

## **Быстрый алгоритм удаления шума из речевого сигнала при высоком уровне акустического шума**

Я.Ю. Кульков

*Муромский Институт (филиал) Владимирского Государственного Университета  
602264, г.Муром Владимирской обл., ул. Орловская, 23,  
тел.:(849234) 7-72-73, E-mail: eivt@mivlgu.ru*

При передаче речевого сигнала происходит неизбежная потеря информации. Хотя речевой сигнал обладает определенной избыточностью, однако различные шумы, искажения и реверберационные помехи могут привести к настолько значительной потере информации, что это сделает невозможным понимание смысла речи. Следует отметить, что "слышимость" и "разборчивость речи" - это разные понятия. Речь может звучать очень громко и быть прекрасно слышна, но быть при этом совершенно неразборчивой (например, в залах вокзалов, аэропортов и др.). Для оценки разборчивости речи разрабатываются специальные методы, отличные от оценок ее громкости, и разработкой этих методов занимаются крупные международные организации: ISO, AES, IEC и др.

Среди многочисленных факторов, влияющих на разборчивость речи, прежде всего можно выделить маскирование другими звуками, в том числе шумами. Шумы могут создаваться вентиляцией, внешними проникновениями, шумами аппаратуры, публикой, электронной аппаратурой и др.

Процент потери разборчивости зависит, прежде всего, от отношения уровня речевого сигнала к уровню шума (S/N), которое должно быть выше определенного уровня, чтобы можно было понять смысловое содержание речи. Степень маскировки шумом будет зависеть от отношения S/N и от спектрального состава шума. Процент словесной разборчивости будет больше 80% только при отношении  $S/N > 12$  дБ.

Если шум узкополосный, то степень маскирования речи и потеря разборчивости зависят от частотной полосы, то более "опасными", чем высокочастотные (1800:2500 Гц) шумы, являются низкочастотные шумы (135:400 Гц).

В настоящей работе предложена адаптивная система снижения шума, которая уменьшает нежелательное присутствие закодированного фонового шума, при этом оказывая минимальное отрицательное влияние на качество закодированной речи и минимально увеличивается потребление ресурсов процессора цифровых сигналов. Способ, предложенный в данной статье, увеличивает разборчивость речи в аудиосигнале, имеющем цифровой вид, за счет пропуска кадров аудиосигнала, преобразованного в цифровой вид, через схему фильтра. Схема фильтра функционирует как регулируемый высокочастотный фильтр, который фильтрует участок цифрового сигнала в области низких звуковых частот и пропускает участок цифрового сигнала, попадающий в области более высоких частот. За счет адаптивной регулировки схемы фильтра и отбора ее частотной характеристики ограничивается объем отфильтрованной речи и в результате имеется минимальное отрицательное влияние фильтрации на разборчивость речи, полученной при передаче ее каналу связи.

Схема фильтра вычисляет оценки шума для кадров аудиосигналов, имеющих цифровой вид. Оценки шума соответствуют количеству фонового шума в кадрах цифровых аудиосигналов. Оценки шума возрастают по мере возрастания относительной величины фонового шума к уровню речи в низкочастотном диапазоне речи. Схема управления фильтром использует оценки шума для регулировки схемы фильтра так, чтобы она фильтровала большие части низкочастотного диапазона речи при увеличении относительной величины фонового шума к уровню речи в низкочастотном диапазоне речи. Когда фоновый шум отсутствует, то никакая часть речевого сигнала не фильтруется. Когда имеется более высокий уровень фонового шума, то выделяются большие части шума и речевой информации. Поскольку шум преимущественно соответствует низкочастотному диапазону и только относительно небольшая часть разборчивой речи попадает в этот низкочастотный диапазон, разборчивость аудиосигнала в целом может быть улучшена, если при увеличении оценок шума будет увеличиваться отфильтровываемая часть низкочастотной энергии.

Энергия кадра, определенная с помощью устройства оценки энергии кадра, извлекается из ОЗУ. Далее принимается решение: превышает ли оценка энергии кадра сумму величины извлеченной оценки шума и заданной величины речевого порога,  $E_f > (\text{оценка шума} + \text{речевой порог})$ .

Величина речевого порога может быть фиксированной величиной, определенной эмпирически; она должна быть больше, чем изменения энергии за короткие промежутки времени (флуктуации энергии) типичного фонового шума, и может быть, например, установлена в 9 дБ. Кроме того, величина речевого порога может быть адаптивно модифицирована, чтобы она отражала изменяющиеся условия речи, например, когда говорящий попадает в более шумную или более тихую обстановку. Если оценка энергии кадра превышает сумму в уравнении 2, то в блоке устанавливается признак, что речь существует. Если детектор речевой составляющей выявляет, что речь существует, тогда устройство оценки шума обходится и извлекается (из ОЗУ) оценка шума, вычисленная для предыдущего кадра в аудиосигнале, преобразованном в цифровой вид, и она используется в качестве текущей оценки шума. В противном случае, если оценка энергии кадра меньше, чем сумма в уравнении 2, тогда признак речи возвращается в исходное состояние.

Если речь не выявлена, тогда устройство оценки шума выполняет последовательность операций по обновлению оценки шума. В последовательности операций по оценке шума, выполняемой устройством оценки шума, дельта ( $\Delta$ ) разности/ошибки определяется в блоке между энергией шума кадра, полученной устройством оценки энергии кадра, и оценкой шума, ранее вычисленной устройством оценки шума, в соответствии со следующим уравнением:

В блоке принятия решения определяется, превышает ли  $\Delta$  нуль. Если  $\Delta$  - отрицательная, что имеет место при высоких величинах оценки шума, тогда оценка шума повторно вычисляется как предыдущая оценка шума +  $\Delta/2$

Поскольку  $\Delta$  - отрицательная величина, то это приводит к понижающей корректировке оценки шума. Относительно большой размер шага  $\Delta/2$ , выбирается для того, чтобы резко (быстро) осуществлять коррекцию для понижающихся шумовых уровней. Однако, если энергия кадра превышает оценку шума, обеспечивая  $\Delta$  большую, чем нуль, тогда оценка шума обновляется, увеличиваясь на значение  $\Delta/256$

Поскольку  $\Delta$  - положительная величина, то оценка шума должна быть увеличена. Однако небольшой размер шага,  $\Delta/256$  (по сравнению с  $\Delta/2$ ), выбирается для постепенного увеличения оценки шума и обеспечения, в значительной степени, невосприимчивости к проходящему шуму.

Оценка шума, вычисленная для текущего кадра, подается на селектор фильтра. Селектор фильтра обращается к справочной таблице и использует текущую оценку шума для отбора величины управления фильтром. Схема фильтра настраивается в зависимости от выбранной величины управления фильтром так, чтобы ее частотная характеристика обеспечивала увеличение количества фильтрованного шума при увеличении оценки шума и фонового шума. выборки, хранящиеся в ОЗУ, затем пропускаются через настроенную схему фильтра, чтобы она фильтровала выборки для устранения шума.