

ОБ ИЗБИРАТЕЛЬНОМ ПРЕОБРАЗОВАНИИ ЧАСТОТНЫХ КОМПОНЕНТ РЕЧЕВЫХ СИГНАЛОВ В ЗАДАЧЕ СЖАТИЯ

Жиляков Е.Г., Болдышев А.В., Курлов А.В., Фирсова А.А.,
Эсауленко А.В.

*Национальный исследовательский университет
"Белгородский государственный университет", г. Белгород*

Постановка проблемы. Одной из особенностей звуков русской речи является сосредоточенность энергии в достаточно узких частотных диапазонах, суммарная ширина которых гораздо меньше частоты дискретизации. Эта особенность может быть использована в различных направлениях области обработки речевых сообщений, в том числе в задаче сжатия речевых данных. Для этого необходимо точно определить, в каком количестве частотных интервалов сосредоточена основная доля энергии. Чтобы определить количество частотных интервалов, в которых сосредоточена основная доля энергии, можно воспользоваться понятием частотной концентрации, которая определяется минимальным количеством частотных интервалов, в которых сосредоточена заданная доля энергии отрезка речевого сигнала. На основании сведений о количестве и расположении частотных интервалов, в которых сосредоточена заданная доля энергии, можно осуществить сжатие речевых данных за счет хранения только составляющих речевого сигнала, соответствующих этим частотным интервалам.

Целью доклада является анализ эффективности сжатия речевых данных за счет использования предлагаемого метода субполосного частотного анализа/синтеза.

Сжатие исходных речевых данных осуществляется за счет хранения вектора значений размерностью равной минимальному количеству ненулевых собственных значений. При этом важным условием является минимизация погрешности восстановления исходного отрезка речевых данных, т.е. обеспечение высокого качества воспроизведения исходного речевого сообщения.

Проведенные вычислительные эксперименты показали высокую эффективность разработанного метода с позиции сжатия речевых данных при сохранении приемлемого для пользователя качества воспроизведения. Предлагаемый подход к сжатию речевых сигналов в среднем позволяет достичь степень сжатия около 6 – 7 раз без существенной потери качества воспроизведения.