

6. Potapov A.A. Metody obrabotki signalov i polej na osnove teorii fraktalov [Methods for processing signals and fields based on the theory of fractals]. *Trudy Pervoj Vserossijskoj NK «Metody i sredstva obrabotki informacii»*, Moscow, MGU Publ., 2003, pp. 559-565.
7. Morozov A.D. *Vvedenie v teoriyu fraktalov* [Introduction to the theory of fractals]. Moscow – Izhevsk, Institut komp'yuternyh issledovanij, 2002, pp. 82-108.
8. Asmuth S., Bloom J. A modular approach to key safeguarding. *Information Theory, IEEE Transactions on*, 1983, vol. 29, no. 2, pp. 208-210. doi: 10.1109/TIT.1983.1056651.
9. Shamir A. How to share a secret. *Communications of the ACM*, New York City: ACM, 1979, vol. 22, no. 11, pp. 612-613.
10. Blakley G.R. *Safeguarding cryptographic keys. Proceedings of the 1979 AFIPS National Computer Conference*, Monval, NJ, USA, AFIPS Press, 1979, pp. 313-317.
11. Mignotte M. How to Share a Secret. *Lecture Notes in Computer Science*, 1983, vol. 149, pp. 371-375. doi: 10.1007/3-540-39466-4_27.

Received 17.01.2017

УДК 621.395

МЕТОД СЖАТИЯ РЕЧЕВЫХ ДАННЫХ БЕЗ ПАУЗ НА ОСНОВЕ ОПТИМАЛЬНОГО КВАНТОВАНИЯ ПО УРОВНЮ КОЭФФИЦИЕНТОВ РАЗЛОЖЕНИЯ ОТРЕЗКОВ РЕЧЕВЫХ СИГНАЛОВ ПО СОБСТВЕННЫМ ВЕКТОРАМ СУБПОЛОСНЫХ МАТРИЦ

Жиляков Е.Г., Белов С.П., Белов Ал.С., Белов Ан.С., Медведева А.А.

Белгородский государственный национальный исследовательский университет, Белгород, РФ

В работе предложен метод сжатия речевых данных без пауз на основе оптимального квантования по уровню коэффициентов разложения отрезков речевых сигналов по собственным векторам субполосных матриц из m -информационных частотных интервалов с применением кодовых книг квазиоптимальных квантователей, применение которого позволяет в зависимости от величины разрядности исходных отсчетов указанных сигналов обеспечить коэффициент сжатия до 12 раз.

Ключевые слова: отрезок речевого сигнала, речевые данные, распределение энергии, субполосная матрица, собственные векторы субполосной матрицы, информационные частотные интервалы, кодовые книги квазиоптимальных квантователей, коэффициент сжатия

Постановка задачи

Проблема уменьшения объемов битовых представлений речевых данных при их хранении и передаче рассматривается в работах многих авторов, особенно специалистов в области телекоммуникаций, что подтверждается результатами анализа научно-технической литературы [1-4].

При этом отмечают два основных аспекта: необходимость обнаружения с последующим их кодированием пауз, возникающих между отдельными словами и в режиме диалога занимающих до 60% длительности исходных звукозаписей, и сокращение объемов битовых представлений собственно звуков речи без пауз. Существующие методы сжатия звуков речи без пауз с использованием грубого квантования по уровню основываются на психоакустической модели, что приводит к необходимости применения так называемых субполосных преобразований отрезков (векто-

ров) отсчетов речевых сигналов, позволяющих получить другие векторы, подвекторы которых отражают частотные свойства исходного вектора в выбранных диапазонах оси частот. Именно компоненты этих подвекторов подвергаются квантованию по уровню с различными шагами, чем достигается учет частотно-избирательных свойств человеческого слуха. В настоящее время для субполосного преобразования принято использовать процедуру прореживания выходных последовательностей КИХ-фильтров (фильтров с конечной импульсной характеристикой), настроенных на соответствующие участки оси частот. Такая процедура субполосного преобразования не является оптимальной в смысле минимума погрешностей аппроксимации спектров исходных векторов в выбранных частотных диапазонах, что приводит к увеличению погрешностей восстановления данных по квантованным значениям

и, как следствие, к ухудшению качества воспроизводимой речи.

В связи с этим авторами предлагается метод сжатия речевых данных без пауз, созданный с применением разработанного математического аппарата на основе собственных векторов субполосных матриц, позволяющего адекватно сформулировать вариационные условия и решить оптимизационные задачи обработки речевых данных.

Математические основы метода

Осуществляется обработка отдельных отрезков (векторов отсчетов) сигнала речи:

$$\vec{x} = (x_1, \dots, x_N)^T \quad (1)$$

в соответствии с выбранным равномерным разбиением полосы нормированных частот $0 \leq \omega \leq \pi$ на R интервалов V вида

$$\begin{aligned} V_{r2} - V_{r1} &= \pi / R; V_{r1+1} = V_{r2}; \\ r &= 1; 2 \dots R - 1 \end{aligned} \quad (2)$$

одинаковой ширины. В основе сжатия данных используется свойство концентрации энергии речевых сигналов в малой доле частотной полосы, что позволяет использовать аппроксимацию

$$\hat{\vec{x}} = a \sum_{r \in R_m} \vec{x}_r, \quad (3)$$

где

$$\vec{x}_r = A_r \vec{x}, \quad (4)$$

коэффициент a при сумме выбирается из условия $\|\hat{\vec{x}}_r\|^2 = m \|\vec{x}\|^2$, что дает

$$a = \frac{\sqrt{m} \|\vec{x}\|}{\left\| \sum_{r \in R_m} \vec{x}_r \right\|}, \quad (5)$$

где R_m – множество частотных интервалов минимальной суммарной ширины, для которых выполняются условия

$$\sum_{r \in R_m} P_r(\vec{x}) \cong m \|\vec{x}\|^2; \quad (6)$$

$$0,85 \leq m \leq 0,98; \quad (7)$$

где m – множество информационных частотных интервалов, а A_r – субполосная матрица, определяемая элементами:

$$A_r = \{a_{i,k}^r\}, i, k = 1, \dots, N;$$

$$a_{ik}^r = \frac{\sin(V_{2,r}(i-k)) - \sin(V_{1,r}(i-k))}{\pi(i-k)}; i \neq k.$$

Выше было указано, что соответствующая некоторому частотному интервалу компонента вида (4) допускает представление

$$\vec{x}_r = \sum_{i=1}^{J_r} \lambda_i^r \alpha_{ir} \vec{q}_i^r; \alpha_{ir} = (\vec{x}, \vec{q}_i^r), \quad (8)$$

где λ_i^r – собственные числа собственных векторов \vec{q}_i^r субполосной матрицы, принимающие значения $0 < \lambda_i^r \leq 1$. Подстановка (8) в (3) дает разложение по набору собственных векторов:

$$\hat{\vec{x}} = \sum_{r \in R_m} \sum_{i=1}^J \beta_{ir} \vec{q}_i^r, \quad (9)$$

где

$$\beta_{ir} = a \lambda_i^r \alpha_{ir}, i = 1, \dots, J_r. \quad (10)$$

Так как наборы собственных векторов \vec{q}_i^r предполагаются известными, то для восстановления исходного отрезка достаточно сохранять информацию о соответствующих коэффициентах разложения. Проведенные исследования показали, что мощность множества частотных интервалов ($\text{int } R_m$) почти для всех звуков русской речи удовлетворяет соотношению

$$\text{int } R_m \approx 0,3R. \quad (11)$$

Поэтому с учетом равенства $J \cong N / R$ получаем коэффициент сжатия за счет использования аппроксимации (3) (по количеству сохраняемых чисел):

$$CH = N / (\text{int } R_m J) \approx 3. \quad (12)$$

Следующий шаг заключается в применении к коэффициентам разложения квантования по уровню с малым их количеством. В общем виде процедура квантования описывается следующим образом, если выполняется условие:

$$\beta_{ir} \in \hat{O}_m = [\varphi_{m-1}, \varphi_m], \quad (13)$$

то положить

$$\beta_{ir}^* = d_m, m = 1, \dots, K, \quad (14)$$

где K – количество используемых уровней квантования.

Проблема заключается в оптимальном выборе границ отрезков в (13) и значений d_m в (14) в смысле минимизации погрешностей аппроксимации исходных данных квантованными значениями:

$$\varepsilon_i^2 = \sum_{m=1}^K \sum_{\beta_{ir} \in S_m} (\beta_{ir} - d_m)^2, (i=1, \dots, J), r \in R_m, \quad (15)$$

где S_m – множество значений β_{ir} , удовлетворяющих условию (13). В результате проведенных исследований было показано, что при заданных отрезках \hat{O}_m в (13) минимум правой части (15) достигается на множестве уровней квантования, равных соответствующим средним значениям:

$$d_m = \sum_{\beta_{ir} \in S_m} (\beta_{ir} / \text{int } S_m), m = 1; 2 \dots K \quad (16)$$

где $\text{int } S_m$ – мощность множества S_m (число попадающих в них значений β_{ir}).

Введем положительную неубывающую последовательность:

$$0 \leq z_k < z_{k+1}, k = 1; 2 \dots NK - 1, \quad (17)$$

$$NK = J \cdot \text{int } R_m, \quad (18)$$

причем

$$\begin{aligned} z_k &\in \{|\beta_{ir}| \} / \gamma, i = 1; 2 \dots J; r \in R_m \\ \gamma &= \max |\beta_{ir}|, \forall i \text{ и } r \in R_m. \end{aligned} \quad (19)$$

Показано, что выполнение условий

$$\sum_{m=1}^K I_m \bar{z}_m^2 = \max, \quad (20)$$

$$\sum_{m=1}^K I_m = NK, \quad (21)$$

где

$$\begin{aligned} \bar{z}_m &= \sum_{i=1}^{L_m} z_{L_{m-1}+i} / I_m, \\ L_m &= \sum_{i=1}^{m-1} I_i, L_0 = 0, \end{aligned} \quad (22)$$

а также выбор уровней квантования в виде

$$\hat{d}_m = \bar{z}_m, m = 1; 2 \dots K \quad (23)$$

дает минимум погрешности аппроксимации z_k квантованными значениями

$$z_k^* = \hat{d}_m, \quad (24)$$

когда выполняется условие

$$z_{L_{m-1}+1} \leq z_k \leq z_{L_{m-1}+I_m}. \quad (25)$$

Реально вместо операции (24) следует использовать кодирование

$$\text{cod} z_k = \log_2 m, \quad (26)$$

имея в виду, что номера уровней квантования целесообразно обозначать двоичными числами разрядности p , так что

$$K = 2^p. \quad (27)$$

Таким образом, количество уровней квантования целесообразно выбирать из множества (2; 4; 8 ...). В соответствии с этим в процессе проведения исследований был разработан алгоритм решения задачи (20), (21) с последовательным делением подпоследовательностей на две части, каждая из которых удовлетворяет этим условиям со своими параметрами I_m и d_m (так как деление на две последовательности любой длины несложно реализуется последовательным перебором).

Использование стандартизованных последовательностей вида (17)-(19) позволяет не хранить значения уровней (23), а для восстановления данных использовать уровни из заранее сформированной кодовой книги, удовлетворяющей условию

$$\sum_{m=1}^K I_m \cdot (\bar{z}_m - d_m^l)^2 = \min \forall D_l, \quad (28)$$

где

$$D_l = \{d_1^l, \dots, d_K^l\}, d_1^l < d_2^l < \dots < d_K^l. \quad (29)$$

Такие кодовые книги сформированы при $K = 2; 4; 8$ с учетом всех звуков русской речи с усреднением по множеству дикторов.

Для иллюстрации работоспособности и эффективности разработанных метода и алгоритма были проведены экспериментальные исследования, которые показали, что разборчивость речи сохраняется уже при $K = 2$. Таким образом, с учетом необходимости сохранения знакового разряда и значения γ достигаемый максимальный коэффициент сжатия может быть равен

$$CH_{\max} = 12N / (N + 12), \quad (30)$$

(в предположении 8-разрядности исходных отсчетов). То есть если N достаточно велико, то

$$CH_{\max} \cong 12. \quad (31)$$

Выводы

В результате проведенных исследований было установлено, что предлагаемый метод сжатия речевых данных без пауз на основе оптимального квантования по уровню коэффициентов разложения отрезков речевых сигналов по собственным векторам субполосных матриц из m -информационных частотных интервалов с применением кодовых книг квазиоптимальных квантователей позволяет в зависимости от величины разрядности исходных отсчетов указанных сигналов обеспечить коэффициент сжатия до 12 раз.

Необходимо также отметить, что с учетом сжатия пауз в речи, которые могут составлять более 60% от продолжительности диалога, общий коэффициент сжатия как за счет обнаружения и кодирования пауз, так и за счет квантования по уровню коэффициентов разложения отрезков речевых сигналов по собственным векторам субпо-

лосных матриц из m -информационных частотных интервалов может достичь величин 20- 25 раз.

Исследования частично финансировались в рамках гранта РФФИ №15-07-01463.

Литература

1. Сергиенко В.С., Баринов В.В. Сжатие данных, речи, звука и изображений в телекоммуникационных системах. М.: Радио Софт, 2009. – 360 с.
2. Сжатие данных в системах сбора и передачи информации. Под ред. В.А. Свириденко. М.: Радио и связь, 1985. – 184 с.
3. Сэломон Д. Сжатие данных, изображений и звука. М.: ТЕХНОСФЕРА, 2004. – 368 с.
4. Цифровая обработка и передача речи. Под ред. О.И. Шелухина. М.: Радио и связь, 2000. – 456 с.

Получено 25.03.2016

Жиляков Евгений Георгиевич, д.т.н., профессор, заведующий Кафедрой информационно-телекоммуникационных систем и технологий (ИТСТ) Белгородского государственного национального исследовательского университета (БелГУ). Тел. (8-472) 230-13-92. E-mail: zhilyakov@bsu.edu.ru

Белов Сергей Павлович, д.т.н., профессор Кафедры ИТСТ БелГУ. Тел. 8-980-323-61-04. E-mail: belov@bsu.edu.ru.

Белов Александр Сергеевич, к.т.н., доцент Кафедры ИТСТ БелГУ. Тел. 8-919-436-86-85. E-mail: belov_as@bsu.edu.ru

Белов Андрей Сергеевич, заведующий лабораторией Кафедры ИТСТ БелГУ. Тел. 8-980-323-61-14.

Медведева Александра Александровна, к.т.н., старший преподаватель Кафедры ИТСТ БелГУ. Тел. 8-960-637-55-23. E-mail: medvedeva_aa@bsu.edu.ru

METHOD OF VOICE DATA COMPRESSION WITHOUT A PAUSE BASED ON THE OPTIMAL LEVEL QUANTIZATION COEFFICIENTS OF THE EXPANSION PIECES OF SPEECH SIGNALS FOR EIGENVECTORS SUBBAND MATRIX

*Zhilyakov E.G., Belov S.P., Belov A.S., Belov A.S., Medvedeva A.A.
Belgorod State National Research University, Belgorod, Russian Federation
E-mail: belov@bsu.edu.ru*

Existing techniques of speech sounds compression without pauses, using a coarse quantization level based on a psychoacoustic model, which results in the need for so-called subband transforms segments (vectors) samples of speech signals, allowing them to obtain the other vectors subvectors which reflect the frequency properties of the input vector in the selected ranges frequency axis. That component of the sub-vectors is quantized by the level of different steps, thus, achieving the accounting frequency-selective properties of human hearing. Currently, for a subband it is common to use transformation process of output sequences of decimation FIR filters (finite impulse response) that are tuned to respective portions of frequency axis. This procedure of subband transformation is not optimal in the sense of minimizing the approximation error of spectra initial vectors in selected frequency ranges, which leads to increase of errors in data recovery in the quantized values and, as consequence, to degradation in the quality of reproduced speech. In this regard, the authors proposed a speech data compression method without pauses, created with the use of mathematical apparatus, developed based on eigenvectors subband matrix, allowing to adequately formulate variational conditions and solve the optimization problem of processing speech data.

The proposed method of speech compression without pauses allows providing compression ratios up to 12 times, depending on the starting bit of mentioned data samples.

Keywords: speech segment, speech data, the power distribution, the subband matrix, subband eigenvectors of subband matrix, information frequency intervals, codebook quasi-optimal quantizers, compression ratio

DOI: 10.18469/ikt.2017.15.1.02

Zhilyakov Evgeny Georgiyevich, Belgorod State National Research University, 85 Pobedy str., Belgorod, 308015, Russian Federation; the Head of Department of Information-telecommunication systems and technologies, Doctor of Technical Science, Professor. Tel.: +74722301392. E-mail: zhilyakov@bsu.edu.ru

Belov Sergey Pavlovich, Belgorod State National Research University, 85 Pobedy str., Belgorod, 308015, Russian Federation; Professor of the Department Information-telecommunication systems and technologies, Doctor of Technical Science, Professor. Tel.: +79803236104. E-mail: belov@bsu.edu.ru

Belov Alexander Sergeevich, Belgorod State National Research University, 85 Pobedy str., Belgorod, 308015, Russian Federation; Assistant Professor of the Department Information-telecommunication systems and technologies, PhD in Technical Sciences, Assistant Professor. Tel.: +79194368685. E-mail: belov_as@bsu.edu.ru.

Belov Andrey Sergeevich, Belgorod State National Research University, 85 Pobedy str., Belgorod, 308015, Russian Federation; the Chief of Laboratory of Department Information-telecommunication systems and technologies. Tel.: +79803236114.

Medvedeva Alexandra Alexandrovna, Belgorod State National Research University, 85 Pobedy str., Belgorod, 308015, Russian Federation; Senior Lecturer of the Department Information-telecommunication systems and technologies. Tel.: +79606375523. E-mail: medvedeva_aa@bsu.edu.ru

References

1. Sergienko V.S., Barinov V.V. *Szhatie dannyh, rechi, zvuka i izobrazhenij v telekommunikacionnyh sistemah* [Compression of data, speech, sound and images in telecommunication systems]. Moscow, Radio Soft Publ., 2009. 360 p.
2. Sviridenko V.A. *Szhatie dannyh v sistemah sbora i peredachi informacii* [Data compression in systems of information collection and transmission]. Moscow, Radio i svjaz Publ., 1985. 184 p.
3. Sjelomon D. *Szhatie dannyh, izobrazhenij i zvuka* [Compressing of data, images and sound]. Moscow, Tehnosfera Publ., 2004. 368 p.
4. Sheluhina O.I. *Cifrovaja obrabotka i peredacha rechi* [Digital processing and voice transmission]. Moscow, Radio i svjaz Publ., 2000. 456 p.

Received 25.03.2016

ТЕХНОЛОГИИ ТЕЛЕКОММУНИКАЦИЙ

УДК 535.3; 621.373.826; 681.7; 53.082.5

РАЗРАБОТКА МЕТОДИКИ ФОРМИРОВАНИЯ ПРЕЦИЗИОННЫХ МАКРОДЕФЕКТОВ В СТРУКТУРЕ КВАРЦЕВЫХ ВОЛОКОННЫХ СВЕТОВОДОВ

Андреев В.А.¹, Бурдин А.В.¹, Бурдин В.А.¹, Василец А.А.², Гаврюшин С.А.¹, Евтушенко А.С.¹,
Казаков В.С.¹, Морозов О.Г.², Севрук Н.Л.¹, Соколов Е.Д.¹

¹Поволжский государственный университет телекоммуникаций и информатики, Самара, РФ
²Казанский национальный исследовательский технический университет им. А.Н. Туполева, Казань, РФ
E-mail: bourdine@yandex.ru

В работе предложена методика формирования прецизионных макроструктурных дефектов типа «перетяжка» и «бочка» в кварцевых многомодовых оптических волокнах с помощью штатного комплекта полевого сварочного аппарата и его модифицированного соответствующим образом программного обеспечения. Приведены результаты ее экспериментальной апробации. Разработаны практические рекомендации по выбору установок параметров