

Поиск ближайшего соседа с вычислением расстояния по взвешенной евклидовой метрике в задачах сжатия речевых сигналов

*Ткаченко А.Н., кандидат технических наук, доцент
Грийо Тукало О.Ф.*

Рассматривается задача повышения производительности компьютерных систем для обработки мультимедийной, в частности, речевой информации за счет уменьшения времени поиска ближайшего вектора в словаре большого размера. Разработанный подход базируется на использовании *kd*-деревьев и сочетает преимущества методов быстрого поиска ближайшего соседа с применением взвешенной евклидовой метрики. Достигнутое снижение вычислительной сложности делает возможной реализацию предложенного метода на процессорах с ограниченной производительностью.

• сжатие речевых сигналов • линейные спектральные частоты
• векторная квантизация • словари • *kd*-дерево • спектральное искажение • взвешенная евклидова метрика

The problem of increasing the performance of computer systems for multimedia processing, in particular voice data, by reducing the nearest vector quantization time in a large dictionary. The approach based on the *kd*-tree is developed that combines the advantages of both: fast vector search and using the weighted Euclidean metric. The achieved reduction in computational complexity makes it possible to implement the proposed method on processors with limited capacity.

• *compressing speech signals* • *linear spectral frequencies* • *vector quantization* • *dictionaries* • *kd-tree* • *spectral distortion* • *weighted Euclidean metric*

ВВЕДЕНИЕ

Благодаря бурному развитию методов записи и хранения данных, объемы информации, передаваемые и обрабатываемые в компьютерных системах, в последнее время существенно выросли. В настоящее время Интернет объединяет сотни миллионов серверов, на которых размещены миллиарды различных сайтов и отдельных файлов с разнообразной информацией. Следует отметить, что в современных сетях основной объем трафика приходится на мультимедийную, в частности, аудио- и речевую информацию. Объемы данных столь значительны, что человек не в состоянии проанализировать их самостоятельно, поэтому необходимость автоматизации



процессов анализа, в частности поиска данных, вполне очевидна. В настоящее время поиск имеет большое значение, он является основой для решения широкого круга задач, а именно: информационный поиск, сжатие данных, распознавание и классификация образов, кодирование изображений, звуковых данных и т. д. Необходимость обеспечения функционирования в реальном масштабе времени накладывает жесткие требования к быстродействию микропроцессорных устройств, входящих в состав систем обработки речевых сигналов. Таким образом, существует проблема недостаточной производительности компьютерных систем, предназначенных для обработки мультимедийной информации. Повышение скорости поиска данных позволяет существенно повысить эффективность работы таких систем.

Задача поиска ближайшего соседа заключается в нахождении среди множества элементов, расположенных в многомерном метрическом пространстве, элементов, близких к заданному согласно некоторой функции близости. Уменьшение вычислительной сложности поиска ближайшего соседа в словарях большого размера рассматривалось в [1-3]. Применение приведенных в этих работах методов позволяет сократить количество операций и, соответственно, время поиска в 20...30 раз. Однако практическая ценность этих методов существенно ограничивается невозможностью применения взвешенной евклидовой метрики (ВЭМ) [4], поскольку все они базируются на предварительной обработке данных в словарях, когда значения весов еще неизвестны.

Целью работы является повышение производительности функционирования компьютерной системы за счет уменьшения вычислительной сложности поиска данных, что достигается благодаря реализации нового подхода, который сочетает преимущества методов быстрого поиска ближайшего соседа с применением ВЭМ. Суть его заключается в том, что на первом этапе на основе быстрого поиска по kd -дереву по невзвешенной евклидовой метрике (НЭМ) отбираются кандидаты, из числа которых на втором этапе с использованием ВЭМ отбирается один, ближайший к входному с заранее заданной вероятностью p .

РАЗРАБОТКА МЕТОДА БЫСТРОГО ПОИСКА БЛИЖАЙШЕГО ВЕКТОРА ПО ВЭМ НА ОСНОВЕ kd -ДЕРЕВА

Пусть $K = \{y_1, y_2, \dots, y_k\}$ – множество векторов, содержащихся в словаре. Суть разработанной двухэтапной стратегии поиска, схематически представленной на рис.1, состоит в том, что:

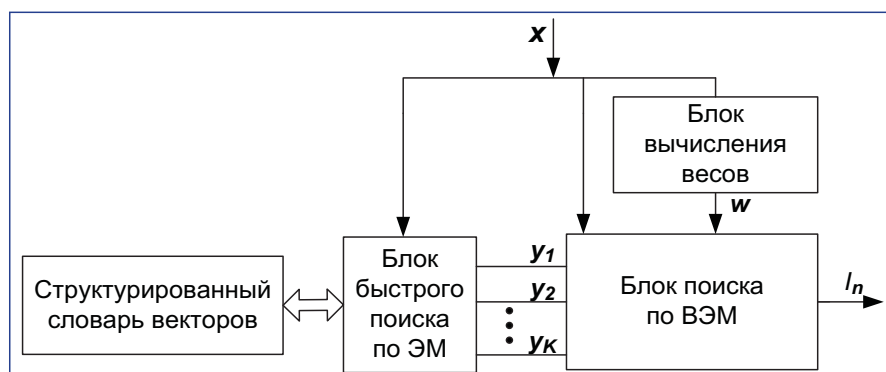


Рис. 1. Структурная схема двухэтапной стратегии поиска векторов в словаре

1. На первом этапе выполняется так называемый быстрый поиск в упорядоченном определенным образом словаре, в процессе которого по НЭМ выбирается множество $C \subset K$ векторов (кандидатов), упорядоченных по возрастанию расстояния к входному вектору X .

$$C = \{y_1, y_2, \dots, y_t\} \mid |C| = t, t \leq k. \quad (1)$$

2. На втором этапе с использованием ВЭМ с отобранного множества кандидатов выбирается один вектор, ближайший к входному (с заданной вероятностью p).

Поскольку по ВЭМ размерности имеют разный вес, вектор из словаря, который является ближайшим по НЭМ, может оказаться не лучшим при учете весов. Такая ситуация для двумерного случая показана на рис. 2:

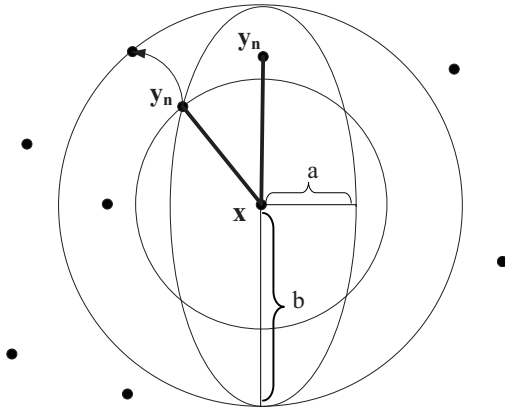


Рис. 2. Поиск ближайшего соседа по ВЭМ в двумерном пространстве

Как видно из рисунка, точка y_n (вектор в двумерном пространстве) является ближайшей по НЭМ, но в результате использования ВЭМ ближайшей оказывается точка y'_n , поскольку для нахождения вектора ближайшего к входному по ВЭМ, в процессе поиска достаточно охватить множество кандидатов, ограниченное кругом радиуса $r_{max} = \max(a, b)$, что детально рассматривалось в [5].

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

1. Вероятность того, что текущий вектор есть ближайшим по ВЭМ, уменьшается с ростом расстояния по НЭМ от вектора-кандидата к входному. Таким образом, значительная часть вычислений расстояний необходима лишь для того, чтобы убедиться в отсутствии лучшего вектора.
2. Пропуск в некоторых случаях ближайших векторов по ВЭМ не приводит к заметному увеличению спектрального искажения [4].

Объясняется это тем, что вместо ближайшего вектора, как правило, выбирается вектор, достаточно близкий к входному. Так, экспериментальные данные [3], показывают, что пропуск ближайших векторов в 5 – 10% фреймов увеличивает спектральное искажение лишь на 0,01 – 0,02 дБ.

Таким образом, нет необходимости вычислять расстояние по ВЭМ ко всем векторам, попадающим в круг (гипершар) радиусом r_{max} , а, задав некоторое значение вероятности p^* , следует найти величину r^* , определяющую радиус поиска, в пределах которого с вероятностью $p \geq p^*$ находится вектор, ближайший к входному по ВЭМ. Для этого необходимо найти зависимость $r = f(p, r_{max})$ или $p = f(r/r_{max})$.

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:

$$p(r) = \frac{V^{(M)}(r)}{V^{(M)}_{el}}, p(r) = \begin{cases} 0, & r = 0 \\ 1, & r \geq r_{max} \end{cases} \quad (2)$$

Геометрическая интерпретация приведенной зависимости для двумерного случая показана на рис. 3:

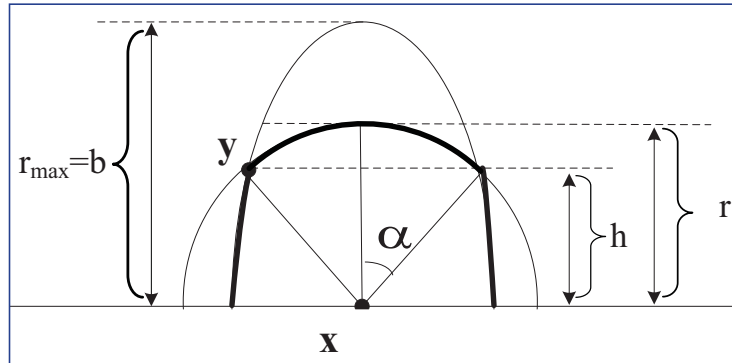


Рис. 3. Геометрическая интерпретация зависимости вероятности от расстояния

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

$$V^{(M)}(h) = V_{el}^{(M)}(h) + V_s^{(M)}(h) - V_c^{(M)}(h), \quad (3)$$

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

$$V_{el}^{(M)}(h) = \int_0^h V_{el}^{(M-1)}(y) dy, \quad (4)$$

При $h = b$, где b – максимальная полуось эллипса, формула (4) позволит получить полный объем полуэллипса.

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

$$V_s^{(M)}(h) = \frac{2\pi r^2}{M} \cdot V_s^{(M-2)} - \frac{(M-1) \cdot r \cdot \sin^{M-3} \alpha \cdot \cos \alpha}{M \cdot (M-2)} \cdot V_b^{(M-1)}, \quad (5)$$

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 .

$$V_c^{(M)}(h) = \frac{V_b^{(M-1)}}{M} \cdot h. \quad (6)$$

Для случая $M = 5$ (разбиение 10-мерного вектора линейных спектральных частот (LSF) на два подвектора при раздельном квантовании) формулы (4) – (6) принимают вид:

$$V_{el}^{(5)}(h) = \int_0^h \pi^2 abcd \left(1 - \frac{y^2}{e^2}\right) dy = \frac{\pi^2 abcd}{2} \left(h - \frac{2}{3e^2} \cdot h^3 + \frac{1}{5e^4} \cdot h^5\right); \quad (7)$$

$$V_s^{(5)} = \frac{4}{15} \pi^2 abcde, \quad (8)$$

$$V_s^{(5)}(h) = \frac{2 \cdot \pi^2 \cdot r^5}{15} \cdot \left(2 - 3 \frac{h}{r} + \left(\frac{h}{r}\right)^3\right), \quad (9)$$

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

где $h = \sqrt{\frac{e^2 \cdot (r^2 - a^2)}{e^2 - a^2}} = r \sqrt{\frac{1 - (\frac{a}{r})^2}{1 - (\frac{a}{e})^2}}$; $a < b < c < d < e$ – полуоси эллипсоида.

Подстановка (7) – (10) в формулы (2) и (3) позволяет получить $p = f(r, a, b, c, d, e)$, однако из-за высокой вычислительной сложности данная зависимость имеет только теоретическое значение. На практике предельные радиусы поиска ближайшего вектора по заданному значению вероятности p^* удобно определить для областей, показанных на рис. 4:

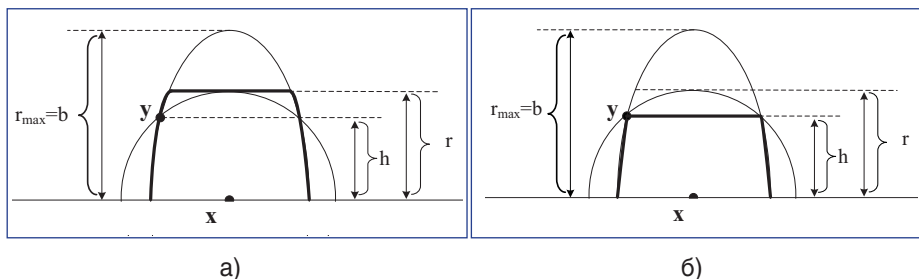


Рис. 4. Геометрическая интерпретация зависимости вероятности от расстояния согласно формул (11) и (12)

В этом случае получаем явную зависимость $p = f(r, r_{max})$, заданную формулами (11), (12):

$$p(r) = \frac{V^{(5)}(r)}{V_{el}^{(5)}} = \frac{15}{8} \cdot \frac{r}{r_{max}} - \frac{5}{4} \cdot \left(\frac{r}{r_{max}}\right)^3 + \frac{3}{8} \cdot \left(\frac{r}{r_{max}}\right)^5 \quad (11)$$

$$p(h) = \frac{V^{(5)}(h)}{V_{el}^{(5)}} = \frac{15}{8} \cdot \frac{h}{r_{max}} - \frac{5}{4} \cdot \left(\frac{h}{r_{max}}\right)^3 + \frac{3}{8} \cdot \left(\frac{h}{r_{max}}\right)^5 \quad (12)$$

Так, используя соотношение (11), можно найти:

$$\left. \frac{r}{r_{max}} \right|_{p^*=0.95} = 0,707, \left. \frac{r}{r_{max}} \right|_{p^*=0.9} = 0,621, \left. \frac{r}{r_{max}} \right|_{p^*=0.85} = 0,506.$$

Это означает, что, например, для определения ближайшего вектора в словаре по ВЭМ с вероятностью $p^* = 0,95$ достаточно пройти лишь 70% расстояния и т.д. Формулы (11) и (12) определяют соответственно верхнюю и нижнюю границу вероятности p^* , что следует из их геометрической интерпретации (см. рис. 4).

Метод поиска вектора из словаря, ближайшего к входному по ВЭМ из числа кандидатов $C \subset K$, упорядоченных по возрастанию расстояния r на втором этапе с заданной вероятностью $p^* < 1$ реализуется следующим образом:

1. Определяются весовые коэффициенты входного вектора $w_k \geq 1$, $k = \overline{1, M}$.
2. Инициализация: $i = 1$, $r_{max} = INF$.
3. По формулам вычисления вероятности (11) или (12) вычисляется значение $\frac{r}{r_{max}}$ для заданного значения вероятности p^* .
4. Для i -го вектора из списка кандидатов $i = \overline{1, |K|}$ вычисляется расстояние по ВЭМ $r_i^{(WE)}$:
$$r_i^{(WE)} = \sum_{j=1}^M [w_j(x_j - y_j)]^2; \quad (13)$$
5. Если $\frac{r_i^{(E)}}{r_{max}^{(E)}} \geq p^*$, где $r_i^{(E)} = \sum_{j=1}^M r_j^{(E)}$, $r_j^{(E)} = (x_j - y_j)^2$; поиск завершается, $|C|=i$. Если $r_i^{(WE)} < r_{max}$, присваивается $r_{max} = r_i^{(WE)}$;
6. $i = i + 1$, переход к п. 4.



Для выполнения быстрого поиска векторов в словаре на первом этапе векторы было упорядочено на основе kd -дерева (k – мерное дерева). Kd -дерево – это бинарное дерево (БД), в котором каждая вершина задает разбиение пространства на два подпространства некоторой плоскостью, проходящей через нее [6]. В kd -дереве кроме корневой присутствуют два типа вершин: терминальные и нетерминальные (узлы).

Математически задача поиска ближайшего соседа по kd -дереву формулируется так: дан некоторый M -мерный вектор; необходимо найти вершину kd -дерева v' , чтобы выполнялось условие:

$$r^E(v', v) = \min \{r^E(v_i, v)\}, i = (\overline{1, n}), v_i \in \mathbf{V}, \quad (15)$$

Использование kd -деревьев позволяет существенно уменьшить количество необходимых измерений, необходимых для поиска ближайшего соседнего вектора (в среднем $\log_2 n$ вместо n при полном поиске). Вместе с тем растёт погрешность квантования и, соответственно, спектральное искажение, поскольку поиск по kd – дереву не гарантирует нахождения действительно ближайшего вектора по формуле (15) [6]. В работе [7] предложена усовершенствованная процедура поиска по kd -дереву, впервые введенная при кодировании изображений, которую предлагается применить для поиска ближайшего вектора в словаре в процессе кодирования речи с учетом специфики речевых сигналов. Иллюстрация поиска ближайшего вектора в словаре, упорядоченного на основе kd -дерева, для двумерного случая показана на рис. 5.

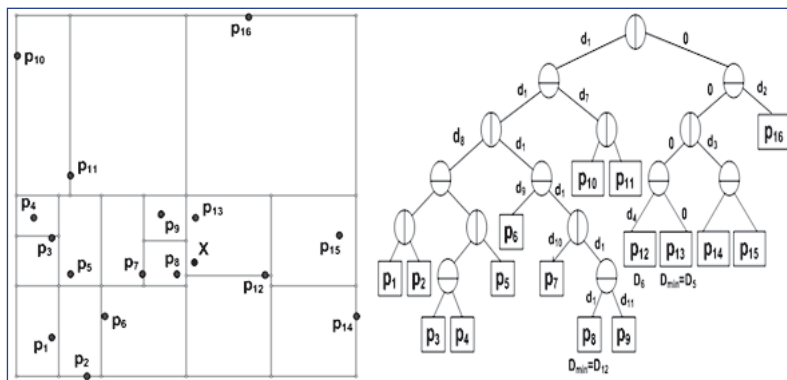


Рис. 5. Иллюстрация поиска ближайшего вектора по kd -дереву для двумерного случая

Чтобы обеспечить нахождение ближайшего вектора, поиск, кроме прямой фазы поиска (спуска по дереву), должен иметь также обратную. Во время прямого поиска фиксируются все расстояния к узлам d_i . Прямая фаза завершается вычислением расстояния к соответствующей терминальной вершине $D_k = D_{min}$. После этого начинается обратная фаза поиска, при этом вычисляются расстояния D_k только к тем вершинам дерева, которые могут обеспечить выполнение $D_k < D_{min}$. Если это условие выполняется, считают, что $D_{min} = D_k$ [6].

Приведенная процедура поиска обеспечивает получение не одного ближайшего вектора, а некоторого множества векторов, упорядоченных по возрастанию расстояния согласно формуле (1). Это возможно за счет того, что данные о расстояниях к уже пройденным терминальным вершинам сохраняются, и одновременно осуществляется их упорядочивание по возрастанию расстояния до входного вектора. Благодаря этому дополнительное нахождение нескольких ближайших векторов не требует больших вычислительных затрат.

ЭКСПЕРИМЕНТАЛЬНЫЕ РЕЗУЛЬТАТЫ

Для экспериментального исследования двухэтапной стратегии поиска использовалась общедоступная часть англоязычного акустического корпуса TIMIT. Тренировочная последовательность состояла из 90 000 векторов LSF, полученных на основе модели линейного прогнозирования десятого порядка [8]. Тестовая последовательность состояла из 15 000 векторов LSF, отличных от векторов тренировочной последовательности. Длина фрейма составляла 20 мс. Расстояние измерялось по ВЭМ с использованием весов разрядов, вычисленных по спектральной чувствительности [9]. Размер словаря составил 4096 векторов размерности $M = 5$.

На рис. 6 приведены результаты проверки соответствия экспериментальным результатам предложенных математических соотношений (3) – (5) для оценки вероятности нахождения ближайшего вектора в словаре по ВЭМ на расстоянии r . Результаты усреднялись для 10 720 фреймов.

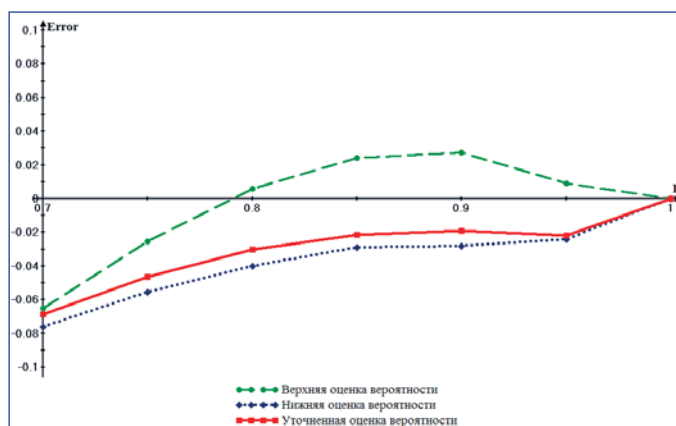


Рис. 6. Погрешность оценивания вероятности

На рис. 6 погрешность определяется по формуле:

$$Error = p_{expr} - p_{teor}$$

где p_{teor} – значения вероятностей, вычисленные на основе предложенных математических соотношений (3) – (5); p_{expr} – значения вероятностей, полученные экспериментально.

Как видно из приведенного рисунка, для больших значений вероятности, которые являются наиболее важными с точки зрения практического применения, все теоретические оценки в достаточной степени совпадают с результатами, полученными экспериментально (погрешность меньше 0,02).

Таблица 1

Производительность двухэтапного поиска ближайшего вектора по ВЭМ на основе kd-дерева

Заданное значение вероятности выбора ближайшего вектора в словаре, p^*		1	0,95	0,9	0,85
Спектральное искажение, SD (дБ)		1,18	1,19	1,20	1,21
К-во векторов из словаря, к которым необходимо вычислить расстояние, $ C = C1 + C2 $	Словарь 1: $ C1 $	11,9	2,7	1,8	1,4
	Словарь 2: $ C2 $	37,9	12,9	8,2	5,8
Общее количество операций во время поиска ($M=5 \times 2$, $n=4096$)	I этап	6870	3270	2070	1690
	II этап	1096	343	220	158



Эффективность двухэтапной стратегии поиска ближайшего вектора с заданной вероятностью оценивалась по спектральному искажению [4] и количеству операций, выполняемых в процессе поиска. Полученные результаты приведены в табл. 1.

В табл. 1 количество выполненных при поиске операций N оценивалась следующим образом:

- полный поиск по ВЭМ (вычисление расстояния до всех векторов в словаре согласно (13)):

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

- 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 НЭМ вектора в словаре на основе kd-дерева; N_1'' – количество операций во время поиска дополнительно $|K-1|$ векторов в словаре по kd-дереву;

$N_2 = (4 \cdot M + 2) \cdot |C|$, $|C| = |C_1| + |C_2|$ – количество операций при выборе ближайшего по ВЭМ вектора из словаря среди $|C|$ кандидатов.

ВЫВОДЫ

Предложенная в статье двухэтапная стратегия поиска векторов в словарях больших размеров позволяет совместить преимущества метода быстрого поиска ближайшего соседа на основе kd-дерева с применением взвешенной евклидовой метрики. Для вероятности выбора ближайшего соседа $p^* = 0,95$ количество операций уменьшается в 45,5 раз по сравнению с полным перебором, что позволяет реализовать данный метод на процессорах с ограниченной производительностью.

ЛИТЕРАТУРА

1. Agrell E. Spectral coding by fast vector quantization / E. Agrell // Proc. IEEE Workshop on Speech Coding for Telecommunications. – Sainte-Adèle, Québec, Canada, 1993. – P. 61 – 62.
2. Arya S. Algorithms for fast vector quantization / S. Arya and D. M. Mount // In J. A. Storer and M. Cohn, editors, Proc. of DCC '93: Data Compression Conference, IEEE Press. – 1993. – P. 381–390. – ISBN:0-89871-329-3.
3. Zhou J., Shoham Y., Akansu A. Simple Fast Vector Quantization of the Line Spectral Frequencies // Image Compression and Encryption Technologies. – 2001. – Vol. 4551. – P. 274–282.
4. Paliwal K.K. Efficient vector quantization of LPC parameters at 24 bits/frame / K. K. Paliwal, B. S. Atal. // IEEE Transaction on Speech and Audio Processing. – 1993. – No. 2, vol. 1. – P. 3–14.
5. Ткаченко О.М. Двоетапна стратегія пошуку в векторних кодових книгах для ущільнення мовлення / О.М. Ткаченко, О.Ф. Грійо Тукало // Вісник Вінницького політехнічного інституту. – Вінниця: ВНТУ, 2011. – №3. – С.194-201. – ISSN 1997-9266.
6. Ткаченко О.М. Пошук векторів у кодових книгах при ущільненні мовлення на основі бінарного дерева / О.М. Ткаченко, О.Ф. Грійо Тукало // Інформаційні технології та комп'ютерна інженерія. – 2011. – №1. – С.38-44. – ISSN 1999-9941.

Ткаченко А.Н., Грийю Тукало О.Ф.

Поиск ближайшего соседа с вычислением расстояния по взвешенной евклидовой метрике...

7. *Arya S.* Algorithms for fast vector quantization / S. Arya and D.M. Mount // In J.A. Storer and M. Cohn, editors, Proc. of DCC '93: Data Compression Conference, IEEE Press. – 1993. – P. 381–390. – ISBN:0-89871-329-3.
8. *Chu W. C.* Speech Coding Algorithms: Foundation and Evolution of Standardized Coders / Wai C. Chu // NY.: John Wiley & Sons, Inc. – 2003 – 558 p. – ISBN 0-471-37312-5.
9. *Hai Le Vu.* Efficient Distance Measure for Quantization of LSF and Its Karhunen–Loeve Transformed Parameters / Hai Le Vu and Laszlo Lois // IEEE Transactions on speech and audio processing. — Nov. 2000. — No. 6, vol. 8.

Сведения об авторах:

Ткаченко Александр Николаевич,

кандидат технических наук, доцент, Винницкий национальный технический университет (ВНТУ), кафедра вычислительной техники (Украина, г. Винница). Область научных интересов: обработка речевых сигналов, системы сжатия речи, низкоскоростное кодирование речи. E-mail: alextk1960@gmail.com

Грийю Тукало Оксана Франсисковна,

аспирантка, ВНТУ, кафедра вычислительной техники (Украина, г. Винница.). Область научных интересов: обработка речевых и аудиоданных, быстрый поиск (поиск музыкальных фрагментов в базах данных мультимедиа), программирование. E-mail: xxmargoxh@gmail.com