

Теоретические проблемы обработки и распознавания сигналов и изображений

УДК 621.39

А.Н. Ткаченко, О.Ф. Грийо Тукало, А.В. Дзись

Метод дельта-квантования параметров речевого сигнала с быстрым поиском ближайшего вектора в кодовой книге

Предложен метод векторного дельта-квантования LSF -параметров речевого сигнала с предсказанием следующего значения, что позволяет уменьшить динамический диапазон входных величин и соответственно ошибку квантования. Разработаны итерационная процедура создания кодовых книг для реализации предложенного метода и двухэтапный метод поиска ближайшего вектора по взвешенной евклидовой метрике на основе kd -дерева, позволяющий уменьшить сложность вычислений без потери качества воспроизведения речевого сигнала.

The paper proposes a vector delta-quantization method of LSF -parameters of speech signal with forecasting the next value, that reduces the dynamic range of input variables and therefore a quantization error. The iterative procedure of codebook design is developed for the proposed method. A two-stage search method of the nearest vector by weighted Euclidean metrics, based on kd -tree, is proposed, that achieves reducing of computational complexity without losing a recovery quality of speech signal.

Запропоновано метод векторного дельта-квантування LSF -параметрів мовленнєвого сигналу з прогнозуванням наступного значення, що дозволяє зменшити динамічний діапазон входних величин і відповідно похибку квантування. Розроблено ітераційну процедуру створення кодових книг для реалізації запропонованого методу та двоетапний метод пошуку найближчого вектора за зваженою евклідовою метрикою на основі kd -дерева, що дозволяє досягти зменшення складності обчислень без втрати якості відтворення мовленнєвого сигналу.

Введение. Замеченная Г. Муром закономерность возрастания производительности процессоров вдвое каждые два года подтверждается практикой на протяжении последних 40 лет [1]. Приблизительно такими же темпами возростали и объемы мультимедийной информации, передающейся и хранящейся в компьютерных системах. Однако наличие фундаментальных ограничений заставило в 2007 г. Г. Мура признать, что закон вскоре перестанет действовать. В то же время аналогичных ограничений на возрастание объемов информации в настоящее время не появилось.

Таким образом, существует проблема недостаточной производительности компьютерных систем, предназначенных для обработки мультимедийной, в частности голосовой информации. Уменьшение объемов данных, необходимых для корректного воспроизведения цифрового речевого сигнала, позволяет существенно повысить эффективность работы таких систем.

Сегодня самую высокую степень сжатия обеспечивают параметрические методы с последующим квантованием полученных параметров. На этапе параметризации сигнал разбивают на кадры (фреймы) одинаковой длины, как правило, по 10–30 мс, и для каждого из них вычисляют некоторый вектор параметров. В современных системах сжатия речевых сигналов стандартом де-факто стало вычисление параметров на основе модели линейного предсказания десятого порядка. На этапе квантования полученный вектор заменяют ближайшим ко входному представителем репрезентативной совокупности параметров, хранящихся в кодовых книгах (КК). В работе [2] было показано:

- В качестве параметров квантования целесообразно использовать линейные спектральные частоты (LSF), вычисляемые по значениям коэффициентов линейного предсказания (LPC).
- Оптимальное квантование полного LSF -вектора сложное в вычислительном отношении, поэтому его разбивают на два или три подвек-

тора, каждый из которых квантуется раздельно. Альтернативным вариантом служит каскадное (многоэтапное) квантование полного вектора при помощи КК меньшего размера, при котором на каждом последующем этапе квантуется ошибка, оставшаяся после квантования на предыдущем этапе.

- Ошибку квантования для i -го фрейма оценивают по спектральному искажению SD_i :

$$SD_i^2 = \frac{1}{F_s} \int_0^{F_s} [10 \log_{10}(P_i(f)) - 10 \log_{10}(\hat{P}_i(f))]^2 df,$$

где $P_i(f) = 1/|A_i(\exp(j2\pi f / F_s))|^2$, $\hat{P}_i(f) = 1/|\hat{A}_i(\exp(j2\pi f / F_s))|^2$ – спектральные плотности начального и квантованного векторов для i -го фрейма, $A_i(z)$, $\hat{A}_i(z)$ – начальный и квантованный LPC -полиномы, соответствующие i -му фрейму, F_s – частота дискретизации.

- Для выполнения условий транспарентности среднее значение ошибки квантования SD_{av} по всем фреймам должно составлять примерно 1 дБ; число фреймов, для которых $SD > 2\delta B$, не должно превышать двух процентов; фреймов, для которых $SD > 4\delta B$, не должно быть вообще.

- Указанные условия достигаются при разбиении вектора LSF -параметров на два подвектора и использовании для квантования входного вектора 24 бит на один фрейм.

- При выборе вектора из КК \mathbf{y} , ближайшего ко входному вектору \mathbf{x} , для вычисления расстояния $d^{(WE)}(\mathbf{x}, \mathbf{y})$ целесообразно использовать взвешенную Евклидову метрику (ВЕМ):

$$d^{(WE)}(\mathbf{x}, \mathbf{y}) = (\mathbf{x} - \mathbf{y})^T \mathbf{W}(\mathbf{x} - \mathbf{y}),$$

где \mathbf{W} – положительно определенная диагональная матрица весов.

Применение ВЕМ обеспечивает выигрыш в $SD \approx 0,1\delta B$ по отношению к невзвешенной Евклидовой метрике (ЕМ).

Дальнейшему уменьшению объема данных для описания спектральной информации посвящены работы [3–5]. Однако задача уменьшения битовой скорости, необходимой для передачи речевой информации, при условии сохра-

нения транспарентности квантования, а также выполнения существующих ограничений по объемам памяти, сложности вычислений и задержки кодирования, остается актуальной.

В статье предлагается метод сжатия речевой информации, позволяющий уменьшить объемы данных для описания спектральной информации до 20 бит на один фрейм при сохранении приемлемого качества воспроизводимого речевого сигнала.

Метод векторного дельта-квантования с предсказанием

Любой метод сжатия основывается на использовании избыточности, присущей речевому сигналу. Эта избыточность сохраняется и после перехода к параметрическому описанию сигнала. Количественной мерой избыточности служит коэффициент корреляции. В табл. 1 приведены значения коэффициента корреляции между отдельными параметрами LSF в пределах фрейма, полученные для тренировочной последовательности из 90000 векторов.

Таблица 1. Коэффициенты корреляции LSF в пределах фрейма

№	1	2	3	4	5	6	7	8	9	10
1	1,00	0,45	-0,05	-0,07	-0,17	-0,27	-0,28	-0,28	-0,24	-0,10
2	0,45	1,00	0,57	0,23	0,11	-0,02	0,03	0,06	-0,04	-0,01
3	-0,05	0,57	1,00	0,56	0,32	0,39	0,38	0,38	0,19	0,06
4	-0,07	0,23	0,56	1,00	0,58	0,46	0,44	0,26	0,19	0,03
5	-0,17	0,11	0,32	0,58	1,00	0,62	0,42	0,24	0,01	0,01
6	0,27	-0,02	0,39	0,46	0,62	1,00	0,62	0,45	0,20	0,01
7	-0,28	0,03	0,38	0,44	0,42	0,62	1,00	0,63	0,25	0,22
8	-0,28	0,06	0,38	0,26	0,24	0,45	0,63	1,00	0,58	0,30
9	0,24	0,04	0,19	0,19	0,01	0,20	0,25	0,58	1,00	0,56
10	-0,15	-0,01	0,06	0,03	0,01	0,01	0,22	0,30	0,56	1,00

В табл. 2 приведены значения коэффициента корреляции между соответствующими LSF -параметрами для последовательных фреймов.

Таблица 2. Коэффициенты корреляции LSF для соседних фреймов

№	1	2	3	4	5	6	7	8	9	10
	0,68	0,73	0,76	0,81	0,86	0,84	0,82	0,81	0,76	0,75

Как видно из таблиц, наблюдается довольно высокая корреляция как между отдельными параметрами в пределах фрейма, так и между компонентами вектора LSF для соседних фрей-

мов. Возможны следующие варианты использования этих двух видов корреляции:

- переход от скалярного квантования (СК) параметров к векторному квантованию (ВК);

- применение методов предсказания с последующим удалением предсказанных значений с целью уменьшения динамического диапазона входных величин.

Существует принципиальная возможность применения данных подходов для использования как межфреймовой корреляции, так и корреляции в пределах фрейма. Однако СК LSF -параметров с предсказанием в пределах фрейма значительно уступает по эффективности ВК, что выражается в увеличении спектрального искажения [3]. В то же время квантование LSF -параметров для последовательных фреймов связано с внесением дополнительной задержки, что недопустимо для некоторых приложений. Таким образом, целесообразным представляется вариант ВК параметров LSF для отдельных фреймов с предсказанием значений для последующих фреймов. Именно этот подход, названный авторами векторным дельта-квантованием с предсказанием (ВДКП), предлагается в данной статье.

В основе ВК при помощи КК лежит простая идея: вместо передачи действительного значения параметров входного речевого сигнала вычисляют и передают индекс ближайшего ко входному вектору из КК. При этом в КК хранятся значения параметров, полученные в результате кластеризации тренировочной последовательности входных векторов LSF . В отличие от этого, по методу ВДКП в КК хранятся значения параметров, полученные в результате кластеризации разницы (дельты) e_i действительного I_i и предсказанного значения \tilde{I}_i LSF . При условии удачного предсказания динамический диапазон входных величин, использующихся для создания КК, и, соответственно, ошибка кластеризации должны уменьшиться, что, в свою очередь, приведет к уменьшению спектрального искажения в процессе квантования.

Таким образом, в процессе создания КК входная последовательность будет состоять из век-

торов $e_i = I_i - \tilde{I}_i$. При использовании авторегрессионной модели линейного предсказания первого порядка предсказанное значение $\tilde{I}_i(m)$ m -го компонента LSF для i -го фрейма вычисляется по формуле:

$$\tilde{I}_i(m) = \alpha(m)\hat{I}_{i-1}(m) + \beta(m), \quad (1)$$

где параметры $\alpha(m)$ и $\beta(m)$ вычисляются, исходя из условия минимизации среднеквадратической ошибки предсказания:

$$\alpha(m) = \frac{COV(I_i(m), \hat{I}_{i-1}(m))}{VAR(I_i(m))}, \quad (2)$$

$$\beta(m) = E(I_i(m)) - \alpha(m)E(\hat{I}_{i-1}(m)), \quad (3)$$

где E , COV , VAR обозначают соответственно математическое ожидание, ковариацию и дисперсию, $\hat{I}_{i-1}(m)$ – квантованное значение m -го компонента вектора LSF на $i-1$ -м фрейме.

Процедура квантования LSF -параметров по методу ВДКП представлена на рис. 1.

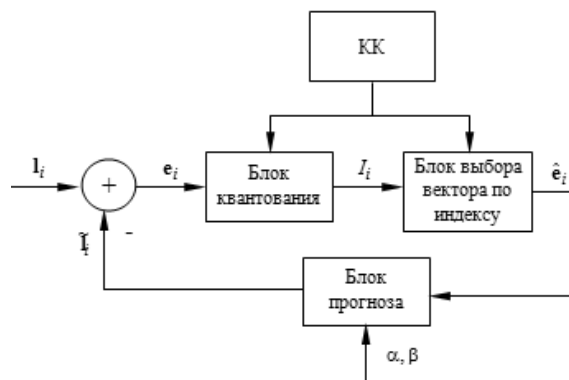


Рис. 1. Структурная схема квантования по методу ВДКП

Создание кодовых книг по методу ВДКП

Параметры α , и β вычисляются для тренировочной последовательности векторов LSF . Это означает невозможность непосредственного применения формул (2) и (3) в процессе создания КК, поскольку значения $\hat{I}_{i-1}(m)$ вычисляются только в процессе квантования, которое, в свою очередь, может происходить лишь при наличии КК.

Таким образом, практическое применение ВДКП требует реализации итерационной процедуры создания КК.

Шаг 1. Для $i=1,2,\dots,N$, $m=1,2,\dots,M$ присвоить $\hat{l}_{i-1}(m) = l_{i-1}(m)$.

Шаг 2. По формулам (2) и (3) вычислить параметры α, β .

Шаг 3. По формуле (1) вычислить предсказываемые значения параметров $LSF \tilde{\mathbf{I}}_i$.

Шаг 4. Вычислить значения векторов ошибки $\mathbf{e}_i = \mathbf{I}_i - \tilde{\mathbf{I}}_i$.

Шаг 5. Для полученных векторов ошибки \mathbf{e}_i по методу k -средних создать КК \mathbf{Y} .

Шаг 6. Используя параметры α, β и созданную КК \mathbf{Y} провести квантование векторов тренировочной последовательности \mathbf{I}_i и получить квантованные значения векторов $\hat{\mathbf{I}}_i$.

Шаг 7. Вычислить ошибку квантования. Если она изменилась на достаточно малую величину в сравнении с предыдущей итерацией, закончить процедуру. Если нет, возвратиться к шагу 2.

Таким образом, процедура создания КК состоит из трех этапов: вычисления параметров предсказания, квантования, а также непосредственно кластеризации. Создание начальной КК целесообразно проводить по усовершенствованному методу k -средних [6]. На последующих итерациях можно применять классический алгоритм k -средних, используя имеющиеся векторы \mathbf{Y} для инициализации центроидов. Схематически процедура создания КК представлена на рис. 2:



Рис. 2. Структурная схема создания КК для метода ВДКП

Рассмотренная процедура создания КК может использоваться как для отдельного, так и для

многоэтапного ВК. В последнем случае процедура создания КК может потребовать дополнительных итераций, обусловленных необходимостью согласования между собой отдельных частей КК [5]. Однако уменьшение спектрального искажения в процессе квантования есть достаточным обоснованием дополнительных временных затрат на стадии создания КК, выполняющейся один раз на подготовительном этапе.

Быстрый поиск ближайшего вектора с использованием kd -дерева

Уменьшение вычислительной сложности поиска ближайшего соседа в КК рассматривалось в работах [7–9]. Однако практическая ценность этих методов существенно ограничивается невозможностью применения ВЕМ, поскольку все они основываются на предварительной обработке КК, когда значения весов еще неизвестны. Для сокращения времени поиска ближайшего вектора предлагается подход, сочетающий преимущества методов быстрого поиска ближайшего соседа с применением ВЕМ.

Пусть \mathbf{K} – множество векторов, содержащихся в КК, $|\mathbf{K}| = k$. Суть разработанной двухэтапной стратегии поиска, схема которой приведена на рис. 3, заключается в следующем.

• На первом этапе выполняется так называемый быстрый поиск в упорядоченной определенным образом КК, в процессе которого по ЭМ отбирается множество $\mathbf{C} \subset \mathbf{K}$ векторов (кандидатов), упорядоченных по возрастанию расстояния до входного вектора:

$$\mathbf{C} = \{\mathbf{y}_1, \mathbf{y}_2, \dots, \mathbf{y}_t\}, |\mathbf{C}| = t, t \leq k.$$

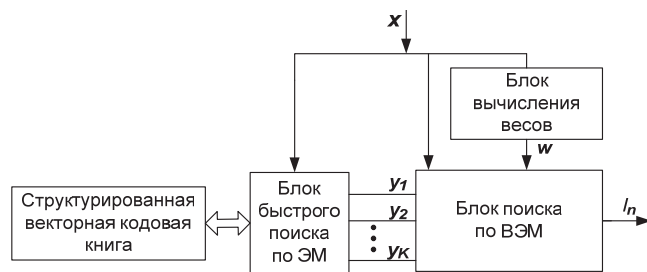


Рис. 3. Структурная схема двухэтапной стратегии поиска векторов в КК

• На втором этапе с использованием ВЕМ из отобранного множества кандидатов выби-

рается один вектор, ближайший к входному (с заданной вероятностью p).

Поскольку по ВЕМ размерности имеют разный вес, вектор КК, являющийся ближайшим по ЕМ, может оказаться не лучшим при учете весов. Такая ситуация для двухмерного случая показана на рис. 4:

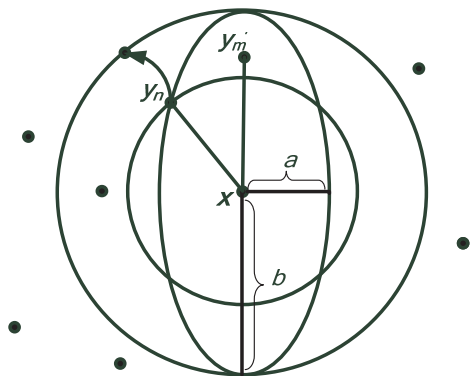


Рис. 4. Поиск ближайшего соседа по ВЕМ в двухмерном случае

Как видно из рисунка, точка y_n (вектор в двухмерном пространстве) является ближайшей по ЕМ, но в результате использования ВЕМ ближайшей оказывается точка y_m , поскольку

$$\frac{a}{b} = \frac{w_{\min}^2}{w_{\max}^2}. \quad (4)$$

Согласно (4), для нахождения вектора, ближайшего к входному по ВЕМ, в процессе поиска достаточно охватить множество кандидатов, ограниченное кругом радиуса $b = r_{\max}$, что детально было рассмотрено в [10].

Дополнительное уменьшение времени поиска можно получить, отказавшись от требования обязательного нахождения на втором этапе вектора, ближайшего к входному по ВЕМ. Предпосылки к этому такие:

- чаще всего вектор, ближайший к входному по ВЕМ, находится в числе первых векторов-кандидатов, определенных по ЕМ и упорядоченных по возрастанию расстояния до входного. Таким образом, значительная часть вычислений расстояний необходима лишь для того, чтобы убедиться в отсутствии лучшего вектора;

- пропуск в некоторых случаях ближайших векторов по ВЕМ не приводит к заметному

увеличению спектрального искажения. Объясняется это тем, что вместо ближайшего вектора, как правило, выбирается вектор, достаточно близкий к входному. Так, экспериментальные данные [9] показывают, что пропуск ближайших векторов в 5–10 процентов фреймов увеличивает спектральное искажение лишь на 0,01 – 0,02 дБ.

Таким образом, нет необходимости вычислять расстояние по ВЕМ до всех векторов, попадающих в круг (гипершар) радиусом r_{\max} , а, задавшись некоторым значением вероятности p_{\lim} , следует найти величину $\lim = \gamma/\gamma_{\max}$, определяющую радиус поиска $r \leq r_{\max}$, в пределах которого с вероятностью $p \geq p_{\lim}$ находится вектор, ближайший к входному по ВЕМ. Для этого необходимо найти зависимость $r = f(p, r_{\max})$ либо $p = f(\gamma/\gamma_{\max})$.

Вероятность нахождения ближайшего вектора на расстоянии r можно найти как отношение части объема M -мерного эллипсоида, ограниченного гипершаром радиуса r , ко всему объему M -мерного эллипсоида:

$$p(r) = \frac{V^{(M)}(r)}{V^{(M)}_{el}}, \quad p(r) = \begin{cases} 0, & r = 0 \\ 1, & r \geq r_{\max} \end{cases}. \quad (5)$$

Геометрическая интерпретация приведенной зависимости для двухмерного случая проиллюстрирована на рис. 5.

Часть объема M -мерного эллипсоида, ограниченного гипершаром радиуса r , можно найти как:

$$V^{(M)}(h) = V_{el}^{(M)}(h) + (V_s^{(M)}(h) - V_c^{(M)}(h)), \quad (6)$$

где $V_{el}^{(M)}(h)$ – объем части полуэллипса в M -мерном пространстве, ограниченного гиперплоскостью, проходящей на расстоянии h ;

$$V_{el}^{(M)}(h) = \int_0^h V_{el}^{(M-1)}(y) dy, \quad (7)$$

при $h = b$, где b – максимальная полуось эллипса, формула (7) даст полный объем полуэллипса.

$V_s^{(M)}(h)$ – объем M -мерного сектора радиусом r и углом 2α ;

$$V_s^{(M)}(h) = \frac{2\pi r^2}{M} \cdot V_s^{(M-2)} - \frac{(M-1) \cdot r \cdot \sin^{M-3} \alpha \cdot \cos \alpha}{M \cdot (M-2)} \cdot V_b^{(M-1)}, \quad (8)$$

где $V_b^{(M-1)} = \frac{2\pi r^2}{M} \cdot V_b^{(M-3)}$ – объем гипершара размерности $(M-1)$; $V_c^{(M)}(h)$ – объем M -мерного конуса высотой h .

$$V_c^{(M)}(h) = \frac{V_b^{(M-1)} \cdot h}{M}. \quad (9)$$

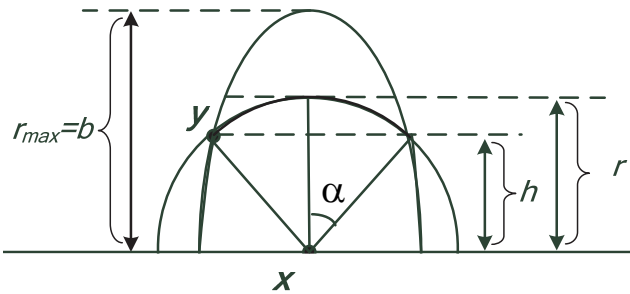


Рис. 5. Геометрическая интерпретация зависимости вероятности от расстояния

Для случая $M=5$ (разбиение 10-мерного вектора LSF на два подвектора) формулы (7) – (9) принимают вид:

$$V_{el}^{(5)}(h) = \int_0^h \pi^2 abcd \left(1 - \frac{y^2}{e^2}\right) dy = \quad (10)$$

$$= \frac{\pi^2 abcd}{2} \left(h - \frac{2}{3e^2} \cdot h^3 + \frac{1}{5e^4} \cdot h^5 \right);$$

$$V_{el}^{(5)} = \frac{4}{15} \pi^2 abcde; \quad (11)$$

$$V_s^{(5)}(h) = \frac{2 \cdot \pi^2 \cdot r^5}{15} \cdot \left(2 - 3 \cdot \frac{h}{r} + \left(\frac{h}{r}\right)^3 \right); \quad (12)$$

$$V_c^{(5)}(h) = \frac{\pi^2 \cdot (r^2 - h^2)^2 \cdot h}{10}, \quad (13)$$

$$\text{где } h = \sqrt{\frac{e^2 \cdot (r^2 - a^2)}{e^2 - a^2}} = r \sqrt{\frac{1 - \left(\frac{a}{r}\right)^2}{1 - \left(\frac{a}{e}\right)^2}};$$

$a < b < c < d < e$ – полуоси эллипсоида.

Подстановка (10) – (13) в формулы (5) и (6) позволяет получить $p = f(r, a, b, c, d, e)$, однако ввиду высокой вычислительной сложности данная зависимость имеет лишь теоретическое значение. На практике для определения предельных радиусов поиска ближайшего вектора по заданным значениям вероятности p удобно воспользоваться следующими формулами:

$$p(r) = \frac{V^{(5)}(r)}{V_{el}^{(5)}} = \frac{15}{8} \cdot \frac{r}{r_{\max}} - \frac{5}{4} \cdot \left(\frac{r}{r_{\max}}\right)^3 + \frac{3}{8} \cdot \left(\frac{r}{r_{\max}}\right)^5, \quad (14)$$

$$p(h) = \frac{V^{(5)}(h)}{V_{el}^{(5)}} = \frac{15}{8} \cdot \frac{h}{r_{\max}} - \frac{5}{4} \cdot \left(\frac{h}{r_{\max}}\right)^3 + \frac{3}{8} \cdot \left(\frac{h}{r_{\max}}\right)^5. \quad (15)$$

Так, используя соотношение (14), можно найти:

$$\frac{r}{r_{\max}} \Big|_{p=0,95} = 0,707, \quad \frac{r}{r_{\max}} \Big|_{p=0,9} = 0,621,$$

$$\frac{r}{r_{\max}} \Big|_{p=0,85} = 0,506, \dots$$

Это означает, что для определения ближайшего вектора КК по ВЕМ с вероятностью $p = 0,95$ достаточно пройти лишь 70 процентов расстояния и т.д.

Формулы (14) и (15) определяют соответственно верхнюю и нижнюю границу вероятности p_{\lim} , что следует из их геометрической интерпретации, приведенной на рис. 6.

Метод поиска вектора КК, ближайшего к входному x по ВЕМ из числа кандидатов $S \subset K$, упорядоченных по возрастанию расстояния r на втором этапе с заданной вероятностью $p < 1$ реализуется следующим образом:

1. Определяют весовые коэффициенты входного вектора $w_k \geq 1, k = \overline{1, M}$.

2. Инициализация: $i = 1, r_{\max} = INF$.

3. По формулам вычисления вероятности (14) либо (15) вычисляют значение r/r_{\max} для заданного значения вероятности p_{\lim} .

4. Для i -го вектора из списка кандидатов $i = \overline{1, |K|}$ вычисляют расстояние по ВЕМ $r_i^{(WE)}$

$$r_i^{(WE)} = \sum_{j=1}^M [w_j(x_j - y_j)]^2. \quad (16)$$

5. Если $\frac{r_i^{(E)}}{r_{\max}} \geq \lim$, где $r_i^{(E)} = \sum_{j=1}^M r_j^{(E)}$, $r_j^{(E)} = (x_j - y_j)^2$; поиск завершают, $|C| = i$. Если $r_i^{(WE)} < r_{\max}$, присваивают $r_{\max} = r_i^{(WE)}$.

6. $i = i + 1$, переход к п. 4.

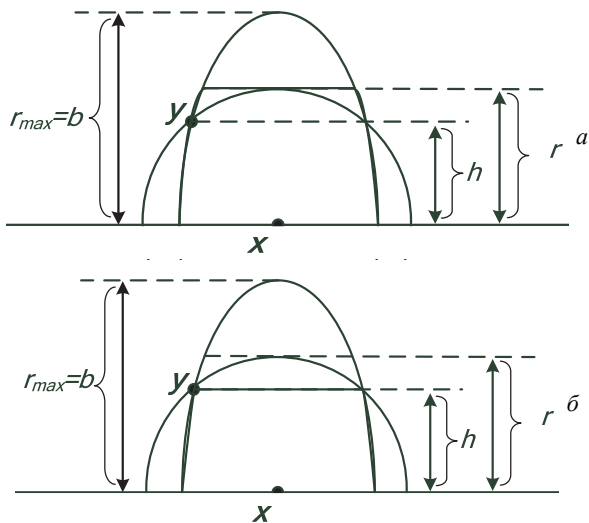


Рис. 6. Геометрическая интерпретация зависимости вероятности от расстояния согласно формулам (14) и (15)

Данная схема поиска может быть использована как при отдельном, так и при многоэтапном квантовании. В последнем случае на первых стадиях для нахождения нескольких кандидатов используется процедура поиска ближайшего вектора на основе kd -дерева [11]. На заключительной стадии для поиска ближайшего вектора может использоваться двухэтапная стратегия.

Экспериментальные результаты

Для экспериментального исследования разработанного метода ВДКП использована общедоступная часть англоязычного акустического корпуса *TIMIT*. Тренировочная последовательность состояла из 90 тыс. векторов LSF , полученных на основе модели линейного предсказания десятого порядка. Тестовая последовательность состояла из 15 тыс. векторов LSF ,

отличающихся от векторов тренировочной последовательности. Длина фрейма составляла 20 мс. Расстояние измерялось по ВЕМ с использованием весов разрядов, вычисленных по спектральной чувствительности [4].

В качестве базы для сравнения выбрана схема отдельного векторного квантования (РВК) с разбиением 10-мерного вектора LSF на два подвектора размерности $M = 5$. Такое же разбиение использовалось при отдельном векторном дельта-квантовании с предсказанием (РВДКП). При использовании многоэтапного квантования процесс кодирования происходил в два этапа, размеры обеих КК были одинаковы. Моделировались два варианта поиска: последовательный, при котором для поиска на втором этапе использовался один ближайший вектор, полученный на первом этапе, и поиск по дереву, при котором на втором этапе использовались N_1 векторов, полученных на первом этапе. Соответствующие схемы обозначены как многоэтапное векторное дельта-квантование с предсказанием (МЭВДКП) с последовательным поиском (ПП) и МЭВДКП с древовидным поиском (ДП).

В табл. 3 приведены результаты, полученные при использовании для квантования входного вектора 24, 22 и 20 бит на один фрейм. Производительность квантования оценивалась по спектральному искажению.

Таблица 3. Производительность квантования спектральной информации

Схема квантования	24 бита/фрейм			22 бита/фрейм			20 бит/фрейм		
	SD, дБ	% превышений		SD, дБ	% превышений		SD, дБ	% превышений	
		SD > 2 дБ	SD > 4 дБ		SD > 2 дБ	SD > 4 дБ		SD > 2 дБ	SD > 4 дБ
РВК	1,18	1,96	0,00	1,34	6,28	0,06	1,51	15,23	0,16
РВДКП	0,93	1,45	0,00	1,06	3,56	0,03	1,22	7,11	0,10
МЭВДКП с ПП	0,94	1,45	0,00	1,06	3,57	0,03	1,23	7,13	0,11
МЭВДКП с ДП, $N_1 = 10$	0,92	1,27	0,00	1,01	2,23	0,01	1,10	3,63	0,05
МЭВДКП с ДП, $N_1 = 30$	0,90	1,12	0,00	0,98	2,04	0,00	1,08	3,42	0,03

Как видно из табл. 3, условия прозрачности полностью выполняются при использо-

вании для описания спектральной информации 22 бит на один фрейм сигнала и частично 20 бит/фрейм.

Вычислительная сложность метода ВДКП практически совпадает с РВК. Лишь при использовании МЭВДКП с ДП количество операций увеличивается, что обусловлено особенностями схемы поиска, а не метода ВДКП.

На рис. 7 представлена ошибка квантования как функция от номера итерации в процессе создания КК. Ошибка вычислялась как значение расстояния между входным вектором и ближайшим вектором КК и усреднялась по всем фреймам. Уже на пятой итерации ошибка квантования изменяется лишь на 0,7 процента.

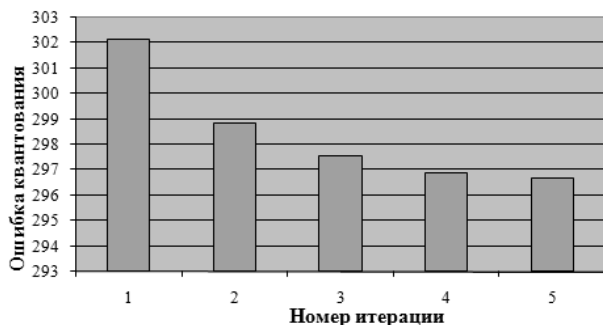


Рис. 7. Зависимость ошибки квантования от номера итерации

Оценка эффективности двухэтапной стратегии поиска ближайшего вектора с заданной вероятностью осуществлялась по спектральному искажению и количеству операций, выполняемых в процессе поиска. Полученные результаты приведены в табл. 4:

Таблица 4. Производительность двухэтапного поиска ближайшего вектора по ВЕМ на основе *kd*-дерева для РВК (24 бит/фрейм)

Заданное значение вероятности выбора ближайшего вектора КК, p	1	0,95	0,9	0,85	
Спектральное искажение, SD (дБ)	1,18	1,19	1,20	1,21	
К-во векторов КК, к которым необходимо вычислить расстояние, $ C = C_1 + C_2 $	КК1: $ C_1 $	11,9	2,7	1,8	1,4
	КК2: $ C_2 $	37,9	12,9	8,2	5,8
Общее количество операций при поиске ($M = 5^2$, $n = 4096$)	I этап	6870	3270	2070	1690
	II этап	1096	343	220	158

В табл. 4 количество совершенных при поиске операций N оценивалось следующим образом:

– полный поиск по ВЕМ (вычисление расстояния ко всем векторам КК согласно (16)):

$$N = 4 \cdot M \cdot n = 4 \cdot 10 \cdot 4096 = 163840 \text{ (операций);}$$

– двухэтапный поиск на основе *kd*-дерева:

$$N = N_1 + N_2 = (N'_1 + N''_1) + N_2,$$

где N'_1 – количество операций при поиске одного ближайшего по ЕМ вектора КК на основе *kd*-дерева; N''_1 – количество операций при поиске дополнительно $|K - 1|$ векторов КК по *kd*-дереву; $N_2 = (4 \cdot M + 2) \cdot |C|$, $|C| = |C_1| + |C_2|$ – количество операций при выборе ближайшего по ВЕМ вектора КК среди $|C|$ кандидатов.

Таким образом, снижение вычислительной сложности делает возможной реализацию предложенных методов на процессорах с ограниченной производительностью.

Заключение. Разработанный метод ВДКП позволяет повысить производительность квантования параметров речевого сигнала благодаря уменьшению динамического диапазона входных векторов, используемых для создания КК и, соответственно, уменьшению ошибки кластеризации. Это достигается за счет усложнения подготовительного этапа, требующего итерационной процедуры создания КК. Экспериментальная проверка разработанного метода показала, что условия транспарентности выполняются для МЭВДКП с поиском по дереву при использовании для квантования 20 бит на один фрейм речевого сигнала.

Предложена двухэтапная стратегия быстрого поиска векторов в кодовых книгах, позволяющая уменьшить количество операций, которое для вероятностей $p = 1$ и $p = 0,95$ составило 4,9 и 2,2 процентов соответственно в сравнении с полным перебором.

1. <http://www.computerra.ru/xterra/homo/21659/>
2. Paliwal K.K., Atal B.S. Efficient vector quantization of LPC parameters at 24 bits/frame // IEEE Transaction on Speech and Audio Processing. – 1993. – 1, N 2. – P. 3–14.
3. Lahouti F., Khandani A.K. Quantization of LSF parameters using a trellis modeling // Ibid. – 2003. – 11, N 5. – P. 400–412.

4. *Hai Le Vu., Laszlo Lois* Efficient Distance Measure for Quantization of LSF and Its Karhunen–Loeve Transformed Parameters // IEEE Transactions on speech and audio processing. – Nov. 2000. – 8, N 6. – P. 744–746.
 5. *Chu W.C.* Speech Coding Algorithms: Foundation and Evolution of Standardized Coders // NY: John Wiley & Sons, Inc, 2003. – 558 p.
 6. *Метод* кластеризації на основі послідовного запуску k -середніх з обчисленням відстаней до активних центроїдів / Ткаченко О.М., Біліченко Н.О., Грійо Тукало О.Ф. та ін. // Реєстрація, зберігання і обробка даних. – 2012. – № 1, Т. 14. – С. 25–34.
 7. *Agrell E.* Spectral coding by fast vector quantization // Proc. IEEE Workshop on Speech Coding for Telecommunications. – Sainte-Adèle, Québec, Canada, 1993. – P. 61–62.
 8. *Arya S., Mount D.M.* Algorithms for fast vector quantization / Ed. J.A. Storer, M. Cohnors // Proc. of DCC '93: Data Compression Conf., IEEE Press. – 1993. – P. 381–390.
 9. *Zhou J., Shoham Y., Akansu A.* Simple Fast Vector Quantization of the Line Spectral Frequencies // Image Compression and Encryption Technologies. – 2001. – 4551. – P. 274–282.
 10. *Ткаченко О.М., Грійо Тукало О.Ф.* Двоетапна стратегія пошуку в векторних кодових книгах для ущільнення мовлення // Вісн. Вінницького політехн. інту. – Вінниця: ВНТУ, 2011. – № 3. – С. 194–201.
 11. *Ткаченко О.М., Грійо Тукало О.Ф.* Пошук векторів у кодових книгах при ущільненні мовлення на основі бінарного дерева // Інформаційні технології та комп'ютерна інженерія. – 2011. – № 1. – С. 38–44.
- E-mail: AlexTk1960@gmail.com, xxmargox@gmail.com, alexdz47@gmail.com*
© А.Н. Ткаченко, О.Ф. Грійо Тукало, А.В. Дзись, 2013



**Ответственные за организацию и авторскую подготовку статей
по материалам конференций УкрОбраз
Васильева Н.Б. и Сажок Н.Н.**