

Е.И. ПРОХОРЕНКО, канд. техн. наук, доц. БелГУ (г. Белгород, Россия),
А.А. ЧЕРНОМОРЕЦ, канд. техн. наук, доц. БелГУ (г. Белгород, Россия),
А.В. ЭСАУЛЕНКО, БелГУ (г. Белгород, Россия)

РАЗРАБОТКА АЛГОРИТМА СЖАТИЯ РЕЧЕВЫХ ДАННЫХ С ИСПОЛЬЗОВАНИЕМ СУБПОЛОСНОГО ПРЕОБРАЗОВАНИЯ

Рассмотрена новая технология сжатия речевых данных с использованием субполосного преобразования, разработан и применен алгоритм сжатия речевых данных с использованием субполосного преобразования с учетом отличительных особенностей речевых данных (значительное количество пауз в сигнале, а также сосредоточение энергии сигнала в относительно узких частотных интервалах). Ил.: 4. Табл.: 1. Библиогр.: 10 назв.

Ключевые слова: сжатие речевых данных, субполосное преобразование, паузы в сигнале.

Постановка проблемы. На сегодняшний день актуальной задачей является создание и развитие новых информационных технологий, которые позволили бы ускорить процесс обмена информацией в обществе. Существенную долю этого обмена составляют речевые данные. Кроме того, зачастую возникает необходимость в хранении речевых данных на жестких носителях, например при записи аудиоконференций, выступлений лекторов, театральных постановок, радиоархивов и т.п. В настоящее время пользователями для сжатия речи, в основном, используются кодеки, предназначенные для сжатия звуковых данных, которые не учитывают специфические особенности речевых сигналов, такие как значительное количество пауз в сигнале, сосредоточение энергии сигнала в относительно узких частотных интервалах и так далее.

Таким образом, одной из важных задач реализации в информационно-телекоммуникационных системах речевого обмена является создание эффективных методов сжатия полученных на этапе регистрации речевых данных с возможностью дальнейшего воспроизведения исходных речевых сообщений с приемлемым качеством на выходе системы передачи информации.

Анализ литературы. Анализ работ многих авторов, посвященных данной тематике, например [1 – 5], показывает, что на сегодняшний день наиболее часто в литературе рассматривается метод сжатия речевых данных, основанный на использовании КИХ-фильтров. Недостатком такого подхода является наличие эффекта элайзинга, то есть наложения частот, и как следствие, ухудшение качества воспроизводимой речи.

Метод оптимального субполосного преобразования [6 – 10], позволяющий избежать недостатков КИХ-фильтрации, учитывает особенности речевого сигнала и позволяет осуществить избирательную обработку составляющих

сигнала, энергия которых сосредоточена в заданных частотных интервалах. Метод субполосного преобразования на основе КИХ-фильтрации получает все большее распространение в алгоритмах цифровой обработки речи.

Цель статьи – исследование и практическое применение метода оптимального субполосного преобразования для сжатия речевых данных.

Уменьшение объемов битовых представлений речевых сигналов основано на методе субполосного преобразования.

Процедура прореживания выходных последовательностей КИХ-фильтров 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. Поэтому целесообразно использовать метод обратимого субполосного преобразования, оптимального в смысле минимума погрешностей аппроксимации спектров исходного вектора в выбранных частотных диапазонах [1 – 4, 6].

Сущность субполосного преобразования состоит в том, что для отрезков сигнала вычисляется вектор $\vec{y} = (\vec{y}_1, \vec{y}_2, \dots, \vec{y}_R)$, состоящий из подвекторов $\vec{y}_r = (y_{1r}, y_{2r}, \dots, y_{Jr})$, которые отражают частотные свойства исходного сигнала в некотором частотном интервале. При этом в данном случае ось частот разбивается на R равновеликих частотных интервалов:

$$V_r = [-v_r, -v_{r-1}) \cup [v_{r-1}, v_r), \quad v_{r-1} < v_r \leq \pi; \quad v_0 = 0; \quad v_r = \pi; \quad r = 1, \dots, R.$$

Субполосное преобразование осуществляется по формуле:

$$\vec{y} = \mathbf{A} \mathbf{A} \vec{x}, \tag{1}$$

где $\mathbf{A} \mathbf{A}$ – блочная матрица вида:

$$\mathbf{A} \mathbf{A} = \begin{bmatrix} \sqrt{L_{11}} Q_{11}^T \\ \dots \\ \sqrt{L_{1R}} Q_{1R}^T \end{bmatrix},$$

где $Q_{1r} = \{\vec{q}_1^r, \vec{q}_2^r, \dots, \vec{q}_J^r\}$ – матрица собственных векторов матрицы A_r , $L_{1r} = \text{diag}(\lambda_1, \dots, \lambda_J)$ – вектор собственных чисел, J – количество собственных чисел матрицы A_r .

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

$$a_{ik}^r = \begin{cases} \frac{\sin[\nu_r(i-k)] - \sin[\nu_{r-1}(i-k)]}{\pi(i-k)}, & \text{если } i \neq k, \\ \frac{\nu_r - \nu_{r-1}}{\pi}, & \text{если } i = k. \end{cases}$$

Если в качестве вектора исходных значений \bar{x} используется отрезок речевого сигнала длительностью N , тогда вектор субполосного преобразования вычисляется по формуле (1).

С использованием полученных значений подвекторов субполосного преобразования вычисляется распределение энергии сигнала по заданным частотным интервалам (рис.1):

$$P_r = \sum_{i=1}^J (y_{ir})^2, \quad r = 1, \dots, R. \quad (2)$$

Полученные значения энергии упорядочиваются по возрастанию (рис. 2), при этом сохраняются номера частотных интервалов. Выбираются значения энергий для тех частотных интервалов, в которых сосредоточена некоторая доля энергии, например, 95%.

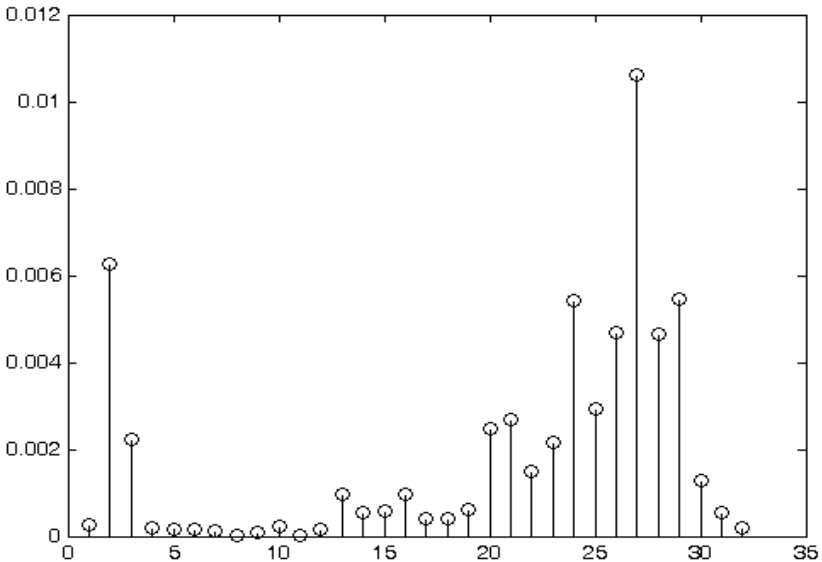


Рис. 1. Распределение энергии, сосредоточенной в $R = 32$ частотных интервалах ($N = 256$)

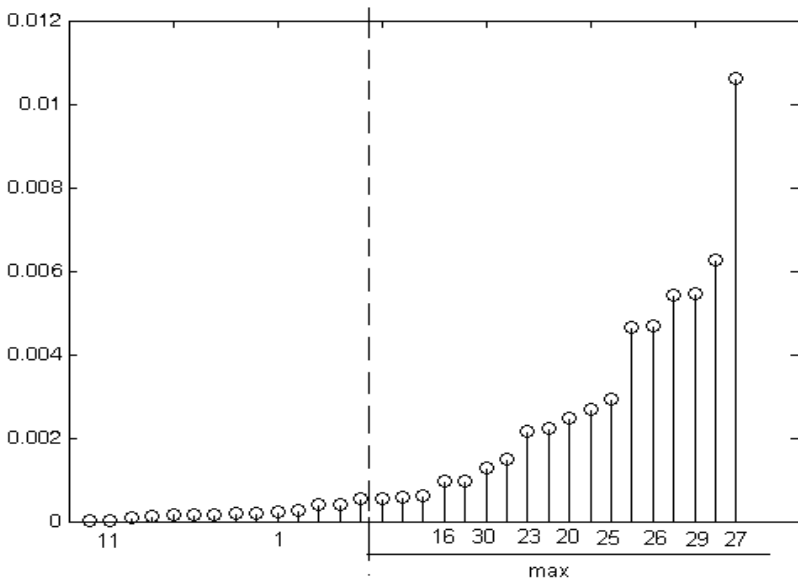


Рис. 2. Распределение энергии по интервалам ($R = 32$, $N = 256$) при сортировке по возрастанию

Значения подвекторов субполосного преобразования квантуются по уровню и записываются в информационный блок совместно со служебной информацией, содержащей номера частотных интервалов.

Алгоритм сжатия представлен на рис. 3.

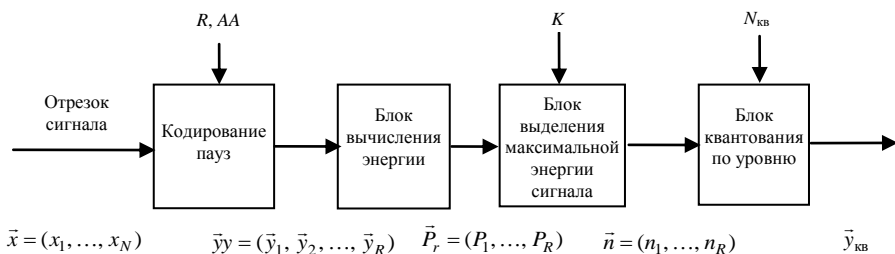


Рис. 3. Функциональная схема системы сжатия речевых данных на основе оптимального субполосного преобразования

Для восстановления сигнала применяют алгоритм, который может быть описан функциональной схемой, представленной на рис. 4.

Из информационного блока осуществляется считывание квантованных значений векторов субполосного преобразования и их восстановление.

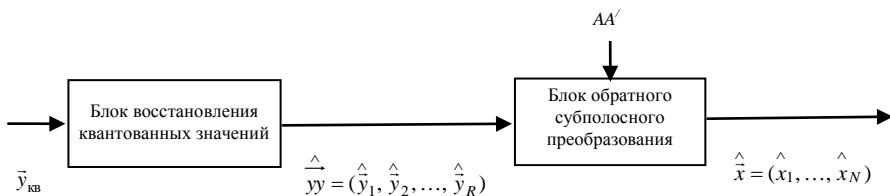


Рис. 4. Функциональная схема системы восстановления речевых данных на основе оптимального субполосного преобразования

Обратное субполосное преобразование осуществляется по формуле:

$$\hat{x} = AA' \hat{y} \quad (3)$$

Для определения параметров субполосного преобразования – длины анализируемого отрезка N , количества частотных интервалов R было проведено большое количество вычислительных экспериментов. В качестве предварительной обработки были обнаружены и закодированы паузы. Таким образом процедуре сжатия подвергается собственно звуковой сигнал.

В ходе экспериментов для различных значений длин отрезков речевых данных $N = 64, 128, 256, 512$ отсчетов, диапазон частот разбивается на $R = 8, 16, 32$ интервалов.

Некоторые результаты приведены в таблице, в которой используются следующие обозначения:

N – длина отрезков речевых данных; K – степень сжатия за счет субполосного преобразования, которая определяется по формуле:

$$K = V_1/V_2, \quad (4)$$

где V_1 – объем исходного файла; V_2 – объем сжатого файла.

Таблица

Степень сжатия речевых данных при различных N

N	K (при $R = 8$)	K (при $R = 16$)	K (при $R = 32$)
64	1,52	1,18	1,22
128	1,81	1,71	1,54
256	1,97	2,38	2,64
512	2,09	2,48	2,80

Необходимо отметить, что в таблице приведены результаты сжатия речевых данных только за счет субполосного преобразования, без учета сжатия за счет удаления пауз и квантования сигнала. Таким образом, на основе полученных результатов можно говорить о высокой эффективности применения данного подхода к сжатию данных, т.к. он позволяет существенно сократить объем речевых данных при их передаче и хранении, при сохранении относительно высокого качества воспроизведения восстановленного сигнала.

При воспроизведении восстановленных звуковых файлов сохраняется качество на уровне телефонного (разборчивость речи, узнаваемость голоса, сохранение тембра голоса).

Выводы. В данной работе представлен алгоритм сжатия речевых данных с использованием субполосного преобразования с учетом отличительных особенностей речевых данных. В результате работы алгоритма достигается сжатие данных в 1,5 – 3 раза в зависимости от выбора длины анализируемого отрезка N и количества частотных интервалов R , при сохранении качества воспроизведения восстановленных данных.

Разработанный алгоритм сжатия речевых данных с использованием субполосного преобразования можно применять для сжатия речевых данных, для их передачи по каналам связи и хранения в ЭВМ. Практическое применение алгоритма сжатия речевых данных с использованием субполосного преобразования позволит ускорить процесс обмена информацией в обществе, а также уменьшить объем памяти, необходимой для хранения речевых данных.

Список литературы: 1. *Сэломон М.* Сжатие данных, изображений и звука / *М. Сэломон*. – М.: Техносфера, 2006. – 368 с. 2. *Михайлов В.Г.* Измерение параметров речи / *В.Г. Михайлов*. – М.: Радио и связь, 1987. – 168 с. 3. *Сергиенко А.Б.* Цифровая обработка сигналов / *А.Б. Сергиенко*. – СПб.: Питер, 2005. – 182 с. 4. *Шелухин О.И.* Цифровая обработка и передача речи / *О.И. Шелухин, Н.Ф. Лукьянцев*. – М.: Радио и связь, 2000. – 208 с. 5. *Орищенко В.И.* Сжатие данных в системах сбора и передачи информации / *В.И. Орищенко, В.Г. Санников, В.А. Свириденко*; под ред. В.А. Свириденко. – М.: Радио и связь, 1985. – 184 с. 6. *Жиляков Е.Г.* Методы обработки речевых данных в информационно-телекоммуникационных системах на основе частотных представлений: монография / *Е.Г. Жиляков, С.П. Белов, Е.И. Прохоренко*. – Белгород: Изд-во БелГУ, 2007. – 136 с. 7. *Жиляков Е.Г.* О субполосном преобразовании звуковых сигналов / *Е.Г. Жиляков, С.П. Белов, Е.И. Прохоренко* // Труды Российского научно-технического общества радиотехники, электроники и связи имени А.С. Попова. – М., 2006. – Выпуск VIII-1. – С. 167–169. – (Серия: цифровая обработка сигналов и ее применение). 8. *Жиляков Е.Г.* Частотный анализ речевых сигналов / *Е.Г. Жиляков* // Научные ведомости Белгородского государственного университета. – Белгород, 2006. – № 2 (31). – Вып. 3. – С. 201–208. 9. *Жиляков Е.Г.* О сжатии речевых сигналов / *Е.Г. Жиляков, С.П. Белов* // Вестник НТУ "Харьковский политехнический институт". Тематический выпуск: Информатика и моделирование. – Харьков: НТУ "ХПИ". – 2005. – № 56. – С. 32–40. 10. *Прохоренко Е.И.* Метод обнаружения пауз в речевых сигналах / *Е.И. Прохоренко* // Вестник Московской академии рынка труда и информационных технологий. – М., 2006. – № 4 (26). – С.13–20.

Статья представлена д.т.н. проф. БелГУ Жиляковым Е.Г.

УДК 621.391

Розробка алгоритму стиснення мовних даних з використанням субсмугового перетворення / Прохоренко Е.І., Черноморец А.А., Есауленко А.В. // Вісник НТУ "ХПІ". Тематичний випуск: Інформатика і моделювання. – Харків: НТУ "ХПІ". – 2009. – № 43. – С. 160 – 166.

Розглянута нова технологія стиснення мовних даних з використанням субсмугового перетворення, розроблений і застосований алгоритм стиснення мовних даних з використанням субсмугового перетворення з урахуванням відмітних особливостей мовних даних (значна кількість пауз в сигналі, а також зосередження енергії сигналу в щодо вузьких частотних інтервалах). Іл.: 4. Табл.: 1. Бібліогр.: 10 назв.

Ключові слова: стиснення мовних даних, субсмугове перетворення, паузи в сигналі.

UDC 621.391

Development of the algorithm of the speech data compression with use of subband transformation / Prokhorenko E.I., Chernomoretc A.A., Esaulenko A.V. // Herald of the National Technical University "KhPI". Subject issue: Information Science and Modelling. – Kharkov: NTU "KhPI". – 2009. – № 43. – P. 160 – 166.

Is considered new technology of the speech data compression with the use of subband transformation, designed and applied algorithm of the speech data compression with the use of subband transformation with provision of discriminating speech data particularities (the great a number of the pauses in the signal, as well as concentration of the signal energy in comparatively narrow frequency interval). Figs.: 4. Table: 1. Refs.: 10 titles.

Key words: speech data compression, subband transformation, pauses in signal.

Поступила в редакцію 10.10.2009