

Подавление широкополосного шума: история и новые разработки

Алексей Лукин

Эта статья написана по материалам лекции "Broadband Noise Reduction: Theory and Applications", прочитанной автором на 125-м конгрессе AES в октябре 2008 года.

Уменьшение широкополосного шума в аудиосигнале – одна из задач, с которыми звукорежиссеры сталкивались с самых ранних времен. Уже в начале 40-х годов в радиоприемниках широко использовались простейшие гейты для подавления помех в отсутствие сигнала. В этой статье рассказывается об эволюции средств и способов подавления шума на протяжении последних 60 лет и о современных системах шумоподавления, основанных на методе спектрального вычитания.

Рассматриваться здесь будут лишь так называемые аддитивные стационарные шумы. Стационарность означает, что такие свойства шума, как мощность и спектр, не меняются во времени. Аддитивность означает, что шум накладывается на чистый сигнал и не вносит в него других видов искажений – модуляций, нелинейных искажений. Многие ровные и однородные во времени шумы удовлетворяют этим условиям. Например, шум магнитной ленты, гул наводок электросети, шипение микрофонного предусилителя, шум вентиляции в студии звукозаписи – все эти шумы приблизительно стационарны и аддитивны.

С другой стороны, такие шумы, как щелчки, выпадения, нелинейные искажения, скрипы и шорохи, не являются стационарными и не будут рассматриваться в этой статье. Для них существуют свои способы реставрации.

Заметим, что рассматриваемые здесь шумы не обязательно должны быть широкополосными. Например, свист наводки от строчного генератора ЭЛТ-видеомонитора (тон с частотой порядка 16 кГц) или гул фона от электросети (тон 50 Гц с гармониками) тоже являются стационарными и аддитивными помехами и допускают подавление рассматриваемыми здесь методами.

Гейт (пороговый шумоподавитель)

Простейший способ подавления шума, применяющийся еще с 40-х годов, – это однополосный гейт (gate). Гейт является устройством динамической обработки (см. статью М. Чернецкого «Устройства динамической обработки сигналов», «Звукорежиссер» 3/1999), подавляющим сигналы, уровень которых ниже заданного порога, и пропускающим остальные сигналы без изменений (рис. 1). В первых радиоприемниках порог срабатывания (threshold) был фиксирован и настроен на подавление радишума в паузах между передачами.

В более современных гейтах порог срабатывания задается пользователем вручную и существенно влияет на качество результата. При слишком высоком пороге тихие сигналы, сравнимые по уровню с шумом, будут подавляться гейтом. При слишком низком пороге случайные всплески шума будут периодически открывать гейт и прорываться в результирующий сигнал.

Для уменьшения эффектов резкого открытия и закрытия гейта всплесками шума существуют схемы замедления работы гейта во времени. Такие параметры,

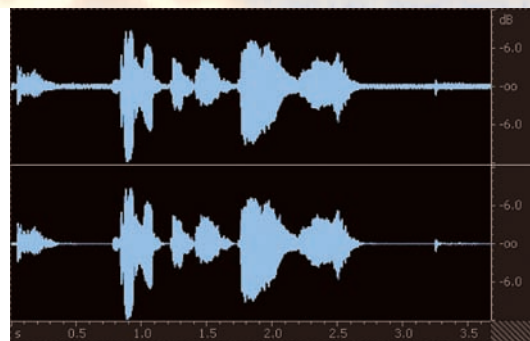


Рис. 1. Сверху – зашумленный сигнал, снизу – обработанный гейтом

как время атаки и время восстановления (attack time, release time), задают скорость реакции гейта на изменения уровня входного сигнала и помогают сгладить процессы открытия и закрытия.

В более сложных устройствах гейты могут приобретать черты экспандеров – похожих устройств динамической обработки, уменьшающих уровень слабых сигналов. Степень подавления шума может зависеть от уровня входного сигнала, а передаточная характеристика может иметь «мягкий порог» (soft knee).

Гейт – самое простое устройство для подавления шума. Простота же является и его основным достоинством. К недостаткам можно отнести эффект модуляции шума сигналами низкой амплитуды и подавление полезных сигналов низкой амплитуды. Фактически гейты подавляют шум только в паузах, а во время звучания программы пользы от них нет, и это является существенным недостатком при малом соотношении сигнал/шум.

Dynaural Noise Suppressor

В 1946 году компанией Н.Н. Scott был представлен первый специализированный прибор для шумоподавления, ставший впоследствии коммерчески успешным. Он назывался Dynaural Noise Suppressor и представлял собой небольшой прибор на трех лампах, реализующий полосовой фильтр. Особенностью прибора является адаптация частот среза полосового фильтра к ширине полосы аудиосигнала. Основным применением системы было улучшение звучания грампластинок (на 78 об/мин) – главного носителя записи в те времена. От системы требовалась возможность подавлять не только стационарные шумы, но и треск винила.

Принцип действия Dynaural Noise Suppressor по современным меркам достаточно прост. Анализируя мощность сигнала в области средних частот, прибор изменяет частоту среза НЧ-фильтра с крутизной 20 дБ/окт в пределах от 2,5 до 12 кГц (рис. 2). Предполагается, что диапазон аудиосигнала тем шире, чем больше в сигнале среднечастотной энергии. Это приблизительно верно для многих оркестровых записей: наиболее широкий спектр у оркестра во время

ударов тарелок или при игре форте, что, как правило, означает и значительную среднечастотную энергию.

Похожим образом в пределах от 70 до 250 Гц изменяется частота среза ВЧ-фильтра с крутизной 10 дБ/окт. Однако при этом уже анализируется мощность сигнала в области средне-низких частот.

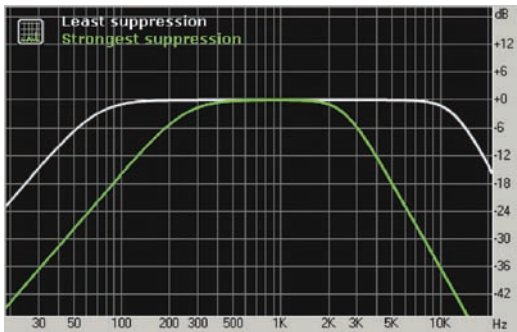


Рис. 2. Частотные характеристики прибора Dynaural Noise Suppressor при минимальном и максимальном подавлении шума

Существенным преимуществом прибора по сравнению с обычным гейтом является подавление шума во время звучания сигнала и адаптация к изменениям сигнала. В то же время неточность определения ширины спектра сигнала приводила к тому, что тембр звука мог приглушаться в результате избыточной фильтрации.

Система Dolby и многополосный гейт

В 1965 году была разработана система Dolby A для подавления шума магнитной ленты. Отличие систем Dolby от других, рассматриваемых в этой статье, в том, что она является двусторонней (double-ended). Сигнал перед записью на ленту подвергается обработке, которая минимизирует влияние шума ленты, а при воспроизведении сигнал «декодируется» для подавления возникших шумов. Остальные же системы, рассматриваемые в этой статье, являются односторонними (single-ended), т.е. работают с уже зашумленным сигналом без доступа к исходному.

Рассмотрим сначала устройство декодера системы Dolby A. Он представляет собой 4-полосный экспандер

с передаточной характеристикой специального вида (рис. 3, справа). Когда уровень сигнала превышает верхний порог, сигнал пропускается без изменений. Когда уровень сигнала находится между нижним и верхним порогом, осуществляется экспандирование (расширение динамического диапазона) с отношением 1:2. Когда уровень сигнала меньше нижнего порога, сигнал ослабляется на фиксированную величину – 10 дБ. Такая передаточная характеристика называется билинейной.

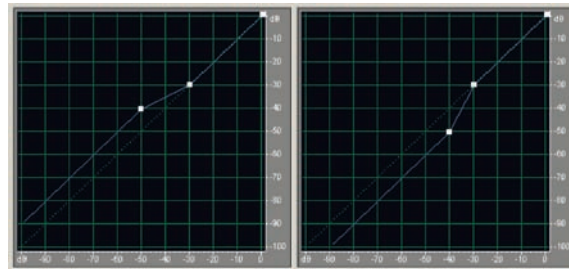


Рис. 3. Передаточные характеристики компрессора и экспандера

Кодер системы Dolby A осуществляет компрандирование (изменение динамического диапазона) симметрично декодеру: он усиливает сигналы низкого уровня на величину до 10 дБ (рис. 3, слева). Его передаточная характеристика является симметричным отражением характеристики декодера относительно оси $y = x$.

Таким образом, система Dolby A сжимает динамический диапазон сигнала перед записью на ленту и восстанавливает его при воспроизведении. При этом в процессе экспандирования уменьшается уровень шума ленты. Это позволяет достичь более эффективного шумоподавления по сравнению с односторонними системами, не имеющими доступа к исходному «чистому» сигналу. Среди недостатков системы Dolby можно отметить не всегда точное восстановление динамического диапазона исходного сигнала в декодере. Даже при отсутствии шума ленты чрезвычайно трудно достичь полной обратимости процесса компрессии кодера из-за наличия баллистики (времени срабатывания) в работе компрессора и экспандера. Добавление шума ленты де-

Реклама

лает эту задачу еще труднее. К счастью, билинейная передаточная характеристика в значительной степени упрощает процесс восстановления динамики, так как сигналы низкой амплитуды имеют фиксированный коэффициент усиления и «не участвуют» в изменении динамического диапазона. (Более подробно о системах Dolby можно прочитать в цикле статей К. Гендри «Системы шумоподавления», «Звукорежиссер» 6...8/2004).

Многими пользователями было замечено, что декодер систем Dolby можно применять для подавления шума даже в записях, которые не были закодированы по системе Dolby. При этом ощущалась потеря низкоуровневых звуков, особенно на высоких частотах, но шум подавлялся столь же эффективно. Фактически, шумоподавление в данном случае выполняет многополосный экспандер.

Первый из известных многополосных гейтов (multiband gate) появился в 1974 году и назывался Garrard Music Recovery Module. Это был 4-полосный экспандер-гейт, в котором степень подавления каждой полосы зависела не только от мощности сигнала в этой полосе, но и от распределения мощностей сигнала в других частотных полосах.

Благодаря наличию нескольких частотных полос Garrard Music Recovery Module мог выборочно подавлять только те частотные полосы, которые в данный момент содержат шум. К сожалению, возможности аналоговой электроники сильно ограничивали рост числа полос. Снять это ограничение удалось с переходом на цифровую обработку звука.

Метод спектрального вычитания

С приходом цифровой обработки сигналов в звукорежиссуру и с увеличением производительности компьютеров стало возможным реализовывать системы шумоподавления с практически неограниченным числом частотных полос. В конце 70-х – начале 80-х годов появились первые научные работы, описывающие метод спектрального вычитания (spectral subtraction) для подавления стационарных аддитивных шумов. А в 1988...1990 годах появились и первые системы, позволяющие выполнять этот алгоритм в реальном времени – Sonic Solutions NoNOISE и CEDAR.

Суть алгоритма достаточно проста. Входной сигнал раскладывается в спектр, точнее – в спектрограмму, то есть существует зависимость амплитуды от частоты и времени. Далее пользователь вручную указывает в звуке фрагмент чистого шума, по которому алгоритм вычисляет спектр шума (рис. 4). Известно, что спектр от суммы сигнала и шума равен сумме спектров сигнала и шума. Поэтому для очистки сигнала от шума осуществляется вычитание амплитудного спектра шума из амплитудного

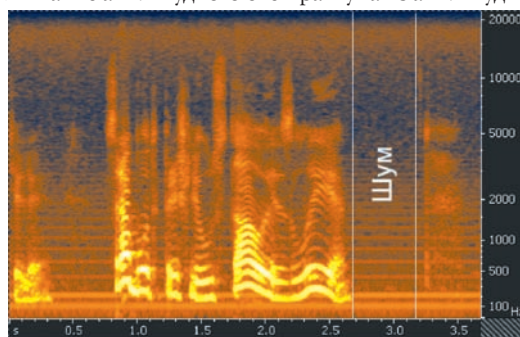


Рис. 4. Спектрограмма речевого сигнала с отмеченной областью шума

спектра аудиосигнала в каждый момент времени (рис. 5). Фазовый спектр аудиосигнала при этом остается неизменным, так как в силу случайности шума нет алгоритма, позволяющего «очистить» и фазовый спектр. После такого вычитания по очищенной спектрограмме сигнала производится синтез результирующего сигнала.

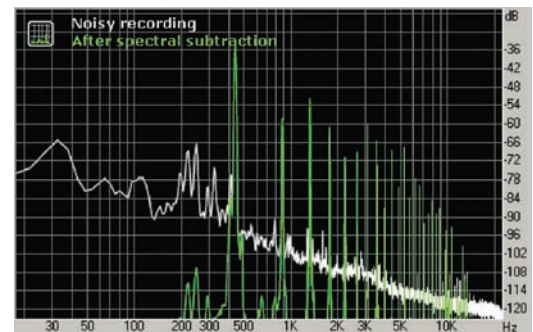


Рис. 5. Зашумленный звук гитарной струны и результат спектрального вычитания (пики зеленой кривой закрывают аналогичные пики белой кривой)

Описанный алгоритм имеет близкое родство с многополосным гейтом. Действительно, спектрограмму сигнала можно трактовать как набор горизонтальных строк, каждая из которых является отдельным субполосным сигналом. И над каждой такой частотной полосой производится операция вычитания из ее амплитуды некоторой константы – уровня шума в этой полосе. Когда уровень сигнала в полосе велик, вычитание небольшой константы практически не меняет сигнала в полосе. Когда уровень сигнала близок к уровню шума, уменьшение уровня сигнала на величину уровня шума фактически эквивалентно применению гейта с «мягким порогом». Поэтому весь процесс спектрального вычитания можно рассматривать как многополосный гейт с большим числом полос и «мягким порогом».

Существует множество различных реализаций метода спектрального вычитания. Одни могут отличаться типами передаточной характеристики у индивидуальных гейтов (правилом «вычитания» спектра шума из спектра сигнала), числом частотных полос, способами борьбы с различными артефактами. Поскольку системы на основе спектрального вычитания являются наиболее современными и распространенными, рассмотрим эти аспекты подробнее.

Частотное разрешение

Типичное число частотных полос в алгоритме спектрального вычитания – порядка тысячи. Это означает, что ширина полос составляет порядка 20 Гц – это позволяет достаточно точно разбивать гармоники многих сигналов по отдельным частотным полосам. А это, в свою очередь, позволяет более надежно отделить полосы, содержащие сигнал, от шумовых полос и буквально отфильтровывать шум между гармоник сигнала (см. рис. 5)

Увеличение числа полос имеет и негативный эффект. Вместе с увеличением частотного разрешения спектрального анализа ухудшается его временное разрешение. Это связано с тем, что импульсы соответствующих разделительных фильтров становятся длиннее по времени и захватывают информацию о более длинных участках сигнала. Это приводит к тому, что спектрограмма размывается по времени и теряется возмож-

ность точно локализовать начало и конец звучания тех или иных событий в аудиосигнале. Это затрудняет подавление шума вблизи транзиентов – резких всплесков энергии аудиосигнала, соответствующих атакам нот или ударам перкуSSIONНЫХ инструментов. В результате вокруг транзиентов появляются области слабо подавленного шума, а сами транзиенты могут размываться по времени и терять четкость.

Для достижения наилучшего компромисса между частотным и временным разрешением спектрограммы большинство алгоритмов фиксируют число частотных полос на некоторой средней величине (порядка 1000). Однако существуют и более сложные алгоритмы, осуществляющие т.н. мультиразрешающий (multiresolution) анализ, адаптивно варьирующий разрешение анализа фонограммы для достижения наилучшей четкости спектрограммы на каждом ее участке. Таким образом они достигают высокого временного разрешения вблизи транзиентов и высокого частотного разрешения для тональных сигналов. Пример такого алгоритма есть в программе iZotope RX (алгоритм C).

Музыкальный шум

Одним из основных артефактов (нежелательных эффектов), остающихся от работы алгоритмов спектрального вычитания, является эффект «музыкального шума». Когда шумовой сигнал пропускается через гейт, случайные всплески сигнала могут вызвать кратковременные открытия гейта. В случае многополосного гейта в результирующем сигнале получаются случайные по времени и частоте кратковременные всплески узкополосного шума. На слух они воспринимаются как льющаяся вода или металлическое позвякивание и называются «музыкальным шумом» (musical noise). Этот неприятный артефакт часто более заметен, чем исходный ровный шум, так как слуху легче адаптироваться к ровному шуму и отключиться от него.

Если рассмотреть спектрограмму «идеального» белого шума (рис. 6, 7), то можно увидеть, что она не является константой. Спектрограмма сама является статистически случайной и колеблется вокруг некоторой средней спектральной плотности шума. Это приводит к тому, что в процессе спектрального вычитания отдельные точки на спектрограмме подавляются слабее других. Они и будут образовывать «музыкальный шум» (рис. 8).

Существует несколько способов борьбы с этим неприятным на слух явлением. Простейший способ – завышение оценки спектра шума, в результате чего шум будет подавляться сильнее и эффект музыкального шума возникать не будет. Однако вместе с этим будут сильнее подавляться низкоуровневые компоненты полезного сигнала, что негативно скажется на тембре.

Другой способ – неполное подавление шума – позволяет оставить в сигнале часть подавляемого шума для маскировки возникающего музыкального шума. Этот способ не устраняет музыкальный шум, а лишь делает его менее заметным при слабом шумоподавлении.

Наиболее популярный способ борьбы с музыкальным шумом – использование сглаженных оценок спектра сигнала. Сглаживание обычно осуществляется простым рекурсивным фильтром вдоль оси времени. Если рассматривать метод спектрального вычитания как многополосный гейт, то такое сглаживание аналогично введению времени атаки/восстановления в работу гейта. Это приводит к замедлению реакции гейта на транзиенты низкого уровня, а также образованию эхообразных шлейфов неподдавленного шума после спадов энергии сигнала (рис. 9).

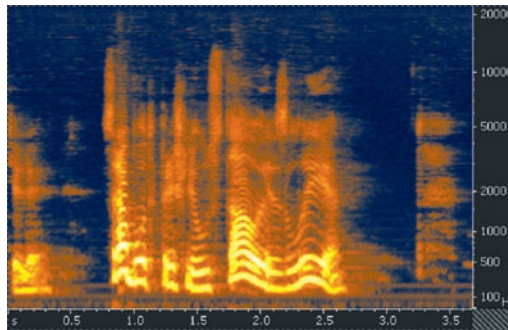


Рис. 9. Результат подавления «музыкального шума» сглаживанием по времени для сигнала, показанного на рис. 4

Более сложные методы подавления музыкального шума могут также производить сглаживание по частоте. Наиболее сложные алгоритмы сглаживания спектрограмм, заимствованные из области обработки изображений, реализованы в программе iZotope RX. Они анализируют двумерную структуру спектрограммы, пытаясь выявить признаки полезного сигнала, потерянного в шуме. Это позволяет устранить «шумовые эхо» после спадов энергии сигнала (рис. 10).

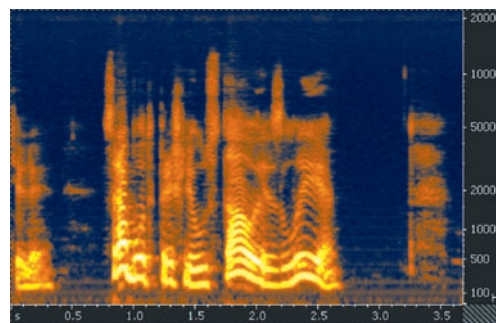


Рис. 10. Результат подавления «музыкального шума» двумерным адаптивным сглаживанием

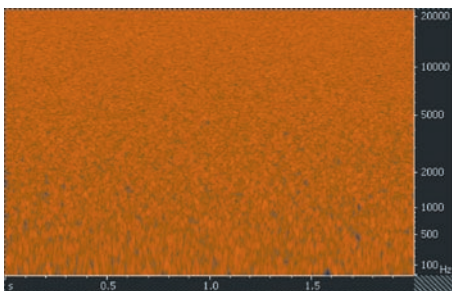


Рис. 6. Спектрограмма белого шума, видны случайные флуктуации энергии спектра

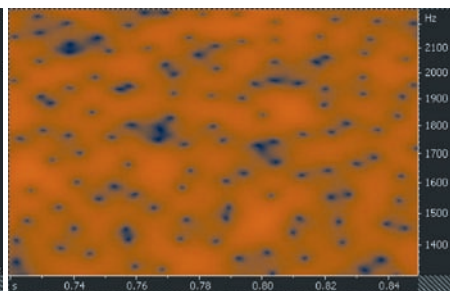


Рис. 7. Спектрограмма белого шума с увеличением

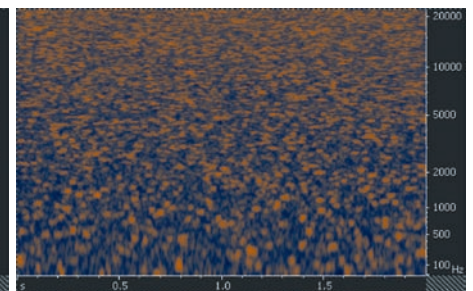


Рис. 8. Артефакт «музыкальный шум» после шумоподавления сигнала, показанного на рис. 6

Психоакустические модели

Во время звучания музыкального материала на фоне шума наблюдается явление психоакустической маскировки (см. статью И. Алдошиной «Основы психоакустики: слуховая маскировка», «Звукорежиссер» 2/2000). В результате этого многие области шума на частотно-временной плоскости становятся неслышимыми. В таких областях проводить подавление шума необязательно и даже нежелательно, ведь любое излишнее вмешательство в сигнал может его исказить.

Поэтому некоторые системы шумоподавления включают в себя психоакустическую модель, строящую пороги маскировки для результирующего аудиосигнала, и руководствуются этими порогами при выборе силы подавления шума на разных частотах.

Подавление различных видов шумов

Шумы, с которыми приходится сталкиваться звукорежиссеру, могут быть разных типов: тональными или случайными, строго стационарными или немного меняющимися во времени. Наиболее качественные системы шумоподавления способны эффективно подстраиваться под различные типы шума.

Так, тональные шумы обычно имеют достаточно стабильный спектр, не требующий сглаживания, но они часто нуждаются в более глубоком подавлении, чем широкополосные шумы, т.к. они могут вносить неприятную окраску.

Случайные шумы требуют подавления артефакта «музыкальный шум» с помощью сглаживания спектра.

Если шум нестабилен во времени, то нужны специальные алгоритмы адаптации, позволяющие оценить спектр шума в каждый момент времени и отличить его от спектра полезного сигнала.

Некоторые системы шумоподавления, например TC Electronic BackDrop и iZotope RX, разделяют шум на несколько составляющих, таких, как тональная и случайная, и позволяют независимо настроить параметры их подавления (рис. 11).



Рис. 11. Пример подавления лишь тональной части шума в программе iZotope RX

Бытует мнение, что гул наводок от электросети или другие тональные помехи лучше всего вырезать из сигнала режекторным фильтром или серией фильтров, настроенных на гармоники помехи. Однако при наличии системы шумоподавления, способной подавлять только тональный шум, более эффективным способом становится спектральное вычитание. Ведь режекторный фильтр подавляет и полезный сигнал, присутствующий на тех же частотах, что и помеха, а также вносит в сигнал существенные фазовые сдвиги или пред-эхо. А качественная

реализация метода спектрального вычитания будет и подавлять помехи, и пропускать без изменений полезный сигнал на тех же частотах.

Заключение

Системы шумоподавления играли важную роль на протяжении всей истории звукозаписи и принимали участие в становлении многих звуковых форматов.

Так, система Dynaural Noise Suppressor впервые сделала возможным проигрывание грампластинок по радио. Ранее считалось, что щелчки и другие помехи пластинок делают их непригодными для радио, где до этого транслировались лишь «живые» концерты и передачи. С появлением этой системы шумоподавления американские радиостанции стало быстро переходить на трансляцию заранее записанной музыки – к неудовольствию многих слушателей.

Появившиеся в 1963 году компакт-кассеты не пользовались большой популярностью из-за высокого уровня шума и ненадежности механики первых магнитофонов. Однако создание в начале 70-х годов бытовых систем шумоподавления Dolby B привело к быстрому распространению кассетного формата.

В свою очередь, развитие систем шумоподавления на основе спектрального вычитания сделало возможным переиздание многих записей со старых носителей на CD. Не всегда результаты такого ремастеринга оправдывают ожидания слушателей. Многие из них привыкли к звучанию и шумам несовершенного оригинального носителя, и чистое выхолощенное звучание материала после реставрации производит впечатление пустоты, бездушности, оторванности от эпохи. Нередко причиной такого негативного впечатления становятся и инженеры реставрации, неумеренно или неаккуратно использующие средства шумоподавления.

Для всех систем подавления широкополосного шума особенно актуален неизбежный компромисс между глубиной подавления шума и деградацией полезного сигнала. Такие низкоуровневые компоненты сигнала, как реверберация, мелкие щелчки, высокочастотная перкуссия, первыми страдают от излишнего шумоподавления. Другой бич систем спектрального вычитания – появление артефактов в виде «музыкального шума», методы борьбы с которым были рассмотрены в этой статье.

Оценка качества системы шумоподавления – процесс достаточно субъективный. С одной стороны, необходимо оценить целостность передачи исходного сигнала, с другой – степень подавления шума и возникновение нежелательных артефактов. Для разных типов материала оптимальными могут оказаться различные системы шумоподавления: одни из них могут быть «заточены» под речь и низкие соотношения сигнал/шум, другие – под музыкальный сигнал с достаточно высоким соотношением сигнал/шум. Одни могут оказаться «толерантными» по отношению к нестационарным или тональным шумам, другие – быть рассчитаны в первую очередь на белый шум.

Чтобы познакомить читателей со звуковыми примерами работы различных систем шумоподавления, мы организовали веб-страничку с примерами обработки одного короткого речевого фрагмента с комбинацией тонального и широкополосного шумов (<http://audio.rightmark.org/lukin/restoration/spectrum.us.htm>). Конечно, одного короткого фрагмента недостаточно для полноценного сравнения, однако из него можно получить представление о возможных артефактах. Мы приглашаем читателей принять участие в дискуссии о расширении теста другими аудиофрагментами и системами шумоподавления. ●