

Министерство образования Российской Федерации  
Ульяновский государственный технический университет

КОМПЬЮТЕРНЫЕ ТЕХНОЛОГИИ  
В ОБЛАСТИ ИНФОКОММУНИКАЦИЙ

Сборник лабораторных работ  
Часть 2 «Исследование речевых сигналов»

Составители: В.Н. Горохин

Ульяновск 2015

УДК 621.391 (076)

ББК 32я7

Т 33

Рецензент канд. тех. наук, доцент кафедры САПР Кадеев Д.Н.

Одобрено секцией методических пособий научно–методического совета университета.

Т 33 Компьютерные технологии в области инфокоммуникаций: Сборник лабораторных работ. Часть 2 «Исследование речевых сигналов»/ сост. В.Н. Горохин. – Ульяновск : УлГТУ, 2015. – 100 с.

Сборник лабораторных работ разработан в соответствии с программой курса «Компьютерные технологии в области инфокоммуникаций» и предназначен для магистров и студентов радиотехнического факультета специальности «Сети связи и системы коммутации», но может использоваться и студентами других специальностей. Лабораторные работы посвящены исследованию речевых сигналов в среде программных продуктов математического профиля.

Сборник подготовлен на кафедре «Телекоммуникации».

УДК 621.391 (076)

ББК 32я7

© Оформление. УлГТУ, 2009

## СОДЕРЖАНИЕ

ВВЕДЕНИЕ .....	4
<b>Лабораторная работа № 1</b>	
ОБРАБОТКА РЕЧЕВЫХ СИГНАЛОВ В СРЕДЕ MATHCAD .....	5
<b>Лабораторная работа № 2</b>	
МОДЕЛИРОВАНИЕ ФОРМАНТНЫХ ВОКОДЕРОВ .....	22
<b>Лабораторная работа № 3</b>	
ИССЛЕДОВАНИЕ КВАНТОВАНИЯ АМПЛИТУД В СПЕКТРЕ РЕЧЕВОГО СИГНАЛА ФОРМАНТНОГО ВОКОДЕРА.....	37
<b>Лабораторная работа № 4</b>	
ИССЛЕДОВАНИЕ КВАНТОВАНИЯ ФАЗ В СПЕКТРЕ РЕЧЕВОГО СИГНАЛА ФОРМАНТНОГО ВОКОДЕРА .....	52
<b>Лабораторная работа № 5</b>	
ИССЛЕДОВАНИЕ МОДЕЛИ ЛИНЕЙНОГО ПРЕДСКАЗАНИЯ РЕЧЕВОГО СИГНАЛА .....	62
<b>Лабораторная работа № 6</b>	
СОЗДАНИЕ И ИССЛЕДОВАНИЕ СПЕКТРОГРАММ РЕЧЕВЫХ СИГНАЛОВ .....	67
<b>Лабораторная работа №7</b>	
РАЗРАБОТКА СПОСОБА РАСПОЗНАВАНИЯ РЕЧИ НА ОСНОВЕ СПЕКТРОГРАММ .....	74
<b>Лабораторная работа № 8</b>	
РАЗРАБОТКА СПОСОБА РАСПОЗНАВАНИЯ РЕЧИ НА ОСНОВЕ МОДЕЛИ ЛИНЕЙНОГО ПРЕДСКАЗАНИЯ .....	82
<b>Лабораторная работа № 9</b>	
ИССЛЕДОВАНИЕ ЗВУКОВЫХ СИГНАЛОВ МУЗЫКАЛЬНЫХ ИНСТРУМЕНТОВ .....	86
БИБЛИОГРАФИЧЕСКИЙ СПИСОК.....	88

## *ВВЕДЕНИЕ*

Речевые сигналы широко используются в различных отраслях человеческой деятельности, особенно в современных системах связи и коммуникаций. Исследование их параметров и создание эффективных способов и средств анализа и синтеза являются актуальными и настоящее время.

Сборник лабораторных работ разработан в соответствии с программой курса «Компьютерные технологии в области инфокоммуникаций» и предназначен для магистров и студентов радиотехнического факультета специальности «Сети связи и системы коммутации», но может использоваться и студентами других специальностей. Лабораторные работы посвящены исследованию речевых сигналов в среде программных продуктов математического профиля.

Сборник подготовлен на кафедре «Телекоммуникации».

## *Лабораторная работа №1*

### *ОБРАБОТКА РЕЧЕВЫХ СИГНАЛОВ В СРЕДЕ MATHCAD*

**Цель работы:** Освоение основных навыков создания речевых сигналов с заданными параметрами и правилами их обработка в среде MathCAD.

#### **Программные среды для проведения лабораторных работ**

Студент самостоятельно осваивает навыки работы в программах CoolEdit и MathCAD.

#### **Задание к лабораторной работе**

Студент создает с помощью программы CoolEdit файлы речевых сигналов . Их параметры:

Длительность – 1..2 сек.

Частота дискретизации - 8 кГц.

Разрядность - 8 (1 байт).

Моно.

Файл должен быть записан в текстовом формате. Любым текстовым редактором у него убирается заголовок.

В качестве примера показан исходный файл S1.txt

SAMPLES: 26112

BITSPERSAMPLE: 8

CHANNELS: 1

SAMPLERATE: 8000

NORMALIZED: FALSE

134

131

131

130

.

.

.

127

127

127

127

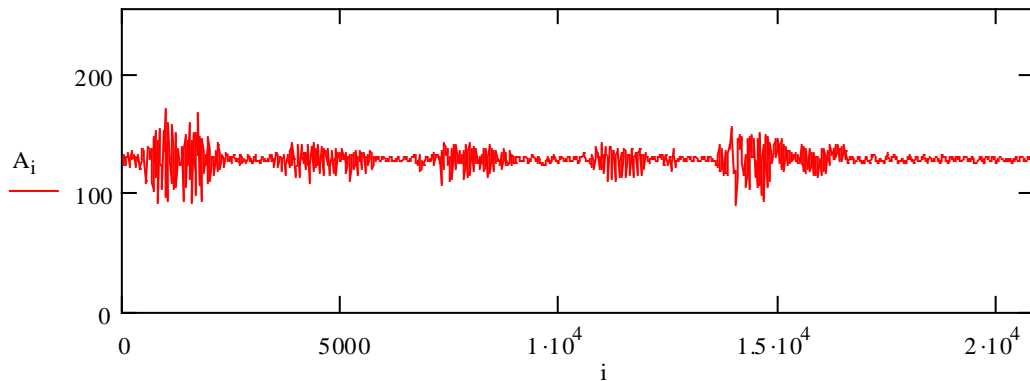
Файл без заголовка S.txt

134

131

131  
 130  
 .  
 .  
 .  
 127  
 127  
 127  
 127

Вид звукового сигнала в среде CoolEdit



Загрузка в массив «А» последовательность отсчетов речевого сигнала с частотой дискретизации 8 кГц - для канала ТЧ из файла «s1-11.txt» в среде MathCAD:

$A := \text{READPRN}("s1-11.txt")$

Разбить последовательность отсчетов звукового сигнала – массив «А» на одинаковые интервалы квазистационарности, например, по 16 мс или по 128 отсчетов для частоты дискретизации 8 кГц. Считаем, что на этих интервалах процесс можно считать достаточно стационарным.

$$j := 0.. \text{floor}\left(\frac{\text{length}(A)}{128}\right) - 1 \quad i := 0.. 127 \quad M_{i,j} := A_{i+128 \cdot j} - 128$$

При этом необходимо вычесть постоянную составляющую (оператор - 128).

Получаем следующую матрицу «М» отсчетов:

$M =$

	0	1	2	3	4	5	6	7	8	9	10
0	-4	2	-1	-5	7	12	-3	-18	-39	0	23
1	-4	0	0	-1	3	7	-17	-25	-39	-4	19
2	-4	-1	1	2	-3	3	-24	-29	-36	-4	16
3	-4	-2	3	5	-8	1	-22	-29	-28	-3	16
4	-3	-4	5	7	-10	-1	-14	-18	-14	-5	16
5	-3	-4	6	7	-11	1	-5	3	8	-9	15
6	-3	-5	7	7	-14	5	5	22	33	-13	12
7	-3	-5	8	6	-16	9	12	31	45	-17	9
8	-3	-6	10	5	-17	12	17	32	39	-22	7

Вычислить быстрое преобразование Фурье (БПФ) для каждого интервала по 128 отсчетов, получится матрица «Mf»:

$$j := 0.. \text{cols}(M) - 1$$

$$Mf^{<j>} := \text{FFT}(M^{<j>})$$

Mf =

	0	1	2
0	1.117	-0.547	0.898
1	1.274+0.822i	-0.885-0.482i	0.104+0.366i
2	-1.405+1.606i	1.305+1.196i	-0.766-2.118i
3	0.543+0.923i	-0.927+1.533i	-1.611-1.851i
4	-0.474+0.365i	0.248-0.394i	-0.282-0.244i
5	-0.204+0.261i	1.154+0.338i	0.61-1.098i
6	-0.017+0.223i	0.089+1.068i	-0.612-1.362i
7	-0.219+0.182i	-0.822-0.594i	0.225+0.233i
8	0.116+0.17i	-0.794-0.068i	-0.336-0.444i
9	-0.12+0.278i	0.281-0.03i	0.042+0.273i

Вычислить обратное быстрое преобразование Фурье

$$Ml^{<j>} := \text{IFFT}(Mf^{<j>}) + 128$$

Получить матрицу восстановленного сигнала «Ml»:

Ml =

	0	1	2	3	4	5
0	124.606	131.139	121.277	128.497	140.212	133.471
1	123.535	130.983	122.864	128.386	137.195	125.945
2	122.987	130.545	124.938	129.659	133.723	120.434
3	123.082	130.042	127.119	131.84	130.181	118.056
4	123.612	129.571	128.97	134.04	127.371	119.145
5	124.208	129.034	130.293	135.634	125.947	123.245
6	124.587	128.236	131.225	136.563	125.997	129.286
7	124.704	127.138	132.046	137.019	127.136	135.885
8	124.692	126.033	132.871	136.9	128.813	141.7

Записываем полученную матрицу в выходной текстовый файл, используя дополнительный массив «Anew»:

$$j := 0.. \text{cols}(Ml) - 1$$

$$i := 0.. \text{rows}(Ml) - 1$$

$$Anew_{i+128,j} := \text{round}(Ml_{i,j})$$

$$\text{WRITEPRN}("s\_out.txt") := Anew$$

Затем этот файл «s\_out.txt» можно прослушать через программу Cool Editor и сравнить его с исходным, они должны быть одинаковые.

После этого необходимо модифицировать исходный звуковой файл, изменяя его спектр. После вычисления быстрого преобразования Фурье, произвести обнуление фаз его спектральных составляющих (мнимые приравнять к нулю, а реальные к модулю). Это соответствует сжатию звукового сигнала в 2 раза. Оценить качество звучания полученного сигнала.

### **Содержание отчета**

Отчет по проделанной лабораторной работе должен быть оформлен в виде презентации и должен содержать:

- основные формулы;
- операторы;
- текст программы;
- демонстрацию звуковых файлов.

### **Контрольные вопросы**



## ***Лабораторная работа № 2***

### **МОДЕЛИРОВАНИЕ ФОРМАНТНЫХ ВОКОДЕРОВ**

**Цель работы:** Изучение и исследование формантной модели речевых сигналов.

#### **Краткие теоретические сведения**

В спектре речевого сигнала на интервалах квазистационарности наблюдаются выраженные максимумы мощности на определенных частотах. Эти частоты называются формантами.

#### **Задание к лабораторной работе**

Произвести исследование звуковых сигналов по количеству оставляемых формант. Определить оптимальное значение формант для приемлемого качества звучания. При этом все остальные отсчеты в спектре сигнала обнуляются. Подсчитать какова при этом получается степень сжатия речевого сигнала.

#### **Содержание отчета**

Отчет по проделанной лабораторной работе должен быть оформлен в виде презентации и должен содержать:

- основные формулы;
- операторы;
- текст программы;
- демонстрацию звуковых файлов.

#### **Контрольные вопросы**

**Лабораторная работа № 3**  
**ИССЛЕДОВАНИЕ КВАНТОВАНИЯ АМПЛИТУД В СПЕКТРЕ** **РЕ-**  
**ЧЕВОГО СИГНАЛА ФОРМАНТНОГО ВОКОДЕРА**

**Цель работы:** Исследовать влияние квантования амплитудного спектра речевого сигнала на качество формантного вокодера.

**Краткие теоретические сведения**

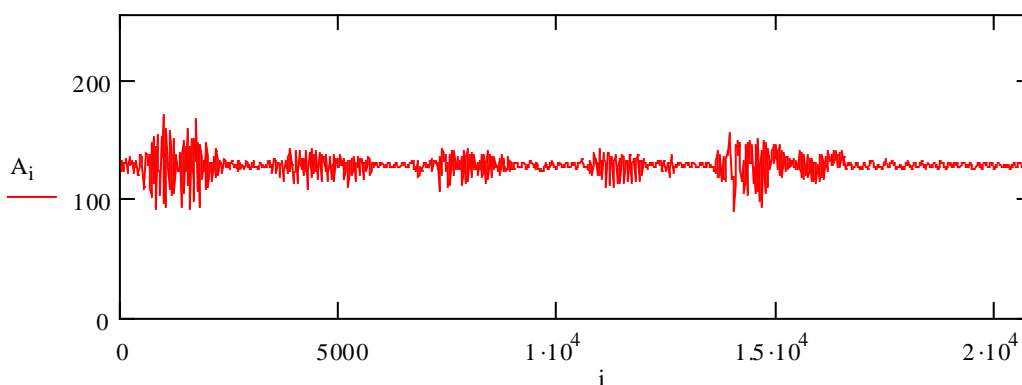
В формантных вокодерах восстановление речевого сигнала производится по информации о формантных максимумах. Параметрами, передаваемыми по каналу связи и позволяющими с достаточным приближением синтезировать картину текущего спектра, являются здесь сигналы о средних частотах и уровнях: формант.

В данной работе исследуется зависимость разрядности квантования амплитуды и фазы спектра речевого сигнала от качества передаваемого сигнала на выходе вокодера. Также определяется оптимальный вариант разрядности квантования для входного сигнала ТЧ 64 Кбит/с.

**Задание к лабораторной работе**

Загрузить в массив последовательность отсчетов речевого сигнала с частотой дискретизации 8 кГц - для канала ТЧ:

```
A := READPRN("s1-11.txt")
```



Разбить эту последовательность на одинаковые интервалы по 16 мс или по 128 отсчетов, предполагая, что на этих интервалах процесс достаточно стационарный.

$$j := 0.. \text{floor}\left(\frac{\text{length}(A)}{128}\right) - 1 \quad i := 0.. 127 \quad M_{i,j} := A_{i+128 \cdot j} - 128$$

Получается следующая матрица отсчетов:

$$M =$$

	0	1	2	3	4	5	6	7	8	9	10
0	-4	2	-1	-5	7	12	-3	-18	-39	0	23
1	-4	0	0	-1	3	7	-17	-25	-39	-4	19
2	-4	-1	1	2	-3	3	-24	-29	-36	-4	16
3	-4	-2	3	5	-8	1	-22	-29	-28	-3	16
4	-3	-4	5	7	-10	-1	-14	-18	-14	-5	16
5	-3	-4	6	7	-11	1	-5	3	8	-9	15
6	-3	-5	7	7	-14	5	5	22	33	-13	12
7	-3	-5	8	6	-16	9	12	31	45	-17	9
8	-3	-6	10	5	-17	12	17	32	39	-22	7

Вычислить быстрое преобразование Фурье (БПФ) для каждого интервала по 128 отсчетов и также получить следующую матрицу:

$$j := 0.. \text{cols}(M) - 1$$

$$Mf^{\langle j \rangle} := \text{FFT}(M^{\langle j \rangle})$$

$$Mf =$$

	0	1	2
0	1.117	-0.547	0.898
1	1.274+0.822i	-0.885-0.482i	0.104+0.366i
2	-1.405+1.606i	1.305+1.196i	-0.766-2.118i
3	0.543+0.923i	-0.927+1.533i	-1.611-1.851i
4	-0.474+0.365i	0.248-0.394i	-0.282-0.244i
5	-0.204+0.261i	1.154+0.338i	0.61-1.098i
6	-0.017+0.223i	0.089+1.068i	-0.612-1.362i
7	-0.219+0.182i	-0.822-0.594i	0.225+0.233i
8	0.116+0.17i	-0.794-0.068i	-0.336-0.444i
9	-0.12+0.278i	0.281-0.03i	0.042+0.273i

Произведем квантование амплитудных значений спектра сигнала с разрядностью квантования равной 4 - 8:

$$\text{razm} := 5$$

$$i := 0.. 64$$

$$Mf\_md_{i,j} := \sqrt{(\text{Re}(Mf_{i,j}))^2 + (\text{Im}(Mf_{i,j}))^2}$$

$$Mf\_mod_{i,j} := \text{floor}\left(\frac{Mf\_md_{i,j}}{\max(Mf\_md^{\langle j \rangle}) + 10^{-10}} \cdot 2^{\text{razm}}\right)$$

Получим следующие значения квантованных амплитуд РС:

$$Mf\_mod =$$

	0	1	2	3	4	5	6	7	8	9
0	16	9	11	26	3	0	0	3	7	1
1	22	17	4	14	11	4	2	2	5	5
2	31	31	29	21	22	22	22	23	15	17
3	16	31	31	24	12	11	10	10	6	6
4	8	8	4	17	15	13	14	14	11	14
5	4	21	16	14	15	8	8	9	4	9
6	3	19	19	8	4	8	8	8	7	5
7	4	18	4	31	22	17	16	14	6	9
8	3	14	7	12	2	4	1	1	8	7

Произведем выделение максимальных по значению формантов. Количество примем равным 5.

Задаем начальные значения отсчетов промежуточных матриц:

$$Mf\_mod\_2_{i,j} := 0 \quad Mf\_mod\_3_{i,j} := 0 \quad Mf\_mod\_4_{i,j} := 0$$

$$i := 1..63$$

Количество оставляемых формантов:  $ff := 5$

Производим выделение формантов так, чтобы каждый отсчет был больше, чем два его соседних, следовательно, он будет максимальным.

$$Mf\_mod\_2_{i,j} := \text{if}(Mf\_mod_{i-1,j} < Mf\_mod_{i,j} > Mf\_mod_{i+1,j}, Mf\_mod_{i,j}, 0)$$

Производим сортировку полученных формантов:

$$Mf\_mod\_3_{\langle j \rangle} := \text{reverse}(\text{sort}(Mf\_mod\_2_{\langle j \rangle}))$$

Выделяем те отсчеты, которые больше заданного минимального порога  $Mf\_mod\_3_{ff}$ , который равен предельному значению отсчета, отстоящего от максимального на количество максимальных формантов (в данном случае 5).

$$Mf\_mod\_4_{i,j} := \text{if}(Mf\_mod\_3_{ff,j} < Mf\_mod\_2_{i,j}, Mf\_mod\_2_{i,j}, 0)$$

Записываем полученную матрицу в исходную  $Mf\_mod$ :

$$i := 0..64 \quad Mf\_mod_{i,j} := Mf\_mod\_4_{i,j}$$

$$Mf\_mod =$$

	0	1	2	3	4	5	6	7	8	9	10	11
0	0	0	0	0	0	0	0	0	0	0	0	0
1	0	0	0	0	0	0	0	0	0	0	0	0
2	31	0	0	0	22	22	22	23	15	17	17	20
3	0	0	31	24	0	0	0	0	0	0	0	0
4	0	0	0	0	0	13	14	14	11	14	14	10
5	0	21	0	0	0	0	0	0	0	0	0	0
6	0	0	19	0	0	0	0	0	0	0	0	0

Также производим квантование фазового спектра сигнала с разрядностью квантования равной 1 - 8:

$$\text{razf} := 3$$

$$\text{Mf}_{\phi i,j} := \begin{cases} \frac{\pi}{2} \cdot \text{sign}(\text{Im}(\text{Mf}_{i,j})) & \text{if } \text{Re}(\text{Mf}_{i,j}) = 0 \\ \text{atan}\left(\frac{\text{Im}(\text{Mf}_{i,j})}{\text{Re}(\text{Mf}_{i,j})}\right) & \text{if } \text{Re}(\text{Mf}_{i,j}) > 0 \\ \text{atan}\left(\frac{\text{Im}(\text{Mf}_{i,j})}{\text{Re}(\text{Mf}_{i,j})}\right) + \pi & \text{if } \text{Re}(\text{Mf}_{i,j}) < 0 \end{cases}$$

$$\text{Mf}_{\phi u_{i,j}} := \text{Mf}_{\phi i,j} - \min(\text{Mf}_{\phi})$$

$$\text{Mf}_{\phi 1u_{i,j}} := \text{floor}\left(\frac{\text{Mf}_{\phi u_{i,j}}}{\max(\text{Mf}_{\phi}^{\langle j \rangle}) + 10^{-10}} \cdot 2^{\text{razf}}\right)$$

В результате получаем матрицу квантованных коэффициентов фаз:

	0	1	2	3	4	5	6	7	8	9
0	2	6	2	2	6	6	2	2	2	6
1	2	6	3	1	6	4	3	1	1	6
2	5	2	7	4	3	1	7	6	5	4
3	3	4	7	2	5	2	7	5	2	6
4	5	0	6	5	2	6	3	7	6	3
5	4	2	0	1	3	6	2	6	0	4
6	4	3	7	5	0	2	5	0	6	1
7	5	6	3	0	1	2	4	6	0	2
8	3	6	7	3	2	5	7	2	6	7

Проводим восстановление сигнала и обратное быстрое преобразование Фурье (ОБПФ):

$$\text{Mf}_{\text{mod}_{i,j}} := \text{Mf}_{\text{mod}_{i,j}} \cdot \frac{\max(\text{Mf}_{\text{md}}^{\langle j \rangle})}{2^{\text{razm}}}$$

$$\text{Mf}_{\phi 1_{i,j}} := \text{Mf}_{\phi 1u_{i,j}} \cdot \frac{\max(\text{Mf}_{\phi}^{\langle j \rangle})}{2^{\text{razf}}} + \min(\text{Mf}_{\phi})$$

$$\text{Mf}_{1_{i,j}} := \text{Mf}_{\text{mod}_{i,j}} \cdot e^{i \cdot \text{Mf}_{\phi 1_{i,j}}}$$

$$j := 0.. \text{cols}(\text{Mf1}) - 1 \quad i := 0.. 64$$

$$\text{Mf}^{\langle j \rangle} := \text{IFFT}(\text{Mf1}^{\langle j \rangle}) + 128$$

Получаем матрицу восстановленного сигнала:

	0	1	2	3	4	5
0	124.606	131.139	121.277	128.497	140.212	133.471
1	123.535	130.983	122.864	128.386	137.195	125.945
2	122.987	130.545	124.938	129.659	133.723	120.434
3	123.082	130.042	127.119	131.84	130.181	118.056
4	123.612	129.571	128.97	134.04	127.371	119.145
5	124.208	129.034	130.293	135.634	125.947	123.245
6	124.587	128.236	131.225	136.563	125.997	129.286
7	124.704	127.138	132.046	137.019	127.136	135.885
8	124.692	126.033	132.871	136.9	128.813	141.7

И записываем ее в выходной текстовый файл:

```
j := 0.. cols(MI) - 1          i := 0.. rows(MI) - 1
Anewi+128.j := round(MIi,j)
WRITEPRN("s_out.txt") := Anew
```

Затем этот файл несложно можно прослушать через программу Cool Editor и оценить качество и разборчивость передаваемого речевого сообщения.

Произвести исследование при различной разрядности амплитудного спектра сигнала.

Оценить минимальное значение разрядности при достаточном качестве речевого сигнала.

Определить коэффициент сжатия и битовую скорость при этом при данном методе.

### Содержание отчета

Отчет по проделанной лабораторной работе должен быть оформлен в виде презентации и должен содержать:

- основные формулы;
- операторы;
- текст программы;
- демонстрацию звуковых файлов.

## **Контрольные вопросы**

## **Лабораторная работа № 4**

### **ИССЛЕДОВАНИЕ КВАНТОВАНИЯ ФАЗ В СПЕКТРЕ РЕЧЕВОГО СИГНАЛА ФОРМАНТНОГО ВОКОДЕРА**

**Цель работы:** Изучение и освоение управления периферийным устройством ввода с клавиатуры, подключенным к AVR микроконтроллеру фирмы Atmel, а также дальнейшее изучение системы команд и системы прерываний.

#### **Задание к лабораторной работе**

##### **Порядок выполнения лабораторной работы**

Необходимо выполнить лабораторную работу №3 изменяя значение количества разрядов при квантовании фазового спектра сигнала от 1 до 8.

Оценить минимальное значение разрядности квантования фазового спектра при достаточном качестве речевого сигнала.

Определить коэффициент сжатия и битовую скорость при этом при данном методе.

#### **Содержание отчета**

Отчет по проделанной лабораторной работе должен быть оформлен в виде презентации и должен содержать:

- основные формулы;
- операторы;
- текст программы;
- демонстрацию звуковых файлов.

#### **Контрольные вопросы**



## Лабораторная работа № 5

### ИССЛЕДОВАНИЕ МОДЕЛИ ЛИНЕЙНОГО ПРОГНОЗИРОВАНИЯ РЕЧЕВОГО СИГНАЛА

**Цель работы:** Исследование модели линейного предсказания речевого сигнала.

#### Задание к лабораторной работе

Проведено исследование линейного предсказания в программе Mathcad, определить ошибки предсказания, а также получить оптимальное решение при передаче речевого сигнала с использованием метода линейного предсказания.

#### Краткое теоретическое введение

Идея линейного предсказания (ЛП) такова: сформировать из нескольких отсчетов речи линейную комбинацию, наиболее точно аппроксимирующую следующий отсчет:

$$\hat{S}(n) = \sum_{i=1}^p a_i \cdot S(n-i) \quad (1)$$

Для этого сигнал речи сегментируют на блоки, обрабатывают окном и для каждого блока вычисляют автокорреляционную функцию (АКФ).

Оказывается, что такая АКФ содержит исчерпывающую информацию о формантном спектре речи на данном сегменте. На втором шаге находят решение системы линейных уравнений относительно коэффициентов предсказания - тех самых, что нужны для формирования синтезирующего фильтра. Фактически найденные коэффициенты задают спектральную модель голосового аппарата человека, и чем выше порядок ЛП ( $P$  в формуле 1), тем точнее модель.

Пропустив исходную речь через фильтр, обратный полученному, мы получаем функцию, близкую к сигналу возбуждения. В ее спектре формантные области сглажены, а спектральный характер напоминает белый шум. Поэтому обратный фильтр также называют отбеливающим.

На третьем шаге параметры ЛП квантуются, после чего передаются на приемную сторону. А ошибка предсказания (разностный сигнал),

$$R(n) = S(n) - \hat{S}(n) \quad (2)$$

обозначенный в формуле (2) через  $R(n)$ , обрабатывается с целью выделения существенных для человеческого восприятия параметров. Например, в простейших кодеках ЛП передается период  $OT$  и мощность возбуждающих импульсов. В кодеке RPE-LTP таких импульсов уже несколько, а в кодеках CELP (code-excited linear prediction - ЛП с кодовым возбуждением) вместо передачи

самых импульсов подбирается наиболее подходящая запись из набора хранящихся в заранее сформированной кодовой книге. Тогда, вместо самой последовательности можно на прием передать ее порядковый номер в кодовой книге. Однако существуют и другие подходы.

В любом случае по каналу связи вместо самой речи передают так или иначе выделенные и квантованные параметры предсказания, интервал и усиление ОТ, параметры возбуждения. В декодере ЛП по принятым параметрам восстанавливают сигнал возбуждения, пропускают его через синтезирующий фильтр и так восстанавливают речь.

### Порядок выполнения лабораторной работы

Загрузить в массив последовательность отсчетов речевого сигнала с частотой дискретизации 8 кГц т.е. для канала ТЧ:

`A := READPRN("s1-11.txt")`

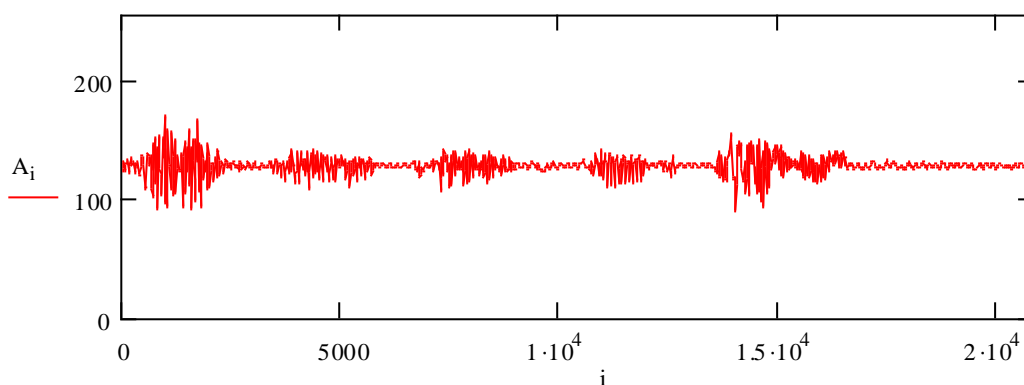


Рис.1. Диаграмма последовательности отсчетов РС

Разбить эту последовательность на одинаковые интервалы по 16 мс или по 128 отсчетов, предполагая, что на этих интервалах процесс достаточно стационарный.

$$j := 0.. \text{floor}\left(\frac{\text{length}(A)}{128}\right) - 1 \quad i := 0.. 127 \quad M_{i,j} := A_{i+128 \cdot j} - 128$$

Получить следующую матрицу отсчетов:

$M =$

	0	1	2	3	4	5	6	7	8	9	10
0	-4	2	-1	-5	7	12	-3	-18	-39	0	23
1	-4	0	0	-1	3	7	-17	-25	-39	-4	19
2	-4	-1	1	2	-3	3	-24	-29	-36	-4	16
3	-4	-2	3	5	-8	1	-22	-29	-28	-3	16
4	-3	-4	5	7	-10	-1	-14	-18	-14	-5	16
5	-3	-4	6	7	-11	1	-5	3	8	-9	15
6	-3	-5	7	7	-14	5	5	22	33	-13	12
7	-3	-5	8	6	-16	9	12	31	45	-17	9
8	-3	-6	10	5	-17	12	17	32	39	-22	7

В данном случае будем использовать 4 порядок модели линейного предсказания

$$j := 0.. \text{cols}(M) - 1 \quad k := 0, 1.. 4$$

Соответственно рассчитаем 5 коэффициентов корреляции, необходимых для нахождения коэффициентов линейного предсказания 4 - го порядка

$$R_{k,j} := \sum_{i=0}^{127-k} M_{i+k,j} \cdot M_{i,j}$$

$$R =$$

	0	1	2	3	4	5
0	$2.495 \cdot 10^3$	$3.254 \cdot 10^3$	$4.221 \cdot 10^3$	$8.924 \cdot 10^3$	$2.502 \cdot 10^4$	$3.773 \cdot 10^4$
1	$2.451 \cdot 10^3$	$3.128 \cdot 10^3$	$4.085 \cdot 10^3$	$8.425 \cdot 10^3$	$2.327 \cdot 10^4$	$3.485 \cdot 10^4$
2	$2.377 \cdot 10^3$	$2.839 \cdot 10^3$	$3.824 \cdot 10^3$	$7.216 \cdot 10^3$	$1.893 \cdot 10^4$	$2.741 \cdot 10^4$
3	$2.286 \cdot 10^3$	$2.455 \cdot 10^3$	$3.467 \cdot 10^3$	$5.548 \cdot 10^3$	$1.321 \cdot 10^4$	$1.743 \cdot 10^4$
4	$2.181 \cdot 10^3$	$2.039 \cdot 10^3$	$3.061 \cdot 10^3$	$3.668 \cdot 10^3$	$7.152 \cdot 10^3$	$6.924 \cdot 10^3$

Исходя из матрицы, найдем коэффициенты линейного предсказания:

$$a^{(j)} := \text{lsolve} \left[ \begin{pmatrix} R_{0,j} & R_{1,j} & R_{2,j} & R_{3,j} \\ R_{1,j} & R_{0,j} & R_{1,j} & R_{2,j} \\ R_{2,j} & R_{1,j} & R_{0,j} & R_{1,j} \\ R_{3,j} & R_{2,j} & R_{1,j} & R_{0,j} \end{pmatrix}, \begin{pmatrix} R_{1,j} \\ R_{2,j} \\ R_{3,j} \\ R_{4,j} \end{pmatrix} \right]$$

$$a =$$

	0	1	2	3	4	5
0	1.28	1.6	1.317	1.591	1.758	1.947
1	-0.232	-0.597	-0.15	-0.613	-0.99	-1.366
2	0.026	-0.136	-0.206	-0.075	0.136	0.37
3	-0.103	0.071	-0.021	-0.011	-0.019	-0.065

$$n := 0, 1.. 3 \quad M1_{n,j} := M_{n,j}$$

Для восстановления сигнала на приемной стороне необходимо передать первые 4 отсчета из каждой исходной выборки речевого сигнала:

$$M1 =$$

	0	1	2	3	4	5	6	7	8	9	10
0	-4	2	-1	-5	7	12	-3	-18	-39	0	23
1	-4	0	0	-1	3	7	-17	-25	-39	-4	19
2	-4	-1	1	2	-3	3	-24	-29	-36	-4	16
3	-4	-2	3	5	-8	1	-22	-29	-28	-3	16

Произведем анализ моделирования для случая, когда восстановление отсчетов речевого сигнала производится только исходя из первых четырех переданных отсчетов и четырех коэффициентов, т.е. без передачи невязок

$i := 4, 5.. 127$

$M1_{i,j} := \text{round}(a_{0,j} \cdot M1_{i-1,j} + a_{1,j} \cdot M1_{i-2,j} + a_{2,j} \cdot M1_{i-3,j} + a_{3,j} \cdot M1_{i-4,j})$

Отсчеты восстановленного сигнала будут выглядеть следующим образом:

	0	1	2	3	4	5	6	7	8
0	-4	2	-1	-5	7	12	-3	-18	-39
1	-4	0	0	-1	3	7	-17	-25	-39
2	-4	-1	1	2	-3	3	-24	-29	-36
3	-4	-2	3	5	-8	1	-22	-29	-28
4	-4	-2	4	7	-11	0	-15	-24	-16
5	-4	-2	5	8	-12	-1	-6	-16	-3
6	-4	-2	5	8	-11	-2	2	-7	8
7	-4	-2	5	7	-9	-3	8	1	15
8	-4	-2	5	6	-6	-3	11	7	17

$j := 0.. \text{cols}(M1) - 1$

$i := 0.. \text{rows}(M1) - 1$

$M1\_new_{i,j} := \text{round}(M1_{i,j})$

$Anew_{i+128,j} := \text{round}(M1_{i,j} + 128)$

$\text{WRITEPRN}("s\_out.txt") := Anew$

В результате ошибка предсказания будет выглядеть следующим образом:

$ii := 0, 1.. \text{length}(Anew) - 1$

$An_{ii} := Anew_{ii} - A_i$

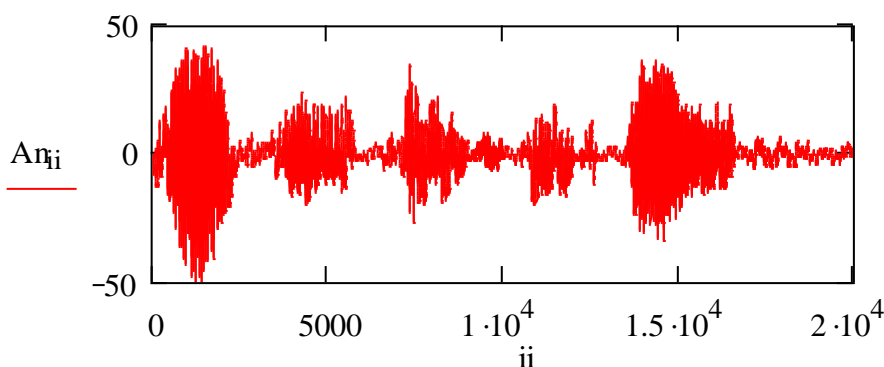


Рис.2. Диаграмма ошибки предсказания без передачи невязок.

Как видно из полученного графика, а также при прослушивании полученного восстановленного сигнала, что ошибки предсказания очень большие и речь практически неразборчива, что означает то, что необходимо передавать невязки.

Производим восстановление сигнала с учетом передачи на приемную сторону невязок размерностью в 1 бит, т.е. будем передавать только знаки невязок

$$i := 4, 5.. 127$$

Зададим знаковую функцию в виде:

$\text{sign}(x) := \text{if}(x < 0, -1, 1)$  т.е она будет принимать только два значения - 1 бит.

$$M1_{i,j} := \text{round}(a_{0,j} \cdot M1_{i-1,j} + a_{1,j} \cdot M1_{i-2,j} + a_{2,j} \cdot M1_{i-3,j} + a_{3,j} \cdot M1_{i-4,j}) + \text{sign}[M1_{i,j} - (a_{0,j} \cdot M1_{i-1,j} + a_{1,j} \cdot M1_{i-2,j} + a_{2,j} \cdot M1_{i-3,j} + a_{3,j} \cdot M1_{i-4,j})]$$

Отсчеты восстановленного сигнала будут выглядеть следующим образом:

	0	1	2	3	4	5	6	7	8
0	-4	2	-1	-5	7	12	-3	-18	-39
1	-4	0	0	-1	3	7	-17	-25	-39
2	-4	-1	1	2	-3	3	-24	-29	-36
3	-4	-2	3	5	-8	1	-22	-29	-28
4	-3	-3	5	8	-10	-1	-14	-23	-15
5	-4	-4	7	9	-11	-2	-5	-13	-1
6	-3	-5	7	8	-11	-1	4	-2	11
7	-4	-4	8	6	-11	1	11	8	19
8	-3	-4	9	5	-11	4	15	15	22

$$j := 0.. \text{cols}(M1) - 1$$

$$i := 0.. \text{rows}(M1) - 1$$

$$M1_{\text{new}i,j} := \text{round}(M1_{i,j})$$

$$A_{\text{new}i+128,j} := \text{round}(M1_{i,j} + 128)$$

WRITEPRN("s\_out.txt") := Anew

В результате ошибка предсказания будет выглядеть следующим образом:

$$ii := 0, 1.. \text{length}(A_{\text{new}}) - 1$$

$$A_{\text{н}ii} := A_{\text{new}ii} - A_i$$

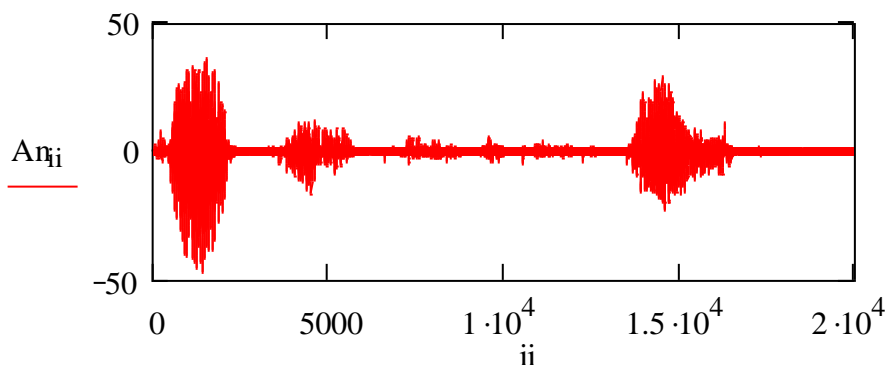


Рис.3. Диаграмма ошибки предсказания с передачей невязок.

В результате получили на приемной стороне речевой сигнал с достаточно разборчивой речью и неплохим качеством. Улучшение сигнала по отношению к предыдущему исследованию также видно из графика ошибок предсказания.

В данной работе произвели исследование модели линейного предсказания на ЭВМ.

В результате исследование показало, что самым оптимальным решением при передаче речевого сигнала с использованием метода линейного предсказания является: передача сигнала с частотой дискретизации 8 кГц, длительность выборки и количество отсчетов – примерно 128 отсчетов или 16 мс, а также количество коэффициентов (порядок модели) равное 4. Улучшение качества от увеличения количества коэффициентов тоже не осуществлялось.

Коэффициент сжатия при этом получается:

При сжатии на выборке 128 байт:

4 байта по 8бит – на коэффициенты

4 байта по 8бит – на начальные отсчеты

128 бит – 4 бит = 124 бит на передачу невязок.

В результате при сжатии передается:  $4*8 + 4*8 + 124 = 188$  бит

Без сжатия:  $128*8 = 1024$  бит

Коэффициент сжатия равен:  $1024/188 = 5,44$

Битовая скорость при этом получается  $64 \text{ кбит/с} / 5,44 = 11,7 \text{ кбит/с}$

Необходимо выполнить все необходимые исследования, используя порядок модели линейного предсказания от 3 до 6, разрядность передаваемых невязок от 1 до 4 бит. Оценить при этом качество восстановленного речевого сигнала, а также коэффициент сжатия и битовую скорость.

### **Содержание отчета**

Отчет по проделанной лабораторной работе должен быть оформлен в виде презентации и должен содержать:

- основные формулы;
- операторы;
- текст программы;
- демонстрацию звуковых файлов.

### **Контрольные вопросы**

## Лабораторная работа № 6

### *СОЗДАНИЕ И ИССЛЕДОВАНИЕ СПЕКТРОГРАММ РЕЧЕВЫХ СИГНАЛОВ*

**Цель работы:** Создание спектрограмм речевых сигналов.

#### **Порядок выполнения лабораторной работы**

Необходимо создать последовательность амплитудных спектров речевых сигналов в виде изображений – спектрограмм. Каждый отрезок квазистационарности речевого сигнала представляется одной спектрограммой – по одной координате номер отсчета в спектре  $K = 0-63$ , по другой координате номер отрезка. Спектрограммы представляются в виде двумерных изображений.

#### Содержание отчета

Отчет по проделанной лабораторной работе должен быть оформлен в виде презентации и должен содержать:

- основные формулы;
- операторы;
- текст программы;
- демонстрацию звуковых файлов.

#### Контрольные вопросы

## *Лабораторная работа № 7*

### *РАЗРАБОТКА СПОСОБА РАСПОЗНАВАНИЯ РЕЧИ НА ОСНОВЕ СПЕКТРОГРАММ*

**Цель работы:** Разработать и исследовать способ распознавания речи на основе спектрограмм.

#### **Порядок выполнения лабораторной работы**

Нужно разработать метод распознавания речевых сигналов на основе их спектрограмм. Метод позволяет выбрать из имеющихся 6 – 10 речевых фраз одну, которая соответствует наибольшей корреляции с произнесенной фразой. Необходимо учитывать небольшой сдвиг при этом речевого сигнала по сравнению с эталоном. Произвести моделирование и оценить вероятностные характеристики распознавания.

#### Содержание отчета

Отчет по проделанной лабораторной работе должен быть оформлен в виде презентации и должен содержать:

- основные формулы;
- операторы;
- текст программы;
- демонстрацию звуковых файлов.

#### Контрольные вопросы



## ***Лабораторная работа № 8***

### **РАЗРАБОТКА СПОСОБА РАСПОЗНАВАНИЯ РЕЧИ НА ОСНОВЕ МОДЕЛИ ЛИНЕЙНОГО ПРЕДСКАЗАНИЯ**

**Цель работы:** Разработать и исследовать способ распознавания речи на основе модели линейного предсказания.

#### **Порядок выполнения лабораторной работы**

Нужно разработать метод распознавания речевых сигналов на основе характеристик метода их линейного предсказания. Метод позволяет выбрать из имеющихся 6 – 10 речевых фраз одну, которая соответствует наибольшей корреляции с произнесенной фразой. Необходимо учитывать небольшой сдвиг при этом речевого сигнала по сравнению с эталоном. Произвести моделирование и оценить вероятностные характеристики распознавания.

#### **Содержание отчета**

Отчет по проделанной лабораторной работе должен быть оформлен в виде презентации и должен содержать:

- основные формулы;
- операторы;
- текст программы;
- демонстрацию звуковых файлов.

#### **Контрольные вопросы**

## ***Лабораторная работа № 9***

### **ИССЛЕДОВАНИЕ ЗВУКОВЫХ СИГНАЛОВ МУЗЫКАЛЬНЫХ ИНСТРУМЕНТОВ**

**Цель работы:** Исследовать параметры одноголосых звуковых сигналов музыкальных инструментов.

#### **Задание к лабораторной работе**

Исследовать параметры звуковых сигналов музыкальных инструментов, применяя к ним различные модели.

По результатам моделирования сравнить эти модели по их адекватности для музыкальных звуков.

#### **Порядок выполнения работы**

Необходимо создать или синтезировать звуковые файлы

Их параметры:

Длительность – 1..2 сек.

Частота дискретизации - 8 кГц.

Разрядность - 8 (1 байт).

Моно.

Файл должен быть записан в текстовом формате. Любым текстовым редактором у него убирается заголовок.

Исследовать параметры этих сигналов, применяя к ним формантную модель и модель линейного предсказания.

Сравнить эти две модели по их адекватности для музыкальных звуков.

#### **Содержание отчёта**

Отчет по проделанной лабораторной работе должен быть оформлен в виде презентации и должен содержать:

- основные формулы;
- операторы;
- текст программы;
- демонстрацию звуковых файлов.

#### **Контрольные вопросы**

## Темы курсовых работ

1. Метод распознавания речи на основе формантной модели (разработка, моделирование и исследование).
2. Метод синтеза речи на основе формантной модели (разработка, моделирование и исследование).
3. Метод распознавания речи на основе модели линейного прогнозирования (разработка, моделирование и исследование).
4. Метод синтеза речи на основе модели линейного прогнозирования (разработка, моделирование и исследование).
5. Метод синтеза музыкальных звуков различных инструментов (разработка, моделирование и исследование).
6. Метод распознавания музыкальных звуков различных инструментов (разработка, моделирование и исследование).
7. Метод создания ключа (пароля) речевым сигналом (разработка, моделирование и исследование).

## *БИБЛИОГРАФИЧЕСКИЙ СПИСОК*

### **Основная литература:**

1. Сергиенко, Александр Борисович. Цифровая обработка сигналов: учебное пособие для студ. вузов, обучающихся по направлению 210300 "Радиотехника" / Сергиенко А. Б. - 3-е изд. - Санкт-Петербург: БХВ-Петербург, 2011. - 756 с.: ил.

### **Дополнительная литература:**

1. Сергиенко, Александр Борисович. Цифровая обработка сигналов: учебное пособие для вузов / Сергиенко А. Б. - 2-е изд. - Санкт-Петербург [и др. ]: Питер, 2006. - 750 с.

2. Основы цифровой обработки сигналов: Учебное пособие для студентов, обучающихся по направлению диплом. Спец. 654401 – Телекоммуникации. БХВ – Петербург, 2005, - 753с.

3. Сергиенко А.Б. Цифровая обработка сигналов: Уч. Пособие для вузов. – Спб Питер, 2006, - 750с.

4. Основы цифровой обработки сигналов: Учебное пособие для студентов, обучающихся по направлению диплом. Спец. 654401 – Телекоммуникации. БХВ – Петербург, 2005, - 753с.

---

**Учебное издание**

**КОМПЬЮТЕРНЫЕ ТЕХНОЛОГИИ  
В ОБЛАСТИ ИНФОКОММУНИКАЦИЙ**

Сборник лабораторных работ  
Часть 2 «Исследование речевых сигналов»

Составитель: **Горохин** Валерий Николаевич

Редактор М. В.Теленкова

Подписано в печать 30.03.2009. Формат 60×84/16.  
Усл. печ. л. 5.81. Уч.–изд. л. 1,20. Тираж 50 экз. Заказ .  
Ульяновский государственный технический университет,  
432027, Ульяновск, Сев. Венец, 32.

Типография УлГТУ, 432027, Ульяновск, Сев. Венец, 32.