

УДК 629.783

И.А.Горбунов (4 курс, каф. РЭСЗИ), Е.А.Попов, к.т.н., доц.

SBR–ТЕХНОЛОГИЯ РАЗДЕЛЕНИЯ СПЕКТРА

Попытки уменьшить скорость передаваемой информации за счет сокращения числа уровней квантования или частоты дискретизации исходных сигналов приводят к снижению качества передаваемой информации, например, ухудшение разборчивости речи.

Одним из возможных способов, позволяющих снизить объемы возникающих потоков цифровых данных без заметной потери качества передаваемых сообщений, является предварительное усечение спектра обрабатываемых сигналов с последующим восстановлением усеченной части на приемной стороне на основе данных, полученных по дополнительному каналу. При этом эффективность такого подхода будет обусловлена как снижением нагрузки на вокодер, так и малым объемом данных в дополнительном канале за счет выявления взаимосвязи между спектральными компонентами. Как показывают многочисленные исследования, между некоторыми спектральными компонентами в верхней и нижней частях спектра звуковых сигналов существует определенная схожесть.

Целью работы является исследование возможности сужения полосы занимаемых частот в системах передачи звуковой информации путем использования технологии разделения спектра SBR (**S**pectral **B**and **R**eplication).

Пусть s_k – цифровые отсчеты сигнала, подлежащие кодированию и последующей передаче в канал. Предлагается следующий алгоритм снижения объема и, соответственно, скорости передаваемой информации (рис. 1).

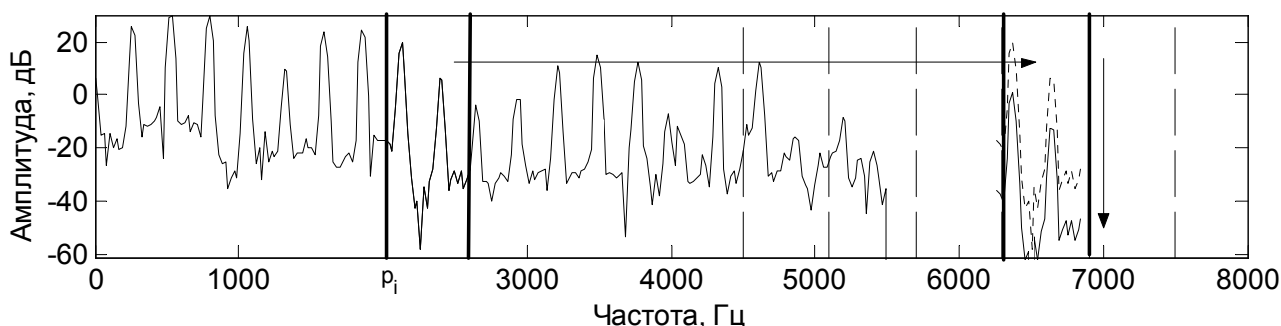


Рис. 1. Иллюстрация работы алгоритма SBR.

1. Исходная последовательность отсчетов $\{s_k\}$ разбивается на сегменты длиной N с последующим дискретное преобразование Фурье каждого сегмента:

$$S(k) = \sum_{n=0}^{N-1} s(n) \cdot \exp\left(-2\pi j \frac{kn}{N}\right), \quad 0 \leq k \leq \frac{N}{2}.$$

2. Получившийся спектр разделяется на области нижних и верхних частот, и поскольку верхняя часть будет отброшена, необходимо иметь информацию для ее последующего восстановления. Для этого область верхних частот разбивается на m полос $[k_i; k_i + l_i - 1]$ ($1 \leq i \leq m$), где i – номер полосы, k_i – начало i -ой полосы, а l_i – ее длина, и вычисляется энергию сигнала E_i в каждой полосе:

$$E_i = \sum_{k=k_i}^{k_i+l_i-1} |S(k)|^2, \quad 1 \leq i \leq m$$

3. Определяется взаимосвязь между различными полосами в верхней и нижней областях спектра по критерию минимума среднеквадратической ошибки, т.е. для каждой полосы производится минимизация выражения $\varepsilon_i^2(p)$ по параметру p :

$$\varepsilon_i^2(p) = \sum_{k=0}^{l_i-1} \left(\sqrt{E_i} \left(\sum_{n=0}^{l_i-1} |S(p+n)|^2 \right)^{-1/2} |S(p+k)| - |S(k_i+k)| \right)^2, \quad 1 \leq i \leq m, \quad p \in [0; k_1 - l_i].$$

Таким образом, формируется 2 семейства параметров $\{E_i\}$ и $\{p_i\}$, отражающих взаимосвязь между нижней и верхней частями спектра.

4. Нижняя часть спектра либо подвергается обратному преобразованию Фурье по $N/4$ точкам, если вокодер работает во временной области, либо непосредственно направляется в вокодер, если он работает в частотной области.

На приемной стороне встает задача восстановить верхнюю часть спектра принятого сигнала $u(n)$, для чего необходимо провести следующие операции.

5. Производится дискретное преобразование Фурье сигнала порядка $N/2$:

$$U(k) = \sum_{n=0}^{N/2-1} u(n) \exp\left(-2\pi j \frac{kn}{N/2}\right), \quad 0 \leq k \leq \frac{N}{4}.$$

6. Получившийся спектр разделяется на полосы, указанные в п. 2, причем все полосы в верхней части изначально пустые.

7. В каждую полосу из области верхних частот переносится соответствующий (по параметру p_i) участок из области нижних частот, при этом производится масштабирование, так чтобы энергия в этой полосе совпадала с энергией исходного сигнала в этой полосе:

$$U(k_i+k) = U(p_i+k) \sqrt{E_i / \sum_{n=0}^{l_i-1} |U(p_i+n)|^2}$$

8. Восстановленный спектр подвергается обратному дискретному преобразованию Фурье порядка N , так что отсчеты восстановленного сигнала s'_k равны

$$s'_k(n) = \frac{2}{N} \sum_{k=0}^{N-1} U'_k \exp\left(2\pi j \frac{kn}{N}\right), \quad \text{где } U'_k = \begin{cases} U_k & \text{для } 0 \leq k \leq \frac{N}{2} \\ (U_{N-k})^* & \text{для } \frac{N}{2} < k < \frac{N}{4} \end{cases}$$

С целью вычисления ошибки восстановления был проведен следующий эксперимент. Исходный сигнал, представляющий собой отрезок гармонического колебания, был подвергнут кодированию, а затем декодирован с использованием технологии SBR. Параметры SBR кодека при этом были $F_D = 22050$ Гц, $m = 10$, $k_1 = 4500$ Гц, $l_i \in [600, 900]$ Гц. При этом среднеквадратическая ошибка восстановления оказалась равной 0,02 при единичной амплитуде исходного сигнала.

Однако, следует заметить, что при использовании сигналов других типов картина может измениться коренным образом. Например, при кодировании сигнала, являющегося суммой шума в полосе от 0 до 8 кГц и колебания с линейной частотной модуляцией от 8 до 16 кГц наблюдаются значительные искажения в области верхних частот спектра. Таким образом, для эффективного применения технологии SBR необходимо, чтобы верхняя часть спектра была похожа на нижнюю, что справедливо для сигналов естественного происхождения, таких как речь или музыка, но может не выполняться для синтетических сигналов.