Министерство образования и науки Российской Федерации

Севастопольский государственный университет

Кафедра ИС

Отчёт

по лабораторной работе №6

дисциплины «Технологии обработки мультимедиа данных»

на тему «Линейный предиктивный анализ речи»

Выполнил:

Ст. гр. ИС/м-21о Дядюшенко С.Е.

Проверил:

Строганов В.А.

Севастополь

2017

1. **Цель работы**

Исследование метода линейного предсказания речи. Определение коэффициентов линейного предсказания.

1. **Постановка задачи**

В ходе выполнения лабораторной работы необходимо:

– Выполнить кратковременный узкополосный и широкополосный анализ вокализованного участка речи.

– Выполнить гомоморфный спектральный анализ этого же участка речи.

– Определить ЛПК-коэффициенты для выбранного участка речи.

– Построить оценку спектра.

– Выполнить сравнение указанных спектров.

– Вычислить ошибку.

– Определить частоту основного тона по графику ошибки.

– По спектру, полученному методом линейного предсказания, определить формантные частоты.

1. **Сценарий MatLab**

В рамках выполнения работы был составлен сценарий MATLAB реализующий построение необходимых сигналов и их спектров, текст которого представлен ниже:

close all;

clear all;

order = 10;

g = 0.01;

filename = 'ee3.wav';

[y, fs] = audioread(filename);

y = y';

k1 = 441;

k2 = 64;

y\_long = y(1 : k1);

y\_short = y(1 : k2);

f\_long = 0 : fs / length(y\_long) : fs / 2 - fs / length(y\_long);

f\_short = 0 : fs / length(y\_short) : fs / 2 - fs / length(y\_short);

s\_long = abs(fft(y\_long));

s\_short = abs(fft(y\_short));

ls\_long = log(s\_long);

ls\_short = log(s\_short);

figure;

plot(f\_long, ls\_long(1 : length(ls\_long) / 2));

grid on;

title('Narrow band log frequency responce');

xlabel('Log F');

ylabel('Amplitude');

figure;

plot(f\_short, ls\_short(1 : length(ls\_short) / 2));

grid on;

title('Wide band log frequency responce');

xlabel('Log F');

ylabel('Amplitude');

s\_y = fft(y);

a\_s\_y = abs(s\_y);

log\_as\_y = log(abs(s\_y));

c\_y = ifft(log\_as\_y);

w = 100;

c\_w = c\_y;

c\_w(w : (length(c\_y) - w)) = 0;

s\_c\_y = fft(c\_y);

s\_c\_w = fft(c\_w);

f = 0 : fs / length(y) : fs / 2 - fs / length(y);

figure

plot(f, s\_c\_w(1 : length(s\_c\_w) / 2));

grid on;

title('Log frequency response');

xlabel('Log F');

ylabel('Amplitude');

a = lpc(y, order);

y\_filtered = filter(g, a, y);

[h, w] = freqz(g, a);

figure;

plot(w \* fs / pi / 2, log(abs(h))');

grid on;

title('Log filter frequency response');

xlabel('Log F');

ylabel('Amplitude');

figure;

plot(f\_long, ls\_long(1 : length(ls\_long) / 2), 'b');

hold on;

plot(f\_short, ls\_short(1 : length(ls\_short) / 2), 'r');

plot(f, s\_c\_w(1 : length(s\_c\_w) / 2), 'g');

plot(w \* fs / pi / 2, log(abs(h)), 'y');

grid on;

title('Log frequency response');

xlabel('Log F');

ylabel('Amplitude');

legend('Narrow band frequency response', 'Wide band frequency response', 'Circumflex of frequency responce', 'Filter frequency response');

t = 0 : 1 / fs : (length(y) - 1) / fs;

figure;

plot(t, y, 'b');

hold on;

plot(t, y\_filtered, 'g');

grid on;

title('Real and predicted signals');

xlabel('Time, sec');

ylabel('Amplitude');

legend('Real signal', 'Predicted signal')

e = y - y\_filtered;

figure;

plot(t, e);

grid on;

title('Difference between real and predicted signals');

xlabel('Time, sec');

ylabel('Amplitude');

1. **Результаты выполнения**

Сигнал вокализованной фонемы был повергнут узкополосному, широкополосному и гомоморфному анализу. Результаты представлены на рисунках 1-3.



Рисунок 1 – Узкополосный спектр вокализованной фонемы



Рисунок 2 – Широкополосный спектр вокализованной фонемы



Рисунок 3 – Результат гомоморфной обработки (огибающая спектра) сигнала

Далее были получены коэффициенты фильтра для линейного предикативного анализа речи. Апмлитудно-частотная характеричтика этого фильтра представлена на рисунке 4.



Рисунок 4 – Амплитудно-частотная характеристика фильтра линейного предикативного анализа речи

На амплитудно-частотной характеристике можно увидеть пики, соответствующие формантным частотам фонемы.

Все полученные амлитудно-частотные характеристики представлены на рисунке 5.



Рисунок 5 –Полученные амплитудно-частотные характеристики

(синий – узкополосный спектр,

красный ­– широкополосный спектр,

зелёный – огибающая спектра, полученная в результате гомоморфного анализа,

жёлтый – амплитудно-частотная характеристики фильтра для линейного предикативного анализа речи)

Далее выполнено построение предсказанного на основе фильтра сигнала. Исходный и предсказанный сигналы представлены на рисунке 6.



Рисунок 6 – Исходный и предсказанный сигналы

(синий – исходный сигнал, зелёный – предсказанный)

Найдена ошибка предсказания. Она представлена на рисунке 7.



Рисунок 7 – Ошибка предсказания

Очевидно, что ошибка предсказания изменяется периодически. Известно также, что она изменяется с частотой основного тона диктора. На основании известных положений соседних пиков была рассчитана частота основного тона:

**Вывод:**

В ходе выполнения данной лабораторной работы исследован метод линейного предсказания речи. Определены коэффициенты линейного предсказания для имеющегося вокализированного сигнала. На его основе построен фильтр. Проведено сравнение его амплитудно-частотной характеристики и спектров, полученных в результате узкополосного, широкополосного и гомоморфного анализов. С помощью полученного фильтра выполнено линейное предсказание сигнала и найдена ошибка. На основе ошибки найдена частота основного тона сигнала.