

УДК 534.87

В.Е. Гай

## АЛГОРИТМ КЛАССИФИКАЦИИ СИГНАЛОВ

Нижегородский государственный технический университет им. Р. Е. Алексеева

Предлагается алгоритм классификации сигналов, принимаемых технической системой, управляемой голосом, на шумовые / не шумовые. Разработанный алгоритм основан на использовании теории активного восприятия, адаптированной к анализу одномерных сигналов. Результаты проведенных экспериментов подтверждают возможность использования предложенного алгоритма для решения поставленной задачи.

*Ключевые слова:* цифровая обработка сигналов, теория активного восприятия, анализ сигналов, классификация цифровых сигналов.

### 1. Постановка задачи

В состав различных технических систем часто входит подсистема голосового управления. Одна из основных задач данной подсистемы – классификация принимаемых сигналов на шумовые и не шумовые. Цель классификации – принятие решения о целесообразности проведения дальнейшего анализа сигнала и выделения из сигнала некоторой команды. Предполагается, что подобную задачу решает человек в процессе информационного взаимодействия с окружающим миром. Достоверно известно, что новорожденные умеют воспринимать звуки и различают их по высоте. Особую чувствительность они проявляют к звукам человеческой речи и предпочитают осмысленную, связную речь произвольному набору слогов. При этом, младенцы способны различать некоторые признаки, не используемые в конкретном языке. Эти утверждения подтверждаются уменьшением  $\alpha$ -ритма на энцефалограмме при воспроизведении речевых сигналов.

Таким образом, предполагается наличие у человека встроенных механизмов опознавания не шумовых сигналов.

В данной работе решается задачи выбора признаков для классификации сигналов. Классификация сигналов на шумовые / не шумовые выполняется на основе теории активного восприятия [1, 2].

Алгоритм классификации должен решать задачу классификации сигнала в реальном масштабе времени. Выбор длительности анализируемого участка сигнала определяется, исходя из условий задачи и используемой системы команд.

На рис. 1 показана структура системы распознавания команд, входящей в состав технической системы.

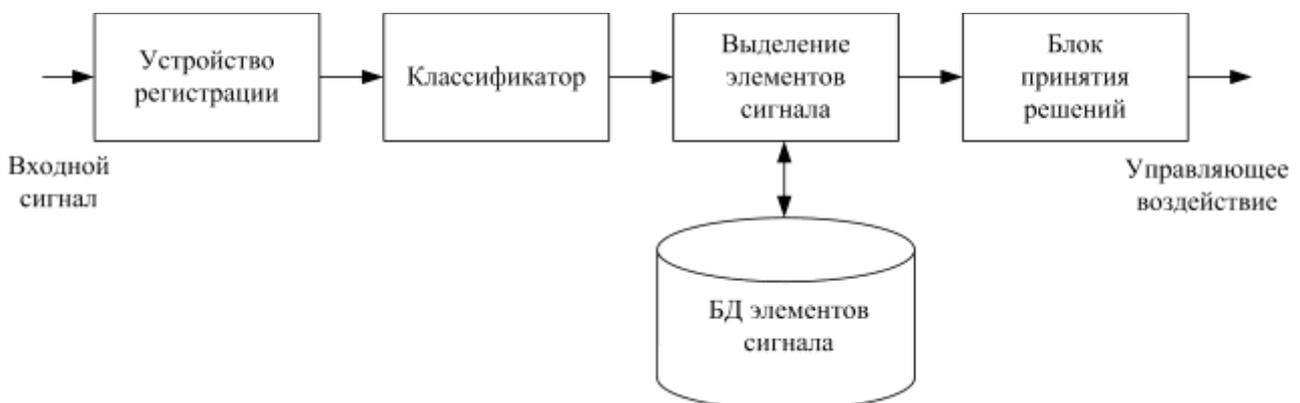


Рис. 1. Система распознавания команд

Использование блока классификации позволяет снизить вычислительную сложность алгоритма анализа сигнала за счёт удаления шумовых фрагментов принимаемого сигнала.

## 2. Предлагаемый алгоритм

Предлагаемый алгоритм классификации сигналов состоит в следующем:

1. Построение спектрального представления исходного сигнала  $f(t)$  с использованием  $U$ -преобразования для одномерных сигналов [1, 2]:

$$S = U(f(t)),$$

где  $S$  – набор из  $M$  мгновенных спектров, где  $M$  – число сегментов сигнала. Каждый спектр представляет собой набор коэффициентов  $S_i = \{s_1, \dots, s_N\}$ ,  $N$  – число фильтров Уолша системы Хармута [3, 4], используемых при формировании спектра (рис. 2). Спектр вычисляется по отрезку сигнала длиной  $len$ .

В слуховой системе человека вычисление мгновенных спектров выполняет основная мембрана [5]. Обычно функциональную модель мембраны представляют в виде набора связанных фильтров. Причём полоса пропускания каждого фильтра изменяется вдоль длины мембраны. Полоса пропускания максимальна в высокочастотной области и минимальна в низкочастотной. Распределение ответов фильтров представляет собой мгновенный спектр сигнала [6].

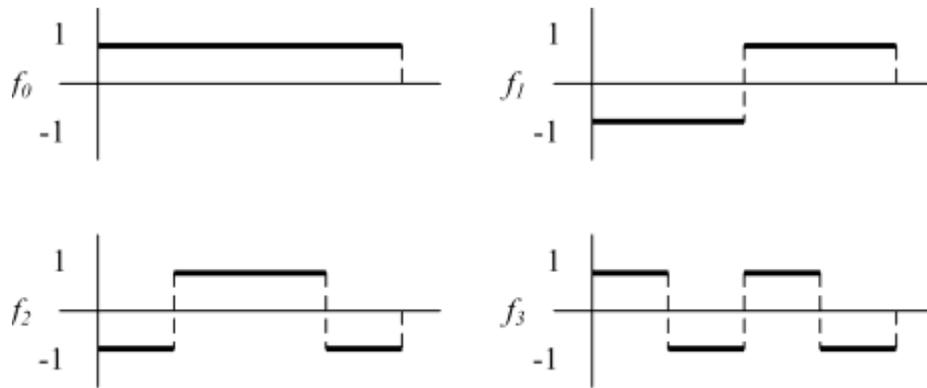


Рис. 2. Фильтры Уолша системы Хармута

2. Поиск похожих цепочек спектров:

а) из  $S$  выбираются ранее неиспользованные цепочки спектров  $L_i$  и  $L_j$  ( $i \neq j$ )  $L_i = \{S_i, \dots, S_{i+c_{len}}\}$ , где  $S_i$  –  $i$ -й спектр набора  $S$ ,  $c_{len}$  – длина цепочки спектров;

б) если абсолютная величина разности между спектрами, входящими в  $L_i$  и  $L_j$ , меньше некоторой пороговой величины ( $thr$ ):

$$\sum_{k=1}^{c_{len}} |L_{ik} - L_{jk}| < thr,$$

считается, что эти две цепочки похожи, после чего цепочка  $L_j$  отмечается как использованная (повторяющаяся) и в дальнейших сравнениях участия не принимает;

в) подсчитывается количество повторяющихся ( $sp_{use}$ ) цепочек спектров сигнала;

3. На основе полученного значения  $sp_{use}$  делается вывод о типе анализируемого сигнала.

## 3. Вычислительный эксперимент

Реализация предложенных алгоритмов выполнена в системе Octave. Тестирование проводилось на ЭВМ с процессором AMD Turion 2 Dual-Core Mobile M500, 2.20 ГГц, 4 Гб оперативной памяти.

При исследовании алгоритма использовались следующие тестовые сигналы (частота дискретизации 16 кГц, глубина кодирования – 16 бит):

1) записи речи 15 дикторов;

- 2) запись трелей японского соловья (как известно [7], птицы используют звуки для обмена информацией);
- 3) музыкальная запись;
- 4) некоррелированный шум с равномерным и нормальным распределением;
- 5) коррелированный шум;
- 6) периодический сигнал (синус с частотой 100 Гц, синус с частотой 50 Гц, амплитудно- и частотно-модулированный синус);
- 7) постоянный во времени сигнал (логическая единица);
- 8) последовательность прямоугольных импульсов.

Значения параметров алгоритма, используемые при проведении эксперимента:

- 1) длина цепочки спектров  $c_{len}$ : 1, 2, 3, 4, 5, 6;
- 2) порог, используемый при сравнении цепочек спектров  $thr$ : 0.1, 0.3, 0.5, 0.7, 1.0;
- 3) длина анализируемого сегмента  $len$  (в отсчётах): 32, 64, 128, 256, в миллисекундах: 2, 4, 8, 16;
- 4) смещение, используемое при выделении сегментов  $shft$ : 16, 32, 64, 128, 256;
- 5) количество используемых фильтров  $N$ : 4, 8;
- 6) размер обрабатываемого сигнала (в отсчётах)  $sig_{len}$ : 50000, 100000, 150000, 200000, 250000 (в секундах: 3.125, 6.25, 9.375, 12.5, 15.625, частота дискретизации 16 кГц).

В табл. 2 и табл. 3 приведены результаты обработки речевого сигнала. Каждая ячейка таблицы содержит количество найденных неповторяющихся / повторяющихся цепочек в сигнале ( $sp_{nuse} / sp_{use}$ ).

Таблица 1

## Результаты работы алгоритма

Параметры алгоритма: $c_{len} = 1, N = 4, thr = 0.1$						
		Длина сигнала в отсчётах				
$len$	$shft$	50000	100000	150000	200000	250000
32	16	1790/1332	3474/2774	4449/4923	5513/6985	6868/8754
32	32	955/606	1898/1266	2504/2182	3089/3160	3820/3991
64	32	1218/342	2454/668	3332/1352	4142/2106	5030/2780
64	64	633/147	1277/284	1736/606	2164/960	2616/1289
128	64	695/83	1403/157	1948/392	2437/685	2934/970
128	128	362/27	732/48	1024/146	1286/275	1548/404
256	128	377/11	757/21	1098/70	1427/133	1752/198
256	256	192/2	384/5	568/16	745/35	918/57

Таблица 2

## Результаты работы алгоритма

Параметры алгоритма: $c_{len} = 1, N = 8, thr = 0.1$						
		Длина сигнала в отсчётах				
$len$	$shft$	50000	100000	150000	200000	250000
32	16	2446/676	5133/1115	7224/2184	8939/3559	11108/4514
32	32	1265/296	2637/487	3737/949	4610/1639	5719/2092
64	32	1519/41	3050/72	4359/325	5482/766	6781/1029
64	64	772/0	1547/14	2218/123	2795/329	3456/449
128	64	778/0	1560/0	2339/1	3109/13	3879/25
128	128	389/0	780/0	1170/0	1559/2	1945/7
256	128	388/0	778/0	1168/0	1560/0	1950/0
256	256	194/0	389/0	584/0	780/0	975/0

На рис. 3 показаны графики изменения числа уникальных спектров в зависимости от длины сигнала (сплошной линией показан график для равномерного шума, пунктирной – для трелей соловья). При вычислении приведённых значений использовались следующие параметры предложенного алгоритма:  $c_{len} = 1$ ,  $N = 4$ ,  $thr = 0.3$ ,  $len = 32$ ,  $shft = 32$ .

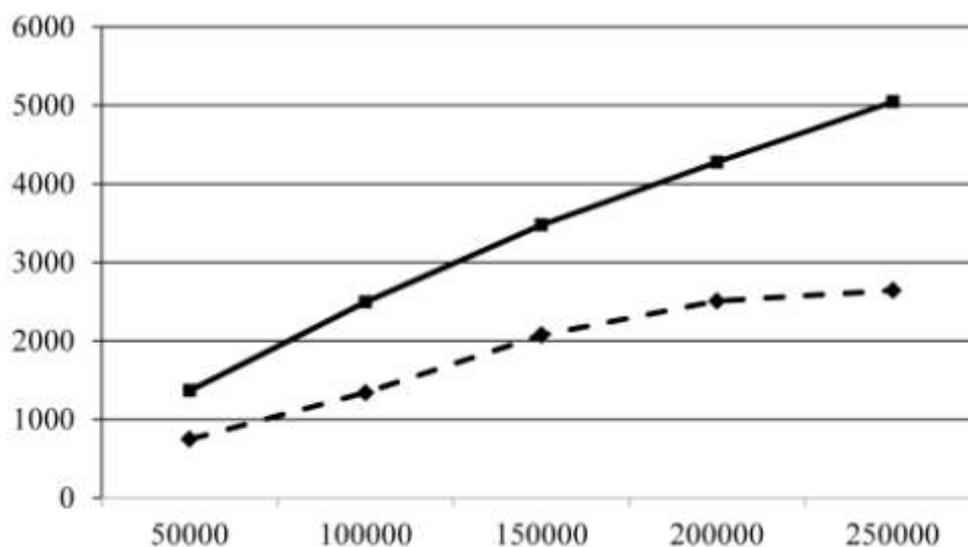


Рис. 3. Зависимость количества уникальных спектров от длины сигнала

Результаты исследования:

1. Записи голосов дикторов, трелей соловья, музыки содержат повторяющиеся цепочки спектров ( $c_{len} \geq 1$ ). Установлено, что для указанных записей с увеличением длительности сигнала возрастает количество уникальных цепочек спектров.

В некоторых случаях (например, при  $N = 8$ ,  $len = 256$ ,  $thr \geq 0.1$ ) количество найденных повторяющихся цепочек близко к нулю. Это объясняется тем, что использование указанных параметров алгоритма приводит к введению достаточно строгих ограничений на подобие цепочек (сравните табл. 1 и табл. 2).

2. Шумовые записи содержат повторяющиеся цепочки спектров при  $c_{len} = 1$ . При  $c_{len} > 1$  повторяющиеся цепочки спектров в шумовых сигналах не обнаружены. Таким образом, при введении более жёстких ограничений на подобие цепочек (например, увеличение числа используемых фильтров) количество повторяющихся цепочек в шумовом сигнале резко уменьшается. Это является характерной особенностью всех исследованных шумовых сигналов.

3. Периодические сигналы включают повторяющиеся цепочки спектров. Количество повторяющихся цепочек с увеличением длительности сигнала остаётся равным некоторой постоянной величине (для определённого набора параметров). Например, синус (частота = 50 Гц, параметры алгоритма:  $c_{len} = 1$ ,  $len = 32$ ,  $shft = 32$ ,  $N = 4$ ,  $thr = 0.1$ ) описывается 10 уникальными цепочками. Периодический сигнал не относится к шумовым, так как несёт информацию о частоте.

4. Постоянный сигнал при любых параметрах алгоритма и любой длине сигнала описывается 1 спектром ( $sp_{noise} = 1$ ). Задача поиска постоянных сигналов может возникать при выделении пауз в речи.

5. Минимальная длина сигнала, необходимая для определения его типа, составляет 0.5 секунды.

6. Длина обрабатываемого сегмента выбрана равной 32 отсчётам, так как слуховая система оценивает изменения в звуковом импульсе, происходящие на интервале времени

меньше 1 мс [5], а при частоте дискретизации 16 кГц 2 мс соответствуют 32 отсчётам. Уменьшать длину анализируемого сегмента непрактично с вычислительной точки зрения, при  $len \gg 32$  определить тип сигнала затруднительно в связи с «грубостью» вычислений.

7. Если требуется определить, является ли не шумовой сигнал периодическим, необходимо выполнить дополнительное исследование: увеличивая длину анализируемого сигнала, зафиксируем количество неповторяющихся спектров ( $sp_{nuse}$ ). Если сигнал периодический, то величина  $sp_{nuse}$  с ростом длины сигнала может изменяться в определённых пределах, но не увеличиваться.

8. При проведении исследований использовано избыточное число значений параметров алгоритма. В результате оказалось, что для определения типа сигнала достаточно использовать значения параметров, приведённые в табл. 3.

9. Время работы алгоритма зависит от длительности обрабатываемого сигнала. Время обработки сигнала длиной 3.125 секунды (50000 отсчётов при частоте дискретизации 16000 кГц) составляет 1 секунду.

Таблица 3

### Параметры алгоритма

№	$len$	$shft$	$N$	$c_{len}$	$thr$	$sig_{len}$
1	32	32	8	1	0.1	50000
2	32	32	4	3	0.1	50000
3	32	32	8	3	0.1	50000

По результатам проведённых экспериментов сформулировано решающее правило для определения типа сигнала: если значение  $sp_{use}$ , вычисленное на основе предложенного алгоритма, для каждого набора параметров, приведённых в табл. 3, больше 10 ( $sp_{use} > 10$ ), то анализируемый сигнал не является шумовым, в противном случае – шумовым (значение 10 подобрано экспериментальным путём).

Таким образом, к классу не шумовых сигналов можно отнести речевой сигнал, музыкальную запись, периодический (в том числе периодическая последовательность прямоугольных импульсов) и постоянный сигналы. Достаточно проблематично различить музыкальную запись и речевой сигнал.

### Вывод

Настоящая работа является попыткой применения теории активного восприятия к решению задачи классификации сигналов на шумовые / не шумовые. Предложенный алгоритм был реализован в системе речевого управления мобильным роботом. Также разработанный алгоритм может найти применение при выделении пауз в речевом сигнале.

Необходимо отметить следующие достоинства предложенного алгоритма: низкая вычислительная сложность, отсутствие настраиваемых параметров, простота реализации.

Дальнейшее направление исследований – оценка с помощью теории активного восприятия влияния шума на сигнал (в том числе и речевой).

### Библиографический список

1. **Утробин, В. А.** Информационные модели системы зрительного восприятия для задач компьютерной обработки изображений: учеб. пособие / В. А. Утробин; НГТУ. – Н. Новгород, 2001. – 234 с.
2. **Утробин, В. А.** Компьютерная обработка изображений. Принятие решений в пространстве эталонов: учеб. пособие / В. А. Утробин; НГТУ. – Н. Новгород, 2004. – 221 с.

3. **Хармут, Х.** Применение методов теории информации в физике: [пер. с англ.] / Х. Хармут. – М.: Мир, 1989. – 344 с.
4. **Хармут, Х.Ф.** Передача информации ортогональными функциями: [пер. с англ. Н. Г. Дядюнова, А. И. Сенина] / Х. Ф. Хармут. – М.: Связь, 1975. – 272 с.
5. Элементы теории биологических анализаторов / под ред. Н. В. Позина. – М.: Наука. Главная редакция физико-математической литературы, 1978. – 360 с.
6. **Харкевич, А. А.** Спектры и анализ / А.А. Харкевич. – М.: Физматгиз, 1952. – 192 с.
7. **Мальчевский, А.С.** Роль звуковой индукции в голосовом поведении птиц / А. С. Мальчевский, А.Б. Пукинский // Русский орнитологический журнал. 2009. Т. 18. Экспресс-выпуск 483. С. 779–801.

*Дата поступления  
в редакцию 28.01.2011*

**V. E. Gai**

### **ALGORITHM OF SIGNAL CLASSIFICATION**

The algorithm of classification of signals accepted by the technical system on noised / not noised is offered. The developed algorithm is based on use of the theory of the active perception adapted for the analysis of one dimensional signals. Results of the spent experiments confirm possibility of use of the offered algorithm for solving the problem.

*Key words:* digital signal processing, theory of active perception, signal analysis, classification of digital signals.