

***Свиридова Ирина Вячеславовна,***

Ассистент кафедры прикладной информатики

и информационных технологий

НИУ «БелГУ» Россия, г. Белгород

***Sviridova Irina Vyacheslavovna,***

Assistant of the Department of Applied Informatics

and information technology

NRU "BelGU" Russia, Belgorod

***Подпругин Александр Ильич,***

Студент магистратуры 1го года обучения НИУ«БелГУ»

Россия, г.Белгород

***Podprugin Alexander Ilyich,***

1st year master's student

NRU "BelGU" Russia, Belgorod

***Гончаров Дмитрий Викторович,***

Ассистент кафедры информационных и робототехнических систем

НИУ «БелГУ» Россия, г. Белгород

***Goncharov Dmitry Viktorovich,***

Assistant of the Department of Information and Robotic Systems

NRU " BelGU" Russia, Belgorod

***Бондаренко Виктория Александровна***

Аспирант 2-го года обучения

института инженерных и цифровых технологий

НИУ «БелГУ» Россия, г. Белгород

***Bondarenko Victoria Alexandrovna***

2nd year postgraduate student

Institute of Engineering and Digital Technologies

NRU "BelGU" Russia, Belgorod

# СЖАТИЕ РЕЧИ С ИСПОЛЬЗОВАНИЕМ СУБПОЛОСНОГО ПРЕОБРАЗОВАНИЯ

## SPEECH COMPRESSION USING SUBBAND TRANSFORM

**Аннотация:** в данной статье описан процесс сжатия речи, а также практическое его применение с использованием субполосного преобразования.

**Ключевые слова:** сжатие речи, алгоритмы, преобразование.

**Abstract:** This article describes the speech compression process, as well as its practical application using subband transform.

**Keywords:** speech compression, algorithms, conversion.

С целью исследования возможности использования субполосного преобразования для очистки речевого сигнала от шумов различного происхождения, были проведены вычислительные эксперименты с реальными речевыми данными.

**Словесное описание алгоритма субполосного преобразования на основе банка КИХ – фильтров.**

1. Задать параметры  $N$  – длина анализируемого отрезка сигнала,  $R$  – количество частотных интервалов.

2. Рассчитать импульсные характеристики для  $R-2$  полосовых КИХ-фильтров, выбирая их полосу пропускания и частоты среза в соответствии с

$$V_k = [-\upsilon_{k2}, -\upsilon_{k1}) \cup [\upsilon_{k1}, \upsilon_{k2}), k = 1, 2, \dots, R$$

$$\sum_{k=1}^R M_k = N, -\upsilon_{M2} = -\pi; \quad \upsilon_{M2} = \pi$$

$$\Omega = \upsilon_{k2} - \upsilon_{k1} = \frac{\pi}{R} = \text{const}, \quad \upsilon_{k1} = (k-1) \frac{\pi}{R}; \quad \upsilon_{k2} = k \frac{\pi}{R}, \quad M_k = M = \frac{N}{R}$$

**Обратное субполосное преобразование (синтез) на основе использования банков КИХ-фильтров**

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.

Сущность субполосного преобразования состоит в том, что для отрезков сигнала вычисляется вектор  $y\vec{u} = (y_1, y_2, \dots, y_R)$ , 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:

$$V_k = [-\nu_{k2}, -\nu_{k1}) \cup [\nu_{k1}, \nu_{k2})$$

Субполосное преобразование осуществляется по формуле:  $y\vec{u} = AA\vec{x}$  (1)

где  $AA$  – блочная матрица вида:

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

$Q$  – матрицы собственных векторов матрицы  $A$ ;  $L$  – вектор собственных чисел,  $J$  – количество собственных чисел матрицы  $A$ .

Если в качестве вектора исходных значений  $x$  используется отрезок речевого сигнала длительностью  $N$ , тогда вектор субполосного преобразования вычисляется по формуле. С использованием полученных значений подвекторов субполосного преобразования вычисляется распределение энергии сигнала по заданным частотным интервалам.

Полученные значения энергии упорядочиваются по возрастанию (рисунок 1), при этом сохраняются номера частотных интервалов. Выбираются значения энергий для тех частотных интервалов, в которых сосредоточена некоторая доля энергии, например, 95%.

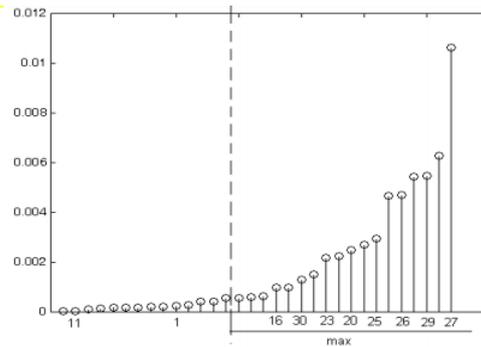
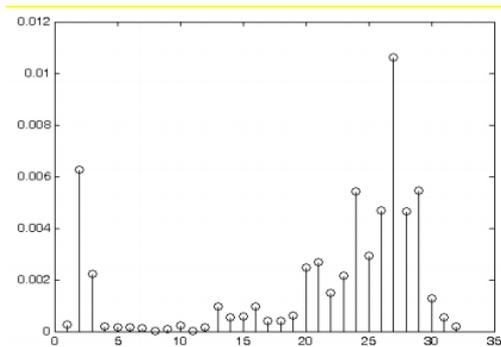


Рисунок 1 - Распределение

Рисунок 2 - Распределение

энергии, сосредоточенной в  $R = 32$  энергии по интервалам ( $R = 32, N = 256$ )

Значения подвекторов субполосного преобразования квантуются по уровню и записываются в информационный блок совместно со служебной информацией, содержащей номера частотных интервалов. Алгоритм сжатия представлен на рисунке 3.

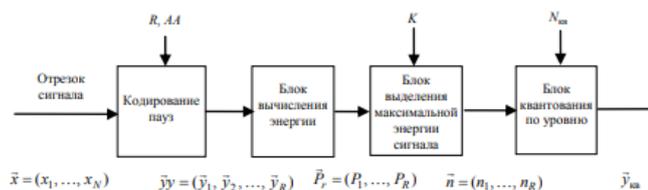


Рисунок 3 - Функциональная схема системы сжатия речевых данных

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:  $K = V1/V2$ , (3) где  $V1$  – объем исходного файла;  $V2$  – объем сжатого файла.

Таблица 1 – 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.

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

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 ИСПОЛЬЗОВАННЫХ ИСТОЧНИКОВ**

1. Жилияков Е.Г., Белов С.П., Прохоренко Е.И. УМК по ОРД в ИТС[текст]/. Е.Г. Жилияков, С.П. Белов, Е.И. Прохоренко: БелГУ, Белгород 2008г.

2. Бернад Скляр, Цифровая связь Теоритические основы и практическое применение. Второе издание: Москва. 2003.