

4. Михелев В.М., Кузнецов К.В., Батищев Д.С., Горопчин Д.А. Алгоритм распознавания нечетких текстовых дубликатов с использованием технологии CUDA. – В сб. Компьютерные науки и технологии. Труды Второй Междунар. НТК. Белгород, 2011, с 112-116.

*Статья поступила 09.12.2011*

**Д.т.н., проф. Е.Г. Жилияков, А.В. Болдышев,  
к.т.н. Е.И. Прохоренко (НИУ «БелГУ»)**

**E.G. Zhilyakov, A.V. Boldyshev, E.I. Prokhorenko**

**АЛГОРИТМ СЖАТИЯ РЕЧЕВЫХ ДАННЫХ  
НА ОСНОВЕ ДВУМЕРНОЙ ОБРАБОТКИ ДАННЫХ**

**SPEECH DATA COMPRESSION ALGORITHM BASED  
ON TWO-DIMENSIONAL DATA PROCESSING**

*В статье изложен новый подход к обработке речевых сигналов, а именно алгоритм двумерной их обработки. Такой подход позволяет сократить затраты ресурсов вычислительных систем на разбиение сигнала на фрагменты и пофрагментную обработку. Приведены результаты вычислительных экспериментов по оценке степени сжатия исходного речевого сообщения, а также коэффициенты корреляции исходного и восстановленного сообщения.*

*In the article describes a new approach to the processing of speech signals, namely the algorithm of two-dimensional processing of speech signals. This approach reduces the overhead of computing systems at signal decomposition into fragments and fragmental processing. The results of computational experiments to assess the degree of compression of the original voice message, as well as the correlation coefficients of the initial and recovered messages.*

*Ключевые слова: сжатие речевых сигналов, двумерная обработка сигналов, информационные системы, телекоммуникационные системы.*

*Key words: compression of speech signals, two-dimensional data processing, information systems, telecommunication systems.*

На сегодня проблема ограниченности ресурсов информационно-телекоммуникационных систем для передачи и хранения речевых данных (пропускная способность канала связи, объем памяти жестких носителей), приводит к необходимости поиска путей их

эффективного использования. Решением этой проблемы занимаются специалисты в области телекоммуникаций, как крупных научно-исследовательских центров, так и подразделений небольших компаний.

В качестве примера можно выделить некоторые области, для которых проблема эффективного использования ресурсов наиболее актуальна: системы передачи экспресс сообщений (Skype, Google Talk, ICQ manager, QIP, Team speak и т.д), системы «голосовой почты», аудиоконференции (звуковая стенография).

В статье предлагается подход к обработке речевых сигналов, заключающийся в представлении исходного речевого сигнала в виде некоего изображения, т.е. не в виде вектора  $f=(f_1, \dots, f_L)^T$ , а в виде матрицы:

$$F = \begin{pmatrix} \bar{x}_1 \\ \dots \\ \bar{x}_M \end{pmatrix}, \quad (1)$$

где:  $L$  – длительность сообщения,

$T$  – операция транспонирования;

$x=(x_1, \dots, x_N)^T$  - отрезок речевого сообщения длительность  $N$  отсчетов;

$M$  – количество отрезков речевого сообщения ( $M=L/N$ ).

При таком представлении происходит обработка всего сигнала сразу, т.е. системе не приходится затрачивать время на разбиение его на фрагменты.

Подобный подход к обработке речевых сигналов направлен на поиск каких-либо общих характеристик отрезков. Выявление общих характеристик различных отрезков может быть использовано непосредственно для решения задачи разработки эффективного метода сжатия.

Итак, полученный сигнал представляет собой некую матрицу:

$$F = \{f_{ik}\}, \text{ где } i = 1, \dots, M; k = 1, \dots, N,$$

где  $f_{ik}$  – элементарное изображение.

После формирования матрицы  $F$  формируется матрица следующего вида:

$$\Phi = FF^T \quad (2)$$

Эта матрица является симметричной и неотрицательно определенной, а, следовательно, обладает полной системой ортонормальных собственных векторов, соответствующих неотрицательным собственным числам [1] и удовлетворяющим условиям:

$$QL = FF^T Q, \quad i = 1, \dots, M, \quad (3)$$

$$Q = \{\bar{q}_1, \bar{q}_2, \dots, \bar{q}_M\},$$

$$L = \text{diag}(\lambda_1, \dots, \lambda_M),$$

$$\lambda_1 \geq \lambda_2 \geq \dots, \lambda_M \geq 0 \quad (4)$$

где:  $Q$  – матрица собственных векторов матрицы  $FF^T$ ,

$L$  – матрица собственных чисел,

$q_i$  – собственный вектор,

$\lambda_i$  – собственное число,

$$\begin{cases} (\bar{q}_i, \bar{q}_k) = 1, & i = k \\ (\bar{q}_i, \bar{q}_k) = 0, & i \neq k \end{cases}, \quad i, k = 1, \dots, M \quad (5)$$

где круглые скобки означают скалярное произведение в виде суммы попарных произведений компонент векторов с одинаковыми индексами.

$i$ -ые строки матрицы собственных векторов отражают общие  $i$ -ых элементов ( $i=1, \dots, M$ ) всех строк матрицы  $F$ . Общность строк матрицы может служить основой для установления идентичности отрезков, этот аспект может быть использован в различных задачах, в т.ч. в задаче сжатия. Через матрицу собственных векторов можно выразить матрицу, столбцы которой отражают общие признаки в каждой строке матрицы  $F$ :

$$\bar{v}_i = F^T \bar{q}_i / \sqrt{\lambda_i} \quad (6)$$

$$\bar{q}_i = F \bar{v}_i / \sqrt{\lambda_i}$$

Причем:

$$\|\bar{v}_i\|^2 * \|\bar{q}_i\|^2 = \lambda_i$$

С помощью этих собственных векторов и чисел исходное изображение, как матрицу, можно записать в виде сингулярного разложения [2], из которого нетрудно получить представление:

$$F = \sum_{i=1}^M \sqrt{\lambda_i} \bar{q}_i \bar{v}_i^T \quad (7)$$

Необходимо отметить, что суммы собственных чисел количественно равны энергии изображения:

$$\sum_{i=1}^M \lambda_i = \sum_{i=1}^M \sum_{k=1}^N f_{i,k}^2 \quad (8)$$

С точки зрения задачи сокращения количества данных, которые необходимо хранить или передавать, можно сформулировать условие нахождения минимального количества собственных значений, которое позволит получить максимальную степень сжатия. При этом важным условием является минимизация погрешности восстановления исходного отрезка речевых данных, т.е. обеспечение высокого качества воспроизведения исходного речевого сообщения.

$$\sum_{i=1}^J \lambda_i / \sum_{i=1}^M \lambda_i \geq c \quad (9)$$

где  $J$  – минимальное количество собственных векторов, необходимое для восстановления исходного сигнала.

Данное условие позволяет определить минимальное количество собственных векторов  $q$ , необходимое для аппроксимации исходных отрезков речевых данных. При этом формула для восстановления исходных значений матрицы  $F$  примет вид:

$$\hat{F} = \sum_{i=1}^J \sqrt{\lambda_i} \vec{q}_i \vec{v}_i^T \quad (10)$$

Таким образом, используя условие (9) осуществляется уменьшение исходного количества данных, необходимых для восстановления речевого сообщения. Степень уменьшения исходных речевых данных (коэффициент сжатия) рассчитывается по следующей формуле:

$$K = \frac{(M * N)}{(M + N) * J} \quad (11)$$

где  $M * N$  – исходное количество речевых данных.

### Вычислительные эксперименты

Вычислительные эксперименты проводились с целью оценки степени сжатия, достигаемой при таком подходе к сжатию. В качестве исходных данных были использованы две звукозаписи: «Мальчик побежал к лагерю», «никогда не проси больше, чем за-

## АЛГОРИТМЫ И ПРОГРАММЫ

служил». Длительность окна анализа была задана 128 отсчетов, порог  $c$  выбирался из диапазона 0.9-0.99. Сначала исходный речевой сигнал обрабатывался алгоритмом обнаружения и кодирования пауз [3], затем полученный сигнал без пауз подвергался сжатию по описанной выше процедуре. Количество отрезков  $M$ , полученное из сигнала без пауз для первой фразы составило 93, а для второй 125. В табл. 1 приведены результаты определения минимального количества собственных векторов согласно условию (9) и получаемый при этом коэффициент сжатия.

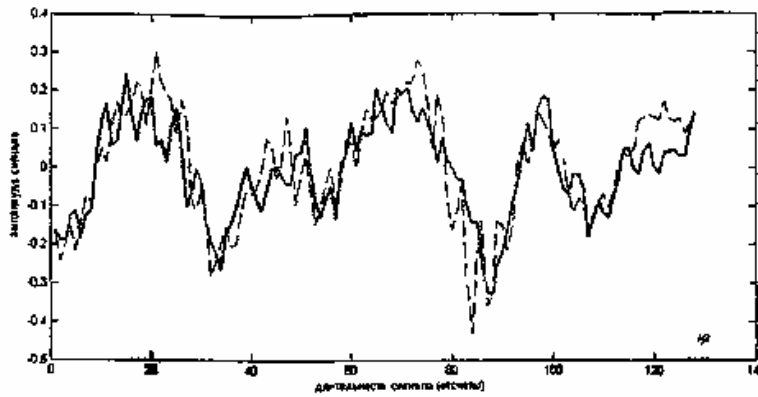
Таблица 1

Оценка степени сжатия исходных речевых данных  
для выбранных речевых сообщений

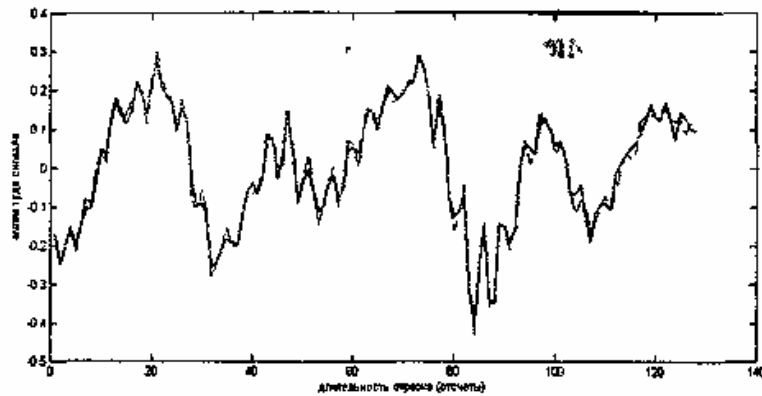
Порог, $c$	«Мальчик побежал к лагерю»)		«Никогда не проси больше, чем заслужил»	
	Минимальное кол-во СВ, $J$	Коэффициент сжатия, $K$	Минимальное кол-во СВ, $J$	Коэффициент сжатия, $K$
<b>0.9</b>	16	3,37	22	2,87
<b>0.91</b>	17	3,17	23	2,75
<b>0.92</b>	17	3,17	25	2,53
<b>0.93</b>	18	2,99	27	2,34
<b>0.94</b>	20	2,69	30	2,11
<b>0.95</b>	21	2,56	32	1,98
<b>0.96</b>	23	2,34	36	1,76
<b>0.97</b>	26	2,07	41	1,54
<b>0.98</b>	30	1,79	48	1,32
<b>0.99</b>	37	1,46	58	1,09

Проведенные вычислительные эксперименты показали, что даже при пороге  $c=0.99$  количество собственных векторов, необходимых для восстановления исходного речевого сообщения составляет около 1/3 от общего количества. Такой подход к обработке речевых сигналов позволяет сократить исходный объем речевых данных в 2-3,5 раза.

На рисунке (рис. 1) приведена форма исходного отрезка сигнала и восстановленного после сжатия.



а)  $c=0.9$



б)  $c=0.9$

Рис. 1

Исходный и восстановленный сигналы (пунктир – исходный сигнал, сплошная линия - восстановленный): а) 18 отрезок, б) 20 отрезок

Как видно из приведенных рисунков маленький порог  $c$  не позволяет получить высокой степени точности восстановления речевого сигнала, что связано с оставлением слишком малого количество собственных векторов. Для достижения приемлемой степени точности восстановления речевого сообщения необходимо использовать порог не менее 0.94-0.95.

### Вывод

Предлагаемый подход к сжатию позволяет сократить количество речевых данных в 2-3,5 раза по сравнению с исходным количеством. Сформулированное условие (9) позволяет варьировать коэффициентом сжатия, однако, как показали проведенные вычислительные эксперименты, порог  $c$  в условии (9) должен быть не менее 0.94.

Исходя из вычисленных коэффициентов корреляции и анализа форм исходного и восстановленного сообщения, можно сделать вывод о достаточно высоком качестве восстанавливаемого сообщения

### Литература

1. Гантмахер, Ф.Р. Теория матриц. М., Физматлит, 2004. 560 с.
2. Хорн Р., Джонсон Ч. Матричный анализ. М., Мир, 1989. 655 с.
3. Прохоренко, Е.И., Болдышев А.В., Фирсова А.А., Эсауленко А.В. Метод оптимального субполосного преобразования в задаче сжатия речевых данных. – «Вопросы радиоэлектроники», сер. ЭВТ, 2010, вып. 1, с. 49-55.

*Работа выполнена при поддержке гранта РФФИ 12-07-00514*

*Статья поступила 09.12.2011*

**Д.т.н., проф. Е.Г. Жилияков, А.В. Курлов**  
(НИУ «БелГУ»)

**E.G. Zhilyakov, A.V. Kurlov**

### **КОМПЬЮТЕРНАЯ КОМПЕНСАЦИЯ ЭХО ИСКАЖЕНИЙ НА ОСНОВЕ РЕШЕНИЯ ИНТЕГРАЛЬНЫХ УРАВНЕНИЙ**

### **COMPUTER ECHO DISTORTION COMPENSATION BASED ON THE SOLUTION OF INTEGRAL EQUATIONS**

*В статье предложена компьютерная процедура компенсации эхо искажений, основанная на решении интегральных уравнений типа свертки. С целью проверки работоспособности проведены вычислительные эксперименты, показывающие влияние частоты дискретизации на погрешность восстановления речевого сигнала.*

*Ключевые слова: речь, компенсация эхо, интегральные уравнения, аппаратная функция.*

*In the article the computer echo distortion compensation procedure, which is based on the solution of integral equations of convolution type. In order to verify that the computational experiments carried out, showing the influence of sampling frequency on the error recovery of the speech signal.*

*Key words: speech, echo distortion, integral equations, apparatus function.*