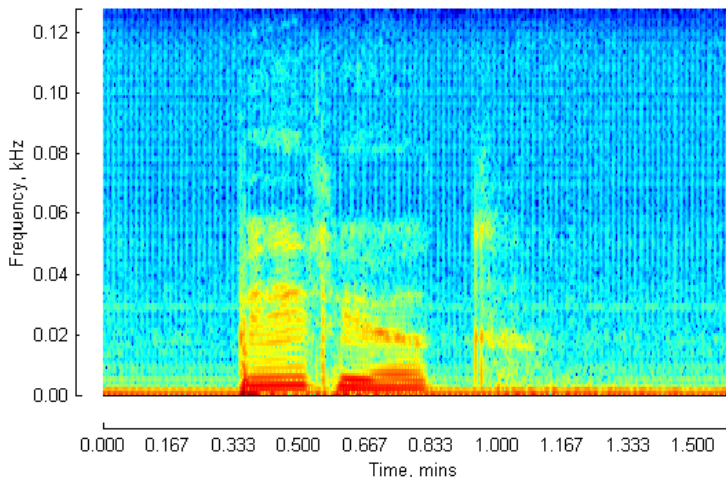
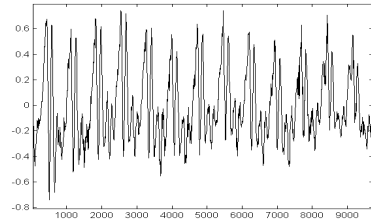


«интеллектуальный дом»



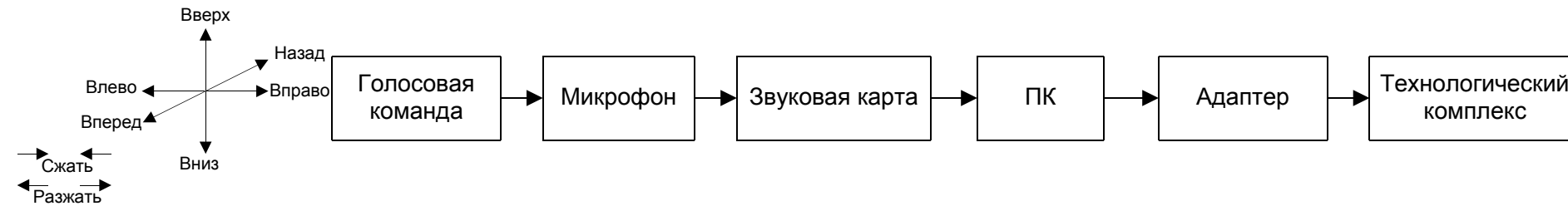
Классификация систем распознавания речи

Основные параметры звука

<p>Спектральные характеристики (зависимость интенсивности сигнала от частоты)</p>	
<ul style="list-style-type: none"> Формантные характеристики (частоты формант, амплитуда, ширина, скорость изменения форманты) Параметры основного тона голоса (среднее значение, дисперсия, статистические моменты, вариации) 	
<p>Число переходов интенсивности сигнала через ноль [6]</p>	<div> $Z = \frac{1}{2} \sum_{k=2}^N \text{sign}(s_k) - \text{sign}(s_{k-1}) ,$ <p>где $\text{sign}(s) = \begin{cases} 1, & s \geq 0; \\ -1, & s < 0 \end{cases}$</p> </div> 

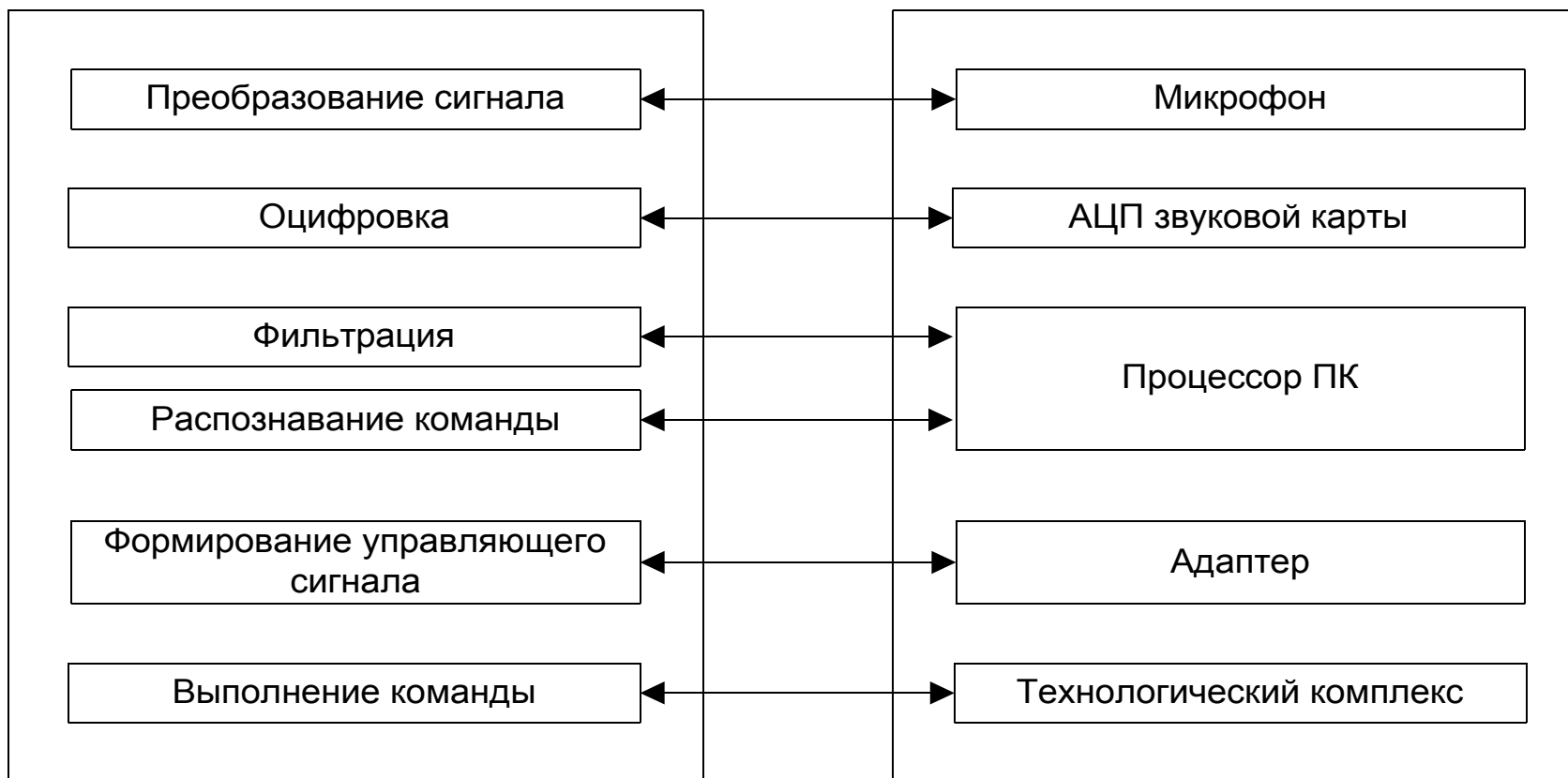
Структурная схема АПК голосового управления

5

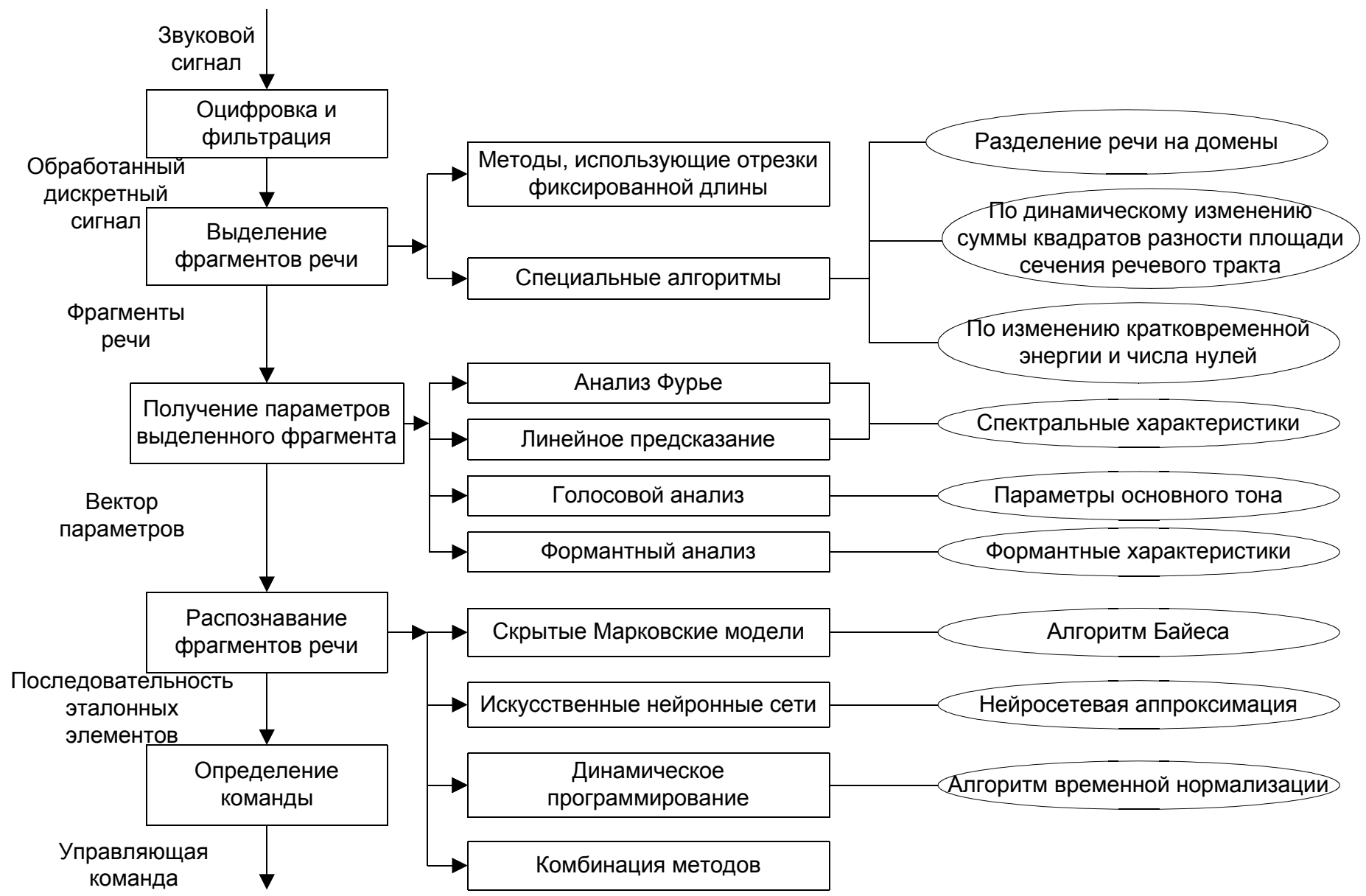


Программное обеспечение

Аппаратное обеспечение

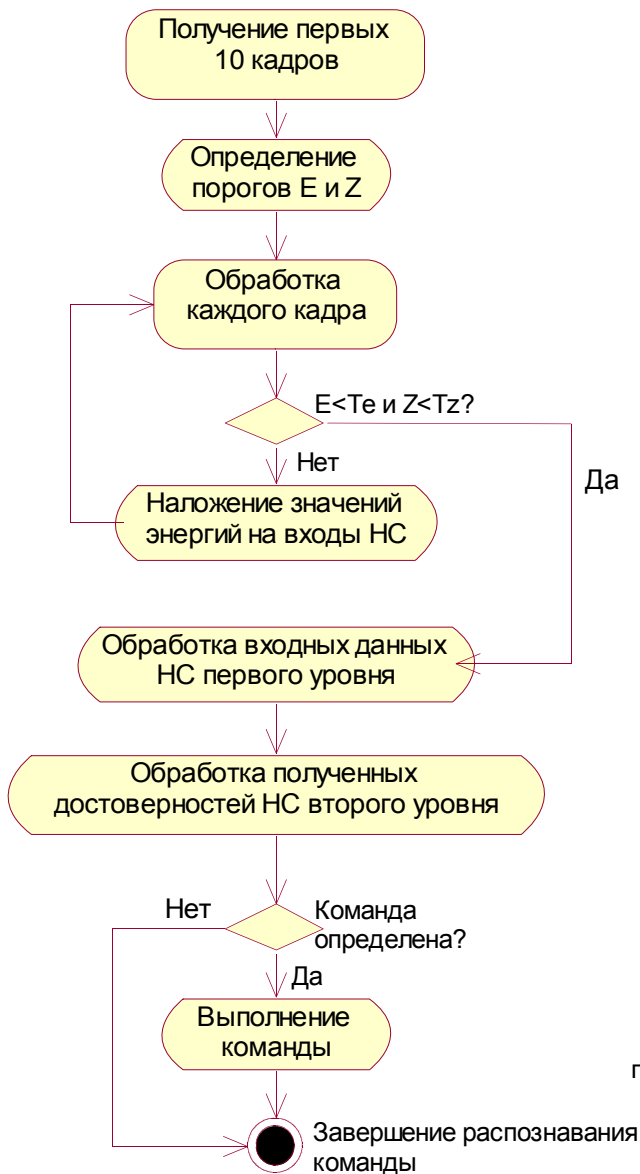


Классификация методов распознавания речи



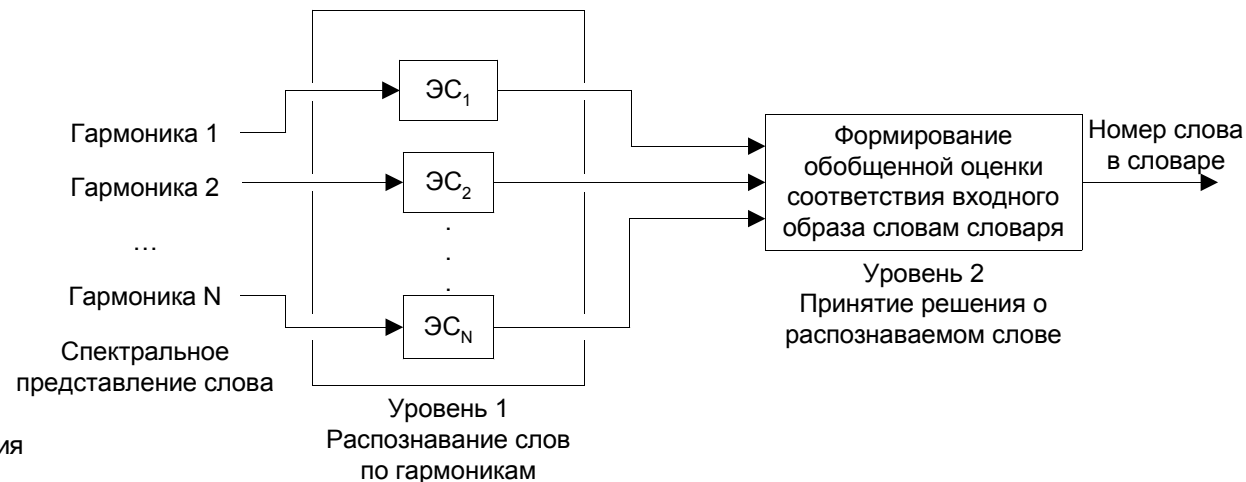
Алгоритм распознавания команды

1. Выделение команды из потока аудиоданных



1. Уровень кратковременной энергии
$$E = \frac{1}{N} \sum_{k=1}^N s_k^2$$
2. Число нулей интенсивности
$$Z = \frac{1}{2} \sum_{k=2}^N |\text{sign}(s_k) - \text{sign}(s_{k-1})|$$
3. Пороговое значение кратковременной энергии
$$T_E = M(E, 10) + 3\sqrt{D(E, 10)} + \frac{1}{400} \max_{1 \leq t \leq L} E_t$$
4. Пороговое значение числа нулей интенсивности
$$T_Z = M(Z, 10) + 3\sqrt{D(Z, 10)} + \frac{1}{20} \max_{1 \leq t \leq L} Z_t$$

2. Схема распознавания команд

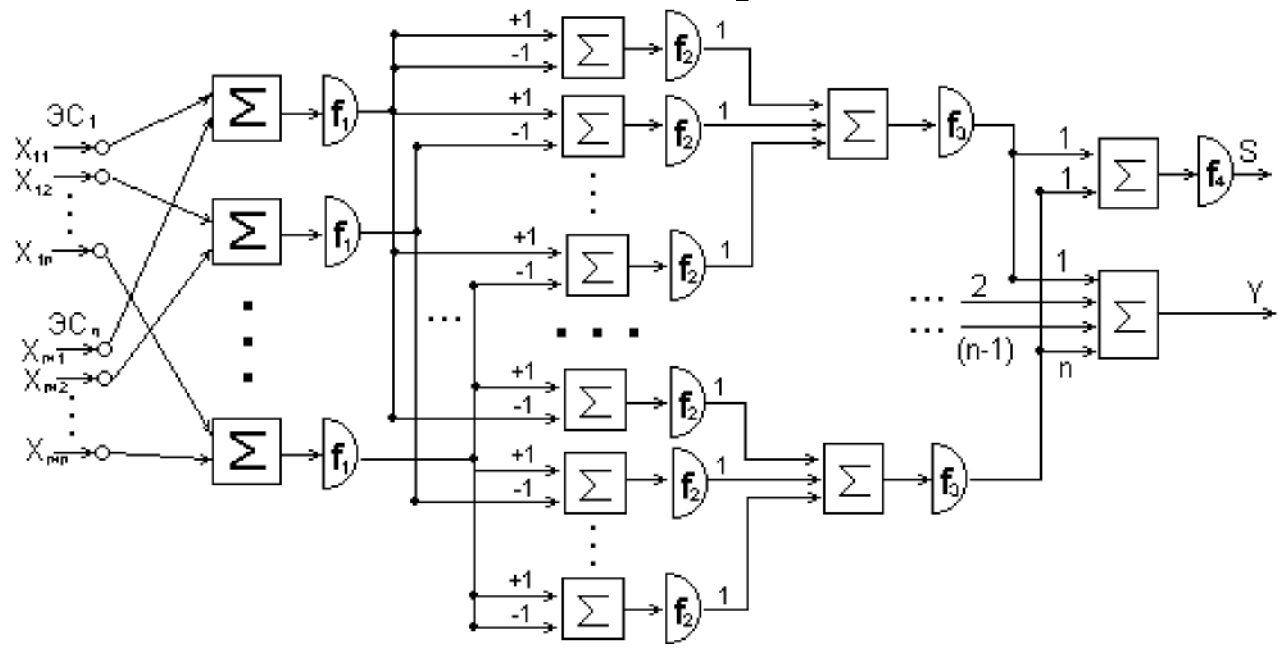


Двухуровневая схема распознавания команд

Уровень 1 (параметры нейросетей)

- входной образ - распознаваемая гармоника;
- выходной сигнал - функция принадлежности гармоники словам словаря;
- желаемый выходной сигнал - вектор размерности словаря из нулей и одной единицы, соответствующей произнесенному слову;
- структура нейросети - трехслойная с полными последовательными связями;
- функция активации $f(g) = 1/(1 + e^{-g})$;
- функция ошибки - отклонение реального выхода от желаемого;
- критерий качества обучения - минимум ошибки по всему обучающему множеству;
- обучение - обратное распространение ошибки.

Уровень 2



$$f_1(g) = \begin{cases} 1, & \text{если } g \geq p \\ 0, & \text{если } g < p \end{cases}$$

$$f_2(g) = \begin{cases} 1, & \text{если } g \geq 0 \\ 0, & \text{если } g < 0 \end{cases}$$

$$f_3(g) = \begin{cases} 1, & \text{если } g \geq n \\ 0, & \text{если } g < n \end{cases}$$

$$f_4(g) = \begin{cases} 1, & \text{если } g > 1 \\ 0, & \text{если } g \leq 1 \end{cases}$$

Функциональное описание системы



Физическая модель АПК голосового управления роботом МП-9С

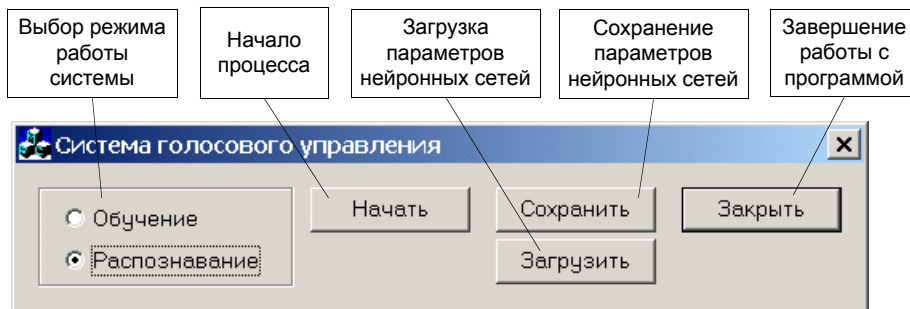
10



Работа с системой голосового управления

Режим обучения

- Добавление/удаление команд управления
- Добавление/удаление обучающих файлов
- Загрузка/сохранение параметров системы (нейронных сетей)
- Обучение системы



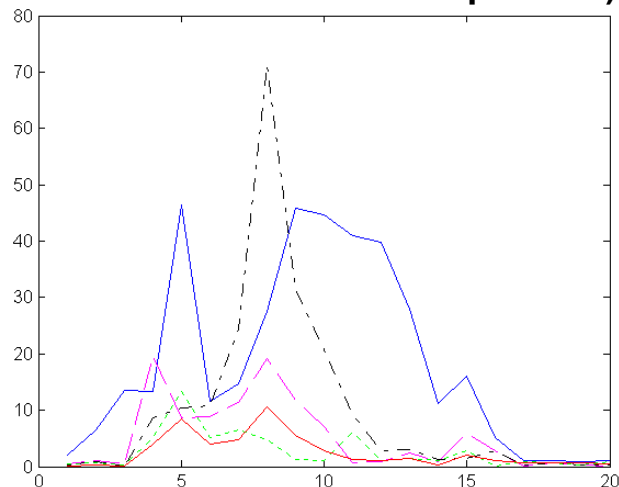
Режим голосового управления

- Загрузка/сохранение параметров системы (нейронных сетей)
- Распознавание и выполнение команд

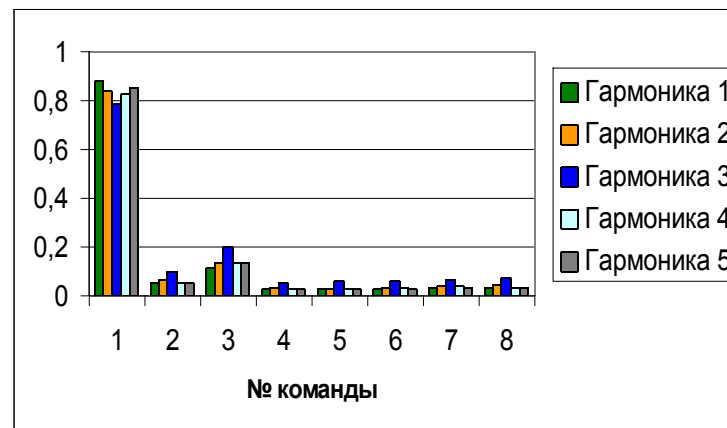
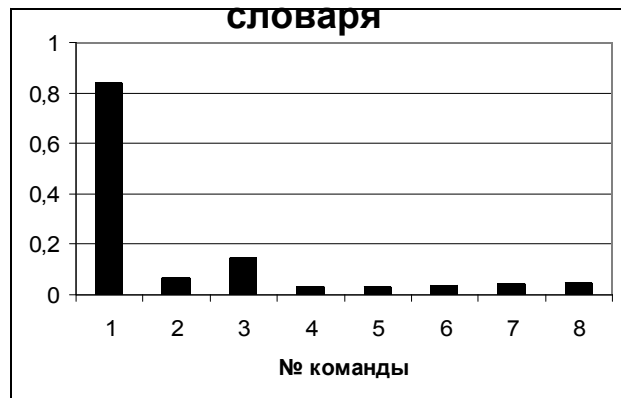
Команда ВВЕРХ

Оценка принадлежности гармоник командам

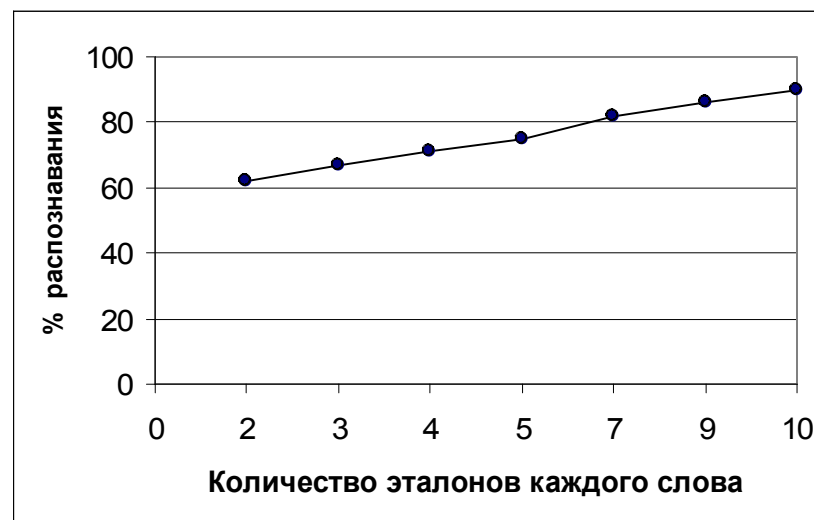
Набор входных данных (значения энергий низкочастотных гармоник)



Интегральная оценка степени соответствия команды словам словаря



Эффективность распознавания



По результатам тестирования был получен процент распознавания 90%, что соответствует требованиям ТЗ

Выводы

- Проведен анализ параметров голосовых команд. В результате анализа выбран набор параметров, необходимый для распознавания голосовых команд управления технологическим комплексом;
- Проведен анализ существующих методов распознавания речи и голосового управления. В результате анализа выбраны наиболее перспективные методы для построения системы голосового управления;
- Разработана модель распознавания голосовых команд управления технологическим комплексом;
- Разработан аппаратно-программный комплекс голосового управления технологическим комплексом;
- Проведена отработка системы на экспериментальном комплексе на базе робота МП-9С.

Апробация

Результаты исследований и разработок были отмечены стипендией Правительства РФ и докладывались на студенческих конференциях:

- Рассказова С.И. «Системы распознавания речи и их использование на радиотехническом предприятии». Сборник трудов V молодежной научно-технической конференции «Наукоемкие технологии и интеллектуальные системы 2003». М: МГТУ им. Н.Э.Баумана, 2003.
- Рассказова С.И. «Некоторые подходы к технологии получения частотных характеристик речи». Сборник трудов VI молодежной научно-технической конференции «Наукоемкие технологии и интеллектуальные системы 2004». М: МГТУ им. Н.Э.Баумана, 2004.