ФЕДЕРАЛЬНОЕ АГЕНТСТВО

ПО ТЕХНИЧЕСКОМУ РЕГУЛИРОВАНИЮ И МЕТРОЛОГИИ

НАЦИОНАЛЬНЫЙ СТАНДАРТ РОССИЙСКОЙ ФЕДЕРАЦИИ

ГОСТ Р 53556.9– 2013

Звуковое вещание цифровое

Кодирование сигналов звукового вещания с сокращением избыточности для передачи по цифровым каналам связи

Часть III

(MPEG–4 audio)

Методы кодирования звука MPEG–1/2 Audio в MPEG–4

Основные технические требования

ISO/IEC 14496-3:2009
(NEQ)

Издание официальное

Москва
Стандартинформ
2014

gost.e-ecolog.ru

Предисловие

1 РАЗРАБОТАН Санкт-Петербургским филиалом Центрального научно-исследовательского института связи «Ленинградское отделение» (ФГУП ЛО ЦНИИС)

2 ВНЕСЕН Техническим комитетом по стандартизации № 480 «Связь»

3 УТВЕРЖДЕН И ВВЕДЕН В ДЕЙСТВИЕ Приказом Федерального агентства по техническому регулированию и метрологии от 22 ноября 2013 г. № 1721-ст

4 Настоящий стандарт разработан с учетом основных нормативных положений международного стандарта ИСО/МЭК 14496-3:2009 «Информационные технологии. Кодирование аудиовизуальных объектов. Часть 3. Звуковое кодирование» (ISO/IEC14496-3:2009 «Information technology - Coding of audio-visual objects - Part 3: Audio» (NEQ)

5 ВВЕДЕН ВПЕРВЫЕ

Правила применения настоящего стандарта установлены в ГОСТ Р 1.0—2012 (раздел 8). Информация об изменениях к настоящему стандарту публикуется в ежегодном (по состоянию на 1 января текущего года) информационном указателе «Национальные стандарты», а официальный текст изменений и поправок — в ежемесячном формационном указателе «Национальные стандарты». В случае пересмотра (замены) или отмены настоящего стандарта соответствующее уведомление будет опубликовано в ближайшем выпуске ежемесячного информационного указателя «Национальные стандарты». Соответствующая информация, уведомление и тексты размещаются также в информационной системе общего пользования — на официальном сайте Федерального агентства по техническому регулированию и метрологии в сети Интернет (gost.ru)

© Стандартинформ, 2014

Настоящий стандарт не может быть полностью или частично воспроизведен, тиражирован и распространен в качестве официального издания без разрешения Федерального агентства по техническому регулированию и метрологии

НАЦИОНАЛЬНЫЙ СТАНДАРТ РОССИЙСКОЙ ФЕДЕРАЦИИ

Звуковое вещание цифровое

КОДИРОВАНИЕ СИГНАЛОВ ЗВУКОВОГО ВЕЩАНИЯ С СОКРАЩЕНИЕМ ИЗБЫТОЧНОСТИ ДЛЯ ПЕРЕДАЧИ ПО ЦИФРОВЫМ КАНАЛАМ СВЯЗИ. ЧАСТЬ III (MPEG–4 AUDIO)

Методы кодирования звука MPEG–1/2 Audio в MPEG–4
Основные технические требования

Sound broadcasting digital.
Coding of signals of sound broadcasting with reduction of redundancy for transfer on digital communication channels. A part III (MPEG–4 audio).
Main positions MPEG–1/2 Audio in MPEG–4

Дата введения – 2014—09—01

1 Область применения

Стандарт *MPEG*–1/2 Аудио в *MPEG*–4 спецификации *MPEG*–4 Аудио определяет использование *MPEG*–1/2 уровня 1, 2 или 3 ориентированным на *MPEG*–4 способом, то есть так, что сигнализация и доступ до обработки на системном уровне идентичны другим типам объектов *MPEG*–4 Аудио.

Чтобы перенести фреймы потока битов *MPEG*–1/2 уровень 1, 2 или 3 в *MPEG*–4, они переформатируются таким образом, что становятся автономными единицами доступа *MPEG*–4. Это облегчает транспортировку по пакетным сетям, произвольный доступ, и возможность редактирования. Автономные единицы доступа, которые используются в системах совместимой транспортировки или формата хранения *MPEG*–4, могут быть переконвертированы в совместимые с *MPEG*–1/2 потоки битов и затем декодированы любым совместимым с *MPEG*–1/2 декодером.

Синтаксис *MPEG*–4 Аудио дополнительно расширяется, чтобы сделать возможными многоканальные конфигурации на базе ГОСТ Р 54711 и ГОСТ Р 54712. Многоканальные конфигурации подобны конфигурациям, определенным для других аудио объектных типов *MPEG*–4 с многоканальными возможностями. Для *MPEG*–1/2 уровня 1 и 2 формат не расширяется. Многоканальный формат для этих уровней описывается в ГОСТ Р 54712.

Разрешенные частоты дискретизации для уровня 3 расширяются для речевого выхода инструментов *FA* и для дублирования *MP* с информацией о форме губ.

Для использования *MPEG*–1/2 уровня 1, 2 или 3 в *MPEG*–4 посредством унаследованного интерфейса *MPEG* 4, используется *ObjectTypeIndication* 0x69 или 0x6*b*.

2 *MPEG_1_2_SpecificConfig*

Синтаксис	Количество битов	Мнемоника
MPEG_1_2_SpecificConfig ()		
{		
extension;*	1	*bslbf*
}		

**extension* должно быть нулем.

3 Отображение канала

Применяются следующие правила:

элементы *single_channel_element () 's* и *lfe_element () 's* представляются монофоническими аудиофреймами;

элементы *channel_pair_element () 's* представляются стереофоническими аудиофреймами;

для уровня 1 и уровня 2 разрешается не больше одного монофонического аудиофрейма, представляющего *single_channel_element* (), или одного стереофонического аудиофрейма, представляющего *channel_pair_element* ().

Издание официальное

4 Формат единицы доступа

4.1 Уровень 1 и 2

Один аудио фрейм отображается непосредственно одной единицей доступа.

4.2 Уровень 3

Одна единица доступа состоит из одного или нескольких элементов *mp3_channel_elements*. Элемент *mp3_channel_element* равняется аудиофрейму уровня 3 со следующими изменениями по сравнению с его определением в ГОСТ Р 54711 или ГОСТ Р 54712:

syncword (12 битов)	Сообщает полную длину *mp3_channel_element* (состоящего из заголовка, *error_check*, дополнительной информация и основных данных) в байтах.
main_data_begin (9/8 бит)	Устанавливается в корректное значение соответствующее, потоку битов *MPEG*–1/2 уровня 3, или обнуляется.
main_data ()	Обычно сохраняется после допонительной информации

Все прочие элементы данных должны быть установлены согласно их спецификации в ГОСТ Р 54711 или ГОСТ Р 54712. Все установки в заголовке должны соответствовать установкам в *AudioSpecificConfig* ().

Все элементы *mp3_channel_elements*, принадлежащие одной и той же метке времени, сохраняются последовательно в одной единице доступа

5 Расширение частоты дискретизации для уровня 3

В этом пункте приводятся спецификации, позволяющие использовать уровень 3 с частотами дискретизации, не определенными в ГОСТ Р 54711 или ГОСТ Р 54712.

Синтаксис и описание потока битов для расширения частот дискретизации ниже, чем определенные в ГОСТ Р 54711-2012, находятся в соответствии с ГОСТ Р 54712-2012 (один фрейм охватывает 576 выборок).

5.1 Скорости передачи

Таблица 1 определяет скорость передачи в зависимости от *bitrate_index* и частоты дискретизации.

Таблица 1 – Скорость передачи в зависимости от *bitrate_index* и частоты дискретизации

bitrate_index	Определенная скорость передачи, Кбит/с		
	8, 11,025, 12 кГц	16, 22,05, 24 кГц (см. ГОСТ Р 54712)	32, 44,1, 48 кГц (см. ГОСТ Р 54711),
'0000'	Запрещено	Запрещено	Запрещено
'0001'	8	8	32
'0010'	16	16	40
'0011'	24	24	48
'0100'	32	32	56
'0101'	40	40	64
'0110'	48	48	80
'0111'	56	56	96
'1000'	64	64	112
'1001'	Запрещено	80	128
'1010'	Запрещено	96	160
'1011'	Запрещено	112	192
'1100'	Запрещено	128	224
'1101'	Запрещено	144	256
'1110'	Запрещено	160	320
'1111'	Запрещено	Запрещено	Запрещено

5.2 Частота дискретизации

В зависимости от частоты дискретизации, сообщенной в *AudioSpecificConfig*, элемент данных *sampling_frequency* в заголовке должен быть установлен, как определено в таблице 2.

Таблица 2 – Установка элемента данных *sampling_frequency* в зависимости от частоты дискретизации, определенной в *AudioSpecificconfig* ()

sampling_frequency	Частота дискретизации
00	11,025 кГц и ее кратные
01	12 кГц и ее кратные
10	8 кГц и ее кратные
11	Зарезервировано

5.3 Дополнение

Дополнение необходимо при частоте дискретизации 11,025 кГц и ее кратных.

5.4 Полосы масштабного коэффициента

Подразделение спектра на scalefactor полосы фиксируется для каждого размера блока и частоты дискретизации и сохраняется в таблицах в кодере и декодере. Таблицы для частот дискретизации, не определенные в ГОСТ Р 54711 или ГОСТ Р 54712, определяются в Приложении А. В соответствии с ГОСТ Р 54711 или ГОСТ Р 54712 масштабный коэффициент для частоты выше самой высокой строки в таблицах является нулем, что означает, что фактический фактор умножения равен 1,0.

Приложение А
(обязательное)

Таблицы полосы масштабного коэффициента

Таблица А.1 – Частота дискретизации 8 кГц, длинные блоки, число линий 576

Полоса масштабного коэффициента	Ширина полосы	*index_of_start*	*index_of_end*
0	12	0	11
1	12	12	23
2	12	24	35
3	12	36	47
4	12	48	59
5	12	60	71
6	16	72	87
7	20	88	107
8	24	108	131
9	28	132	159
10	32	160	191
11	40	192	231
12	48	232	279
13	56	280	335
14	64	336	399
15	76	400	475
16	90	476	565
17	2	566	567
18	2	568	569
19	2	570	571
20	2	572	573

Таблица А.2 – Частота дискретизации 8 кГц, короткие блоки, число линий 192

Полоса масштабного коэффициента	Ширина полосы	*index_of_start*	*index_of_end*
0	8	0	7
1	8	8	15
2	8	16	23
3	12	24	35
4	16	36	51
5	20	52	71
6	24	72	95
7	28	96	123
8	36	124	159
9	2	160	161
10	2	162	163
11	2	164	165

Таблицы А.3 – Частота дискретизации 11,025 кГц, длинные блоки, число линий 576

Полоса масштабного коэффициента	Ширина полосы	*index_of_start*	*index_of_end*
0	6	0	5
1	6	6	11
2	6	12	17
3	6	18	23
4	6	24	29
5	6	30	35
6	8	36	43
7	10	44	53
8	12	54	65
9	14	66	79
10	16	80	95
11	20	96	115
12	24	116	139
13	28	140	167
14	32	168	199
15	38	200	237
16	46	238	283
17	52	284	335
18	60	336	395
19	68	396	463
20	58	464	521

Таблица А.4 – Частота дискретизации 11,025 кГц, короткие блоки, число линий 192

Полоса *scalefactor*	Ширина полосы	*index_of_start*	*index_of_end*
0	4	0	3
1	4	4	7
2	4	8	11
3	6	12	17
4	8	18	25
5	10	26	35
6	12	36	47
7	14	48	61
8	18	62	79
9	24	80	103
10	30	104	133
11	40	134	173

Таблицы А.5 – Частота дискретизации 12 кГц, длинные блоки, число линий 576

Полоса масштабного коэффициента	Ширина полосы	*index_of_start*	*index_of_end*
0	6	0	5
1	6	6	11
2	6	12	17
3	6	18	23
4	6	24	29
5	6	30	35
6	8	36	43
7	10	44	53
8	12	54	65
9	14	66	79
10	16	80	95
11	20	96	115
12	24	116	139
13	28	140	167
14	32	168	199
15	38	200	237
16	46	238	283
17	52	284	335
18	60	336	395
19	68	396	463
20	58	464	521

Таблица А.6 – Частота дискретизации 12 кГц, короткие блоки, число линий 192

Полоса масштабного коэффициента	Ширина полосы	*index_of_start*	*index_of_end*
0	4	0	3
1	4	4	7
2	4	8	11
3	6	12	17
4	8	18	25
5	10	26	35
6	12	36	47
7	14	48	61
8	18	62	79
9	24	80	103
10	30	104	133
11	40	134	173

Приложение Б
(справочное)

Преобразование потоков битов *MPEG*–1/2 уровень 3 в элементы *mp3_channel_elements*

Использование битового накопителя обычно вызывает запуск появления *main_data* ()в прошлом фрейме потока битов. Это следует изменить, перемещая *main_data* () сразу после его дополнительной информации. Каждый результирующий элемент *mp3_channel_element* отображается непосредственно в единицу доступа. Получающийся заголовок и дополнительная информация обозначаются как *H'* и *SI'* соответственно.

Все элементы данных *header ()* должны быть сохранены. Элемент данных *main_data_begin* может быть обнулен. В этом случае *CRC* должно быть пересчитано.

Приложение В
(справочное)

Преобразование элементов *mp3_channel_elements* в потоки битов *MPEG*–1/2 уровень 3

В.1 Обзор

Элементы *mp3_channel_elements*, извлеченные из единицы доступа, должны подвергнуться следующим операциям преобразования, чтобы получить аудиопотоки битов *MPEG*–1/2 уровень 3, соответствующие ГОСТ Р 54711 или ГОСТ Р 54712:

для каждого *mp3_channel_element* на устройство доступа открыть экземпляр декодера или выходной поток;

для каждого *mp3_channel_element* в каждой единице доступа выполнить:

восстановить *syncword* и *IDex;*

скорректировать *bitrate_index;*

орегулировать *main_data_begin;*

пересчитать *crc_word* ;

восстановить кадрирование.

В.2. Сигнализация о частоте дискретизации

Чтобы позволить использовать сигнализацию о частотах дискретизации, не определенных в ГОСТ Р 54711 или ГОСТ Р 54712, используется последний бит *syncword*. Это приводит к следующей модификации синтаксиса:

Синтаксис	Количество битов	Мнемоника
header ()		
{		
syncword;	11	*bslbf*
Idex;	1	*bslbf*
...		

syncword Битовая строка '1111 1111 111'.

IDex Один бит, чтобы указать на расширенный *ID* алгоритма. Имеет значение '0' для частот дискретизации, не определенных в ГОСТ Р 54711 или ГОСТ Р 54712.

Следующая таблица определяет частоту дискретизации в зависимости от значений для *IDex* и *ID*:

IDex	*ID*	Частота дискретизации
0	0	8, 11,025, 12 кГц
1	0	16, 22,05, 24 кГц (см. ГОСТ Р 54712),
1	1	32, 44,1, 48 кГц (см. ГОСТ Р 54711),

В.3. Инструкции по востановлению

Этот процесс реконструкции предоставляет определенные степени свободы:

bitrate_index (чтобы отрегулировать длину фрейма потока битов согласно новым настройкам *bitrate_index*, частоте дискретизации и *padding_bit*, может потребоваться вставка битов (стаффинг)),

1) установить максимально позволенное значение (сигнализация максимально позволенной длины фрейма потока битов).

2) установить ближайшее более высокое значение, которое соответствует длине *mp3_channel_element*.

3) установить ближайшее более высокое значение, которое соответствует длине *mp3_channel_element* минус *main_data_begin* текущего аудиофрейма.

4) *main_data_begin* обнулить.

5) установить *main_data_begin* в значение, указывающее на конец *main_data* предыдущего аудиофрейма.

6) установить корректное значение *main_data_begin* соответствующего потока битов *MPEG*–1/2 уровень 3.

7) расположение наполнения (вставки битов) в конце *main_data*: сохраняет вспомогательные

данные, записанные в прямом направлении, начиная после последней кодовой комбинации Хаффмана.

8) расположение наполнения (вставки битов) в конце последней кодовой комбинации Хаффмана (расположение может быть вычислено, используя *part_2_3_length*): сохраняет вспомогательные данные, записанные в обратном направлении, начиная перед *main_data* следующего фрейма.

9) никакое наполнение не требуется: сохраняет любые вспомогательные данные.

В зависимости от требований скорости передачи и вспомогательной обработки данных, эти возможности могут быть объединены несколькими способами.

Самый простой метод устанавливает максимальную величину скорости передачи. Это препочтительный метод, когда питание существующих декодеров *MPEG*–1/2 уровень 3. *main_data_begin* обнуляется. Биты стаффинга добавляются до или после вспомогательных данных.

Более передовой метод можно получить из этого простого метода, устанавливая *bitrate_index* в ближайшее более высокое значение, которое соответствует длине *mp3_channel_element*. С этой модификацией скорость передачи может быть значительно уменьшена.

Для частот дискретизации до 24 кГц (то есть в случаях, где одна гранула формирует фрейм), размер гранулы может превысить максимальный размер фрейма. Максимальный индекс скорости передачи может не позволить хранить целый фрейм после заголовка (*main_data_begin*=0). Это следует из того факта, что максимальная длина гранулы составляет 960 байтов (7680 битов), но максимальная длина фрейма, основанная на самом высоком индексе скорости передачи, составляет 576 байтов (для 8 кГц), 417 байтов (для 11,025 кГц), 384 байта (для 12 кГц), 720 байтов (для 16 кГц), 522 байта (для 22,05 кГц), 480 байтов (для 24 кГц). В этом случае *main_data_begin* должен быть должным образом скорректирован между нулем и различием между максимальным размером гранулы и максимальной длиной фрейма.

Чтобы избежать необходимости стаффинга и неопределенности, *main_data_begin* устанавливается в значение, указывающее на конец *main_data* предыдущего фрейма. *bitrate_index* устанавливается в ближайшее более высокое значение, которое соответствует длине *mp3_channel_element* минус *main_data_begin* текущего аудиофрейма. Только если *main_data_begin* превысит дозволенное значение, должен быть выполнен стаффинг.

Исходный поток битов уровня 3 отлично может быть восстановлен, если корректное значение *main_data_begin* соответствующего потока битов *MPEG*–1/2 уровень 3 было сохранено.

Приложение Г
(справочное)

Интерфейс унаследованных систем *MPEG*–4 в *MPEG*–1/2 Audio

Г.1 Обзор

Это приложение обеспечивает помощь в использовании *decSpecificInfo* и *accessUnit*, чтобы применять *MPEG*–1/2 уровень 1, 2, 3 и *MPEG* 2 AAC в *MPEG*–4, используя следующие значения *objectTypeIndication*:

0x6*b* (ГОСТ Р 54711)
0x69 (ГОСТ Р 54712)
0x66 (ГОСТ Р 54712 основной профиль)
0x67 (ГОСТ Р 54712 профиль малой сложности)
0x68 (ГОСТ Р 54712 профиль масштабируемой частоты дискретизации)

Г.2 Специальная информация о декодере

В ГОСТ Р 53556.1 *decSpecificInfo* определяется для некоторой информации о декодере носителей информации. Эта специальная информация о декодере составляет непрозрачный контейнер с информацией для специального медиа-декодера (декодера носителей). При ее наличии, она может использоваться для инициализации декодера и априорной реализации наборщика. Нет необходимости определять эту специальную информацию декодера. Ее существование и семантика зависят от значений *DecoderConfigDescriptor.streamType* и *DecoderConfigDescriptor.objectTypeIndication*.

Нехватка доступности любого *decSpecificInfo* приводит к ситуации, когда формат памяти композиции нельзя априорно различить, чтобы реализовать наборщиком. Следовательно декодер определяет формат памяти композиции.

Г.2.1 *MPEG*–2 *AAC*

Для *MPEG*–2 *AAC* определяется *decSpecificInfo*, то есть в случае значений *DecoderConfigDescriptor.objectTypeIndication*, которые обращаются к потокам, удовлетворяющим ГОСТ Р 54712.

В этом случае аудио декодеры получают всю релевантную информацию от этого *decSpecificInfo*, которая состоит из *adif_header* (), и могут переслать формат памяти композиции в память композиции.

Г.2.2 *MPEG*–1 *Audio* и *MPEG*–2 *Audio*

Для *MPEG*–1 *Audio* и *MPEG*–2 *Audio* никакой *decSpecificInfo* не определяется, то есть в случае значений *DecoderConfigDescriptor.objectTypeIndication*, которые относятся к потокам, соответствующим ГОСТ Р 54711 и ГОСТ Р 54712. В этих случаях аудиодекодеры получают всю значимую информацию в элементе *'header()'* их собственного потока битов и могут передать формат памяти композиции в память композиции. Таким образом, динамически внося изменения в выходном формате, необходимые для того, чтобы иметь дело с ним, то есть без элементарного обновления дескриптора потока.

Г.3 Единицы доступа

Фрейм *MPEG*–1/2 уровней 1, 2 или 3 (данные между синхронизирующими словами) или фрейм *MPEG*– 2 *AAC* (*raw_data_block*) могут быть обработаны как единицы доступа аудио не только в контексте ГОСТ Р 54711 и ГОСТ Р 54712, но также и в контексте настоящего стандарта.

При обработке фреймов *MPEG*–1/2 уровней 1, 2, 3 или *MPEG*–2 *AAC,* как единиц *MPEG*–4, единицам доступа присваивается информация синхронизации.

Так как определения единицы аудиодоступа точно не соответствуют между *MPEG*–1/2 и *MPEG*–4, то нужно принимать во внимание некоторые специальные соображения.

В частности для уровня 3 единица аудиодоступа определяется в *MPEG*–1/2, как часть потока битов, которая может быть декодирована только с использованием ранее полученной основной информации, которая не отражает определение единицы аудиодоступа в *MPEG*–4.

Впоследствии некоторые единицы аудиодоступа могут быть не декодированы из-за нехватки некоторой потерянной основной информации в случае перфораций в потоке битов и произвольного доступа. Однако информация синхронизации сохраняется правильно.

В случае, когда считают необходимым иметь лучшее редактирование или возможности вставки перфораций для потоков уровня 3, желательно использовать потоки, закодированные *VBR*.

Существует возможность преобразовать любой существующий поток уровня 3 в поток *VBR:*

однозначно;

полностью совместимый с *MPEG*–1 или *MPEG*–2;

декодируемый любым существующим декодером уровня 1, 2 или 3.

Это можно сделать следующим образом:

main_data () для единственного фрейма помещается непосредственно рядом с его дополнительной информацией. Указатель *main_data_begin* обнуляется. Фрейм за фреймом индексы скорости передачи (*bitrate_index*) увеличиваются до минимального значения, нужного, чтобы получить длину фрейма, которая может разместить исходный заголовок, *error_check*, дополнительную информацию и основные данные. Из-за гранулярности в доступных скоростях передачи обычно эта длина фрейма больше, чем длина заголовка, *error_check*, дополнительная информация и основные данные. В этом случае в конце *main_data* добавляются биты стаффинга, чтобы получить совместимые фреймы.

Библиография

[1] ИСО/МЭК 14496–3:2009 Информационные технологии. Кодирование аудиовизуальных объектов. Часть 3. Звуковое кодирование (ISO/IEC14496–3:2009 *Information technology - Coding of audio-visual objects - Part* 3: *Audio*)

УДК 621.396 : 006.354 ОКС 33.170

Ключевые слова: звуковое вещание, электрические параметры, каналы и тракты, технологии MPEG–кодирования, синтетический звук, масштабирование, защита от ошибок, поток битов расширения, психоакустическая модель

Подписано в печать 05.11.2014. Формат 60x84 1/8.
Усл. печ. л. 1,86. Тираж 36 экз. Зак. 3898.

Подготовлено на основе электронной версии, предоставленной разработчиком стандарта

ФГУП «СТАНДАРТИНФОРМ»

123995 Москва, Гранатный пер., 4.

gost.e-ecolog.ru gost.e-ecolog.ru

gost.e-ecolog.ru