

Б.Ж. Медетов, А.Ж. Наурзбаева, А.Б. Манапбаева

Казахский Национальный университет им. Аль-Фараби, г. Алматы, Казахстан

МЕТОД ОБРАБОТКИ СИГНАЛОВ «АДРЕСНОЕ СОПОСТАВЛЕНИЕ ЯЧЕЕК» ПРИ РЕШЕНИИ ЗАДАЧ РАСПОЗНАВАНИЯ ОБРАЗОВ

Аннотация. В работе представлено описание логики работы алгоритма «адресное сопоставление ячеек», который может быть использован в задачах распознавания образов на этапе принятия решения. Также приведены примеры применения данного метода при обработке реальных сигналов.

Ключевые слова: сигнал, алгоритм, распознавание образов, адресное сопоставление ячеек.

Введение

В большинстве случаев задача по распознаванию образов по сигналам решается трехэтапно [1]:

1. Предварительная обработка;
2. Векторизация;
3. Идентификация и верификация;

На первом этапе обычно проводятся различные подготовительные работы такие, как шумоочистка, нормализация и центрирование сигнала, сегментирование и т.д. В результате выполнения данного этапа из первичного сигнала выделяются ее наиболее информативные части. Например, при обработке речевых сигналов предварительная обработка может понадобиться для определения участков пауз, вокализованных и невокализованных звуков. Также на этом этапе выполняется сегментирование, т.е. нарезание обрабатываемого сигнала на куски определенной длины, называемых фреймами. Необходимость сегментирования сигнала продиктована тем, что сигнал является продуктом определенных динамических процессов, которых можно представить состоящими из множества стационарных процессов. Соответственно, длина фрейма должна быть такая же как длительность этих стационарных процессов. В общем случае, задача точного определения длительности и разграничение стационарных процессов является труднорешаемой и, пожалуй, не существуют никаких универсальных рецептов. Тем не менее, данная задача решается в конкретных случаях, например

при обычном темпе разговорной речи считается, что длительность стационарных участков примерно составляет порядка 20 мс - 30 мс. Соответственно, при обработке речевых сигналов длина фрейма выбирается равной длительности 20 мс - 30мс.

На следующем этапе, т.е. на этапе векторизации, для каждого фрейма сигнала вычисляются соответствующие числовые характеристики. Таким образом, каждый фрейм сигнала будет описываться вектором характеристик. Пусть N -количество характеристик, а M -количество фреймов, тогда обрабатываемому сигналу в пространстве характеристик соответствует матрица размером $M \times N$. В общем случае, количество характеристик является фиксированным, а количество фреймов – нет. В таком случае, в зависимости от длительности сигнала количество фреймов может быть довольно большим, что в свою очередь создает определенные проблемы для скорости принятия идентификационных решений. Во избежания данной проблемы используют так называемые кодовые книги, обладающие компактным размером матрицы. Для построения кодовой книги можно использовать различные методы, например, метод K -средних, метод ближайшего соседа, метод опорных векторов и т.д.

И наконец на этапе идентификации и верификации на основе «сравнения» кодовых книг обрабатываемого и эталонного сигналов выносится решение об их степени соответствия. На сегодняшний день имеются множество различных методов, при-

меняемых для осуществления процесса сравнения двух кодовых книг. Примерами подобных методов являются – Скрытые марковские цепи, Гауссовские смеси и т.д. Фактически во всех этих методах сравнение производится на основе вычисления «расстояний» между элементами кодовых книг. При этом в качестве меры расстояния могут быть использованы следующие его разновидности:

1. Евклидово расстояние;

$$\rho(x, x') = \sqrt{\sum_i^n (x_i - x'_i)^2} \quad (1)$$

2. Квадрат евклидова расстояния;

$$\rho(x, x') = \sum_i^n (x_i - x'_i)^2 \quad (2)$$

3. Расстояние городских кварталов (манхэттонское расстояние);

$$\rho(x, x') = \sum_i^n |x_i - x'_i| \quad (3)$$

4. Расстояние Чебышева;

$$\rho(x, x') = \max(|x_i - x'_i|) \quad (4)$$

5. Степенное расстояние:

$$\rho(x, x') = \sqrt[r]{\sum_i^n (x_i - x'_i)^p} \quad (5)$$

Здесь мы рассмотрели лишь основные, часто применяемые на практике меры вычисления “расстояний” между разными объектами в многомерном пространстве. Существуют и другие подобные меры.

Из формул 1-5 видно, что алгоритмическая сложность вычисления всех этих мер фактически одинакова и составляет порядка $O(N^2)$, где N – количество элементов в кодовой книге. При небольшом количестве N данный алгоритм может работать довольно быстро, но с ростом количества элементов в кодовой книге скорость вычисления расстояний будет сильно падать, что в некоторых практических применениях, особенно где требуется высокая скорость принятия решений, такой подход может не дать нужного результата.

В связи с этим, нами предлагается новый метод «адресное сопоставление ячеек», который как будет показано ниже, имеет алгоритмическую сложность порядка $O(N)$, то есть предлагаемый нами метод, которого можно использовать на этапе идентификации и верификации, будет работать гораздо быстрее, чем методы, основанные на вычислении различных расстояний между элементами кодовых книг.

Описание предлагаемого метода

Основная идея метода «адресное сопоставление ячеек» заключается в следующем. В первую очередь необходимо установить минимально и максимально допустимые значения по всем характеристикам. Например, известно, что частота основного тона, используемая в задачах идентификации дикторов или распознавания речи, меняется в пределах от 50 Гц до 400 Гц. Конечно, в крайне редких случаях у некоторых дикторов данная характеристика может выйти за рамки указанного предела. В подобных случаях значение характеристики необходимо приравнять ближайшей границе, то есть либо максимальному, либо минимальному значению.

В некоторых случаях минимальные и максимальные значения определенных характеристик сигнала можно найти теоретически. Например, энтропия любого информационного сигнала не может быть больше энтропии случайного сигнала, и не может быть меньше нуля.

Итак, после того, как установлена граница возможных значений определенной характеристики, необходимо разделить этот диапазон значений на ячейки. Эту процедуру можно осуществлять следующими способами:

- 1) Линейное деление:

- 1.1) с фиксированным количеством ячеек;

- 1.2) с фиксированным размером ячеек;

- 2) Нелинейное деление;

Пусть у нас имеется набор характеристик $\{x_i\}$, а $\max(x_i)$ и $\min(x_i)$ соответственно максимальное и минимальное значения характеристики x_i . Тогда в случае линейного деления с фиксированным количеством ячеек N_i размер ячеек ΔN_i можно определить по формуле:

$$\Delta N_i = \frac{\max(x_i) - \min(x_i)}{N_i} \quad (6)$$

А в случае линейного деления с фиксированным размером ячеек ΔN_i необходимо найти количество ячеек по формуле:

$$N_i = \frac{\max(x_i) - \min(x_i)}{\Delta N_i} \quad (7)$$

Если принято решение использовать способ нелинейного разделения, то количество ячеек и размер каждого из них задается исследователем по-своему усмотрению. Нелинейное разделение ячеек дает дополнительную возможность для управления над точностью принятия идентификационных решений. Например, в области наиболее вероятных значений характеристики размер ячеек можно сделать малыми, а в области наименее вероятных значений наоборот большими.

В итоге мы имеем многомерное дискретное пространство, размерность которого равна количеству используемых характеристик. В таком случае, каждый фрейм анализируемого сигнала будет отображаться в виде определенной точки в данном пространстве. При этом координаты этой точки могут принимать только дискретные значения, определяемые номерами ячеек, в которых она попадает. Благодаря дискретизации многомерного пространства характеристик для определения координаты точек мы можем использовать еще один сверхбыстрый алгоритм.

Пусть для j -го фрейма сигнала значение характеристики с индексом i равно $x_{i,j}$, тогда номер соответствующей ячейки можно определить по формуле:

$$N_{i,j} = \text{Fix} \left(\frac{x_{i,j} - \min(x_i)}{\Delta N_i} \right) + 1 \quad (8)$$

где оператор Fix означает взятие целой части выражения, находящегося в скобках.

Работу данного алгоритма можно проиллюстрировать с помощью следующего примера. Пусть какая-то характеристика меняется в пределах от 1 до 7. Разделим этот диапазон методом линейного разделения с фиксированным размером ячеек, равным 1, тогда количество ячеек будет равно:

$$N = \frac{7-1}{1} = 6 ;$$

т.е. получаем 6 ячеек, которые графически можно представить так, как показано на рисунке 1.

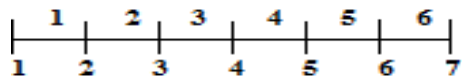


Рисунок 1– Пример разделения на ячейки диапазона значений характеристики.

На рисунке 1 цифры снизу показывают границы, а цифры сверху номера ячеек. Допустим, что у определенного фрейма сигнала значение рассматриваемой характеристики оказалось равным 3,2.

Вычислим номер соответствующей ячейки по формуле (8):

$$n = \text{Fix} \left(\frac{3,2 - 1}{1} \right) + 1 = 3$$

Это означает, что данный фрейм сигнала по этой характеристике попадает в 3-ю ячейку. Таким образом, можно определить номера всех ячеек по каждой характеристике для рассматриваемого фрейма сигнала. В итоге данный фрейм сигнала будет характеризоваться набором номеров (адресов) этих ячеек. Соответственно, в предлагаемом нами методе в качестве элементов кодовой книги можно использовать вектор (набор) адресов ячеек. На этапе идентификации и верификации, то есть при принятии решения, как правило, требуется провести сравнение двух кодовых книг – обрабатываемого и эталонного сигналов. Естественно, по парный поиск элементов кодовой книги при их большом объеме может занять довольно большое время. В связи с этим, для повышения производительности данной операции требуется найти соответствующий алгоритм.

Известно, что одним из самых быстрых алгоритмов поиска является метод бинарного поиска в отсортированном одномерном массиве. В нашей задаче можно провести небольшое преобразование с тем, чтобы привести структуру кодовой книги к подходящему виду для использования метода бинарного поиска. Отметим, что на этапе вычисления номеров ячеек они уже оказываются отсортированными, то есть остается лишь привести структуру ячеек в одномерно-линейный вид. Этот процесс назовем линейризацией кодовой книги и покажем на примере как она осуществляется. Для простоты примем, что у нас имеются всего лишь две характеристики - x_1 и x_2 . Пусть первая характеристика имеет три, а вторая четыре ячейки. Тогда в двумерном пространстве структура кодовой книги будет выглядеть так, как показано на рисунке 2.

x_1	11	12	13	14
	21	22	23	24
	31	32	33	34
	x_2			

Рисунок 2.– Пример структуры кодовой книги в двумерном пространстве.

Как видим из рисунка 2, всего ячеек – 12. В ячейках вышеприведенной таблицы прописаны их двумерные номера. Далее необходимо осуществить переход из двумерного представления номеров ячеек в одномерное представление, согласно которому нумерации ячеек таблицы можно придать следующий линейный вид: 11, 21, 31, 12, 22, 32, 13, 23, 33, 14, 24, 34.

Теперь определим с помощью, какой математической операции или выражения можно осуществить подобную линеаризацию. Пусть I_1 набор ячеек по первой, а I_2 набор ячеек по второй характеристике. Тогда в нашем примере $I_1 = \{1,2,3\}$, а $I_2 = \{1,2,3,4\}$. Обозначая через I набор ячеек линейного представления, получаем $I = \{1,2,3 \dots 12\}$. В таком случае, нетрудно убедиться в том, что элементы набора I можно вычислить по формуле:

$$I = I_1 + (I_2 - 1) * n_1 \quad (9)$$

где n_1 - количество ячеек первой характеристики, в нашем примере оно равно 3. С помощью нижеследующей таблицы можем продемонстрировать логику работы выражения (9):

Последовательность	I_1	I_2	Номера ячеек в I
11	1	1	$1+(1-1)*3=1$
21	2	1	$2+(1-1)*3=2$
31	3	1	$3+(1-1)*3=3$
12	1	2	$1+(2-1)*3=4$
22	2	2	$2+(2-1)*3=5$
32	3	2	$3+(2-1)*3=6$
13	1	3	$1+(3-1)*3=7$
23	2	3	$2+(3-1)*3=8$
33	3	3	$3+(3-1)*3=9$
14	1	4	$1+(4-1)*3=10$

24	2	4	$2+(4-1)*3=11$
34	3	4	$3+(4-1)*3=12$

Выражение (9) легко можно превратить для трехмерного случая:

$$I = I_1 + (I_2 - 1) * n_1 + (I_3 - 1) * n_1 * n_2 \quad (10)$$

А для четырехмерного случая получаем следующую формулу:

$$I = I_1 + (I_2 - 1) * n_1 + (I_3 - 1) * n_1 * n_2 + (I_4 - 1) * n_1 * n_2 * n_3 \quad (11)$$

Анализируя выражения (9), (10) и (11), для любого случая, то есть для произвольного количества характеристик, получаем следующее выражение, осуществляющее линеаризацию кодовой книги:

$$I = I_1 + \sum_{j=2}^m (I_j - 1) * \prod_{k=1}^j n_k \quad (12)$$

где m – количество характеристик.

В итоге для обрабатываемого сигнала в качестве кодовой книги имеем отсортированный одномерный массив ячеек, характеризующий данный сигнал.

Степень сходства двух сигналов определяется следующим образом. Пусть A множество адресов ячеек эталонного, а B множество адресов обрабатываемого сигналов, тогда степень сходства этих сигналов в процентном соотношении определяем по формуле:

$$C = \left(1 - \frac{2 * Cnt(A \cup B)}{Cnt(A \cap B)}\right) * 100\% \quad (13)$$

где операция $Cnt(A \cup B)$ - вычисляет количество одинаковых элементов множеств A и B , соответственно $Cnt(A \cap B)$ - вычисляет общее количество элементов A и B . Алгоритм бинарного поиска используется для вычисления выражения $Cnt(A \cup B)$

Заключение

Впервые метод «адресное сопоставление ячеек» был применен в работах [2] и [3], в которых были достигнуты удовлетворительные результаты. Также в работе [4] приводились результаты сравнительного анализа производительности предлагаемого метода с традиционными методами. В результате данных работ можем заключить, что метод «адресное сопоставление ячеек» работает устойчиво и способен обеспечить вычислительный процесс очень высокой производительностью.

Список литературы

- 1 Горелик А. Л., Скрипкин В. А. Методы распознавания.— 4-е изд.—//М.— 2004.— Высшая школа
- 2 Наурзбаева А.Ж., Медетов Б.Ж.. Выявление предвестников землетрясений по сигналам от сейсмической активности методом информационно-энтропийного анализа.// Вестник КазНУ, серия физическая.— 2012.— том 41, №2.
- 3 Наурзбаева А.Ж., Медетов Б.Ж.. Идентификация звуковых сигналов на

основе нелинейного анализа.// Вестник КазНУ, серия физическая.— 2012.— том 41, №2.

4 Наурзбаева А.Ж., Медетов Б.Ж. Применение нелинейного анализа сигналов в задачах распознавания образов.// 8-ая Международная научная конференция «Хаос и структуры в нелинейных системах. теория и эксперимент»/ – Караганда.– 2012
Принято в печать 11.06.2013

Б.Ж. Медетов, А.Ж. Наурзбаева, А.Б. Манапбаева

Казахский Национальный университет им. Аль-Фараби, г. Алматы, Казахстан

**МЕТОД ОБРАБОТКИ СИГНАЛОВ «АДРЕСНОЕ СОПОСТАВЛЕНИЕ ЯЧЕЕК»
ПРИ РЕШЕНИИ ЗАДАЧ РАСПОЗНОВАНИЯ ОБРАЗОВ**

В работе представлено описание логики работы алгоритма «адресное сопоставление ячеек», который может быть использован в задачах распознавания образов на этапе принятия решения. Также приведены примеры применения данного метода при обработке реальных сигналов.

Ключевые слова: сигнал, алгоритм, распознавание образов, адресное сопоставление ячеек

Б.Ж. Медетов, А.Ж. Наурзбаева, А.Б. Манапбаева

ал-Фараби атындағы Қазақ Ұлттық Университеті, Алматы қ., Қазақстан

**ОБРАЗДАРДЫ ТАНУ ЕСЕПТЕРІН ШЕШУДЕ ҚОЛДАНЫЛАТЫН «ҰЯШЫҚТАР
АДРЕСІН СӘЙКЕСТЕНДІРУ» ТӘСІЛІ**

Осы жұмыста образдарды тану есептерін шығарудағы шешім қабылдау этапында қолдануға қажетті «ұяшықтар адресін сәйкестендіру» тәсілі логикасының қалай жұмыс істейтіндігі сипатталған. Сонымен қатар, аталған тәсілді нақты сигналдарды өңдеуде пайдалану мысалдары келтірілген.

B. Medetov, A. Naurzbaeva, N. Albanbai, A. Manapbaeva

Al-Farabi Kazazakh National Universiti, Almaty, Kazakhstan

**SIGNAL PROCESSING METOD “MATCHING CELLS ADDRESS” IN SOLVING
PROBLEMS OF IMAGE RECOGNITION**

The paper describes the logic of the algorithm "matching cells address " that can be used in pattern recognition problems in the decision-making stage. It also shows examples of application of this method in the processing of real signals.