

УДК 621.395

ЛИВШИЦ М.З.

ВЕКТОРНОЕ КВАНТОВАНИЕ В ШИРОКОПОЛОСНЫХ КОДЕРАХ РЕЧИ

Рассматриваются пути повышения эффективности векторного квантования параметров речевого сигнала. Осуществляется анализ методов построения кодовых книг, описывается модифицированный алгоритм обучения, рассматриваются вопросы применения векторного квантования к широкополосным речевым сигналам.

Векторное квантование (ВК) широко используется как в технике низкоскоростной передачи речевых сигналов, так и при построении акустических моделей систем распознавания речи. ВК представляется как процесс исключения избыточности, при котором эффективно используются четыре взаимосвязанных свойства векторных параметров: линейная зависимость (корреляция), нелинейная зависимость, форма функции плотности вероятности (ФПВ) и многомерность вектора.

Реализация векторного квантования требует применения двухэтапной процедуры. На первом этапе (этапе обучения) происходит формирование кодовой книги векторного квантователя. На втором этапе (этапе квантования) наблюдаемому вектору ставится в соответствие наиболее подходящий вектор кодовой книги. Эффективность функционирования квантователя существенно зависит от качества проведения операции обучения. Как правило, для этой цели используют процедуру кластеризации, известную в литературе как алгоритм К-средних или алгоритм Линдо-Бьюзо-Грэя [1...2]. Данный подход кластерного анализа заключается в сортировке множества векторов обучающей выборки на группы с похожими свойствами – кластеры или ячейки Вороного [1].

Критерием сходства обычно является мера расстояния между вектором обучающей выборки и центром кластера (ячейки). Наибольшее распространение получили такие меры искажений как: среднеквадратическая ошибка (СКО) и взвешенная СКО. В работе [3] была предложена комплексная мера искажения, состоящая из трех составляющих: СКО, направления и скорости изменения вектора. В настоящей работе для исследования влияния параметров векторного квантования использовалось специальное программное обеспечение, с помощью которого были проведены исследования влияния различных мер искажения и методов задания начальных условий на качество обучения. Методы задания начальных условий представлены на рис. 1 – 3.

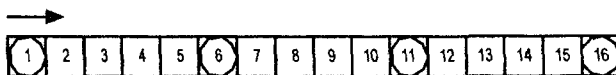


Рис.1. Задания начальных условий по методу "равномерный шаг"

В книгу помещаются векторы обучающего множества, отмеченные окружностью, стрелкой указано направление движения при выборе.

В методе "случайный выбор". Базовые векторы выбираются случайным образом по индексу из массива норм векторов.

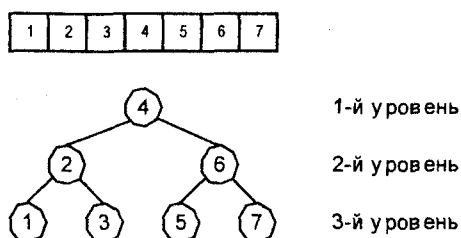


Рис.2. Задание начальных условий по методу "равномерное дерево"

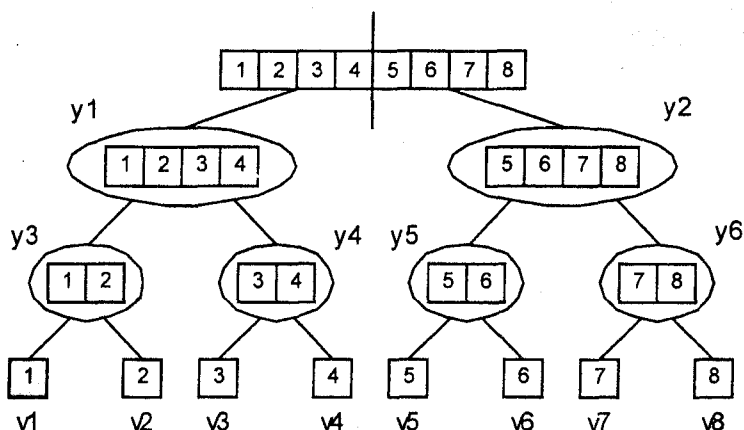


Рис.3. Построение книги для поиска по методу дихотомии
*y1–y6 – центры соответствующих векторов обучающего множества;
 v1–v8 – искомые векторы (восьмиуровневая книга).*

Было выявлено, что предлагаемые в работах [1,2] алгоритмы кластеризации обладают существенным недостатком - сходимостью к локальному оптимуму. При этом применение комплексной меры искажения, предложенной в работе [3] не обеспечило весомых улучшений качества векторного квантования. На основании экспериментальных данных, представленных в табл.1, был предложен модифицированный алгоритм К-средних, обеспечивающий нахождение глобального оптимума для всей обучающей выборки.

Таблица 1

Анализ качества составления кодовых книг для алгоритма К-средних

Длительность вектора, мс	1.25		2.5		5		7.5		10		Среднее по уровням	
	64	128	64	128	64	128	64	128	64	128	64	128
SEGSNR, дБ:												
"Равномерный шаг";	7.9	9.2	6.1	6.8	3.7	4.4	2.9	3.5	2.3	2.9	4.58	5.36
"Равномерное дерево";	7.7	9.1	6.1	6.7	4.0	4.1	2.6	3.2	2.4	2.5	4.56	5.12
"Случайный выбор";	6.8	9.1	6.8	5.8	3.7	3.9	2.4	3.5	1.8	2.4	4.30	4.94
Дихотомия.	6.9	7.2	5.1	5.3	3.0	3.7	2.8	3.2	2.7	2.8	4.10	4.44

Упрощенная схема алгоритма представлена на рис.4.

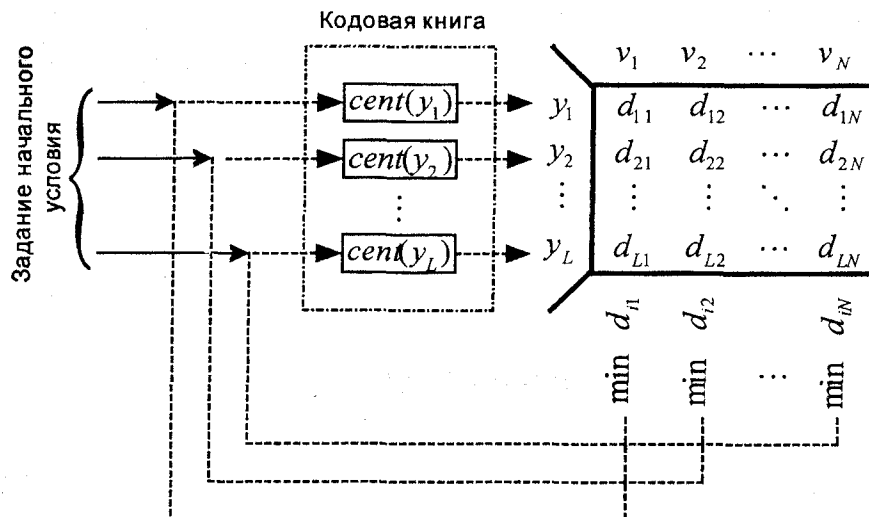


Рис.4. Обучение книги по модифицированному алгоритму К-средних $cent(y_i)$ - центроид i -го кластера (уровень книги y_i); v_1, \dots, v_N - векторы обучающего множества; d_{ij} - расстояние j -го вектора обучающего множества от i -го уровня книги; $\min d_{ij}$ - индекс ближайшего j -го вектора обучающего множества к i -му уровню книги

Сравнение экспериментальных данных для алгоритма, описанного в [1,2], и предложенного в данной работе представлено в табл.1 и табл.2, соответственно. Анализ табл.1 и 2 показывает, что модифицированный алгоритм обеспечивает лучшее качество векторного квантования.

В связи с актуальностью построения широкополосных компрессоров речевых сигналов с частотным диапазоном 50 - 7000Гц для дальнейшего улучшения качества реконструированного сигнала в настоящей работе было предложено использовать многополосные кодовые книги. Исследование влияния параметров многополосных кодовых книг на качество векторного квантования показало, что даже при использовании банка фильтров, не обладающего перфектной реконструкцией, многополосная кодовая книга обеспечивает существенное уменьшение искажения.

Таблица 2

Анализ качества составления кодовых книг для модифицированного алгоритма

Длительность вектора, мс	1.25		2.5		5		7.5		10		Среднее по уровням	
	64	128	64	128	64	128	64	128	64	128	64	128
SEGSNR, дБ	7.9	9.5	7.5	8.5	4.3	5	3	3.5	2.5	3	5.04	5.9

В качестве примера на рис.5 представлена зависимость меры искажения от глубины однополосной и многополосной кодовой книги.

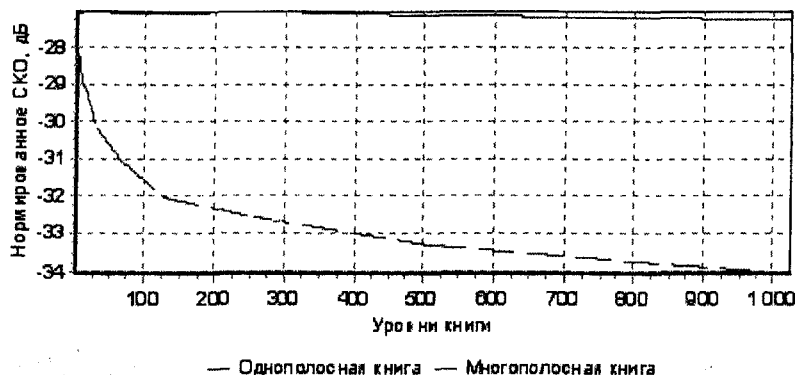


Рис.5. Зависимость СКО от количества уровней в книге для вектора длительностью 5 мс

Таким образом, предлагаемый модифицированный алгоритм обучения (составления) кодовых книг в сочетании с квантованием исходного сигнала в критических полосах обеспечивает существенное улучшение качества реконструированной речи при относительно небольшом увеличении вычислительной сложности.

БИБЛИОГРАФИЧЕСКИЙ СПИСОК

1. Digital Signal Processing: HandBook /edited by Vijay K. Madisetti and Douglas B. Williams, USA: CHAPMAN & HALL/CRCnetBASE, 1999.
2. Макхоул Дж., Рукос С., Гиш Г. Векторное квантование при кодировании речи // ТИИЭР, 1985. – Т.73. – №11. – С.19-61
3. Стукалов Д.Н. Кластеризация параметров речевого сигнала с учетом их динамических характеристик // Доклады 3-й межд. конф. "Цифровая обработка сигналов и ее применения", DSPA-2000, Москва, Россия, т.2., 2000.

Лившиц Михаил Зенадьевич

Студент факультета компьютерного проектирования

Белорусский государственный университет информатики и радиоэлектроники, г.Минск

Тел.: +375 (17) 261-85-10

E-mail: mlivshitz@tut.by