



Программа микроконтроллера для системы распознавания речевых сигналов

Выполнил :
Чуркин Вадим Сергеевич

Научный руководитель:
Загоруйко Николай Григорьевич

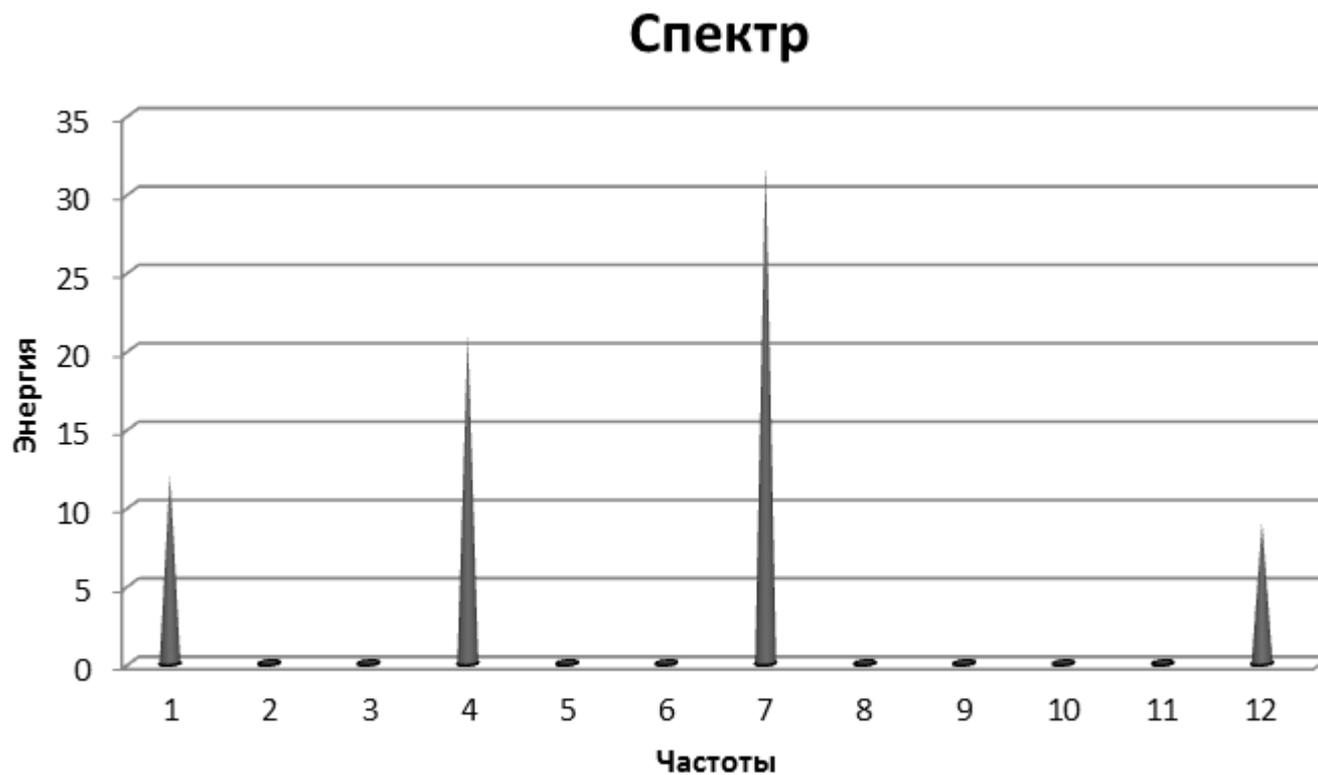


Эффект Маскировки

Эффект маскировки это особенность, присущая слуховой системе человека. Этот эффект используется для более компактного описания речевых сигналов со сложным спектром, за счёт выделения из полного спектрального описания сигнала значений частот - формантных признаков, которые используются в качестве критериев, для принятия фонемного решения.

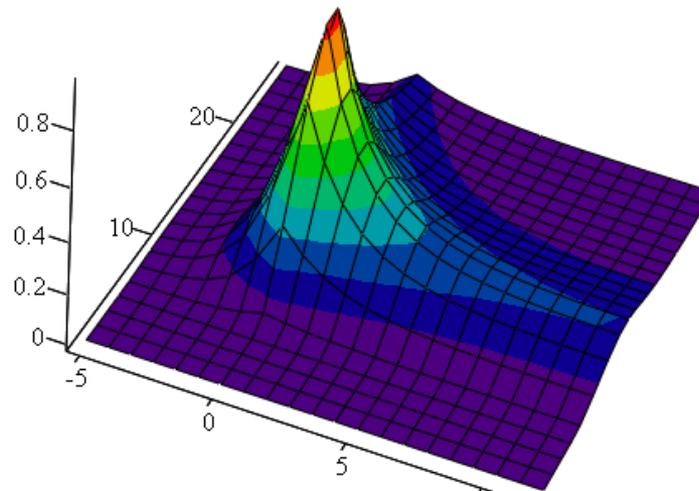
Маскировка по частоте представляет собой скрывание испытательных тонов, если они ниже определённого уровня ослабления в зависимости от амплитуды маскира и частоты расхождения от маскира. Маскировка по времени имеет различие в том, что вместо частоты анализируется время частоты от маскира.

Как это работает:

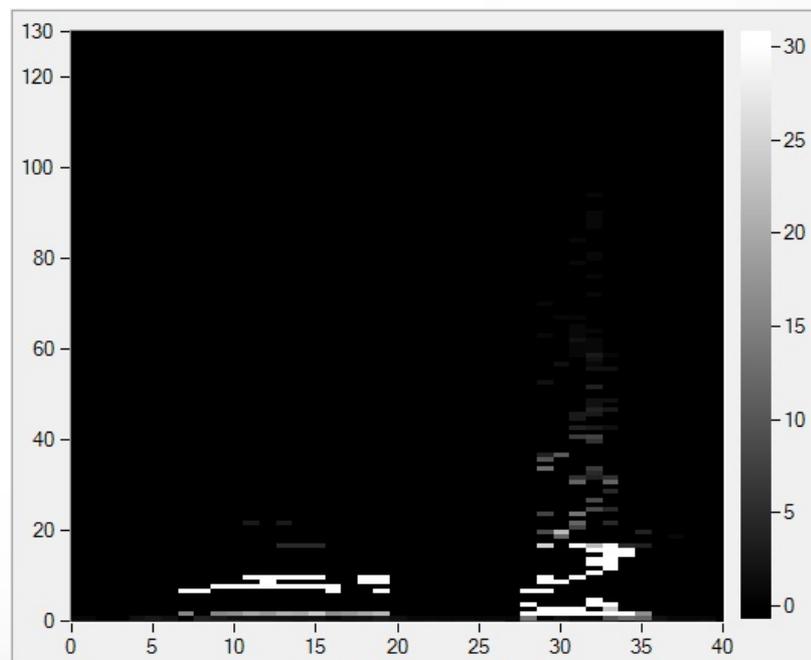
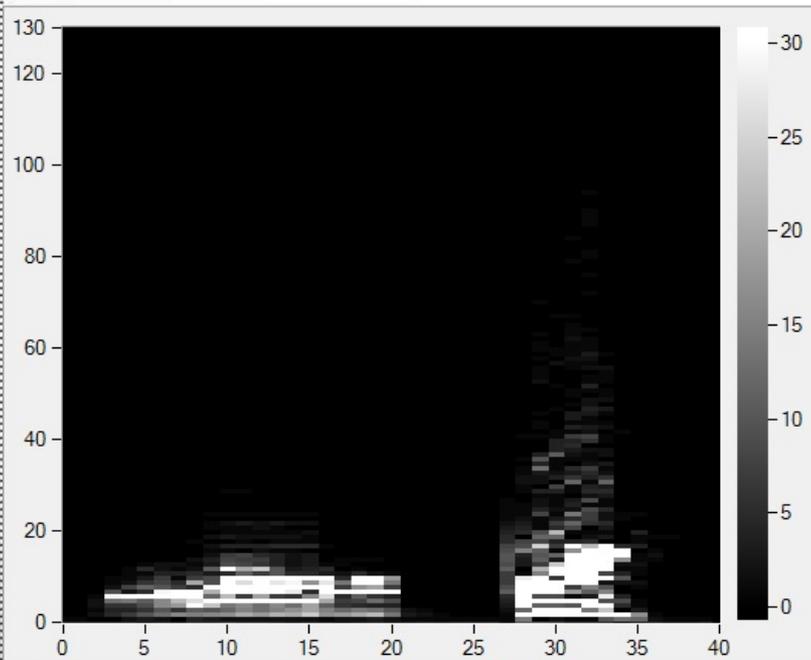


Математическая модель маскировок

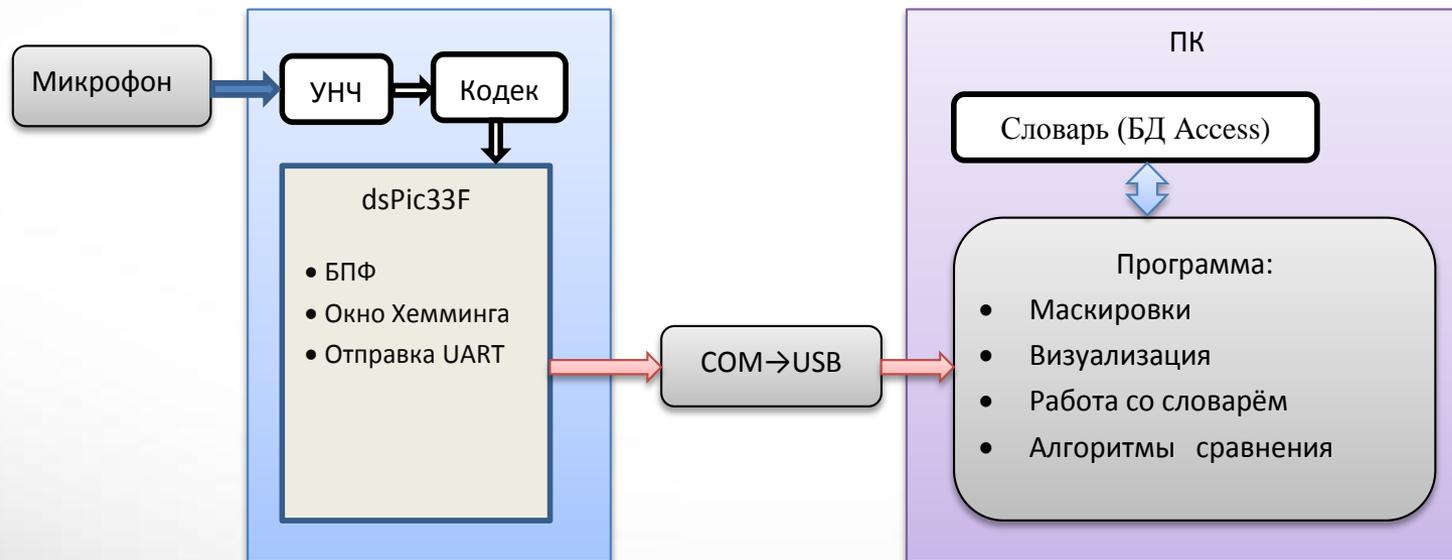
$$\text{masks}(k, j) := \begin{cases} I_m \cdot \exp\left[-1.44 \cdot Q^2 \left(\frac{\omega_m}{\omega(j)} - 1\right)^2\right] \cdot \left(\beta \cdot \exp\left(\frac{k \cdot \Delta t}{\tau}\right)\right) & \text{if } k < 0 \\ I_m \cdot \exp\left[-1.44 \cdot Q^2 \left(\frac{\omega_m}{\omega(j)} - 1\right)^2\right] \cdot \beta \cdot \left[1 - \alpha \cdot \left(1 - \exp\left(\frac{-k \cdot \Delta t}{\tau}\right)\right)\right] & \text{if } k \geq 0 \end{cases}$$



Влияние маскировки



Блок схема преобразований всей системы





Микроконтроллер dsPIC33F

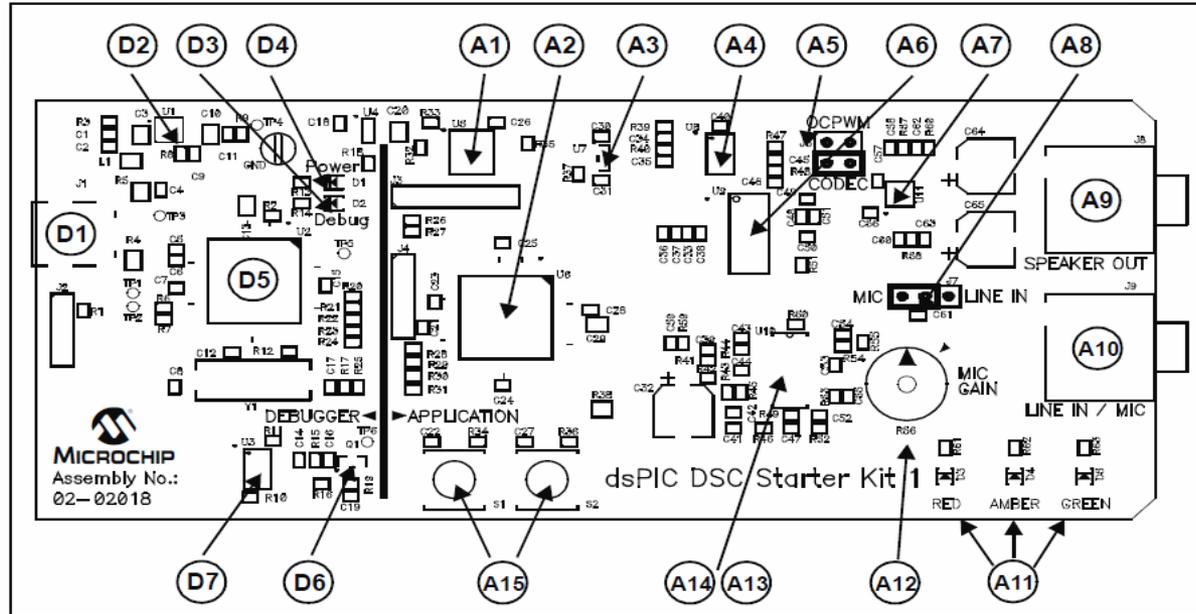
- **Периферия общего использования:**
 - календарь и часы реального времени RTCC
 - аппаратный подсчет CRC
 - расширенная периферия (SPI, I2C, UART (с поддержкой IrDA, LIN), CAN (ECAN))
 - 10-и и 12-битные АЦП
 - компараторы
 - 10-и и 16-битные ЦАП
 - прямой доступ к памяти (DMA)
 - ведущий параллельный порт (PMP)
 - программное переназначение выводов (PPS)
 - многоуровневая система защиты кода (Code Guard)
 - специализированный ШИМ для управления приводом (Motor Control PWM)
 - интерфейс квадратурного энкодера
- **Для импульсных источников питания (SMPS)**
 - Специализированный сверхбыстрый ШИМ с высоким разрешением (SMPS PWM)
 - Специализированные сверхбыстрые АЦП (SMPS ADC)
- **Для работы со звуком:**
 - 12-битный АЦП
 - 16-битный ЦАП
 - специализированный ШИМ (output compare PWM)
 - интерфейс кодирования данных DCI (I2S, AC97)
- **Для управления графическими дисплеями:**
 - ведущий параллельный порт PMP (QVGA)
 - модуль измерения времени заряда CTMU (сенсорные дисплеи touch-screen)

Started kit for dsPIC33F



Лучшее решение для разработки новых программ!

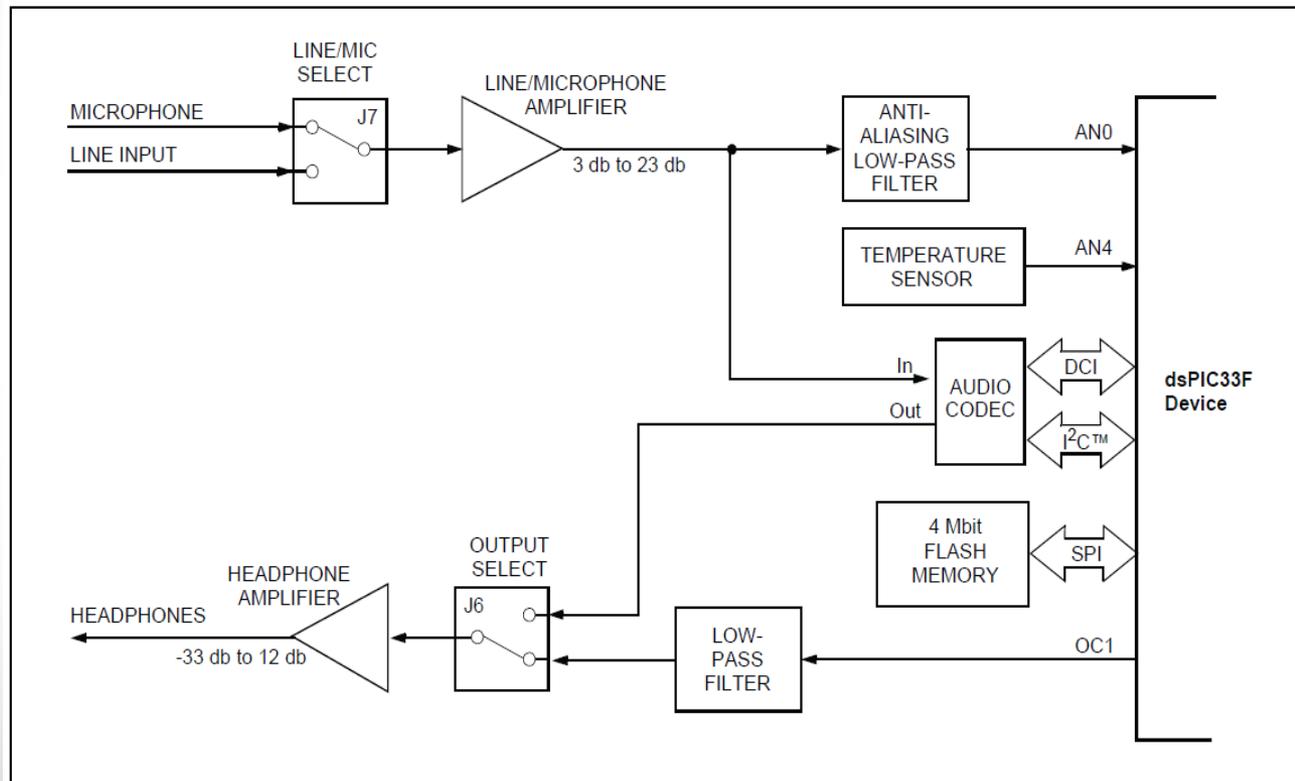
Started kit for dsPIC



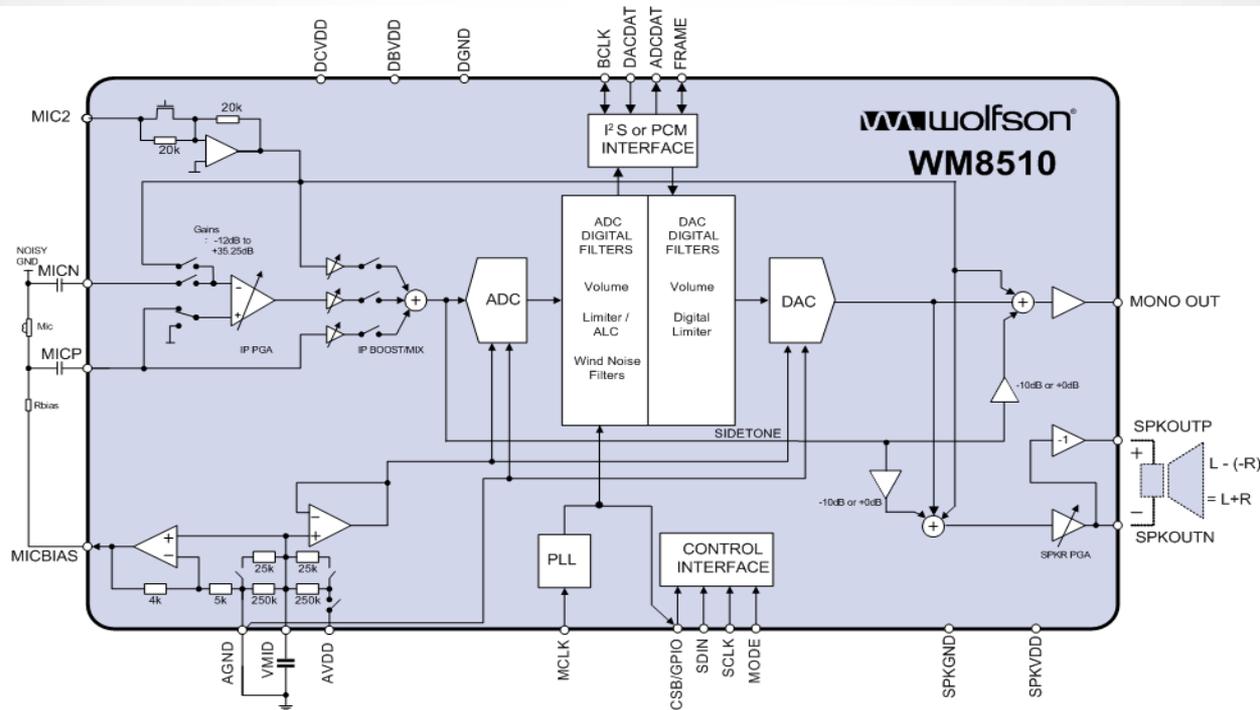
| Ref | Component | Ref | Component |
|-----|---|-----|---|
| D1 | Mini-B USB Connector (J1) | A5 | Output Select Jumper (J6) |
| D2 | MCP1727 (U1) | A6 | Codec (U9) |
| D3 | STATUS LED – Debug (D2) | A7 | Headphone Amplifier (U11) |
| D4 | STATUS LED – System Power (D1) | A8 | Line/Microphone Input Select Jumper (J7) |
| D5 | PIC18F67J50 MCU (U2) | A9 | Headphone Output Jack (J8) |
| D6 | Low Vce Saturation PNP Transistor Switch (Q1) | A10 | Line/Microphone Input Phone Jack (J9) |
| D7 | 25LC010A Serial EEPROM (U3) | A11 | User LEDs (D3,D4 and D5) |
| A1 | Flash Memory (U5) | A12 | Microphone Gain Control (R56) |
| A2 | Digital Signal Control (U6) | A13 | Line/Microphone Pre-Amplifier (U10:A) |
| A3 | Temperature Sensor (U7) | A14 | Anti-Aliasing Low-Pass Filter (U10:B,C,D) |
| A4 | PWM Low Pass Filter (U8:A,B) | A15 | User Switches (S2 and S1) |

D# = Debug components
A# = Audio components

Преобразование аналогово сигнала



Кодек



- Усиление (порядка 50 дБ)
- Фильтрация
- Оцифровка



Функции микроконтроллера

- Получение сегментов. С частотой 16 дискретизации КГц. По теореме Котельникова: сигнал можно восстановить без потерь из его дискретизированного, только если частота дискретизации выше в 2 раза .
- Окно Хемминга - для снижения эффекта «замыливания» спектра, т.к. сегмент конечен.
- Быстрое преобразование Фурье. После преобразования сокращается в 2 раза частота и размер выборки.
- Формирование сигнала в логарифмической шкале, и отправка на ПК.

Передача на ПК



- Скорость передачи : 38400 бод, 19200 бод, 9600 бод.
- Восемь бит в пакете.
- Контроль четности (паритет) – положительный.
- Два стоп бита.
- 3 стробирующих байта со значением 0x55. Характеризуют что начался следующий сегмент.
- Значения в диапазоне от 0x64 до 0xFF, в логарифмической шкале.



Приём данных.

Данные с МК идут непрерывно, программа может их принимать либо игнорировать, отображая текущую спектрограмму в окне, её можно сохранить в словарь или распознать.



Как происходит сравнение

- Тестируемая гистограмма разбивается на частотные области(рекомендованные физиологами).
- Из словаря берется гистограмма разбитая по частотным областям.
- Применяются различные анализы к обоим гистограммам(если требуется)
- Формируется массив схожести 2х гистограмм.
- Алгоритм «Динамо» из массива схожести определяет конечную величину похожести 2х слов.



Разбиение на частотные области

Спектрограмма – набор сегментов.

Сегмент – спектр слова в отдельный момент времени.

Каждый сегмент разбивается на 8 полос(в герцах):
150 – 400 , 400 – 640, 640 – 1040, 1040 – 1520,
1520 – 2200, 2200 – 3040, 3040 – 4200, 4200 – 7000,
7000 – 10000.



Словарь

База данных - файл MS Access.

Поля таблицы:

- Индекс.
- Слово – текстовый вид.
- 2-мерный сериализированный массив спектрограммы.

Преимущества:

- Удобства хранения и переноса(как файл)
- Удобство просмотра записей посредством MS Access.
- Высокая скорость работы со словарем.
- Гибкость структуры(можно вносить доп. информацию)



Мера похожести

- Расстояние между сегментами (Декартово):

$$\rho_{mn}^2 = \sum_{i=1}^p \left(\ln \left(\frac{\tilde{E}_i^{(m)}}{E_0^{(m)}} \right) - \ln \left(\frac{\tilde{E}_i^{(n)}}{E_0^{(n)}} \right) \right)^2$$

$$E_0^{(m)} = \sum_i E_i^{(m)}$$

$$\tilde{E}_i^{(m)} = \sum_{j=li}^{ri} \bar{E}_i^{(m)}$$

- Мера близости:

$$d_{mn} = \frac{\alpha^2}{\alpha^2 + \rho_{mn}^2}$$

Работа алгоритма «Динамо»

- Позволяет распознавать слова с разной длиной.

Массив сравнения – каждый сегмент одного слова, сравнивается с каждым сегментов другого (мера сравнения – производная от Декартового расстояния).

| | | | | | | |
|----|----|----|----|---|----|----|
| | | | | | 18 | 0 |
| 53 | 15 | 1 | 3 | 4 | 3 | 7 |
| | 6 | 10 | 20 | 7 | 10 | 2 |
| | 12 | 8 | 5 | 2 | 8 | 18 |
| | 15 | 15 | 15 | 6 | 6 | 6 |
| | 8 | 2 | 15 | 1 | 4 | 6 |
| | 0 | 0 | 0 | 0 | 0 | 0 |

$$A(i, j) = \max \begin{cases} A(i, j + 1) \\ A(i + 1, j + 1) + d(i, j) \\ A(i + 1, j) \end{cases} \quad i = I, I - 1, \dots, 1; j = J, J - 1, \dots, 1.$$

Пример работы:

The screenshot shows the StarterKid Connect software interface. On the left, there are settings for the COM port (COM1 or COM3), frequency (98400), parity (EVEN), bits per packet (EIGHT), and stop bits (TWO). A central spectrogram displays a signal with a color scale from -130 to -50. A dialog box in the foreground lists the most suitable word and its similarity score, along with other words and their scores.

COM Порт
COM1
COM3
ЗАКРЫТЬ

Частота 98400
Четность EVEN
Бит в пакете EIGHT
Стоп битов TWO

Вывод в файл

Настройки Маскировок Маскировка

Наиболее подходящее слово: Воздух С величиной похожести: 18,3653643830715

| | |
|---------------|---------------------|
| Слово: Воздух | -- 18,3653643830715 |
| Слово: Чешуя | -- 16,3051629591596 |
| Слово: Бубен | -- 15,9539331395261 |
| Слово: Дверь | -- 14,3958827075292 |
| Слово: Шабаш | -- 12,7718315975735 |
| Слово: Хищник | -- 10,4983620105025 |
| Слово: Химия | -- 9,86370416084802 |
| Слово: Свист | -- 8,85495068082523 |
| Слово: Помпа | -- 6,87275123500114 |



Публикации:

- Диплом 3й степени МНСК (выступление).
- Журнал «Альманах современной науки и образования» - принято к публикации.



Что сделано:

- Реализация алгоритма маскировок
- Реализация алгоритма Динамо
- Разработка интерфейса МК- ПК
- Разработка устройства согласования МК-ПК
- Разработка словаря, для хранения слов и механизмов работы со словарем
- Реализация механизмов сравнения слов
- Разработка алгоритма определения длинны слов

A wireframe sphere composed of intersecting lines, rendered in a light blue-green color, is positioned in the upper-left corner of the slide. It is set against a solid black circular background.

Спасибо за внимание!

ВОПРОСЫ?