

«Система распознавания изолированных речевых команд»

Автор: Лукьянчикова А.В.

Руководитель: Гай В.Е.

Цель работы

Разработка и программная реализация
системы распознавания изолированных
речевых команд

Задачи

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

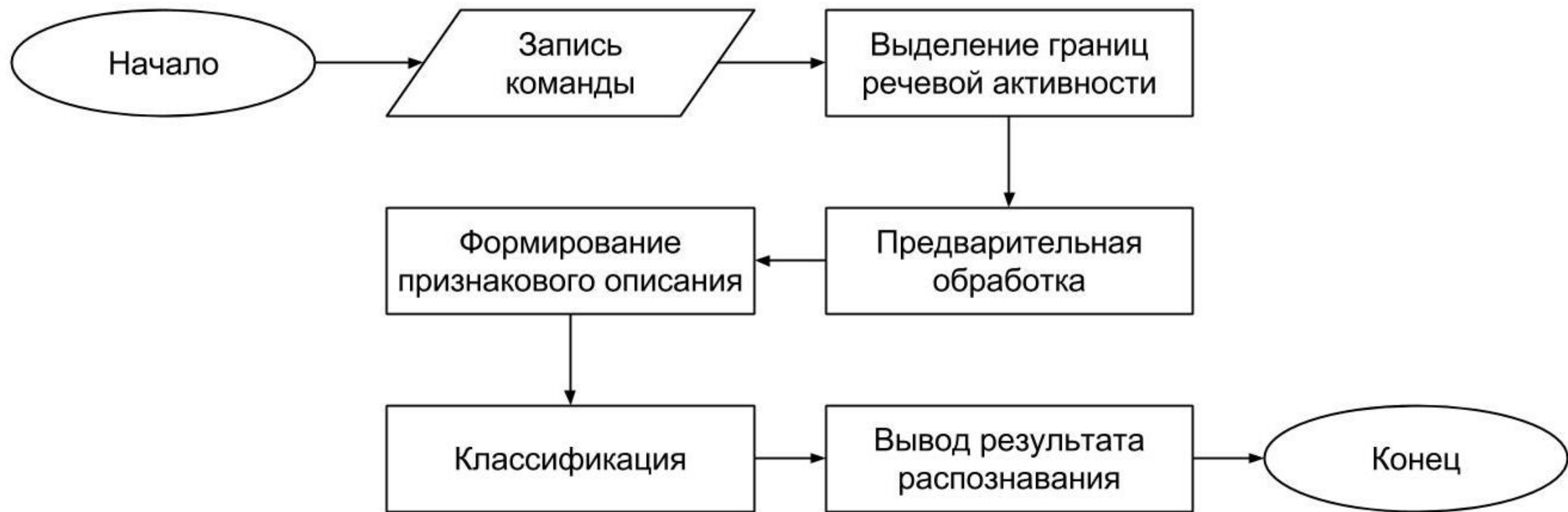
Структура программной системы

- 1) Пользовательский интерфейс;
- 2) Программная логика;
- 3) Данные.

Программная логика

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

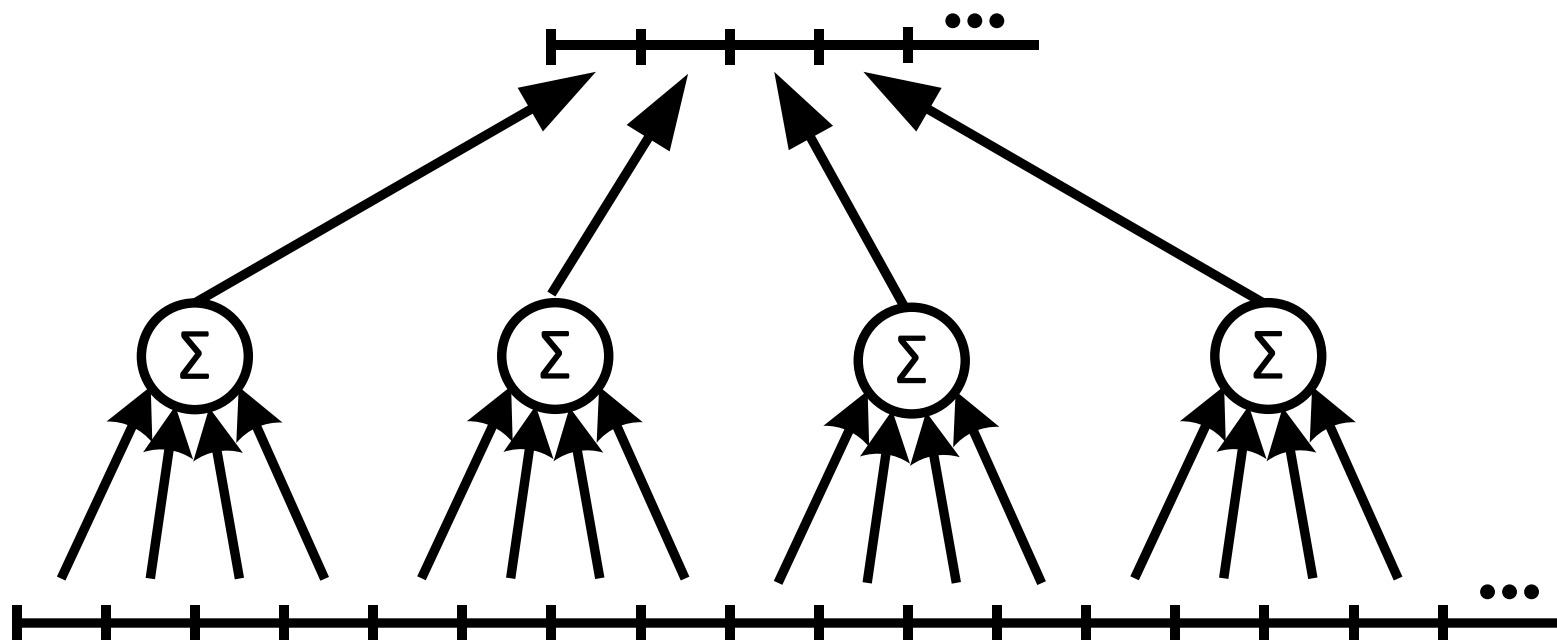
Блок-схема работы алгоритма программной системы



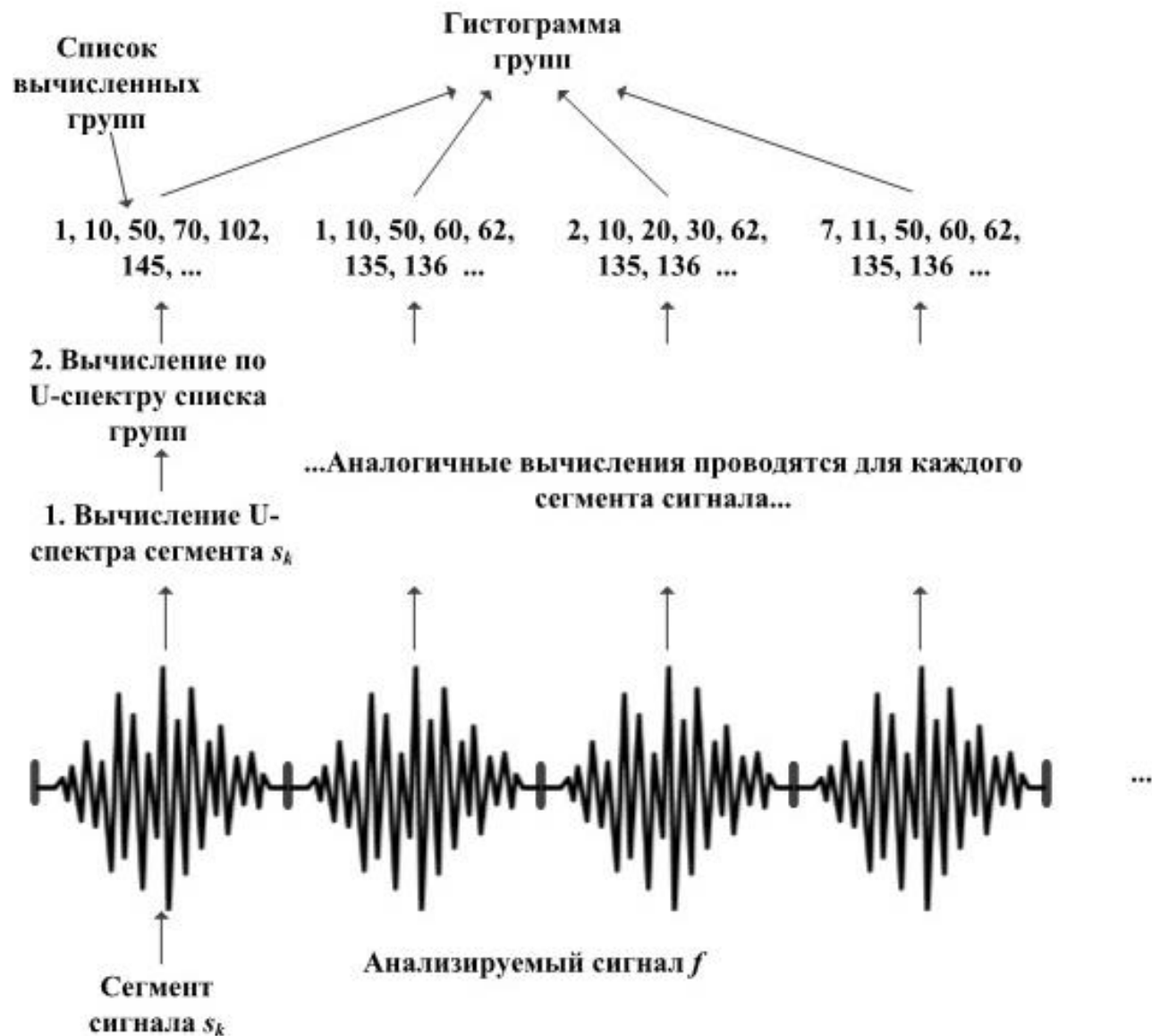
Предварительная обработка

Сигнал g

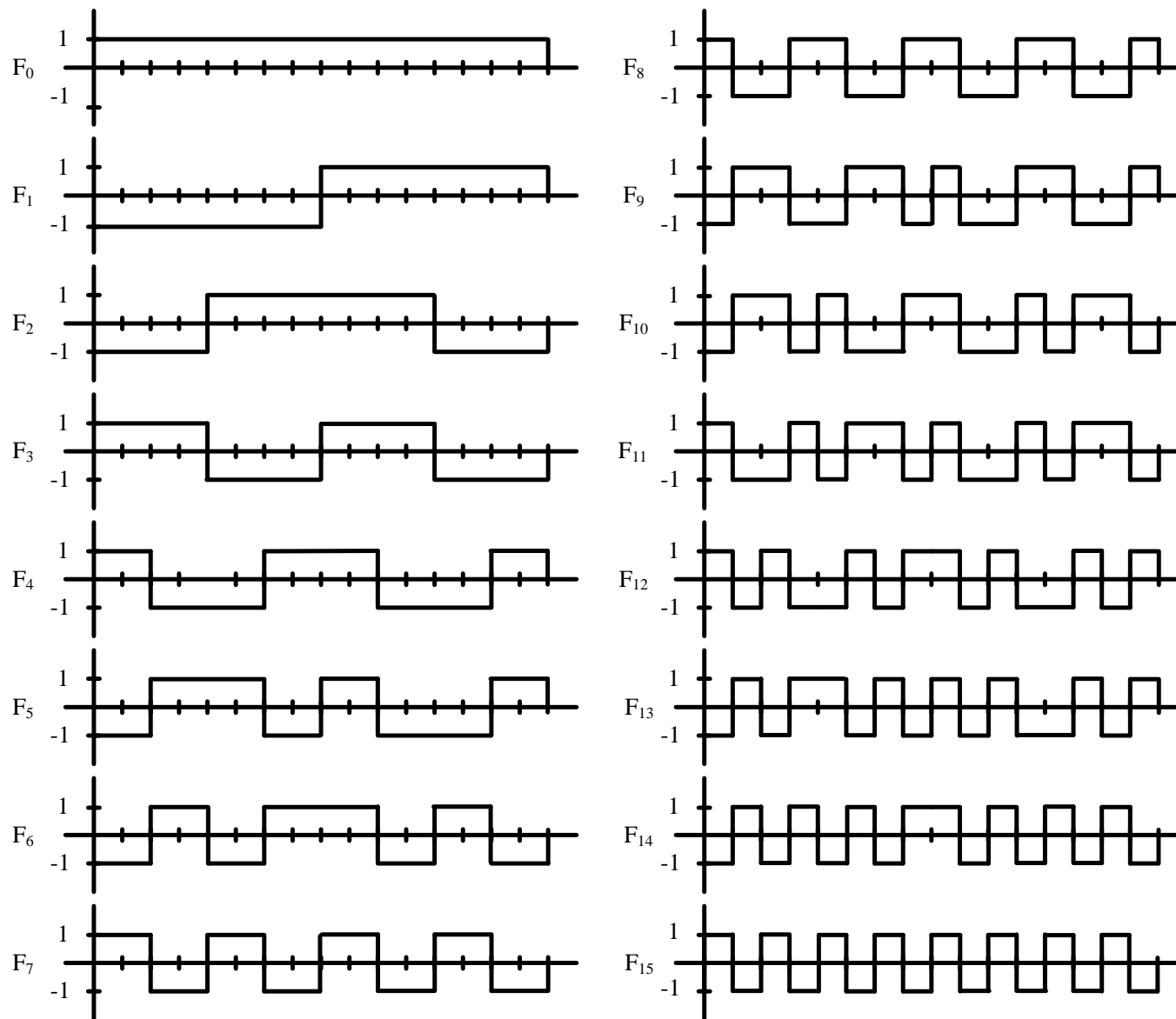
Сигнал f



Формирование признакового описателя



Фильтры

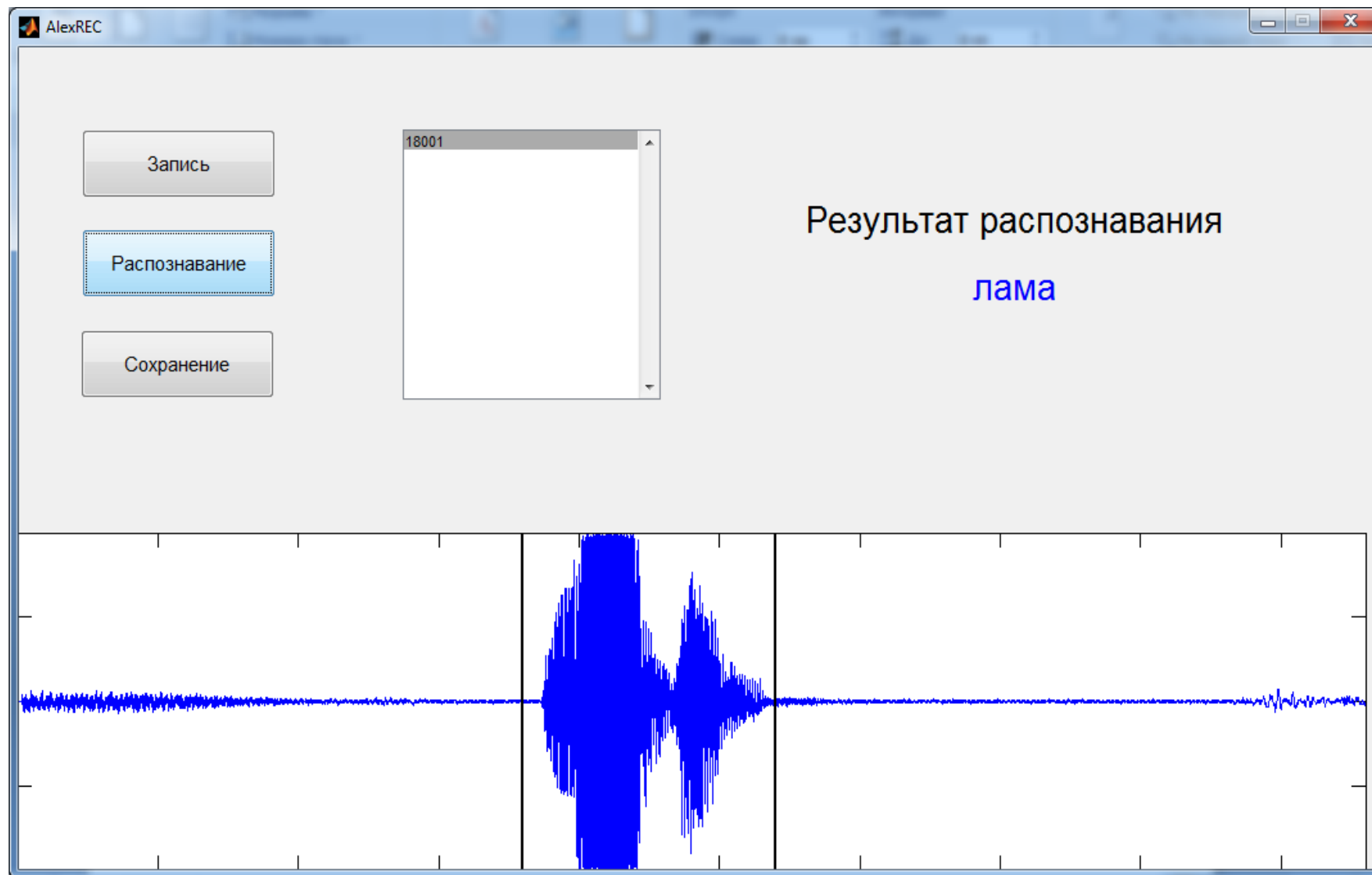


Подсистема классификации

- 1) Классификатор k-ближайших соседей;
- 2) Классификатор на основе метода опорных векторов;
- 3) Классификатор на основе нейронных сетей.



Пользовательский интерфейс



База эталонов

- Полная база эталонов содержит 150 различных слов
- Каждое слово в базе записано в 20 экземплярах
- Всего 3 000 записей
- Параметры записи слова

Формат: *.wav

Длительность (макс.): 1 с

Глубина кодирования: 32 bit

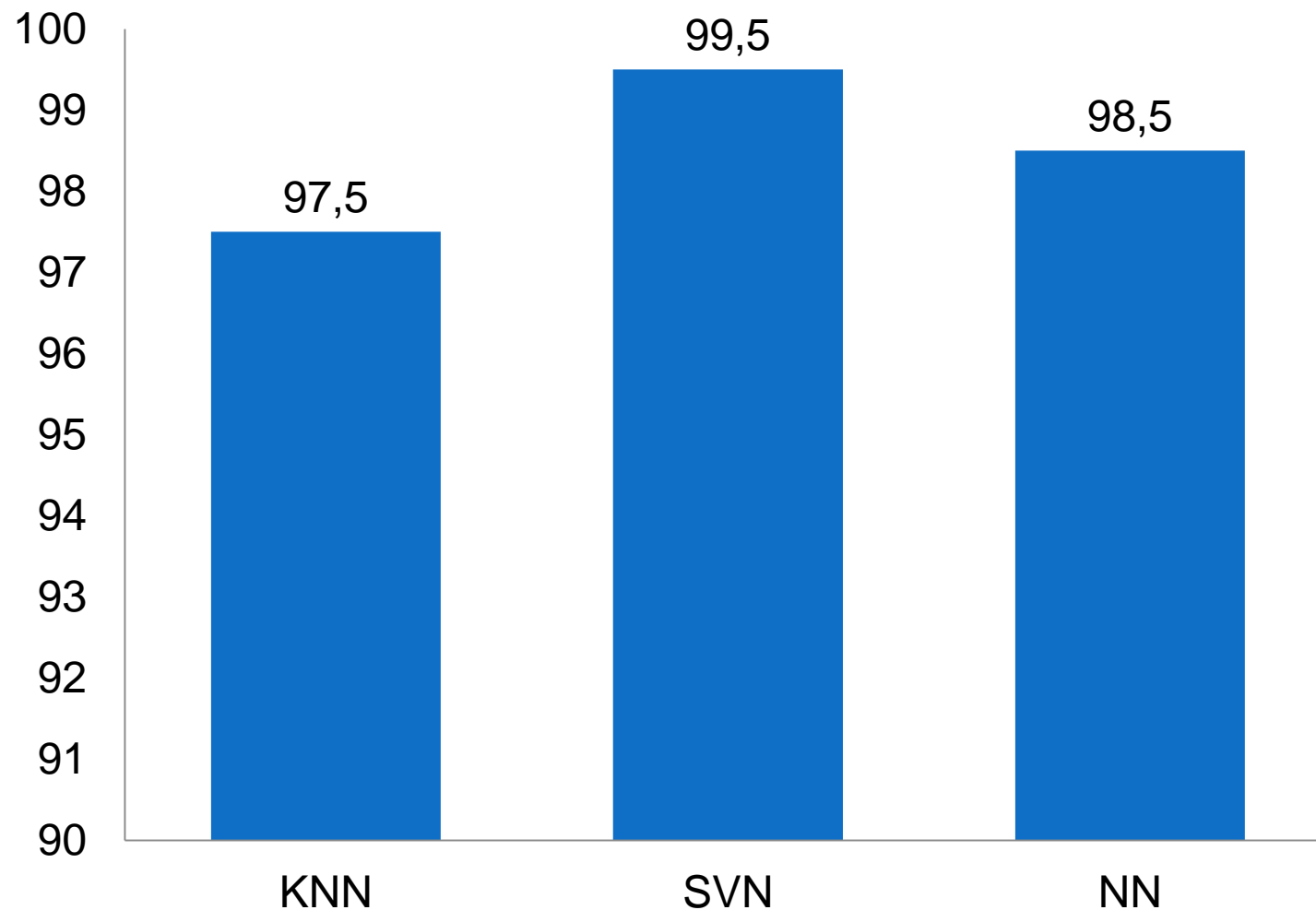
Частота дискретизации: 16 kHz

Тестирование системы

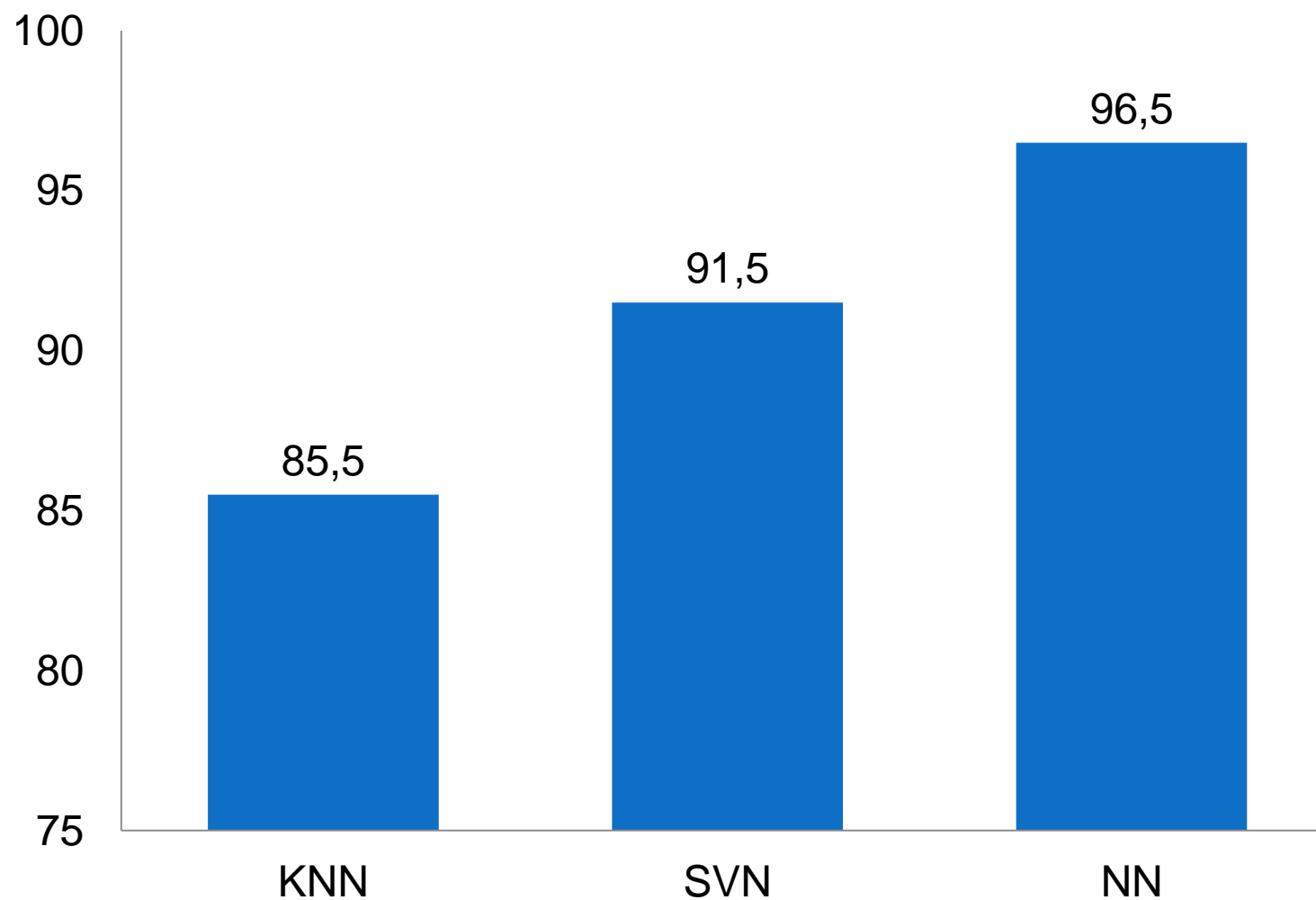
- 1) **Первый набор слов** – слова, различные по своему звучанию: **стол, телефон, почта, лего, дверь, маска, слово, блокнот, картина, газета.**
- 2) **Второй набор слов** – слова, схожие по своему звучанию : **мама, лама, рама, дама, гамма, кресло, тесно, стена, место, весна.**

База эталонов содержит 20 записей каждого слова.

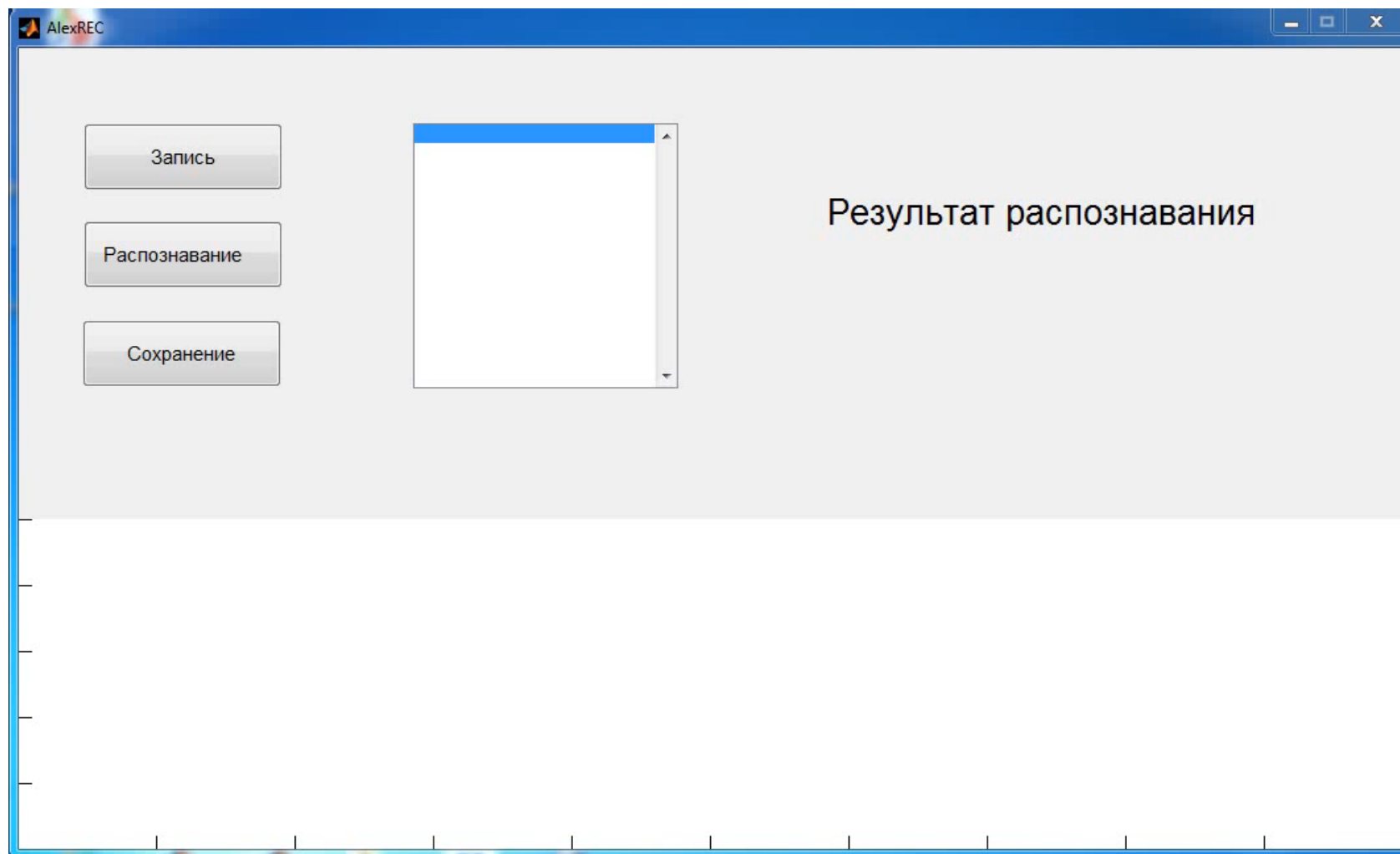
Точность распознавания первого набора слов



Точность распознавания второго набора слов



Демонстрация работы(видео)



Публикации

- А.А. Филяков, А.В. Лукьянчикова, В.Е. Гай.

Аппаратно-программная робототехническая система распознавания изолированных речевых команд. Конференция МСИТ, Томск, 2014.

- А.А. Филяков, А.В. Лукьянчикова, В.Е. Гай.

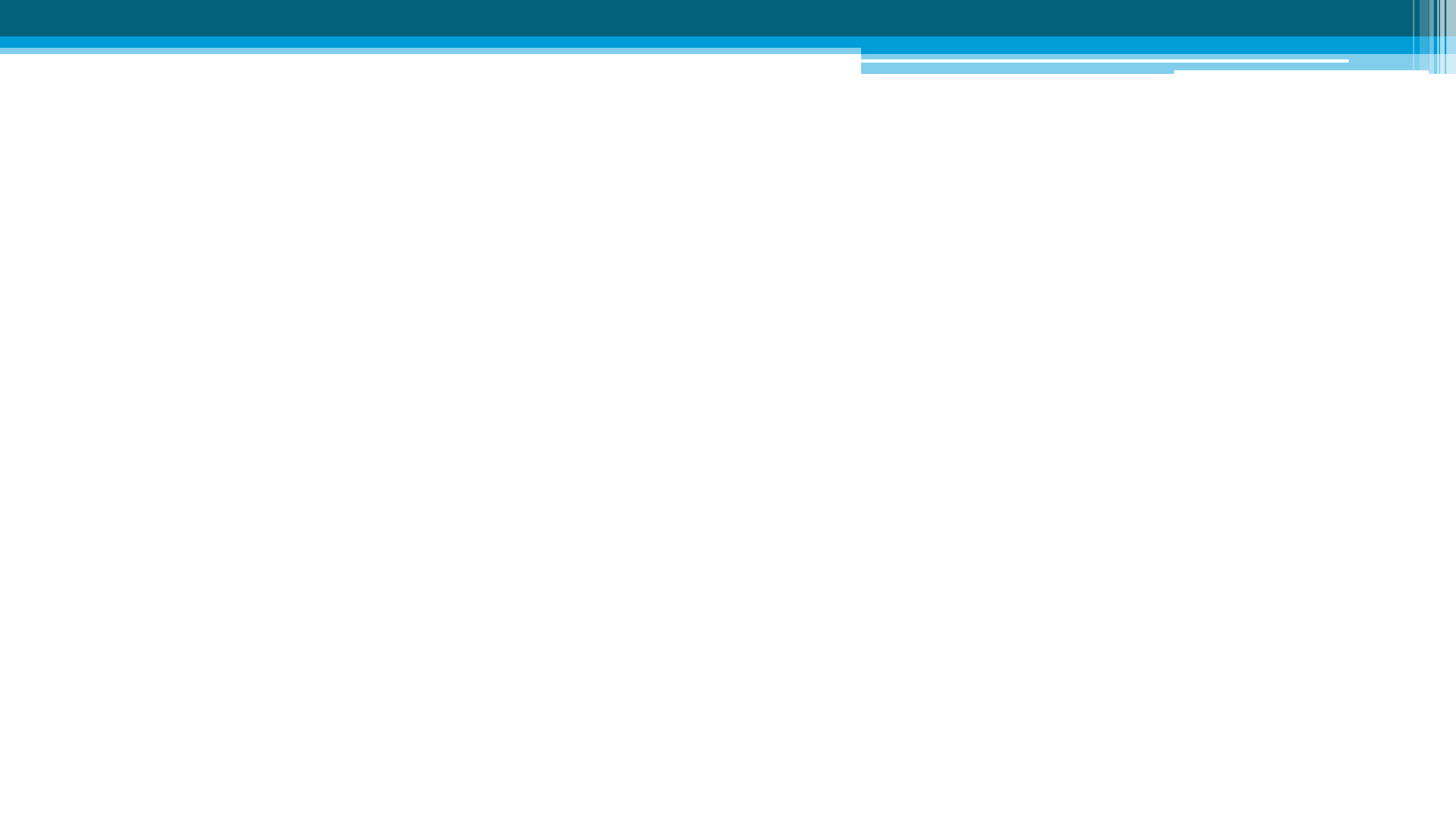
Робототехническая система распознавания и исполнения голосовых команд. Конференция ИСТ-2015, Нижний Новгород

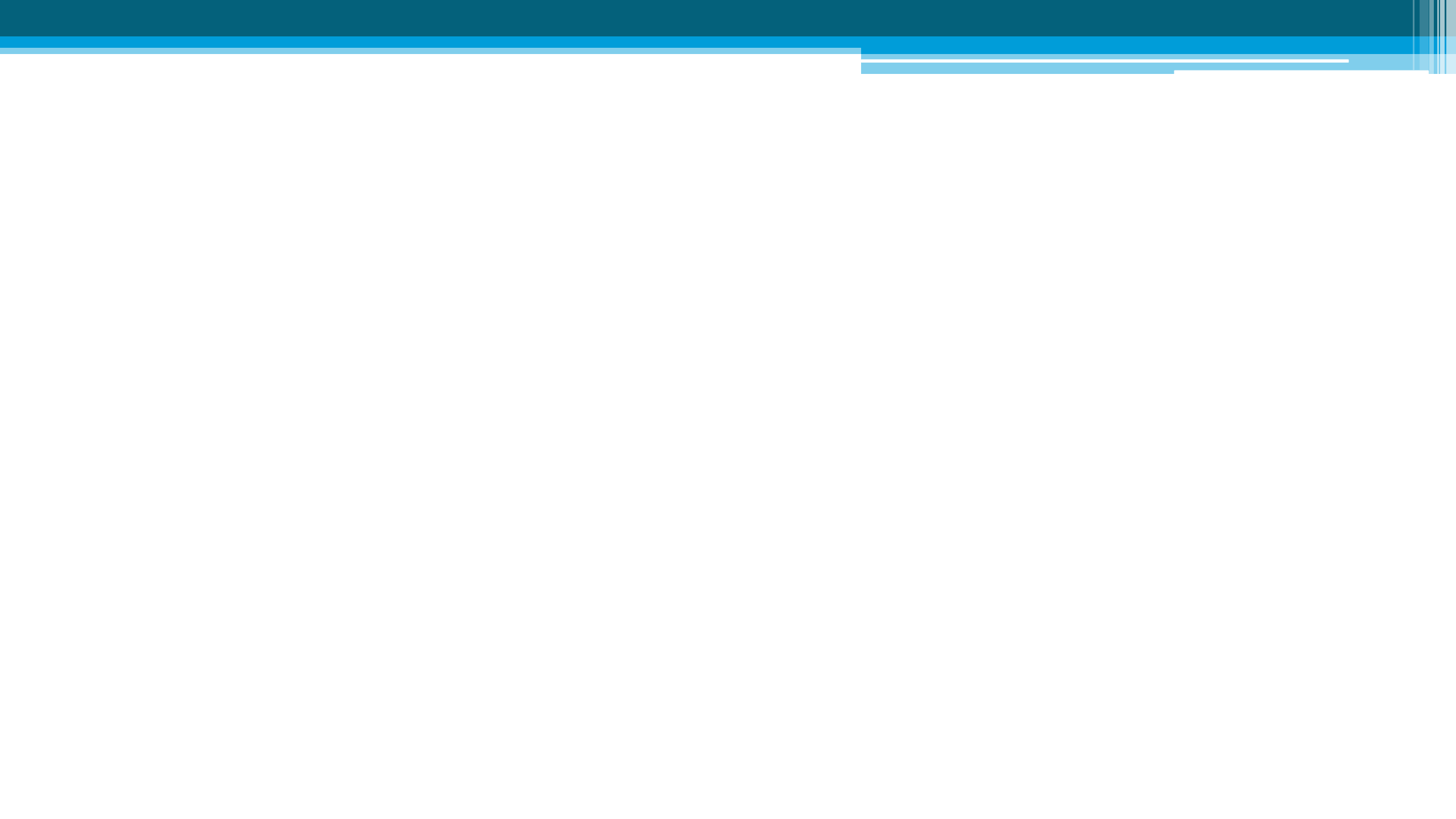
- В. Е. Гай, В. А. Утробин, А. В. Лукьянчикова, И. В. Поляков.

Распознавание изолированных речевых команд с позиций теории активного восприятия (В печати)

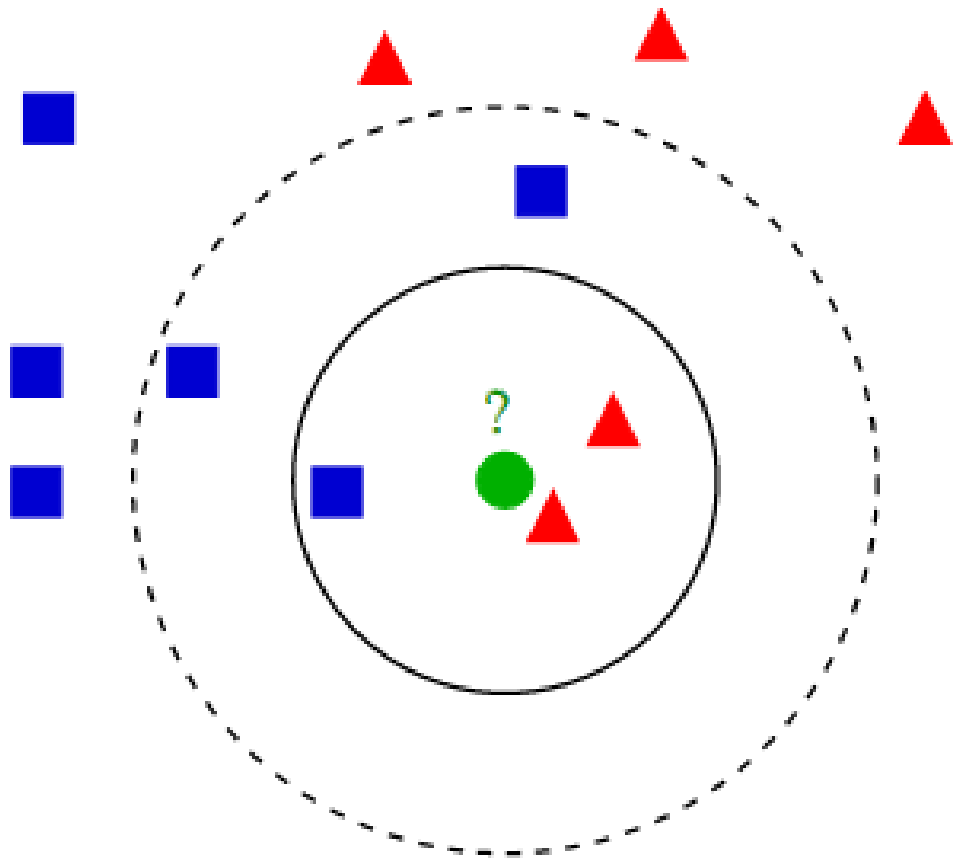
Спасибо за внимание!

A decorative graphic element consisting of several horizontal lines in shades of blue and white, extending across the width of the slide below the main text.





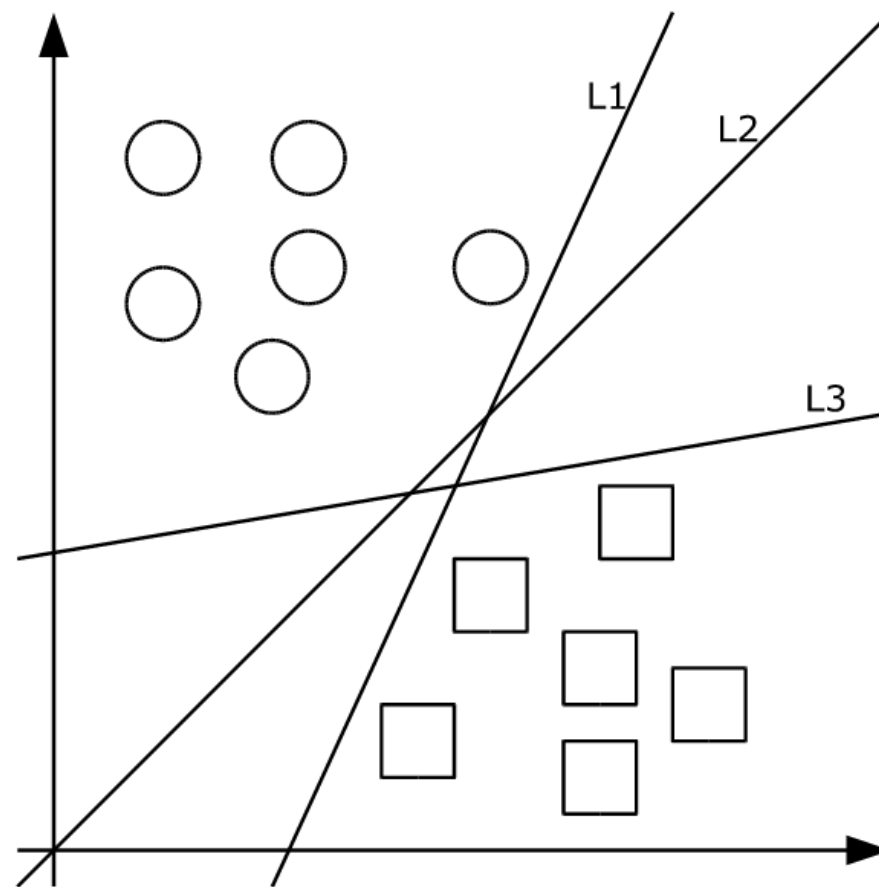
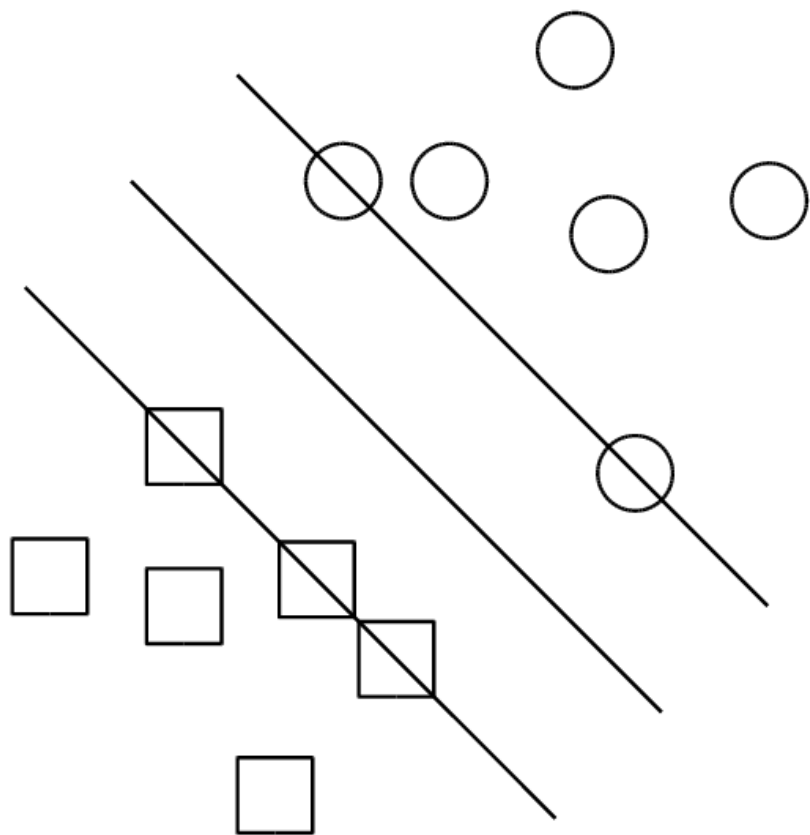
Метод k-ближайших соседей



Пример классификации k ближайших соседей. Тестовый образец (зеленый круг) должен быть классифицирован как синий квадрат (класс 1) или как красный треугольник (класс 2). Если $k = 3$, то она классифицируется как 2-й класс, потому что внутри меньшего круга 2 треугольника и только 1 квадрат. Если $k = 5$, то он будет классифицирован как 1ый класс (3 квадрата против 2ух треугольников внутри большего круга).

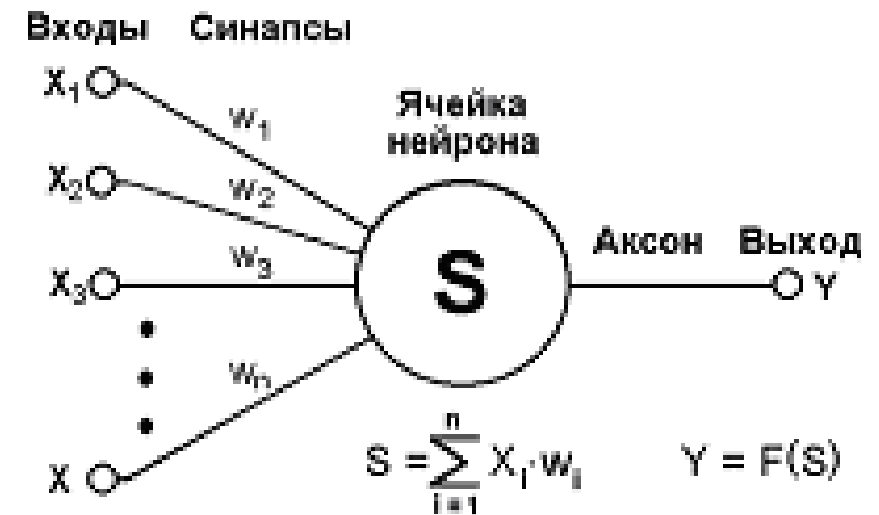
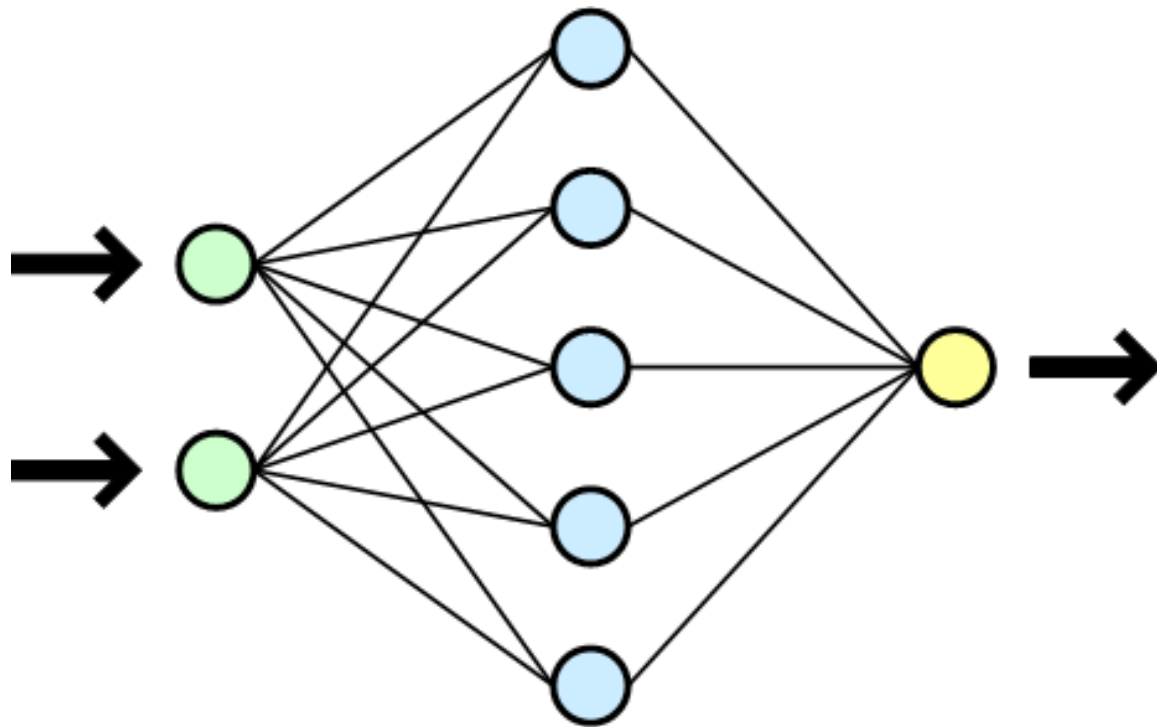
Метод опорных векторов

Оптимальная разделяющая гиперплоскость для метода опорных векторов, построенная на точках из двух классов. Ближайшие к параллельным гиперплоскостям точки называются опорными векторами



Несколько классифицирующих разделяющих прямых (гиперплоскостей). Но только одна достигает оптимального разделения

Нейронная сеть

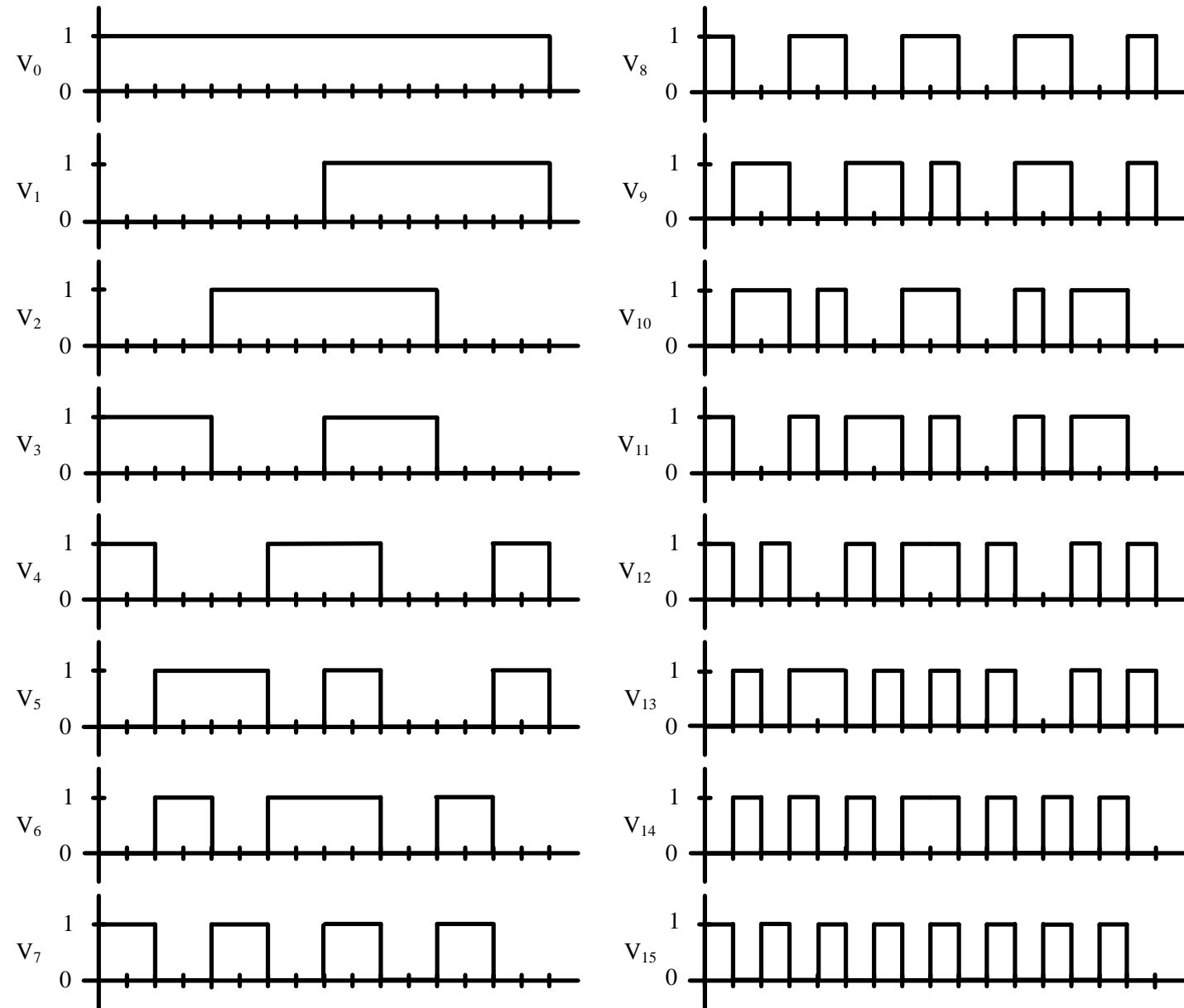


Алгебра групп

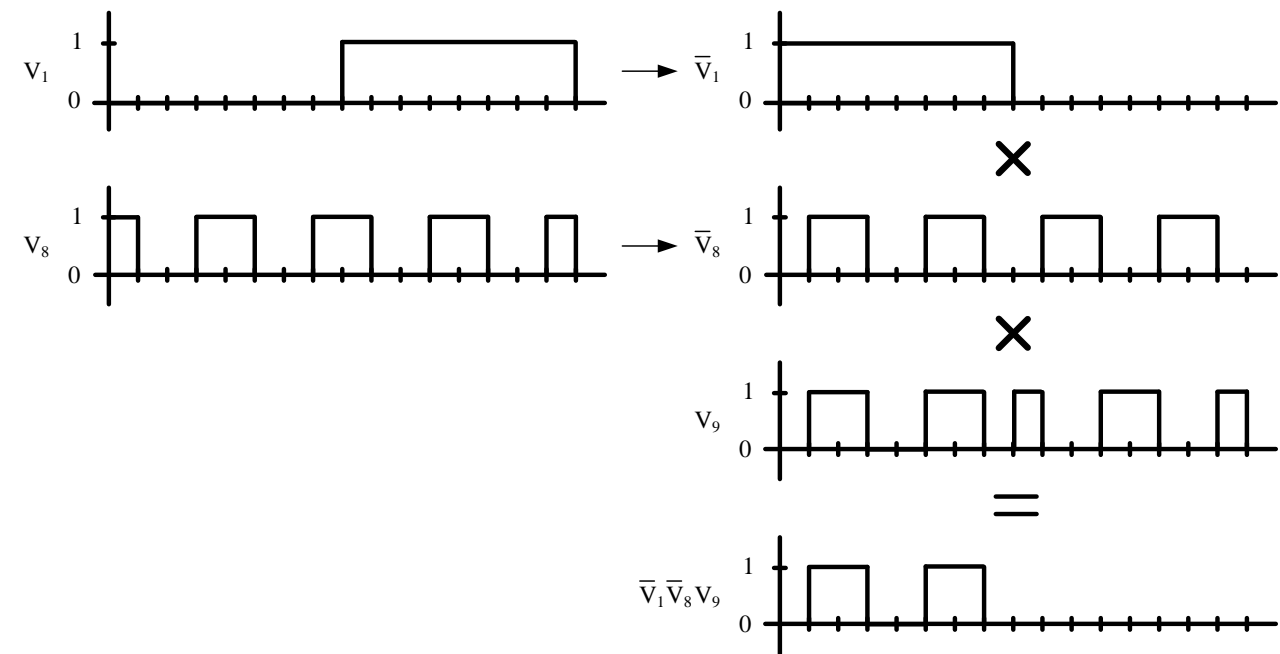
Алгебра групп используется для формирования описания речевого сигнала. В алгебре A_v существуют алгебраические группы (V_i - бинарные операторы):

- 1) P_{ni} (названы полными), образованы на тройках операторов (V_i, V_j, V_k)
- 2) P_{si} (названы замкнутыми), образованы на четверке операторов (V_i, V_j, V_p, V_m)

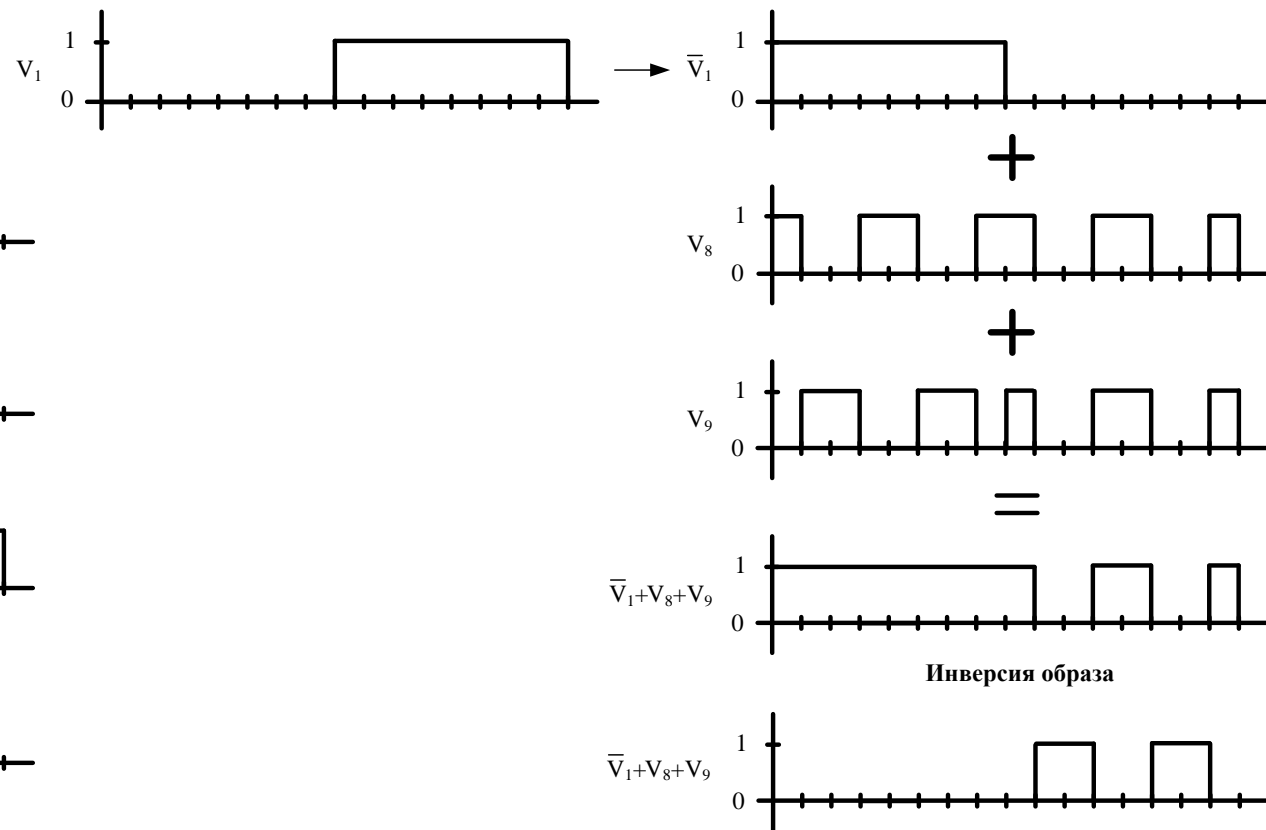
Операторы



Полные группы



Вычисление образа полной группы на операции умножения



Вычисление образа полной группы на операции сложения

Замкнутые группы

