

Звуковой динамический процессор или телевизионный процессор громкости?

Окончание. Начало в №5/2012

Александр Сироткин,

аспирант СПбГУТ им. проф. М.А. Бонч-Бруевича

От теории – к практике!

Итак, в первой части были рассмотрены истоки проблемы. Теперь можно перейти к практическим вариантам ее решения.

Первое, что приходит на ум – изначально готовить все материалы с нормализацией по среднеквадратическому значению громкости (RMS). Для этого даже нет необходимости «портить» исходный файл, так как информация о громкости каждого звукового объекта может сохраняться в его метаданных, и при воспроизведении надо лишь автоматически понижать или повышать уровень, опираясь на значения этих метаданных. Данный вариант подходит для телеканалов, работающих исключительно с записанными материалами, но как быть телеканалам с прямыми эфирами?

До 2006 года у телевизионных вещателей был, по существу, единственный путь – применение для звукового сопровождения телепрограммы аудиопроцессора, осуществляющего динамическую обработку выходного сигнала, подаваемого в эфир. Использовались самые разные динамические процессоры, которые находят широкое применение в звукозаписи, при мастеринге или в радиовещании. Причем для телеканалов, транслирующих прямые эфиры, наилучшим вариантом представляется использование в качестве устройств финальной обработки процессоров, ориентированных на радиовещание, так как там решаются сходные задачи – в радиозэфире чередуются прямые трансляции, музыка, реклама, передачи, выступления ди-джея и т.д.

Казалось бы, вот оно – решение! И чем дороже и именитее прибор динамической обработки, тем звук должен стать более «сочным» и «прокачанным» – как на луч-

ших FM-радиостанциях! Однако не все так радужно, как кажется на первый взгляд.

Главная цель финальной динамической обработки, осуществляемой в FM-процессоре, – сделать звук максимально громким в заданных пределах допустимой девиации, очевидно, чтобы станция звучала громче (по меньшей мере, не тише) конкурента на соседней частоте. И для достижения этой цели могут не браться в расчет прочие параметры. Иллюстрация этого приведена на рис. 7.

Вот что происходит после предельно глубокой финальной обработки динамическим процессором звукового фрагмента, необработанная сигналограмма которого была рассмотрена в ч. 1. Рисунок хорошо иллюстрирует принципиальный подход в обработке, выполняемой динамическим процессором – сделать звук максимально компрессированным и предельно громким. Видно, что к динамическому диапазону рекламного ролика (или крика «ГООООЛ!») были «дотянуты» (то есть экстремально ужаты) все остальные фрагменты. Очевидно, что при

таком подходе звуковые фрагменты действительно будут выровнены, нормализованы, вот только вряд ли мы хотим слышать шум стадиона и гомон болельщиков столь же громко, как и овации, что неизбежно сопровождают залетевший в ворота мяч.

На первый взгляд, логичное решение – ослабить степень динамической обработки сигнала в процессоре. Но так как громкость ролика или оваций остается неизменной (и большой!) при выбранной парадигме на увеличение громкости (напомним – быть громче конкурента), то возвращаемся, в конечном счете, к тому, что имели в начале. И вроде бы трудится в звуковом тракте дорогой именитый прибор финальной обработки, а телезрители продолжают жаловаться на громкую рекламу – воз и ныне там.

Первый серьезный шаг, чтобы изменить сложившуюся ситуацию, был сделан в 2006 году, и западное сообщество аудиоинженеров быстро оценила его как настоящую революцию в мире аудио. Шаг этот – опубликование рекомендации BS.1770 Меж-

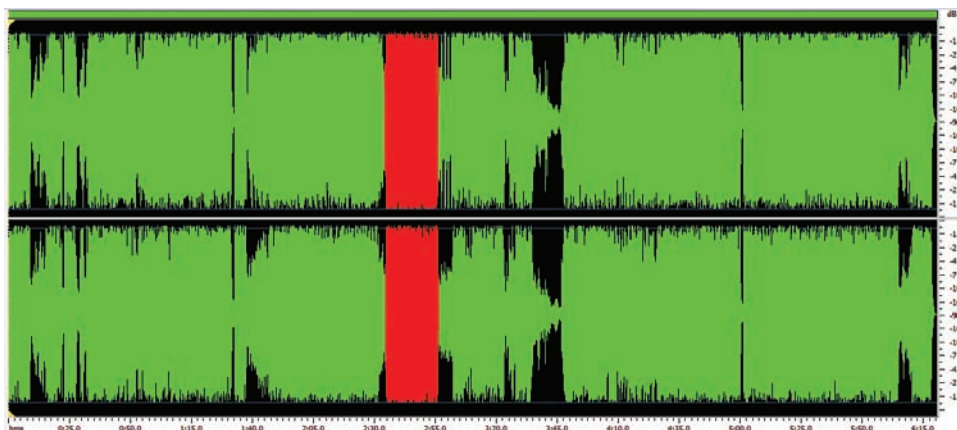



Рис. 7. Финальная динамическая обработка звуковым процессором



Не укротить **ГРОМКОСТЬ?**

У нас **есть решение!**

digiton[®]
systems

+7 812 324 66 42
www.digiton.ru

дународного Союза Электросвязи (секции «Радио»). Вкратце суть «революции» выражается в том, что МСЭ-Р был разработан простой и элегантный способ объективного измерения громкости звукового сигнала, имеющий высокую корреляцию с субъективной оценкой громкости слуховой системой человека. В основе этого метода лежит предложенная в 2004 году модель объективной оценки громкости монофонических сигналов. Опубликованная же в 2006-м рекомендация включила в себя модификацию этой модели, что позволило с равной степенью корреляции между объективным измерением и субъективным восприятием оценивать громкость как монофонических, так и стерео- и многоканальных сигналов (5.1). Другими словами, представленная модель позволяет оперировать сугубо техническими величинами и единицами измерения, которые, в свою очередь, с достаточной степенью точности эквивалентны субъективному восприятию ухом человека.

Таким образом, отпадает необходимость в жесткой нормализации звукового сопровождения телепрограммы, поскольку меняется сама парадигма оценки – теперь стало возможным принимать во внимание не столько электрический уровень сигнала или его амплитуду (в недостаточной степени информативную для оценки субъективно воспринимаемой громкости), а то, насколько громким данный фрагмент звука воспринимается слухом. Измерив таким способом громкость сигнала, можно сравнить его с целевой (желаемой) громкостью и соответствующим образом нормализовать. Это иллюстрируется рис. 8.

На рисунке в условной форме показано, что произойдет с исходным сигналом, который рассматривается на протяжении всей статьи. Видно, что основная часть программы осталась неизменной, тогда как рекламный ролик оказался значительно уменьшен по ам-

плитуде – до такой степени, чтобы его громкость (как она субъективно воспринимается слухом) оказалась выровнена с субъективно воспринимаемой громкостью соседних фрагментов. А так как субъективное восприятие громкости теперь можно измерить технически, задача финальной обработки свелась к нормализации звука в соответствии с заданной величиной объективного уровня громкости. Именно эту задачу и должен выполнять телевизионный процессор громкости.

Итак, зафиксируем одно из ключевых отличий в обработке, осуществляемой звуковым динамическим процессором для радиовещания и телевизионным процессором громкости. Состоит оно в том, что первый старается максимизировать громкость звука, тогда как задачей второго является нормализация различных фрагментов программы, то есть приведение их громкости к единой заданной величине. Причем компрессия исходного сигнала осуществляется только для тех фрагментов звука, которые простым линейным усилением уровня невозможно «дотянуть» до заданной величины громкости, в остальных же случаях динамические свойства сигнала остаются нетронутыми.

Таким образом, при использовании звукового динамического процессора в качестве устройства финальной обработки, когда принудительная компрессия осуществляется в любом случае, могут существенно пострадать художественные характеристики исходного материала (фильма, записанной телепередачи и пр.). Например, в случае задуманного звукорежиссером драматического эффекта за счет выверенного контраста уровней, динамический баланс сигнала этого фрагмента будет уже полностью настроен. Используя специализированный телевизионный процессор громкости, звукорежиссер может не опасаться пренебрежительного отно-

шения к своему труду – художественный баланс уровней будет сохранен и доберется в первоначальной форме до телезрителя. Или, быть может, телеслушателя...

Самое время перейти от творческих сфер к техническим величинам, поскольку исторически сложилось так, что в этом, казалось бы, очевидном вопросе далеко не все до конца прозрачно и ясно.

Единицы и величины: Америка и Европа

В качестве единицы измерения громкости Международным Союзом Электросвязи была введена единица LKFS (Loudness, K weighted, relative to nominal full scale), что расшифровывается как «громкость, взвешенная по кривой К, относительно номинальной полной шкалы» (данная кривая используется в модели измерения, описанной в Рекомендации BS.1770). Принятая единица непосредственно отражает величину громкости, и она эквивалентна децибелу, поскольку увеличение уровня сигнала на 1 дБ приведет к увеличению показателя громкости на 1 LKFS.

В то же самое время в опубликованной Рекомендации BS.1771, описывающей требования к приборам, измеряющим громкость, в качестве единицы относительной громкости используется LU (Loudness Unit, единица громкости), которая отражает смещение относительно опорной величины, выраженной в LKFS.

Вообще говоря, после выхода в свет Рекомендации BS.1770 Европейский Вещательный Союз начал вести активную научную деятельность по вопросам нормализации громкости. В 2010 году эта деятельность ознаменовалась опубликованием Рекомендации EBU R 128, принятой европейскими теле вещателями (в отличие от североамериканских коллег, техническим базисом для которых является Рекомендация МСЭ-Р BS.1770).

Надо сказать, что в Рекомендации EBU R 128 для измерения громкости используется модель, описанная в BS.1770, но дополненная и расширенная. Так, были введены новые дескрипторы (диапазон громкости, максимальный истинный пиковый уровень); введены три шкалы времени при измерении уровня громкости (мгновенная M, кратковременная S, интегральная I); при работе с интегральной шкалой введена функция стробирования, предназначенная для того, чтобы на величину измеряемой интегральной громкости сигнала не влияли длительные интервалы тишины.

Кроме того, интересно отметить некоторую терминологическую неувязку. В обеих упомянутых выше Рекомендациях МСЭ-Р

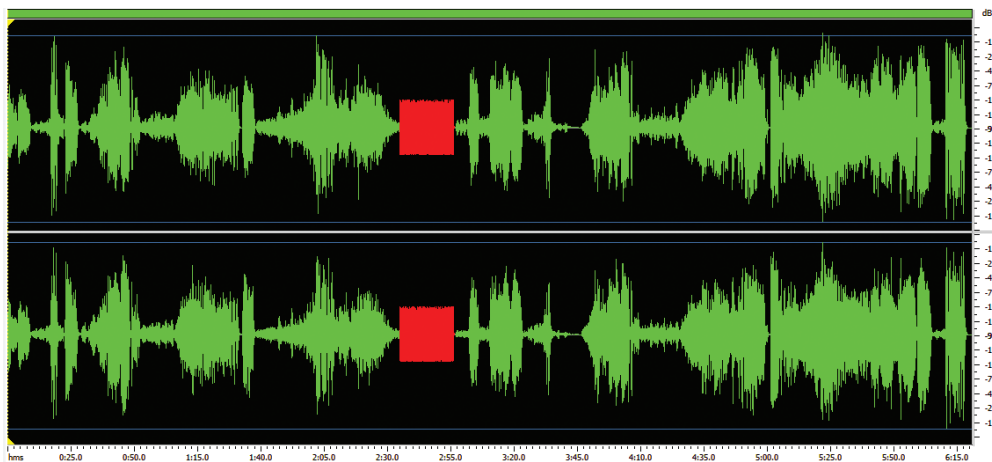


Рис. 8. Финальная динамическая обработка телевизионным процессором громкости

используется термин «громкость», но не «уровень громкости», что не согласуется с принятой практикой именованья для логарифмических величин. На это и ряд других несоответствий было обращено пристальное внимание, и, в частности, была дана рекомендация переименовать единицы LKFS в LUFS, что в большей степени соответствует международному стандарту именованья акустических величин ISO-80000-8.

Поэтому в европейской рекомендации EBU R 128 была введена эквивалентная дБ единица измерения LUFS, определяемая как «единица громкости относительно полной шкалы» (Loudness Unit, referenced to Full Scale). Таким образом, в Европе принято обозначение LUFS, в Соединенных Штатах – LKFS. Фактически же эти величины равнозначны.

В 2009 году североамериканский комитет по разработке перспективных стандартов в области ТВ (Advanced Television Systems Committee, ATSC) выпустил Рекомендацию A/85 (Techniques for Establishing and Maintaining Audio Loudness for Digital Television – «Методы создания и поддержания громкости звука для цифрового телевидения»), в основе которой лежит Рекомендация МСЭ-Р BS.1770.

Исходя из Рекомендации ATSC A/85, громкость сигнала устанавливается равной -24 LKFS с максимальным пиковым уровнем -2 dBFS. В 2010 году в ответ на жалобы телезрителей в США законодательно принимается так называемый CALM Act (Commercial Advertisement Loudness Mitigation Act), по которому каждый телеведущатель, кабельный оператор или любой другой поставщик видеослужб, осуществляющий деятельность на территории США, должен выполнять нормализацию громкости сигнала звукового сопровождения своей телепрограммы, исходя из Рекомендации ATSC A/85.

Тем временем в Европе, в рамках работ по громкости, проводимых Европейским Ведущательным Союзом, организуется экспертная группа PLOUD (поначалу именовавшаяся P/LOUD), которая занимается вопросами стандартизации разработки и внедрения решений на базе приведенной Рекомендации EBU R 128, а также прорабатывает дальнейшее развитие упомянутого алгоритма.

Экспертные оценки, проводимые группой PLOUD, работы которой периодически освещаются в прессе, сформировали следующие требования к параметрам нормализации: уровень громкости сигнала устанавливается равным -23 LUFS (вспомним принятые в США -24 LKFS в полностью эквивалентных величинах) с допустимым разбросом ± 1 LUFS, уровнем истинных пиков -1 dBTP (в США это -2 dBFS). Кроме того, в EBU R 128 введен порог -8 LU для срабатывания функции стробирования. Третья редакция Рекомендации BS.1770-2 также включает в себя функцию стробирования, описанную в EBU R 128, однако в качестве порогового значения срабатывания функции стробирования вводится значение -10 LU.

Как видно, и в этом вопросе европейцы и американцы разошлись, хотя и идут предельно параллельными курсами.

А что у нас? Выводы

В России принят «Федеральный закон от 13.03.2006 N 38-ФЗ», согласно которому «при трансляции рекламы ее звук не может быть громче звука транслируемой программы». При этом никоим образом не разъяснено, что подразумевается под термином «громче», и не дается конструктивных рекомендаций, чем и как эту самую громкость измерять. Как уже было упомянуто выше, все телеканалы соблюдают существующий сегодня регламент относительно громкости звукового сопровождения телевизионных программ, в котором предписана нормализация по пиковым уровням сигнала. Однако это ничуть не приближает к истинной нормализации – выравниванию уровней громкости рекламы и транслируемой программы в отношении к субъективному восприятию слуховой системой человека.

На сегодняшний день уже практически стандартом стал метод измерения громкости, изложенный в Рекомендациях МСЭ-Р BS.1770 и EBU R 128, на базе которого законодательно принимаются технические регламенты в ряде стран, а также сформировалось внушительное сообщество компаний-производителей оборудования. Не за горами коренные изменения и в стане отечественных вещателей.

Помимо метода, представленного в BS.1770 и EBU R 128, в мире существуют и другие алгоритмы измерения громкости, которые в чем-то могут превосходить его. Однако все фирменные (закрытые) алгоритмы, особенно если они рассчитаны на многополосную обработку, существенно сложнее в реализации. Данный же метод дает очень высокие показатели корреляции между объективным измерением и субъективной оценкой при довольно несложной реализации.

Несомненно, что некоторое время спустя и у нас будет разработан и внедрен технически продуманный закон, регламентирующий вопросы нормализации громкости. Будут ли в его основу положены ITU-R BS.1770 и EBU R 128? Наверняка. И тогда все телеведущатели в обязательном порядке станут нормализовать громкость.

Однако уже сейчас есть проверенное на множестве телеканалов в ряде западных стран средство избежать жалоб телезрителей на громкую рекламу, сохранив при этом столь трепетно оберегаемый звукорежиссерами на телевидении художественный баланс уровней громкости. Это средство – телевизионный процессор громкости, который избавит от бесконечного ропота недовольных потребителей и писем из ФАС.

Рекомендацию EBU R 128 на сегодняшний день поддерживают более полутора сотен компаний-производителей оборудования. Автору данной статьи и его коллегам посчастливилось на практике послушать результаты работы различных процессоров громкости – сравнивались исходные и прошедшие обработку сигналы реальных телевизионных программ различных жанров. С уверенностью можно сказать, что наибольших успехов на этом поприще достигла американская компания Linear Acoustic. Произведенные ею телевизионные процессоры громкости используются большинством американских телеканалов и крупнейшими европейскими вещателями – например, Discovery. Технологии от Linear Acoustic в виде готовых модулей применяют в своих решениях такие компании, как Axon, Miranda, Dolby, Cobalt, Ross Video, Snell, Radiant Grid и Wohler.

А вы уже выбрали свое средство экономии нервов?..



Rycote для ПРОФЕССИОНАЛОВ		ВЕТРОЗАЩИТА RYCOTE Мировой лидер в производстве ветрозащит			
		ВЕТРОЗАЩИТЫ МОДУЛЬНЫЕ ПОДВЕСЫ ЦЕПЕЛИНЫ		тел./факс: (495) 737-7440 e-mail: video@tivionica.ru www.rycote.ru	
				COMPANY HOLDING 	