

ФОРМИРОВАНИЕ ТЕСТОВЫХ СИГНАЛОВ ДЛЯ ДЕТЕКТОРА РЕЧИ

Борисевич Дм. А.

Кафедра защиты информации, Белорусский государственный университет информатики и радиоэлектроники
Минск, Республика Беларусь
E-mail: zmicierb@tut.by

Предложен набор тестовых аудиосигналов для детекторов речи с разделением на речевые и неречевые сигналы. В работе проведен анализ тестовых сигналов с помощью следующих информационных параметров: количество пересечений с нулем, стационарность спектра, значение кепстральных коэффициентов. В работе проведен анализ тестовых сигналов в зависимости от соотношения сигнал/шум.

ВВЕДЕНИЕ

Для эффективности работы детекторов речи необходимо проводить сравнительные испытания. Для этого необходимы тестовые сигналы речевого и неречевого характера. Набор тестовых сигналов должен содержать небольшое число аудиосигналов, которые моделируют реальные условия использования прибора. Сигналы должны быть небольшой длительности и иметь одинаковое качество записи.

I. ТЕСТОВЫЕ СИГНАЛЫ

В работе предложен набор тестовых аудиосигналов с разделением на речевые и неречевые сигналы. Неречевые сигналы включают в себя записи тональных сигналов, высокочастотный шум, стук, хлопki, низкочастотный шум. Аудиосигнал моделирующий тональный сигнал содержит образцы тестовых сигналов, которые различаются по амплитуде ($A_1=1$, $A_2=0,5$, $A_3=1,5$) и частоте ($f_1=440$ Гц, $f_2=100$ Гц, $f_3=850$ Гц). Амплитуда отдельных образцов постоянна, частоты тестовых сигналов лежат в среднечастотной области.

Аудиосигнал моделирующий высокочастотные шумы содержит запись звуков писка. Во временной области сигнал характеризуется короткой длительностью и высокой амплитудой на протяжении большей части времени звучания. В спектральной области сигнал характеризуется небольшим числом всплесков в высокочастотной области.

Аудиосигнал моделирующий стук содержит запись хлопков. Во временной области сигнал характеризуется короткой длительностью, в начале звучания высокое значение амплитуды, а затем идет резкое снижение амплитуды до уровня шума. В спектральной области сигнал характеризуется наличием всплесков в области низких частот, основная спектральная мощность расположена в области низких частот.

Аудиосигнал моделирующий низкочастотный шум содержит запись звуков льющейся воды. Во временной области сигнал характеризуется равномерной амплитудой на всем протяжении

со случайными всплесками амплитуды. В спектральной области сигнал характеризуется равномерным распределением спектра по всей спектральной плоскости.

Речевые сигналы включают в себя запись дикторов и речевой коктейль. Записи дикторов включают в себя образцы мужских, женских и детских голосов. Во временной области сигнала встречаются квазигармонические участки с высокой амплитудой, что соответствует гласным звукам, шумоподобные участки с меньшей амплитудой, что соответствует согласным звукам. В спектральной области сигнал сосредоточен в области низких и средних частот.

II. АНАЛИЗ ТЕСТОВЫХ СИГНАЛОВ

Для каждого тестового сигнала представлены параметры во временной и спектральной области, параметры кепстрального спектра. В работе анализируются основные параметры, используемые в детекторах речи: количество пересечений с нулем, стационарность спектра, значение кепстральных коэффициентов.

Расчеты параметров для тестовых сигналов производились в программном приложении MatLAB. В качестве тестовых сигналов брались аудиозаписи длительностью 4-5 секунд с частотой дискретизацией 22050 Гц. Параметры рассчитывались для кадров, которые содержат 600 отчетов, что составляет 327 мс, а затем рассчитывались средние значения параметров для всего сигнала. Также на тестовые сигналы был наложен «белый» шум с различными уровнями сигнал/шум в 20, 15, 10, 6, 0 дБ. Мощность сигнала оценивалась с помощью автокорреляционной функции. Стационарность спектра оценивалась для трех октавных полос в диапазоне от 88 Гц до 704 Гц. Кепстральные коэффициенты рассчитывались с помощью встроенной функции в MatLAB.

Информационный параметр количество пересечений с нулем - характеристика, которая показывает, сколько раз на временном интервале сигнал пересекает нулевой уровень. С учетом параметров тестовых аудиосигналов, диапазон зна-

чений критерия сигнализирующих о наличии речи равен от 3 до 20. Значения информационного параметра стационарность спектра, сигнализирующие о наличии речевых сигналов, не должны равняться 0 и должны иметь широкий диапазон значений. Информационный параметр значение кепстральных коэффициентов детектирует речь путем оценки корреляционных свойств кепстрального вектора. Большое значение информационного параметра (больше 50) является признаком наличия речевого сигнала. Значения информационных параметров полученные для тестовых аудиосигналов представлены в таблицах.

III. СПИСОК ЛИТЕРАТУРЫ

1. Патент US № 6,061,647 Voice activity detector.
2. Патент US № 5,649,055 Voice activity detector for speech signals in variable background noise.
3. Чистович Л.А., Венцов А.В., Гранстрем М.П. и др. Физиология речи. Восприятие речи человеком. В серии "Руководство по физиологии"/ Л., Наука, 1976. 388 с.
4. Патент US № 6,556,967 Voice activity detector.
5. Патент US № 7,191,128 Method and system for distinguishing speech from music in digital audio signal in real time.

Таблица 1 – Значения информационных параметров для тестового сигнала «Тональные сигналы»

Уровень сигнал/шум, дБ	Число пересечений с нулем	Стационарность спектра	Кепстральные коэффициенты
Без помех	10 – 32	1.31 – 131.11	0 – 9
20	5 – 36	6.52 – 123.46	2.01 – 29.21
15	1 – 100	9.18 – 116.69	2.7 – 31.44
10	10 – 154	21.17 – 101.84	4.37 – 27.37
6	40 – 159	24.03 – 99.77	3.85 – 28.92
0	84 – 152	42.12 – 79.62	4.13 – 34.91

Таблица 2 – Значения информационных параметров для тестового сигнала «Высокочастотный шум»

Уровень сигнал/шум, дБ	Число пересечений с нулем	Стационарность спектра	Кепстральные коэффициенты
Без помех	30 - 91	1.8 – 11.5	2 – 16
20	31 – 120	2.92 – 11.07	2.27 – 19.57
15	33 – 120	4.55 – 11.44	2.3 – 21.89
10	30 – 124	6.15 – 13.12	2.67 – 17.32
6	22 – 117	9.11 - 15.73	3.17 – 28.4
0	35 – 118	11.2 – 23.98	3.23 – 21.7

Таблица 3 – Значения информационных параметров для тестового сигнала «Стук»

Уровень сигнал/шум, дБ	Число пересечений с нулем	Стационарность спектра	Кепстральные коэффициенты
Без помех	11 – 89	6.17 – 77.51	4 – 25
20	11 – 111	6.01 – 76.32	3.32 – 40.02
15	13 – 121	7.94 – 75.06	4.33 – 27.6
10	19 – 123	9.63 – 74.64	4.58 – 30.49
6	27 – 118	9.37 – 71.26	4.35 – 28.07
0	56 – 122	16.24 – 71.27	3.52 – 29.4

Таблица 4 – Значения информационных параметров для тестового сигнала «Низкочастотный шум»

Уровень сигнал/шум, дБ	Число пересечений с нулем	Стационарность спектра	Кепстральные коэффициенты
Без помех	45 – 111	0.52 – 0.73	2 – 21
20	68 – 114	0.51 – 0.73	3.63 – 19.26
15	71 – 114	0.52 – 0.75	3.34 – 26.44
10	72 – 108	0.48 – 0.76	3.74 – 27.2
6	85 – 110	0.62 – 0.73	3.6 – 22.42
0	93 – 116	0.57 – 0.92	3.6 – 25.78

Таблица 5 – Значения информационных параметров для тестового сигнала «Речь»

Уровень сигнал/шум, дБ	Число пересечений с нулем	Стационарность спектра	Кепстральные коэффициенты
Без помех	2 – 126	0.85 – 87.03	1 – 52
20	3 – 123	2.08 – 86.31	1.93 – 27.38
15	8 – 134	2.44 – 87.05	2.73 – 33.44
10	8 – 145	3.81 – 86.44	3.52 – 29.75
6	11 – 140	5.68 – 81.01	4.99 – 35.58
0	35 – 144	15.92 – 85.37	4.77 – 31.02