

ВЛИЯНИЕ АЛГОРИТМОВ КОМПРЕССИИ КОНТЕНТ-СЕРВЕРА СИСТЕМЫ DRM НА КАЧЕСТВО ПЕРЕДАВАЕМЫХ ЗВУКОВЫХ ПРОГРАММ

DOI: 10.36724/2072-8735-2021-15-7-4-13

Соколов Сергей Анатольевич,
ООО «ДИГИТОН СИСТЕМС», г. Санкт-Петербург, Россия,
sokoloff@digiton.ru

Ковалгин Юрий Алексеевич,
Санкт-Петербургский государственный университет
телекоммуникаций им. проф. М.А. Бонч-Бруевича,
г. Санкт-Петербург, Россия, kowalgin@sut.ru

Manuscript received 02 April 2021;
Accepted 12 May 2021

Ключевые слова: цифровое радиовещание,
DRM-передатчик, скорость цифрового потока,
качество аудиоконтента

Изучено влияние алгоритмов компрессии кодеков MPEG-4 HE-AAC v.2 и MPEG-4 xHE-AAC на качество программ цифрового радиовещания. В первой части статьи на основе анализа опубликованных работ исследовано качество алгоритмов компрессии стандартов MPEG-4 ISO/IEC 14493-3 (AAC, AAC+SBR, PS, AAC+SBR+PS) и MPEG D Surround, применяемых в кодеках MPEG-4 HE-AAC v.2 и MPEG-4 xHE-AAC систем DAB и DRM. Показано, что их совместное применение позволяет при кодировании высококачественных звуковых сигналов радиовещания с полосой частот 40...15000 Гц уменьшить скорость цифрового потока на выходе кодера до 24...30 кбит/с без появления заметных на слух артефактов. Этот вывод подтверждается во второй части статьи. Здесь изложены результаты, полученные при сравнительной оценке качества воспроизведения звуковых сигналов, переданных по аналоговому ЧМ- и цифровому DRM-трактам. Описано оборудование, использованное при их реализации, методика получения тест-фонограмм для проведения субъективно-статистических экспертиз, приведены шкалы оценки сравниваемых пар звучаний. Результаты экспертиз подтвердили, что при скорости цифрового потока на выходе кодера MPEG-4 xHE-AAC равной 30 кбит/с имеет место слабо заметное различие в качестве звуковых сигналов с полосой частот 40...15000 Гц, передаваемых по аналоговым и цифровым трактам передачи, а также пар звучаний, полученных с выхода цифровых трактов вещательного процессора и DRM-приемника. Показано также преимущество в качестве и в удобстве применения на практике кодека MPEG-4 xHE-AAC по сравнению с его предшественником MPEG-4 HE-AAC v.2, сформулированы на основе накопленного опыта общие закономерности деградации качества, свойственные всем алгоритмам компрессии стандартов MPEG.

Информация об авторах:

Соколов Сергей Анатольевич, генеральный директор ООО ДИГИТОН СИСТЕМС, г. Санкт-Петербург, Россия

Ковалгин Юрий Алексеевич, д.т.н., профессор, Санкт-Петербургский государственный университет телекоммуникаций им. проф. М.А. Бонч-Бруевича, г. Санкт-Петербург, Россия

Для цитирования:

Соколов С.А., Ковалгин Ю.А. Влияние алгоритмов компрессии контент-сервера системы DRM на качество передаваемых звуковых программ // Т-Comm: Телекоммуникации и транспорт. 2021. Том 15. №7. С. 4-13.

For citation:

Sokolov S.A., Kovalgin Yu.A. (2021) Influence of the compression algorithms on the quality of audio programs in DRM digital radio broadcasting system. T-Comm, vol. 15, no.7, pp. 4-13. (in Russian)

Введение

Напомним, что скорость цифрового потока в канале пользователя MSC системы DRM при работе в режиме устойчивости E (диапазон ОВЧ) в зависимости от вида модуляции поднесущих частот OFDM-символа и скорости кода может меняться в пределах от 37,3 до 186,6 кбит/с. Несмотря на большое количество публикаций, посвященных оценке качества алгоритмов компрессии цифровых аудиоданных, работ посвященных оценке их применению на реальных трактах цифрового радиовещания крайне мало. Решению этой проблемы и посвящена данная работа.

Исследование качества алгоритмов компрессии контент-сервера и влияние характеристик DRM-передатчика на число и качество передаваемых звуковых программ и сопутствующей мультимедийной информации выполним в два этапа. На первом этапе, опираясь на данные публикаций, рекомендаций и отчетов ITU-R, определим минимально возможные скорости цифровых потоков на выходе кодера, при которых может быть обеспечено качество звучания эквивалентное ЧМ-радиовещанию при соответствующем выборе характеристик приемно-передающего тракта. На втором этапе подтвердим полученные результаты экспериментально методом субъективно-статистических экспертиз, сравнивая качество воспроизведения звуковых программ при их передаче по реальным радиоканалам ЧМ- и DRM-радиовещания.

Качество алгоритмов компрессии по данным публикаций. В системе DRM для компрессии цифровых аудиоданных используются алгоритмы, относящиеся к семейству стандартов MPEG – это стандарты MPEG-4 ISO/IEC 14496-3 [1] и MPEG D Surround Part 1 ISO/IEC 23003-1 [2]. В системах DAB и DRM разрешено применение двух кодеков – MPEG-4 HE-AAC v.2 и MPEG-4 xHE-AAC [3,4,6,7].

Анализ результатов многочисленных исследований [5,6,7,8,9,10] и накопленный собственный опыт свидетельствуют, что любому алгоритму компрессии стандартов MPEG присущи следующие общие особенности:

- пока установленное значение скорости цифрового потока таково, что доступное для кодирования число битов равно или больше требуемого психоакустической моделью для так называемого прозрачного кодирования, то изменения, вызванные при компрессии в сигнале более грубым квантованием в субполосах кодирования, все еще маскируются полезным сигналом, т. е. лежат ниже порога их слуховой заметности; в этом случае оба сигнала: исходный и прошедший процедуру кодирования в кодеке практически по качеству при воспроизведении практически не отличаются;

- для каждого алгоритма компрессии существует минимально-возможная скорость цифрового потока и максимально-возможная полоса частот, при которых качество звучания сигнала на входе и выходе кодера практически не отличаются. В этом случае требуемое психоакустической моделью число битов, необходимое для прозрачного кодирования, не превышает доступное их число при установленной пользователем скорости цифрового потока, а значит, и возможные артефакты компрессии лежат ниже их слуховой заметности;

- дальнейшее уменьшение скорости цифрового потока кодера сопровождается, прежде всего, понижением верхней

частоты звукового сигнала: биты перестают выделяться в верхние субполосы кодирования. При этом, чем меньше будет установленное значение скорости цифрового потока, тем ниже будет верхняя частота звукового сигнала после кодера; кроме сокращения полосы частот другие артефакты компрессии при слуховом восприятии пока еще не появляются;

- затем, когда доступных для кодирования битов станет слишком мало, появятся явно заметные и другие артефакты, вызванные компрессией; этот порог выбирается для каждого кодера исследователями, исходя из имеющегося у них опыта;

- для каждого алгоритма компрессии с потерями качество кодированного сигнала зависит не только от выбранной скорости цифрового потока, но и от структурных особенностей самого кодируемого сигнала, т.е. для сигналов разных жанров оно может претерпевать вполне заметные колебания. Даже в сигнале одного жанра качество компрессии для одного и того же алгоритма может меняться от одной выборки к другой с изменением структуры сигнала.

Оригиналом при оценке качества алгоритмов компрессии цифровых аудиоданных чаще всего является сигнал студийного тракта. Напомним, что полоса частот студийного тракта составляет 20...20000 Гц, в то время как полоса частот канала звука при FM-радиовещании равна 40...150000 Гц [13], в [14] указано значение 30...15000 Гц, т.е. значительно уже. Это различие по полосе частот канала звука является существенным при компрессии аудиоданных.

Следовательно, если верхняя частота канала звука при компрессии цифровых аудиоданных остается выше значения 15000 Гц, то можно ожидать, что при выбранном алгоритме компрессии, предназначенном для кодирования высококачественных звуковых сигналов, будет обеспечиваться качество близкое или эквивалентное ЧМ-радиовещанию. Это важное предположение, однако, оно требует экспериментальной проверки. Будем учитывать это соображение при оценке возможностей алгоритмов компрессии, рекомендованных для применения в системе DRM.

Обычно для слуховой (экспертной) оценки качества алгоритмов компрессии используют отрывки звучаний разных жанров, представленные на диске EBU SQAM. Они отображены экспертами группы MPEG, исходя из наибольшей слуховой заметности артефактов, вызванных компрессией. Это отрывки студийного качества, представляющие собой натуральные звучания с полным динамическим диапазоном, для каждого из них полоса частот канала звука специально не ограничена, составляет 20...20000 Гц, а внесенные изменения, вызванные предварительной обработкой, лежат существенно ниже их слуховой заметности.

Не все алгоритмы компрессии, имеющиеся в этих кодеках, обеспечивают возможность получения качества близкого к ЧМ-радиовещанию. Контент-сервер системы DRM содержит следующие алгоритмы компрессии аудиоданных [3].

1. Алгоритм MPEG-4 AAC (Advanced Audio Coding), базируется на опыте, накопленном при разработке алгоритма компрессии Layer 3 стандартов MPEG ISO/IEC 11172-3 и 13818-3; применяется для кодирования высококачественных звуковых сигналов сложной структуры. Существует множество публикаций, посвященных субъективной оценке качества этого алгоритма, да к тому же и множество его модификаций. Данные, заслуживающие наибольшего доверия, представлены на рисунке 1.

При скорости цифрового потока 96 кбит/с искажения, вызванные компрессией цифровых аудиоданных, для сигналов студийного качества еще остаются практически незаметными для большинства слушателей.

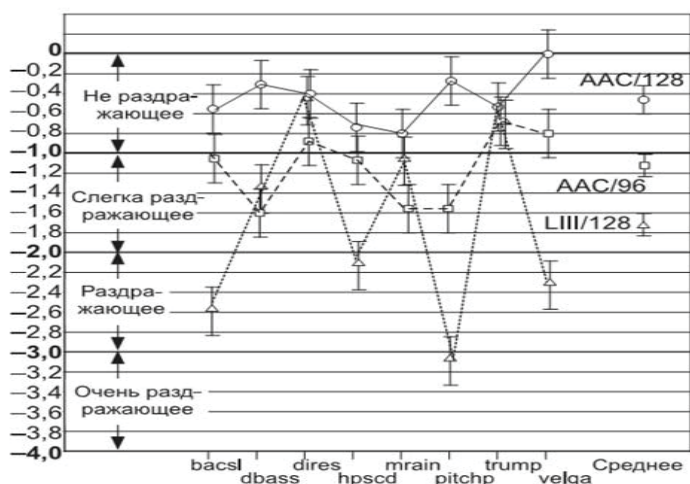


Рис. 1. Результаты слуховой оценки качества алгоритмов компрессии цифровых аудиоданных Layer 3 (LIII) и AAC для групп звучаний с ярко выраженными как спектральными, так и динамическими (ритмическими) неоднородностями (горизонтальная ось), на которых искажения, вызванные компрессией наиболее заметны. Отрывки звучаний взяты с компакт-диска EBU SQAM (студийное качество), [6,7]

При кодировании звукового сигнала с полосой частот 40...15000 Гц (эта полоса частот канала звука имеет место при ЧМ-радиовещании) при скорости цифрового потока 64 кбит/с отличия от оригинала (с полосой частот 40...15000 Гц), вызванные компрессией цифровых аудиоданных, также практически незаметны для слушателей. Субъективная оценка качества в этом случае соответствует градации «Отлично», («Excellent»), по шкале «MUSHRA», [3,6,7].

При скорости цифрового потока около 32 кбит/с искажения, вызванные компрессией цифровых аудиоданных, остаются почти незаметными (слегка раздражающими) для большинства слушателей при кодировании сигнала с полосой частот 40...15000 Гц, оценка качества «Хорошо» («Good») по шкале MUSHRA.

2. Алгоритм MPEG-4 AAC+SBR (SBR - Spectral Band Replication). Алгоритм SBR применяется для кодирования высокочастотной части спектра звукового сигнала с целью дополнительного уменьшения скорости цифрового потока. Обычно используется совместно с алгоритмом компрессии MPEG-4 AAC. При кодировании высокочастотной части спектра звукового сигнала с полосой частот 6000...15000 Гц алгоритм SBR обеспечивает скорость цифрового потока около 2 кбит/с. Данный алгоритм применяется также совместно и с алгоритмами CELP (CELP+SBR) и HVXC (HVXC+SBR). В этих двух последних случаях не обеспечивается качество передачи аудиоконтента эквивалентное ЧМ-радиовещанию). По этой причине эти последние два сочетания при работе системы DRM в диапазоне ОБЧ не применяются.

Совместное применение алгоритмов компрессии AAC+SBR позволяет получить следующие значения верхней частоты канала звука и скорости цифрового потока для звуковых сигналов радиовещания в случае, если [1]:

- кодируется монофонический сигнал, то возможны три значения верхней частоты канала звука: 10875 Гц при скорости цифрового потока 14...18,46 кбит/с; 13125 Гц при скорости цифрового потока 18,48...22,46 кбит/с; 15375 Гц при скорости цифрового потока 22,48...28,46 кбит/с;

- кодируется стереофонический сигнал формата 2/0; возможны два значения верхней частоты звукового сигнала на выходе кодека – это 13125 Гц (при скорости цифрового потока 26,48...28,48 кбит/с) и 15375 Гц (при скорости цифрового потока 28,48 кбит/с). Для звуковых сигналов разных жанров в этом случае наблюдается слабо заметное отличие от оригинала с полосой частот 40...15000 Гц, что соответствует оценке «Хорошо» по шкале MUSHRA (верхняя частота канала звука составляет 15375 Гц).

Результаты тестирования показывают (рис. 2, кодер MPEG-4 HE-AAC), что совместное применение алгоритмов AAC+SBR дает среднюю оценку «Хорошо» («Good» по шкале MUSHRA) при скорости цифрового потока 24 кбит/с. Здесь входной сигнал кодера имеет студийное качество с полосой частот 20...20000 Гц.

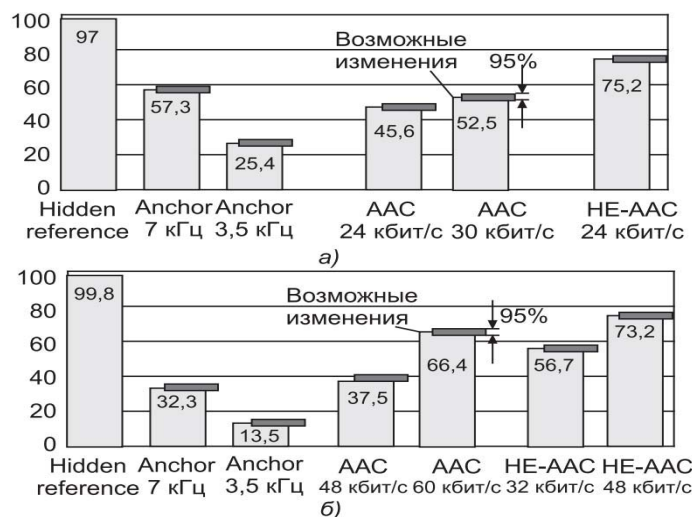


Рис. 2. Результаты тестирования кодеков стандарта MPEG – 4 AAC и MPEG-4 HE-AAC (AAC+ SBR) в шкале MUSHRA для разных значений скорости цифрового потока: а – монофонические испытательные сигналы; б – то же самое, но при стереофоническом тестовом сигнале. Слева – качество исходного сигнала, правее – качество исходного сигнала, ограниченное по полосе значением 7 кГц, правее – тоже самое, но при полосе 3 кГц, далее – оценки качества исходных сигналов на выходе соответствующего кодека. Здесь представлены результаты, усредненные для звуковых сигналов разных жанров [3,9]

3. Алгоритм MPEG-4 PS (Parametric Stereo). Идея параметрического кодирования заключается в выделении пространственных параметров сигналов стереопары, объединении исходной пары сигналов в один монофонический сигнал, последующем кодировании и передаче декодеру пространственных параметров исходного стереосигнала и полу-

ченного при матрицировании монофонического сигнала. Объединение сигналов стереопары в один монофонический сигнал происходит с помощью понижающего матрицирования (Downmix). Выделение пространственных параметров осуществляется в результате анализа субполосных сигналов, выделенные пространственные параметры квантуются, а затем кодируются с учетом психоакустики в параметрическом кодере. Для каждой такой пары идентичных субполос вычисляются обычно три пространственных параметра: разность уровней в дБ (как логарифм отношения энергий соответствующей пары одинаковых по ширине полосы субполосных сигналов левого и правого каналов стереопары), среднее различие по фазе и взаимная корреляция левого и правого субполосных сигналов.

Полученные цифровые потоки передаются по каналу связи к декодеру. На приемной стороне монофонический сигнал декодируется, после чего он используется для реконструкции исходного стереосигнала на основе переданных пространственных параметров. Данный алгоритм компрессии обеспечивает суммарную скорость цифрового потока при кодировании сигналов обычной стереофонии около 28 кбит/с при верхней частоте канала звука 15 000 Гц: 16...20 кбит/с при кодировании полученного монофонического сигнала и 6...8 кбит/с при кодировании пространственных параметров сигналов стереопары.

4. Совместное применение алгоритмов компрессии AAC+SBR+PS позволяет снизить скорость цифрового потока при кодировании стереосигнала формата 2/0 с полосой частот 40...15000 Гц уже до 24 кбит/с, правда, за счет некоторого дополнительного снижения качества по сравнению с оригиналом, полоса частот которого также составляет 40...15000 Гц. Кодек AAC Plus v.2, кодируя стереосигнал формата 2/0, при скорости цифрового потока 24 кбит/с дает такое же качество, как и кодек AAC при скорости цифрового потока 32 кбит/с. При этом отличие от оригинала в обоих случаях не превышает 1 балла по пятибалльной шкале оценки. Это усредненные данные для звуковых сигналов разных жанров [5, 10].

5. MPEG-D Surround; применяется только для кодирования сигналов многоканальной стереофонии (рис.3): скорость цифрового потока при передаче пространственных параметров многоканального сигнала составляет не более (6...8) ... (10...12) кбит/с. При скорости цифрового потока 160 кбит/с кодер MPEG-D Surround дает оценку «Отлично» («Excellent»), кодируя многоканальный сигнал студийного качества формата 5.1 [11]. Более подробные сведения о качестве алгоритмов компрессии при кодировании сигналов многоканальной стереофонии приведены в [16], где основное внимание уделено оценке качества алгоритмов компрессии AAC и HE-AAC (AAC+SBR), а также их модификациям. Как и ранее, используется шкала оценки MUSHRA.

При кодировании звукового сигнала формата 5.1 скорость цифрового потока может быть уменьшена примерно до 64 кбит/с с сохранением оценки «слабо заметное различие» по сравнению с оригиналом, имеющим полосу частот основных каналов до 15000 Гц.

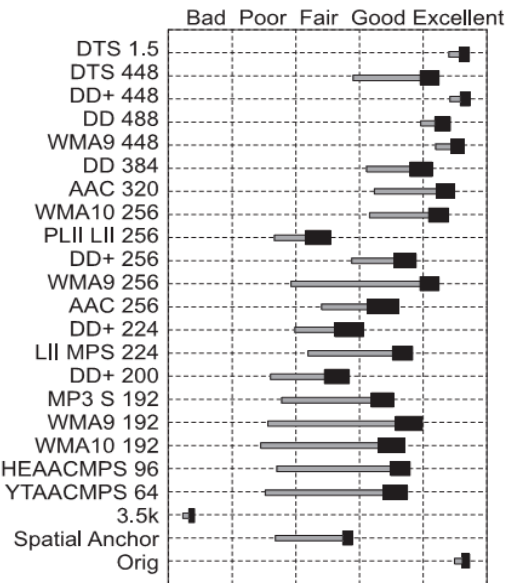


Рис. 3. Качество алгоритмов компрессии цифровых аудиоданных при кодировании многоканального сигнала, усредненные данные при кодировании сигналов студийного качества. Шкала оценки «MUSHRA», [11], [EBU tests of multi-channel audio codecs/ Convention Paper 7052, Вена, май 2007 г.]

6. MPEG-4 CELP (Code Excited Linear Prediction), применяется для кодирования речи, обеспечивает передачу речевых сигналов со скоростью 3, 86...14 кбит/с; в данном алгоритме компрессии полосы частот канала звука составляют 100...3800 Гц и 50...7000 Гц; в сочетании с алгоритмом SBR полоса частот канала звука расширяется до 14 кГц. При работе системы DRM в режиме устойчивости E алгоритм CELP не используется из-за низкого качества.

7. MPEG-4 HVXC (Harmonic Vector eXcitation Coding), предназначен для кодирования речевых и музыкальных сигналов с простой структурой, обеспечивает передачу со скоростью 2...6,56 кбит/с; при частоте дискретизации 8 кГц полоса частот канала звука равна 100...3800 Гц; при использовании совместно с алгоритмом SBR полоса частот канала звука расширяется до 8 кГц; его применение не обеспечивает высококачественную передачу звуковых сигналов с полосой частот 40...15000 Гц. При работе системы DRM в режиме устойчивости E не применяется.

Для большей наглядности алгоритмы компрессии, обеспечивающие передачу высококачественных звуковых сигналов (с полосой частот 40...15000 Гц) при наименьшей скорости цифровых потоков как при ЧМ-радиовещании приведены в таблице 1 и на рисунке 4, где приведены самые последние данные, посвященные сравнительной оценке качества кодеков MPEG-4 HE-AAC v.2 и MPEG-4 xHE-AAC.

Итак, по данным опубликованных работ применение алгоритмов компрессии AAC, AAC+SBR, PS, AAC+SBR+PS при передаче звуковых сигналов формата 2/0 (обычное стерео) является наиболее предпочтительным. В этом случае ожидаемая минимальная возможная скорость цифрового потока при кодировании высококачественного стереофонического сигнала формата 2/0 с полосой частот 40...15000 Гц должна лежать в пределах от 26 до 28 кбит/с, обеспечивая качество звучания как при ЧМ-радиовещании.

Таблица 1

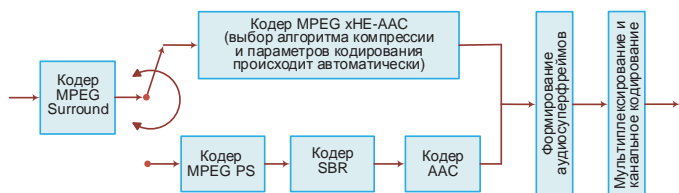
Ожидаемые (по данным публикаций) минимальные значения скорости цифровых потоков на выходе кодера MPEG-4 xHE-AAC, обеспечивающие передачу высококачественных звуковых сигналов с полосой частот 40...15 000 Гц с качеством не хуже, чем при ЧМ-радиовещании

Алгоритм компрессии	Ожидаемое минимальное значение скорости цифрового потока на выходе кодера, кбит/с
AAC	32
AAC+ SBR	26...28
PS	24...2
AAC+SBR+PS	
MPEG D Surround	56...64

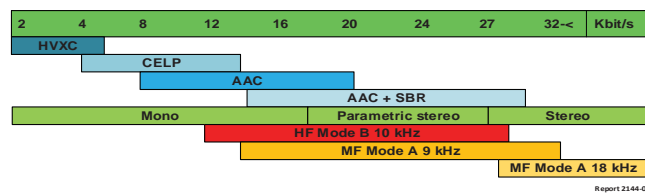
Заметим также, что кодер MPEG-4 xHE-AAC, являясь более поздней моделью, обеспечивает более высокое качество передачи аудиосигнала при существенно более низкой скорости цифрового потока в сравнении с кодером MPEG-4 HE-AAC v.2, рис. 4, в и г [3,4]. Причем, это различие становится наиболее заметным при скорости цифрового потока ниже 32 кбит/с. В тоже время при скорости цифрового потока от 64 кбит/с и выше оба кодера обеспечивают примерно одинаковое качество, соответствующее оценке «Отлично» или «Превосходно» по методу оценки качества звучания «MUSHRA», [4,9].

Большой выбор устанавливаемых параметров в кодере MPEG-4 HE-AAC v.2 затрудняет его практическое применение радиовещателями при отсутствии должного опыта. Кроме того (и это важно для практического применения) качество каждого из имеющихся алгоритмов существенно зависит от текущей структуры выборки. Один и тот же алгоритм обеспечивает разное качество, кодируя разные фрагменты (выборки) одного и того же звукового отрывка. В отличие от этого в кодере MPEG-4 xHE-AAC предусмотрена возможность только установки только скорости цифрового потока. По результатам анализа структуры сигнала выборки (без участия пользователя) автоматически выбирается алгоритм компрессии, обеспечивающий наиболее высокое качество при кодировании. Это важное изменение, позволяющее существенно упростить его эксплуатацию радиовещателями. В данном кодексе сохранены алгоритмы компрессии, реализованные в его предшествующей версии, MPEG-4 HE-AAC v.2, при этом в каждый из них внесены небольшие изменения, позволившие улучшить качество при компрессии.

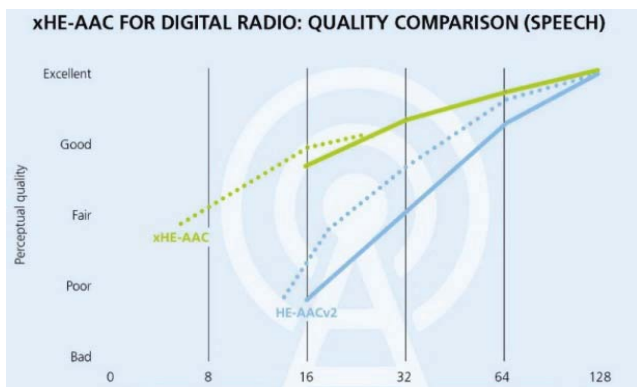
В области компрессии цифровых аудиоданных разработчики достигли по сути дела феноменальных результатов. При изначальной скорости студийных стереосигналов равной 1536 кбит/с или при записи компакт-дисков равной 1411 кбит/с они сумели, используя алгоритмы кодера MPEG-4 xHE-AAC, понизить скорость цифрового потока без появления заметных на слух обычному слушателю артефактов до значений близких к 24...30 кбит/с, кодируя стереосигнал *студийного* качества (более, чем в 50 раз!).



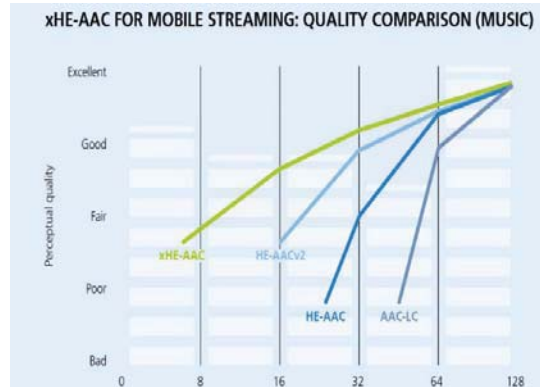
а)



б)



в)



г)

Рис. 4. К оценке качества аналогового ЧМ- и цифрового DRM-радиовещания: а – укрупненная структурная схема кодера системы DRM; б – рекомендуемые скорости цифровых потоков для разных алгоритмов компрессии; в – сравнительная перцепционная оценка качества кодеков MPEG-4 HE-AAC v.2 и MPEG-4 xHE-AAC для речи (слева) и музыки (справа) по данным компании RFmondial GmbH), метод и шкала оценки MUSHRA, [ETSI ES 201 980 V4.1.1 (2014-01); Report ISO/IEC JTC1/SC29/WG11]

Экспериментальная установка и метод сравнительной оценки качества ЧМ- и DRM-радиовещания в полосе частот 87,5...108 МГц диапазона ОВЧ. Структурная схема установки для получения звуковых сигналов, подлежащих слуховой оценке приведена на рис.5.

Звуковой сигнал студийного качества подвергался предварительной обработке в вещательном процессоре Omnia.9 с целью адаптации к характеристикам радиоканала, обеспечения однородности звучания разных типов фонограмм и учета психологических особенностей целевой аудитории слушателей, на которую ориентированы программы радиостанции.

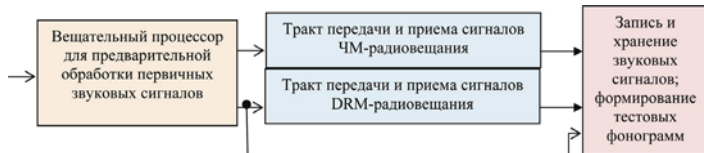


Рис. 5. Укрупненная структурная схема установки для получения звуковых сигналов аналогового и цифрового радиовещания для последующей слуховой оценки их качества

Перечень и характеристики используемого в ней оборудования представлены ниже.

Таблица 2

Оборудование для передачи, записи и формирования испытательных звуковых сигналов, несущая частота ЧМ-канала 95,9 МГц, DRM-канала 95,7 МГц

Название оборудования	Тип оборудования
Вещательный процессор с двумя отдельными каналами обработки звуковых сигналов студийного качества для их последующей подачи на передатчики аналогового (ЧМ) и цифрового (DRM) радиовещания. Канал обработки для УКВ-ЧМ вещания включает в себя стереокодер и лимитер комплексного стерео сигнала.	Omnia.9
Передатчик аналогового радиовещания, ЧМ-модулятор и радиочастотный усилитель	Полус ПТ
Приемник аналогового ЧМ-радиовещания	Marantz ST 251
Передатчик цифрового радиовещания стандарта DRM, включающий контент-сервер, DRM-модулятор и радиочастотный усилитель	Fraunhofer + RFmondial + Полус ПТ
Приемник цифрового DRM-радиовещания	RFmondial
Кодек системы DRM	MPEG-4 xHE-AAC
Устройство записи и хранения испытательных сигналов	Жесткий диск компьютера с профессиональной звуковой картой
Звуковоспроизводящее оборудование	Звуковые колонки Genelec 1029A, головные телефоны Beyer dynamic DT 770 PRO

Примечание: характеристики выбранного оборудования трактов ЧМ- и DRM-радиовещания удовлетворяют требованиям соответствующих стандартов, [12, 13].

Верхняя частота канала звука на выходах вещательного процессора для обоих трактов была установлена равной 17000 Гц. Девиация частоты ЧМ-передатчика составляла +/- 75 кГц, постоянная времени цепи предсказаний равнялась 50 мкс. Иными словами, все параметры передающего и приемного трактов системы ЧМ-радиовещания с пилот-тоном соответствовали ГОСТ Р 51197-97 «Системы стереофонического радиовещания. Основные параметры. Методы измерений» [13]. Текущие измерения параметров качества передатчика ЧМ-радиовещания были выполнены по ГОСТ Р 51741-2001 «Передатчики радиовещательные стационарные диапазона ОВЧ. Основные параметры, технические требования и методы измерений» [14]. При проведении измерений были учтены также требования рекомендаций 450-1 МККР и J.21 МККТТ.

Предсказание в канале ЦРВ системы DRM соответствовало рекомендации 651 МККР, для компрессии аудиоданных использовался кодек MPEG-4 xHE-AAC. Скорость цифрового потока на выходе кодера изменялась в пределах от 12 до 42 кбит/с ступенями через 6 кбит/с. Модуляция для поднесущих частот OFDM-блока системы DRM была выбрана QAM-4, скорость кода составляла 0,5. Вероятность появления цифровых ошибок в приеме-передающем тракте BER составляла менее 10^{-4} , что соответствует условиям комфортного прослушивания и требованиям стандарта [12].

При ЧМ-радиовещании использовался приемник Marantz ST 251, прием цифрового сигнала осуществлялся на контрольный приемник компании RFmondial. Звуковые сигналы снимались с линейных выходов приемников, подавались на профессиональную звуковую карту и далее на ноутбук для записи. Фрагменты записываемых звуковых отрывков формировались из текущей программы радиостанции, включали разные жанры: речь, речь на фоне музыки, музыкальные сигналы разных жанров и с разной динамикой. Длительности записываемых отрывков не превышали 1 минуты. Всего в общей сложности для последующей подготовки тест-фонограмм было отобрано 48 фрагментов звучаний.

База отобранных фрагментов звучаний включала отрывки с выхода приемо-передающих трактов ЧМ- и DRM-радиовещания, а также эти же отрывки с соответствующих выходов вещательного процессора. Полоса частот звуковых сигналов на выходах вещательного процессора для аналогового и цифровых трактов составляла 17000 Гц.

Из полученной базы звучаний были подготовлены тест-фонограммы, представляющие собой следующие друг за другом в случайной последовательности не известной слушателям пары одинаковых по содержанию звучаний, но полученные с разных точек экспериментального аналогового и цифрового трактов.

Для воспроизведения тест-фонограмм использовалось профессиональное оборудование студийного качества. Эксперты располагались на оси симметрии системы воспроизведения в точке оптимального слушания. Прослушивания программ выполнялись в помещениях, близких по характеристикам к условиям типовой жилой комнаты. Дополнительно для прослушивания звучаний использовались профессиональные головные телефоны. Полученные тест-фонограммы предварительно оценивались звукорежиссером радиостанции для выявления возможных дефектов.

Экспертами являлись слушатели, не имеющие дефектов слуха, преимущественно молодые люди, возраст которых не превышал 30-ти лет, в общей сложности число экспертов составило 50 человек. При жестком контроле условий проведения экспертиз для получения достоверных результатов достаточно иметь группу экспертов, состоящую из 20 человек [8]. Каждая пара отрывков в тест-фонограммах повторялась (через неизвестные слушателям промежутки времени) несколько раз, что необходимо для контроля психологического состояния экспертов в момент проведения прослушивания. Пары звучаний предлагались экспертам в неизвестной им случайной последовательности. При обработке результатов экспертиз не учитывались показания тех слушателей, которые давали разные оценки для пар звучаний, содержащих одинаковые отрывки звуковых сигналов.

При проведении сравнительного прослушивания пар отрывков экспертами оценивалась их качество по пятибалльной шкале (табл. 3), рекомендация МСЭ-R BS. 1284-2 «Общие методы субъективной оценки качества звука». Данная шкала оценки качества («отлично» – 5 баллов, «хорошо» – 4 балла, «удовлетворительно» – 3 балла, «неудовлетворительно» – 2 балла, «плохо» – 1 балл) согласована с так называемой непрерывной 100-балльной шкалой оценки качества «MUSHRA» (рис. 6), разделенной также на пять равных частей для соответствия со стандартной 5-балльной шкалой.

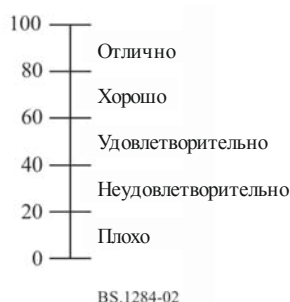


Рис.6. Непрерывная 100-балльная шкала, не содержащая знаков, с пятью словесными понятиями оценки качества звука [рекомендации МСЭ-R BS. 1284-2 и МСЭ-R BS.1534-2]

При обработке экспертопоказаний вычислялась средняя оценка по стандартной методике.

Таблица 3

Шкалы оценки качества и ухудшения качества [8,15,17]

Качество звучания	Ухудшение качества
5 баллов – Отлично	5 баллов – Незаметное
4 балла – Хорошо	4 балла – Заметное, но не раздражающее
3 балла – Удовлетворительно	3 балла – Слегка раздражающее
2 балла – Неудовлетворительно	2 балла – Раздражающее
1 балл – Плохо	1 балл – Очень раздражающее.

Результаты сравнительной оценки качества звучания. Ниже в качестве примера приведены типичные спектрограммы, полученные на выходах вещательного процессора, а также на выходах прямо-передающих трактов сис-

тем DRM (при разных скоростях цифровых потоков на выходе кодера) и ЧМ-радиовещания.

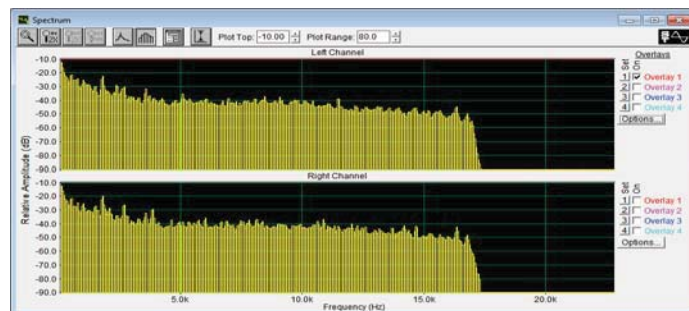


Рис. 7. Спектрограмма звукового сигнала на выходе вещательного процессора

Анализ полученных выше результатов и данных экспертиз позволяет сделать следующее заключение.

1. Поведение кодера MPEG-4 xHE-AAC соответствует общим закономерностям, наблюдаемым в кодеках MPEG с потерями при компрессии цифровых аудиоданных.

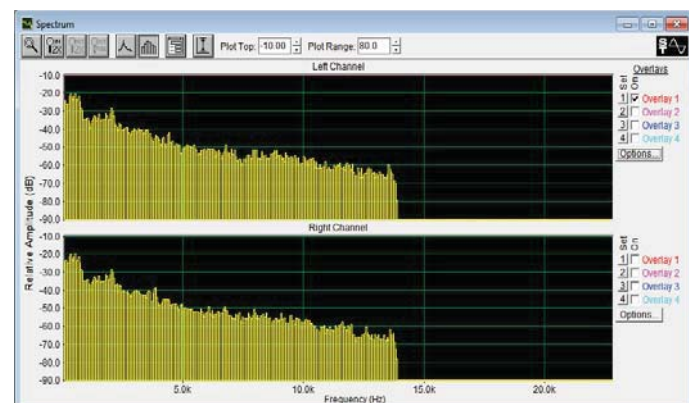


Рис. 8. Спектрограмма звукового сигнала на выходе DRM-приемника: скорость цифрового потока на выходе кодера 12 кбит/с; верхняя частота звукового сигнала составляет около 14 000 Гц

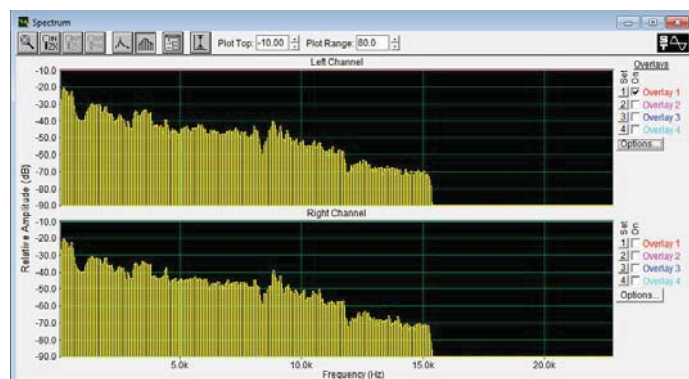


Рис. 9. Спектрограмма звукового сигнала на выходе DRM-приемника (скорость цифрового потока на выходе кодера 18 кбит/с; верхняя частота звукового сигнала составляет чуть выше 15 000 Гц)

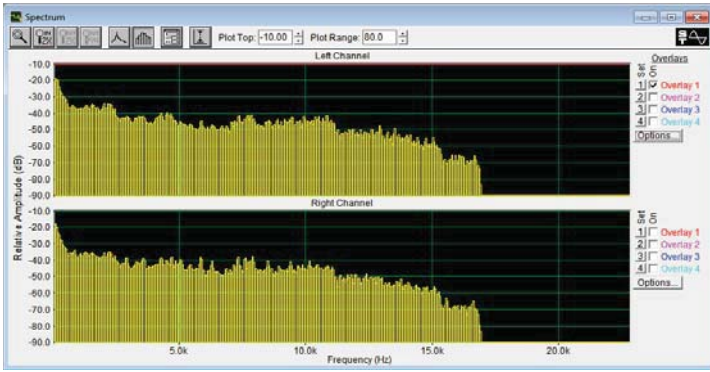


Рис. 10. Спектрограмма звукового сигнала на выходе DRM-приемника: скорость цифрового потока на выходе кодера 24 кбит/с; верхняя частота звукового сигнала составляет около 17 000 Гц, т.е. уже не отличается от верхней частоты входного сигнала

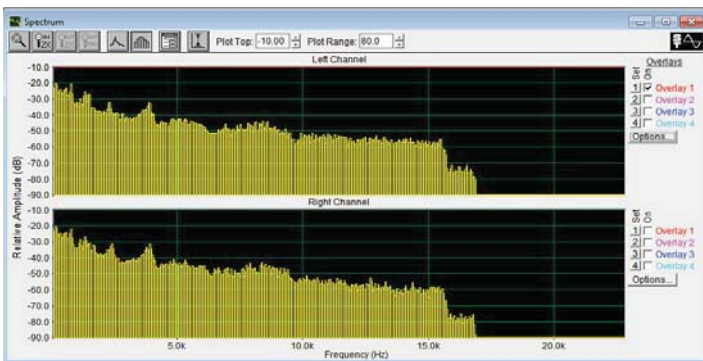


Рис. 11. Спектрограмма звукового сигнала на выходе DRM-приемника (скорость цифрового потока на выходе кодера 30 кбит/с; верхняя частота звукового сигнала составляет около 17 000 Гц, т.е. уже не отличается от верхней частоты входного сигнала)

2. Результаты экспертиз показали, что скорость цифрового потока на выходе кодера MPEG-4 xHE-AAC равная 30 кбит/с – это наименьшее значение, при котором качество принятого сигнала имеет отличие от исходного сигнала, практически незаметное для большинства обычных слушателей при прохождении DRM-тракта. Средняя оценка для обычных слушателей, лежит в интервале «Отлично» по шкале «MUSHRA», составляет 4,25 балла по пятибалльной шкале оценки качества (табл. 3). На рисунке 12 слева отложены баллы пятибалльной шкалы оценки, а справа – значения по шкале MUSHRA. Ширина доверительного интервала для экспертных оценок при всех значениях скорости цифрового потока не превышала 10% при вероятности 0,95. Все же следует отметить, что при значениях скорости цифрового потока менее 30 кбит/с эксперты отмечают более существенную деградацию качества звука в цифровом тракте по сравнению с данными, приведенными на рис.4, в и г. При скорости цифрового потока 12 кбит/с звучание в цифровом тракте по качеству становится неприемлемым: имеет место явно заметное пропадание верхних частот, резкое ухудшение прозрачности и четкости звучания, оно воспринимается как «кашеобразное», глухое, тембрально обедненное. При скорости цифрового потока 18 кбит/с сокращение полосы частот еще уверенно отмечается всеми экспертами, тем-

бральные изменения становятся на слух менее заметными, но четкость и прозрачность звучания по-прежнему не соответствует оригиналу. При скорости цифрового потока 24 кбит/с средняя оценка экспертов составляет 3,52 балла, что по шкале MUSHRA соответствует оценке «Хорошо». Иными словами, ухудшение качества здесь «заметное, но не раздражающее». Правда, следует отметить, что в нашем случае верхняя частота канала звука в DRM-тракте составляла 17000 Гц, что превышает аналогичное значение в канале звука при ЧМ-радиовещании (равное 15000 Гц). При ограничении верхней частоты в канале звука DRM-тракта значением 15000 Гц средняя оценка составляет уже 3,95 балла, табл.3).

3. Результаты проведенных экспертиз хорошо согласуются с данными, представленными компанией Fraunhofer (DRM-xHE-AAC-Demo_v2_20130913). Здесь представлены фрагменты звучаний трех жанров (речь, речь на фоне звучания оркестра, звучание оркестра), полученные на выходе кодера MPEG-4 xHE-AAC для скоростей цифрового потока на выходе кодера 12 кбит/с и 24 кбит/с в сравнении с оригиналом ЧМ-радиовещания. При скорости 24 кбит/с имеет место слабо заметное отличие от оригинала, что соответствует оценке внутри интервала «Отлично» – «Хорошо» по шкале «MUSHRA».

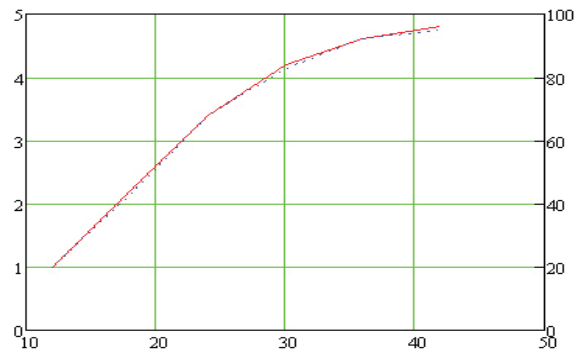


Рис. 12. Результаты субъективно-статистических экспертиз по оценке качества звукового сигнала на выходе приемного тракта системы DRM в функции от скорости цифрового потока кодера MPEG-4 xHE-AAC: усредненные данные для звуковых сигналов разных жанров. По горизонтальной оси отложена скорость цифрового потока, кбит/с; верхняя частота в канале звука 17000 Гц

4. Результаты субъективно-статистических экспертиз позволяют утверждать, что при модуляции QAM-4 и скорости кода $R=1/3$ по каналу DRM можно передать только одну стереопрограмму с качеством не хуже, чем при ЧМ-радиовещании. При модуляции QAM-16 и скорости кода $R=1/2$ число передаваемых звуковых программ может быть увеличено до 4-х. В обоих случаях для передачи сопутствующей и мультимедийной информации максимальная скорость цифрового потока составит около 20 кбит/с. Однако, переход к модуляции QAM-16 требует увеличения мощности DRM-передатчика при сохранении прежнего радиуса зоны покрытия.

Заключение

Для получения окончательных результатов необходимо провести субъективную оценку качества, используя отрывки реальных звучаний с компакт-диска EBU SQAM группы исследователей MPEG, который содержит специально отобранные отрывки реальных звучаний студийного качества, на которых искажения, вызванные компрессией цифровых аудиоданных наиболее заметны. Отрывки с диска EBU SQAM предварительно должны быть обработаны вещательным процессором, при этом верхняя частота канала звука должна быть ограничена значением 15000 Гц.

Следует помнить, что повторная компрессия цифровых аудиоданных, уже ранее подвергнутых этой процедуре недопустима, она может вызвать существенную деградацию качества исходного сигнала.

Литература

1. ISO/IEC FCD 14496-3 Subpart Information Technology-Very Low Bitrate Audio-Visual Coding. Part 3: Audio, 1998-05-10 (ISO/JTC 1/SC 29, N 2203).
2. ISO/IEC 203-1:2007, "Information Technology— MPEG Audio Technologies-Part 1: MPEG Surround," International Standards Organization, Geneva, Switzerland (2007).
3. Ковалгин Ю.А. Цифровое радиовещание: системы и технологии, [монография]. М.: Горячая линия – Телеком, 2021. 580 с.
4. Ковалгин Ю.А., Сантуш Виржилио. Влияние режимов работы DRM-передатчика на качество передачи аудиоконтента в диапазоне низких и средних частот // Труды учебных заведений связи, 2019.
5. Den Brinker A.C., Breebaart J., Ekstrand P., Engdegaerd J., Henn F., Kjoerling K., Oomen W., and Purnhagen H., Review Article. An Overview of the Coding Standard MPEG-4 Audio Amendments 1 and 2: HE-AAC, SSC, and HE-AAC v.2//EURASIP Journal on Audio, Speech, and Music Processing Volume 2009, Article ID 468971, 21 pages.
6. Ковалгин Ю.А. Психоакустика и компрессия цифровых аудиоданных: [монография]. СПб: Издательство СПбГУТ, 2012. 300 с.
7. Ковалгин Ю.А., Вологдин Э.И. Аудиотехника. Учебник для вузов. М.: Горячая линия – Телеком, 2013. 742 с.
8. Рекомендация ITU-R BS.1534-2 (06/2014). Метод субъективной оценки промежуточного уровня качества аудиосистем. Серия BS. Радиовещательная служба (звуковая).
9. Breebaart J., Steven van de Par, Kohlausch A, Schuijers E. Parametric Coding of Stereo Audio // EURASIP Journal on Applied Signal Processing. 2005. Vol. 9. P. 1305–1322.
10. Purnhagen H., Engdegaerd J., Oomen, and Schuijers E. Combining low complexity parametric stereo with High Efficiency AAC//ISO/IEC JTC1/SC29/WG11 MPEG2003/M10385, December 2003.
11. Here J., Kjoerling K., Breebaart J., Faller C., Dish S., Purnhagen H., Koppens J., Hilpert J., Roeden J., Oomen W., Linzmeier K., and Kok Seng Chong. MPEG Surround – The ISO/MPEG Standard for Efficient and Compatible Multichannel Audio Coding // J. Audio Eng. Soc., vol. 56, № 11, 2008, November, pages 932-952.
12. ETSI ES 201 980 v.4.1 (2017-04). Digital Radio Mondiale (DRM); System Specification. ETSI Standard, EBU Operating Eurovision. 196 p. (<http://www.etsi.org/standards-search>).
13. ГОСТ Р 51107-97. Системы стереофонического радиовещания. Основные параметры. Методы измерений. М.: Госстандарт России.
14. ГОСТ Р 51741-2001. Передатчики радиовещательные стационарные диапазона ОВЧ. Основные параметры, технические требования и методы измерений. М.: Госстандарт России
15. Рекомендация МСЭ-R BS. 1284-2 (01/2019). Общие методы субъективной оценки качества звука.
16. Report on the Verification Tests of MPEG-D Surround // ISO/IEC JTC 1/SC 29/WG 11 № 8851, 2007.
17. Рекомендация МСЭ-R BS.1116-3 (02/2015) Методы субъективной оценки небольшого ухудшения качества в звуковых системах Серия BS Радиовещательная служба (звуковая).

INFLUENCE OF THE COMPRESSION ALGORITHMS ON THE QUALITY OF AUDIO PROGRAMS IN DRM DIGITAL RADIO BROADCASTING SYSTEM

Sergei A. Sokolov, OOO Digiton Systems, St. Petersburg, Russia, sokoloff@digiton.ru

Yuri A. Kovalgin, St. Petersburg State University telecommunications them prof. M.A. Bonch-Bruevich, St. Petersburg, Russia, kovalgin@sut.ru

Abstract

This work studied the influence of MPEG-4 HE-AAC v.2 and MPEG-4 xHE-AAC compression algorithms on the reproduction performance of high-quality audio signals. The first part of the article is based on the analysis of earlier publications, studied quality of the compression algorithms for MPEG-4 ISO/IEC 14493-3 (AAC, AAC+SBR, PS, AAC+SBR+PS) and MPEG D Surround standards applied in MPEG-4 HE-AAC v.2 and MPEG-4 xHE-AAC codecs of DAB and DRM systems for digital audio broadcasting, respectively. It has been demonstrated that their use for processing of high-quality audio signals in the frequency range of 40...15000 Hz, allowed to reduce the digital-data rate at the output to 24...30 kbit/s, without observing noticeable phenomena. This conclusion was confirmed in the second part of the article, which described the results of relative comparison of the quality of reproduction of audio signals broadcasted via analogues FM or digital DRM channels. The equipment employed for their realization was described, as well as the methodology to obtain test recordings in order to perform the statistical data analysis. Further, the specific data scales were shown to evaluate the signal pairs being compared. Results of the experiments and their analysis confirmed that, for a digital data rate of 30 kbit/s obtained at the output of the MPEG-4 xHE-AAC coder, the difference in quality was rather weak for audio signals ranging from 40 to 15000 Hz, broadcasted via both analogues and digital channels. In addition, only a marginal difference was noticed for the signal pairs received from the digital output channels of the transmitter processor and DRM receiver. It has been demonstrated that the MPEG-4 xHE-AAC codec, compared to the previous MPEG-4 HE-AAC v.2 version, exhibited clear advantages for practical applications in terms of sound quality. Based on the gained experience, common criteria of deterioration of the data quality were formulated and generalized for all audio compression MPEG standards.

Keywords: digital audio broadcasting, DRM-transmitter, digital data rate, audio-content quality.

References

1. ISO/IEC FCD 14496-3 Subpart 1. Information Technology-Very Low Bitrate Audio-Visual Coding. Part 3: Audio, 1998-05-10 (ISO/JTC 1/SC 29, N 2203).
2. ISO/IEC 23003-1:2007, "Information Technology- MPEG Audio Technologies-Part 1: MPEG Surround," International Standards Organization, Geneva, Switzerland (2007).
3. Kovalgin Y.A. (2021). Digital audio broadcasting: systems and technologies, [monograph]. Moscow: Hot line – Telecom. 580 p.
4. Y.Kovalgin, V. Santos. (2019). Influence of DRM-Transmitter Operation Mode of the Quality of Audio Content Transmission in the Low and Medium Frequency Range. *Proceedings of Telecommunication Universities*. 5(1):56-63 (in Russ) Available from: <https://doi.org/10.31854/1813-324X-2019-5-1-56-63>.
5. Den Brinker A.C., Breebaart J., Ekstrand P., Engdegaard J., Henn F., Kjoerling K., Oomen W., and Purnhagen H., (2009). Review Article. An Overview of the Coding Standard MPEG-4 Audio Amendments 1 and 2: HE-AAC, SSC, and HE-AAC v.2//*EURASIP Journal on Audio, Speech, and Music Processing* Vol. 2009, Article ID 468971, 21 p.
6. Kovalgin Yu. A. (2012). Psychoacoustics and compression (encoding) of digital audio-data: [monograph] . SPb.: SPbSUT. 300 p.
7. Kovalgin Y.A., Vologdin E.I. (2013). *Audiotechnics. Textbook for Universities*. Moscow: Hot line – Telecom. 742 p.
8. Recommendation ITU-R BS.1534-2 (06/2014). Method for the subjective assessment of intermediate quality level of audio systems. BS Series. Broadcasting service (sound).
9. Breebaart J., Steven van de Par, Kohlausch A, Schuijers E. (2005). Parametric Coding of Stereo Audio. *EURASIP Journal on Applied Signal Processing*. Vol. 9. P. 1305-1322.
10. Purnhagen H., Engdegaard J., Oomen, and Schuijers E. Combining low complexity parametric stereo with High Efficiency AAC//ISO/IEC JTC1/SC29/WG11 MPEG2003/M10385, December 2003.
11. Here J., Kjoerling K., Breebaart J., Faller C., Dish S., Purnhagen H., Koppens J., Hilpert J., Roeden J., Oomen W., Linzmeier K., and Kok Seng Chong. (2008). MPEG Surround – The ISO/MPEG Standard for Efficient and Compatible Multichannel Audio Coding//*J. Audio Eng. Soc.*, vol. 56, no. 11, November. P 932-952.
12. ETSI ES 201 980 v.4.1.2 (2017-04). Digital Radio Mondiale (DRM); System Specification. ETSI Standard, EBU Operating Eurovision. - 196 p. (<http://www.etsi.org/standards-search>).
13. GOST P 51 107-97. Stereophonic broadcasting systems. Main parameters. Methods of measurements. M.: Go standard Russia.
14. GOST. P 51741-2001. Broadcasting transmitters, fixed Very High Frequency (VHF). Main parameters, technical requirements and methods measurement. M.: Go standard Russia.
15. Recommendation ITU-R BS.1284-2 (01/2019). General methods for the subjective assessment of sound quality. BS Series. Broadcasting service (sound).
16. Report on the Verification Tests of MPEG-D Surround//ISO/IEC JTC 1/SC 29/WG 11 № 8851, 2007.

Information about authors:

Sergei A. Sokolov, the General Director of OOO Digiton Systems, St. Petersburg,, Russia

Yuri A. Kovalgin, doctor of technical sciences, professor, department of television and metrology of St. Petersburg State University telecommunications them prof. M.A. Bonch-Bruevich, St. Petersburg, Russia