

2. Бутаков Е.А., Островский В.И., Фадеев И.Л. Обработка изображений на ЭВМ. — М.: Радио и связь, 1987. — 240 с.

## АЛГОРИТМЫ ВЕКТОРНОГО КВАНТОВАНИЯ В СИСТЕМЕ ИДЕНТИФИКАЦИИ РЕЧЕВЫХ СИГНАЛОВ.

*Садыхов Р.Х.<sup>\*</sup>, Ракуш В.В.<sup>\*\*</sup>*

*<sup>\*</sup>Белорусский Государственный Университет  
Информатики и Радиоэлектроники.*

*<sup>\*\*</sup>Научно-Исследовательский Институт Криминологии,  
Криминалистики и Судебных Экспертиз.*

### Введение.

Как отмечается в работах по автоматическому распознаванию личности по голосу [1], в системах нового поколения в качестве структурных элементов используются субречевые единицы, образующиеся при переходе от скалярного представления сигнала к векторному пространству признаков. С этой целью разработано несколько разновидностей алгоритмов векторного квантования [2]. Их суть заключается в том, что если  $\bar{x} = [x_1, x_2, \dots, x_N]^T$  — N-мерный параметрический вектор, координаты которого  $\{x_k, 1 \leq k \leq N\}$  — действительные случайные величины, то он может отражаться в N-мерный вектор  $\bar{y}$ , координатами которого также являются действительные случайные числа. Множество  $\bar{Y} = \{\bar{y}_i, 1 \leq i \leq L\}$  называется кодовой книгой. L — размер кодовой книги, а  $\{\bar{y}_i\}$  — множество кодовых векторов (центроидов). Для построения

#### 4. Распознавание образов и анализ изображений

кодовой книги  $M$ -мерное пространство векторов  $\bar{x}$  разбивается на  $L$  областей  $\{c_i, 1 \leq i \leq L\}$ . Каждой области  $c_i$  соответствует вектор-центроид  $\bar{y}_i$ . Если  $\bar{x}_i$  лежит в  $c_i$ , ему присваивается значение кодового вектора  $\bar{y}_i$ . Естественно, что при этом возникает ошибка квантования. Отклонение  $\bar{x}$  от  $\bar{y}$  может быть определено мерой близости  $d(\bar{x}, \bar{y})$ , в качестве которой может использоваться среднеквадратичное отклонение (СКО)

$$d(\bar{x}, \bar{y}) = \frac{1}{N} (\bar{x} - \bar{y})^T (\bar{x} - \bar{y}) = \frac{1}{N} \sum_{k=1}^N (x_k - y_k)^2 \quad (1)$$

$N$  – количество координат в векторе.

Создавая таким образом кодовые книги-эталоны и кодовые книги-тесты, затем сравнивая их, можно верифицировать или идентифицировать личность по голосу [3]. Для систем верификации, работающих по парольной фразе, кодовые книги-эталоны и кодовые книги-тесты имеют одинаковую размерность  $L$ , их сравнение не вызывает каких-либо затруднений. Другое дело, когда распознавание осуществляется по произвольной речи, продолжительность и семантика которой не регламентированы. В этом случае эталонная кодовая книга должна иметь максимальную размерность, а кодовая книга-тест будет иметь переменное значение  $L$ , зависящее от продолжительности и семантики высказывания. В этом случае возникает необходимость решения задачи оптимального квантования входной реализации в соответствии с некоторым критерием качества векторного квантования, оцененным априори.

### Алгоритмы формирования кодовых книг.

В этом разделе мы рассмотрим модификацию алгоритма векторного квантования, применяемую для формирования эталонной и тестовой кодовых книг.

Одним из методов построения кодовой книги является алгоритм  $K$ -средних [2]. Пусть  $K=L$ . В этом случае алгоритм разбивает набор входных векторов  $\bar{x}_i$  на  $L$  областей (кластеров). Обозначим через  $m$  номер итерации, а через  $C_i(m)$  —  $i$ -ый кластер на  $m$ -ой итерации с центроидом  $y_i(m)$ . Тогда алгоритм  $K$ -средних можно разбить по шагам.

*Шаг 1. Начальное разбиение на  $L$  кластеров.*  $m=0$ . Для  $m=0$  разбиение осуществляется с помощью обычного дихотомического алгоритма, результатом работы которого является последовательность векторов-центроидов  $\bar{y}_i(0), 1 \leq i \leq L$ .

*Шаг 2. Классификация.* Входные векторы  $\{\bar{x}_i, 1 \leq i \leq M\}$  разбиваются по кластерам  $C_i$  с помощью правила ближайшего соседа  $\bar{x} \in c_i(m)$  тогда и только тогда, когда  $d[\bar{x}, \bar{y}_i(m)] \leq d[\bar{x}, \bar{y}_j(m)]$  для всех  $j \neq i$ .

*Шаг 3. Коррекция кодового вектора.* Для  $m=m+1$  производится перерасчет центроидов всех кластеров  $\bar{y}_i(m) = \text{cent}(c_i(m)) \quad 1 \leq i \leq L$ .

*Шаг 4. Проверка на окончание процедуры.* При формировании кодовой книги-эталона значение  $L$  заранее известно. В этом случае квантование завершается, если интегральная мера близости

$$D(m) = \frac{1}{L \cdot N} \sum_{i=1}^L \sum_{j=1}^N d_{\min}(\bar{x}_i, \bar{y}_i) \quad (2)$$

на итерации  $m$  не уменьшилось по отношению к  $D(m-1)$ . В противном случае осуществляется переход на шаг 2.

При формировании кодовой книги-теста значение  $L$  заранее неизвестно. Поэтому, чтобы завершить квантование, необходимо  $D(m)$  сравнивать с каким-то пороговым значением ошибки квантования  $P$ . Для оценки порога  $P$  была создана фонетически сбалансированная речевая база данных продолжительностью звучания более 5 минут. Пять дикторов-

#### 4. Распознавание образов и анализ изображений

мужчин начитывали этот текст, который затем вводился в компьютер и служил в качестве эталона.

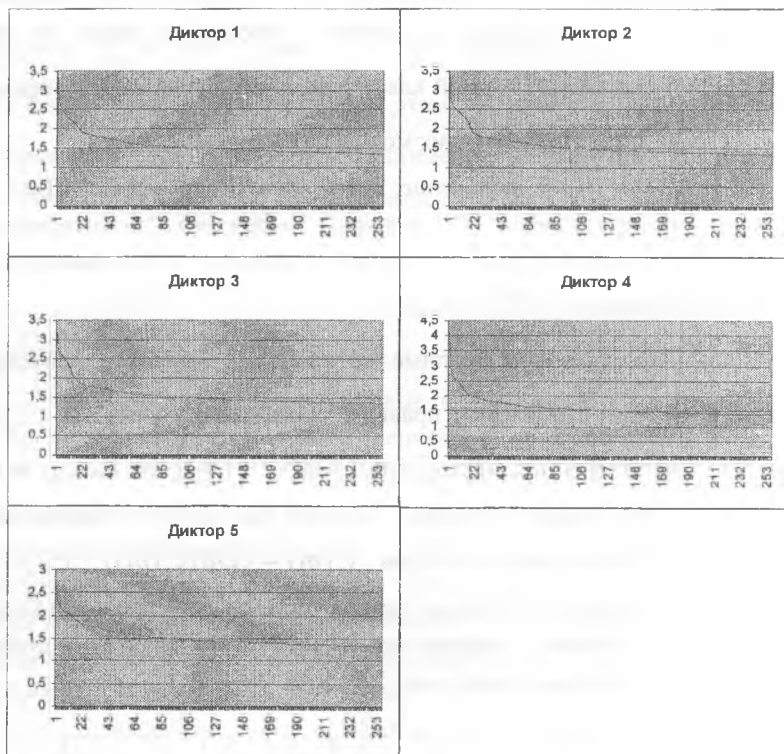


Рис 1. Графики зависимости ошибок квантования от размерностей кодовых книг для 5 дикторов.

После этого рассчитывались кодовые книги для различных размерностей  $L$ , в зависимости от нее определялись ошибки квантования (1). Графики зависимости  $d(\bar{x}, \bar{y})$  от  $L$  для пяти дикторов показаны на рис

1. Из них видно, что после  $L=200$   $d(\bar{x}, \bar{y}) \approx \text{const} = 1,4$ . Окончательно в наших экспериментах  $L$  была выбрана 256.

Таким образом, для формирования тестовой кодовой книги пороговое значение ошибки квантования  $P$  при сравнении на шаге 4 должно быть равным 1,4.

### **Экспериментальная проверка работоспособности алгоритма.**

Для проверки работоспособности алгоритма была разработана специальная программа, прототип системы фоночета голосов потенциальных правонарушителей, количество которых состояло из 5 человек. В качестве эталонов использовались описанные в разделе 2 книги, созданные из фонетически сбалансированных речевых файлов 5-минутной продолжительности.

Во время тестирования каждый из 5 дикторов наговаривал в микрофон произвольный текст произвольной длительности. На экране фиксировались пять мер близости между пятью эталонными и одной тестовой кодовыми книгами.

Результаты таких предварительных экспериментов показали, что, как правило, мера близости между тестовой реализацией и эталоном диктора, который в данный момент произносил фразу, была самой минимальной. Ошибок на базе из 5 дикторов зарегистрировано не было.

### **Выводы.**

Предлагаемый алгоритм структуризации речи со среднеквадратичным критерием качества может быть использован в системах идентификации личности по произвольной речи, при этом размерность кодовой книги должна состоять не менее чем из 200-256 кодовых векторов. К недостаткам данного алгоритма следует отнести невысокое быстродействие работы,

особенно в режиме обучения, которая может продолжаться несколько дней непрерывного счета на компьютере типа Pentium 166.

Таким образом, устранение этого недостатка, по-видимому, должно стать объектом дальнейших исследований в этом направлении.

### **Литература.**

1. Рылов А.С. Некоторые аспекты проблемы автоматического распознавания личности по голосу. Сб. статей Вопросы криминологии, криминалистики и судебной экспертизы. № 11, 1996 г., стр. 162-176.
2. Макхоул Д., Рукос С., Гиш Г. Векторное квантование при кодировании речи. ТИИЭР т.73, № 11, 1985 г., стр. 19-61.
3. Рылов А.С., Сапронович И.З., Францкевич А.В. Система автоматической верификации голоса говорящего. Тез. докладов международной конференции "Информатизация правоохранительных систем". Москва, июль, 1995 г.

УДК 681.3

## **АЛГОРИТМЫ РАСПОЗНАВАНИЯ РУКОПИСНЫХ СИМВОЛОВ НА ОСНОВЕ КОЛЬЦЕВЫХ И СЕКТОРНЫХ МОМЕНТНЫХ ФУНКЦИЙ**

*Садыхов Р.Х., Муравин А.Л.*

*Институт технической кибернетики НАН Беларуси*

Задача распознавания рукописных символов является частным случаем более общей задачи - задачи распознавания графических образов. Основными методами решения поставленных задач являются методы цифровой обработки сигналов (ЦОС).