

Primera edición
30.11.2007

Válida a partir de
01.12.2007

**Televisión digital terrestre — Codificación de
video, audio y multiplexación
Parte 2: Codificación de audio**

Palabras clave: Televisión digital terrestre. Codificación de fuente. AAC. Nivel y perfil.

ICS 33.160.01

ISBN 978-85-07-00882-8

© ABNT 2007

Todos los derechos reservados. A menos que se especifique de otro modo, ninguna parte de esta publicación puede ser reproducida o utilizada por cualquier medio, electrónico o mecánico, incluyendo fotocopia y microfilm, sin permiso por escrito de la ABNT.

ABNT

Av. Treze de Maio, 13 - 28º andar

20031-901 - Rio de Janeiro - RJ

Tel.: + 55 21 3974-2300

Fax: + 55 21 2220-1762

abnt@abnt.org.br

www.abnt.org.br

Impresso en Brasil

Índice

Página

Prefacio.....	iv
1 Alcance	1
2 Referencias normativas	1
3 Términos y definiciones.....	1
4 Abreviaturas.....	2
5 Formato de entrada de audio	3
5.1 Condiciones generales	3
5.2 Principales parámetros	3
5.2.1 Formatos	3
5.2.2 Interfaces.....	3
5.2.3 Niveles de señal de audio	4
5.2.4 Modos o configuraciones multicanal	4
5.2.5 Metadatos	4
6 Servicios de audio y canales auxiliares	5
7 Sistema de codificación de audio	6
8 Procedimientos para compresión y transmisión de audio	6
8.1 Visión general del estándar de codificación	6
8.2 Perfiles y niveles.....	7
8.3 Capa de transporte y multiplexación.....	7
9 Restricciones de los parámetros de codificación de audio.....	8
9.1 Restricciones de los parámetros de codificación de audio para servicio <i>full-seg</i>	8
9.1.1 Modos de codificación de audio	8
9.1.2 Principales parámetros	9
9.1.3 Restricciones operativas referentes a la compatibilidad con receptor estéreo	10
9.2 Restricciones de los parámetros de codificación de audio para servicio <i>one-seg</i>	10
9.2.1 Modos de codificación de audio	10
9.2.2 Principales parámetros	11
Bibliografía	12

Prefacio

La Associação Brasileira de Normas Técnicas (ABNT) es el Fórum Nacional de Normalización. Las Normas Brasileñas, cuyo contenido es responsabilidad de los Comités Brasileños (ABNT/CB), de los Organismos de Normalización Sectorial (ABNT/ONS) y de las Comisiones de Estudios Especiales (ABNT/CEE), son elaboradas por Comisiones de Estudio (CE), formadas por representantes de sus sectores implicados de los que forman parte: productores, consumidores y neutrales (universidades, laboratorios y otros).

Los Documentos Técnicos ABNT se elaboran de acuerdo con las reglas de Directivas ABNT, Parte 2.

La Associação Brasileira de Normas Técnicas (ABNT) llama la atención sobre la posibilidad de que algunos de los elementos de este documento pueden ser objeto de derechos de patente. La ABNT no debe ser considerada responsable por la identificación de cualesquiera derechos de patente.

La ABNT NBR 15602-2 fue elaborada por la Comisión de Estudio Especial de Televisión Digital (ABNT/CEE-00:001.85). El Proyecto circuló en Consulta Nacional según Edicto nº 07, de 29.06.2007 a 28.08.2007, con el número de Proyecto 00:001.85-002/2.

En caso que surja cualquier duda con relación a la interpretación de la versión en español siempre deben prevalecer las prescripciones de la versión en portugués

Esta Norma está basada en los trabajos del Fórum del Sistema Brasileiro de Televisão Digital Terrestre, según establece el Decreto Presidencial nº 5.820, de 29/06/2006.

La ABNT NBR 15602, bajo el título general "Televisión digital terrestre – Codificación de video, audio y multiplexación", tiene la previsión de contener las siguientes partes:

- Parte 1: Codificación de video;
- Parte 2: Codificación de audio;
- Parte 3: Sistemas de multiplexación de señales.

Esta versión en español es equivalente a la versión corregida de la ABNT NBR 15602-2:2007, de 07.04.2008.

Televisión digital terrestre — Codificación de video, audio y multiplexación

Parte 2: Codificación de audio

1 Alcance

Esta parte de la ABNT NBR 15602 especifica los parámetros para las señales de audio y el sistema de codificación y decodificación de sonido a ser utilizado en el sistema brasileño de televisión digital terrestre (SBTVD).

2 Referencias normativas

Los documentos indicados a continuación son indispensables para la aplicación de este documento. Para las referencias fechadas, se aplican solamente las ediciones citadas. Para las referencias sin fecha, se aplican las ediciones más recientes del documento citado (incluyendo enmiendas).

ABNT NBR 15602-3:2007, *Televisión digital terrestre - Codificación de video, audio y multiplexación – Parte 3: Sistemas de multiplexación de señales*

ABNT NBR 15603-2:2007, *Televisión digital terrestre – Multiplexación y servicios de información (SI) — Parte 2: Estructura de datos y definiciones de la información básica de SI*

ISO/IEC 13818-1:2007, *Information technology – Generic coding of moving pictures and associated audio information: Systems*

ISO/IEC 14496-3:2005, *Information technology – Coding of audio-visual objects – Part 3: Audio*

ITU Recommendation BS.775-1, *Multichannel stereophonic sound system with and without accompanying picture*

3 Términos y definiciones

Para los efectos de esta parte de la ABNT NBR 15602, se aplican los siguientes términos y definiciones.

3.1

codificación

proceso de transformación de señales externas en bits que representen tales señales

NOTA La codificación se da, por ejemplo, a través de muestreo, y la información obtenida puede aún ser compactada.

3.2

decodificación

proceso responsable por la recuperación de la señal original a través de los bits recibidos del codificador

NOTA La decodificación puede, eventualmente, realizar también la descompactación de la información recibida.

3.3

downmix

operación para transformar una matriz de n canales y obtener menos de n canales, comúnmente realizada para la conversión de un programa multicanal en estéreo o mono

3.4

LATM/LOAS

mecanismo de transporte definido en el MPEG-4 que utiliza dos capas, una de multiplexación y otra de sincronización

NOTA La capa de multiplexación LATM administra la multiplexación de varios *payloads* de audio (datos de audio) y sus datos de configuración constantes en los elementos de *AudioSpecificConfig()*. La capa de sincronización LOAS especifica una sintaxis para sincronismo automático en el flujo de transporte de audio del MPEG-4.

3.5

nivel

número máximo permitido de canales y de su tasa de muestreo y que indica la complejidad computacional para el decodificador

3.6

receptor *full-seg*

dispositivo capaz de decodificar informaciones de audio, video, datos etc., contenidas en la capa del flujo de transporte de 13 segmentos, destinada al servicio fijo (*indoor*) y móvil

NOTA La clasificación *full-seg* se aplica a los convertidores digitales, también conocidos como *settop box*, y a los receptores de 13 segmentos integrados con pantalla de exhibición, pero no exclusivos a éstos. Este tipo de receptor es capaz de recibir y decodificar señales de televisión digital terrestre de alta definición y, a criterio del fabricante, también recibir y decodificar informaciones transportadas en la capa "A" del flujo de transporte, aplicada para los servicios dirigidos a los receptores portátiles, definidos como *one-seg*.

3.7

receptor *one-seg*

dispositivo que decodifica exclusivamente informaciones de audio, video, datos etc., contenidas en la capa "A" asignada en el segmento central de los trece segmentos

NOTA La clasificación *one-seg* se destina a los receptores del tipo portátil, también conocidos como "*handheld*", especialmente recomendados para pantallas de exhibición de dimensiones reducidas, normalmente hasta 7 pulgadas. Entre los productos clasificados como *one-seg*, están los receptores integrados con teléfono móvil, PDA, *dongle* y televisores portátiles, los cuales son alimentados por una batería interna y, por lo tanto, sin necesariamente requerir una fuente externa de energía, así como aquellos destinados a automóviles. Este tipo de receptor es capaz de recibir y decodificar sólo señales de televisión digital terrestre transportadas en la capa "A" del flujo de transporte y, como consecuencia de ello, únicamente señales de perfil básico, destinadas a los dispositivos portátiles de recepción.

3.8

unidad de acceso de audio

parte de audio de un *elementary stream* que es accesible individualmente

NOTA Para los efectos de esta parte de la ABNT NBR 15602, una unidad de acceso de audio equivale a un *rawdatablock()*.

4 Abreviaturas

Para los efectos de esta parte de la ABNT NBR 15602, se aplican las siguientes abreviaturas:

AAC	Codificación de Audio Avanzada (<i>Advanced Audio Coding</i>)
CPE	<i>Channel Pair Element</i>
HDMI	<i>High-definition Multimedia Interface</i>
LATM	<i>Low Overhead Audio Transport Multiplex</i>
LFE	<i>Low Frequency Enhancement</i>
LOAS	<i>Low Overhead Audio Stream</i>

PCE	<i>Program Configuration Element</i>
PCM	<i>Pulse-code Modulation</i>
PS	<i>Parametric Stereo</i>
PSI	<i>Program Specific Information</i>
SAP	<i>Second Audio Program</i>
SCE	<i>Single Channel Element</i>
SDI	<i>Serial Digital Interface</i>
SBR	<i>Spectral Band Replication</i>
TS	Flujo de Transporte (<i>Transport Stream</i>)

5 Formato de entrada de audio

5.1 Condiciones generales

Las condiciones generales para el formato de entrada de audio deben ser obligatoriamente las siguientes:

- frecuencia de muestreo de la señal de audio: 32 kHz, 44,1 kHz ó 48 kHz;
- configuración de señales estereofónicas y multicanal (es decir, señales compuestas por dos o más señales de audio para lograr una reproducción envolvente o espacial del sonido); la tasa de muestreo para todas las señales debe ser obligatoriamente la misma;
- cuantización de las señales de entrada debe emplear 16 bits ó 20 bits;
- un programa de audio debe tener obligatoriamente por lo menos un canal de audio. El número máximo de canales en el programa debe ser obligatoriamente limitado al número máximo de canales permitidos por la ISO/IEC 14496-3;
- se recomienda que los programas multicanales se preparen de conformidad con la ITU Recommendation BS.775-1;
- los programas de audio en modo multicanal compatibles con los modos previstos en la ITU Recommendation BS.775-1 deben obligatoriamente estar en una de las configuraciones permitidas en la Tabla 3;
- en el caso de transmisión de solamente un programa multicanal sin transmisión de un programa estéreo, el programa multicanal debe obligatoriamente estar en modo 3/2 (5.0 ó 5.1, con o sin adición del canal LFE de enriquecimiento de las bajas frecuencias) para permitir el *downmix* para estéreo.

5.2 Principales parámetros

5.2.1 Formatos

Deben ser obligatoriamente admitidos flujos de bits o archivos conteniendo audio digital no comprimido en formato PCM, como WAVE o AIFF, estéreo y multicanal.

5.2.2 Interfaces

Entre las interfaces (buses) de entrada/salida digital permitidos, deben obligatoriamente estar AES3 (AES/EBU, conteniendo dos canales PCM por flujo de bits), SDI, HD-SDI y HDMI.

5.2.3 Niveles de señal de audio

El nivel de referencia para la intensidad o presión sonora debe ser obligatoriamente igual a 0 dB. La banda dinámica admisible de excursión debe estar limitada obligatoriamente a + 20 dB (headroom) y - 70 dB con respecto a la referencia, correspondiendo a una banda dinámica típica de 90 dB. Es conveniente que los niveles de audio medio estén a - 20 dBFS (0 dB), para permitir homogeneidad en el volumen entre canales distintos. La señal debe acomodar picos de por lo menos 4 veces su potencia media RMS.

5.2.4 Modos o configuraciones multicanal

El modo de transmisión se refiere a la configuración multicanal utilizada, al número de canales disponible en el flujo de bits y a la forma de codificación de esos canales.

El número de canales de audio fuente debe ser obligatoriamente por lo menos uno para una configuración básica, dos para transmisión estándar estéreo típico y cinco canales más un canal de bajas frecuencias (LFE) para transmisión multicanal "5.1" estándar. Las señales fuentes deben ser obligatoriamente preprocesadas y/o combinadas previamente a la entrada del codificador, para que produzcan los canales de transmisión que deben estar presentes obligatoriamente en el flujo de bits.

Una misma programación de audio se puede transmitir en más de un modo, por ejemplo, en estéreo (dos canales) más modo multicanal 3/2 (5.1) simultáneamente, aunque la transmisión simultánea no es obligatoria.

En el caso de la transmisión exclusiva en modo multicanal 3/2 (5.1), los receptores deben ser capaces obligatoriamente de sintetizar la señal estéreo por medio de conversión (*downmixing*), operaciones de replicación, *dematrixing*, combinación y procesamiento de señal en el ámbito funcional del sistema de reproducción de audio del receptor. Los modos multicanal permitidos para codificación y transmisión deben cumplir obligatoriamente lo descrito en 9.1.1.

5.2.5 Metadatos

Datos auxiliares deben contener obligatoriamente informaciones como descripciones de contenido de los programas de audio, parámetros de configuración de los servicios de audio y parámetros de las señales de audio transmitidos en el flujo de bits.

Se pueden admitir como tipos de datos auxiliares:

- a) descripción del contenido de los programas de audio que están siendo transmitidos (por ejemplo, clasificación de programa sonoro, descripción de los objetos de audio mezclados en el contenido, descripción del contenido del canal de audio auxiliar etc.);
- b) modo multicanal;
- c) volumen de referencia para operaciones de ecualización en la reproducción en el terminal de acceso. Datos auxiliares y la descripción de contenido de programas de audio se pueden clasificar en dos niveles.

Un primer nivel debe ser obligatoriamente normativo. Obligatoriamente, ese nivel debe afectar directamente la operación del receptor (decodificación de los flujos de bits) como, por ejemplo, información de cantidad y modo de los canales y perfil y nivel de codificación extraídos directamente de las tablas PSI. Obligatoriamente, los datos en esta categoría deben ser esenciales para la decodificación y reproducción correcta del servicio de audio en el receptor.

Un segundo nivel debe ser obligatoriamente informativo. Ese nivel no debe afectar la decodificación, pero sí traer informaciones sobre los contenidos de los programas de audio asociados a cada PID. Los datos en esta categoría se deben usar obligatoriamente para procesamiento de información sobre los programas en el receptor.

La Tabla 1 resume los tipos de datos auxiliares para audio admisibles en el sistema.

Tabla 1 — Tipos de datos auxiliares para audio

Parámetros	Descripción y uso
<i>matrix_downmix_idx</i>	<p>Descripción: Indicador del coeficiente a ser utilizado en el <i>downmix</i> de multicanal para estéreo. Se debe transmitir obligatoriamente en el flujo de bits como metadato, conforme especificado en ISO/IEC 14496-3</p> <p>Uso: obligatorio cuando se transmite un programa en modo multicanal</p> <p>Cuando el parámetro <i>matrix_mixdown_idx_present</i> esté activo en "1" en el PCE respectivo del programa (PID), se debe utilizar obligatoriamente el sistema de <i>downmix</i> descrito en la ISO/IEC 14496-3:2005, subsección 4.5.1.2.2 y Tabla 4.70</p> <p>Cuando el terminal receptor realice operación de <i>downmix</i>, la salida estéreo analógica siempre debe estar activa obligatoriamente con esta señal</p>
<i>program_ref_level</i>	<p>Descripción: valor representativo de la intensidad media del volumen del programa de audio de largo plazo para todos los canales combinados con relación a la referencia 0 dBFS. Representado en 128 niveles (7 bits), cuantizado en pasos de 0,25 dB, totalizando una excursión de 32 dB con relación al fondo de escala (0 dBFS)</p> <p>Este parámetro debe contener obligatoriamente una descripción informativa del volumen de referencia adoptado por la emisora (0 dB) con relación al fondo de escala (0 dBFS), para efecto de normalización del diálogo y para tornar más confortable el cambio de canales para el usuario</p> <p>Uso: obligatorio. Se recomienda utilizar <i>program_ref_level = 80 (0x50)</i>, que corresponde a un valor indicativo de -20 dBFS como referencia 0 dB, de conformidad con la ISO/IEC 14496-3</p> <p>El parámetro deberá ser transmitido junto a la estructura DRC, conforme ISO/IEC 14496-3:2005, sección 4.5.2.7</p>
<i>Dynamic Range Control (DRC)</i>	<p>Descripción: El control de banda dinámica es especialmente indicado para las transmisiones en modo multicanal y se puede señalar en la forma de metadatos, de conformidad con la ISO/IEC 14496-3:2005, sección 4.5.2.7</p> <p>Uso: en la codificación es opcional, aunque el decodificador debe soportar obligatoriamente esa herramienta. En el caso que la información de DRC no sea enviada por el codificador, el decodificador no debe aplicar la herramienta DRC</p>

6 Servicios de audio y canales auxiliares

Servicios de audio incluyen la transmisión de programas de audio adicionales al programa principal y se consideran obligatoriamente servicios opcionales, con excepción del servicio de descripción de audio (DA), cuya transmisión es obligatoria conforme la legislación vigente.

La transmisión de estos servicios se debe realizar a través de la asignación de canales de audio auxiliares adicionales en programas de audio (PID) distintos, o en el mismo flujo de bits de un mismo PID, respetándose siempre el número máximo de canales permitidos en el flujo de bits por el perfil/nivel de codificación usado.

Canales adicionales al programa principal se pueden utilizar para transmitir audio en otros idiomas (como, por ejemplo, servicio de programa de audio secundario SAP), para transmitir servicios de descripción de audio (DA), para transmitir programas adicionales al programa principal y audio secundario proveniente de otras tomas de sonido (contenido adicional, como, por ejemplo, efectos).

Todos los canales adicionales referentes a servicios de audio auxiliares se deben indicar en forma obligatoria y apropiada utilizando una identificación válida de tipo de componente (*component_type*) en el respectivo descriptor de audio (*audio_component_descriptor*) del programa.

Los canales auxiliares se deben transmitir obligatoriamente en programas distintos (PID distintos), con la debida señalización e identificación de sus canales, para ser seleccionados, decodificados y reproducidos en conjunto con o en sustitución de los canales de audio del programa principal.

El servicio de DA (descripción de audio) consiste comúnmente en un canal monoaural de voz y suministra una descripción de la escena como un subcomponente asociado a un servicio de televisión. El mismo debe auxiliar obligatoriamente en la comprensión del entretenimiento principal (pero no exclusivamente) para telespectadores con problemas visuales.

La transmisión de DA se debe implementar obligatoriamente utilizando al menos uno de los mecanismos siguientes:

- a) como canal auxiliar (monoaural o estéreo) conteniendo la descripción de audio previamente mezclada con un programa de audio principal;
- b) como canal auxiliar conteniendo la descripción de audio en separado, para posterior mezcla con el programa de audio en el receptor.

En ambos casos se recomienda señalar el servicio a través del parámetro *component_type* disponible en el descriptor "audio component descriptor", conforme ABNT NBR 15603-2:2007, Tabla 28.

La habilidad de mezclar uno o más canales suplementarios de descripción con el audio del programa principal puede tener otras aplicaciones, incluyendo comentarios en otros idiomas (multilingüe), interactividad y propósitos educativos.

7 Sistema de codificación de audio

Las señales de audio se deben codificar obligatoriamente por una combinación de codificación transformada en el tiempo y en frecuencia. La transformada en frecuencia debe descomponer la señal de entrada en sus componentes de frecuencia empleando la transformada discreta del coseno (DCT – *Discrete Cosine Transform*) cuando se reduce la cantidad de información, reduciéndose el desvío en frecuencia de cada componente.

Como herramienta de compresión adicional se emplean estimativas psicoacústicas de la señal, donde los códigos se deben estimar obligatoriamente para minimizar la degradación en las bandas de frecuencias perceptibles por el oído humano.

La compresión de audio y los procedimientos de transmisión deben ser obligatoriamente compatibles con la ISO/IEC 14496-3.

El decodificador se debe construir obligatoriamente asumiendo que cualquier estructura válida de la ISO/IEC 13818-1, incluyendo descriptores privados, se puede incluir en el flujo de bits aunque en el inicio de las transmisiones esos descriptores no se utilicen. El decodificador de audio no debe tener en cuenta obligatoriamente estructuras "reservadas" o aquéllas que corresponden a funciones no implementadas por el receptor.

8 Procedimientos para compresión y transmisión de audio

8.1 Visión general del estándar de codificación

La Figura 1 muestra los procedimientos para compresión y transmisión de audio.

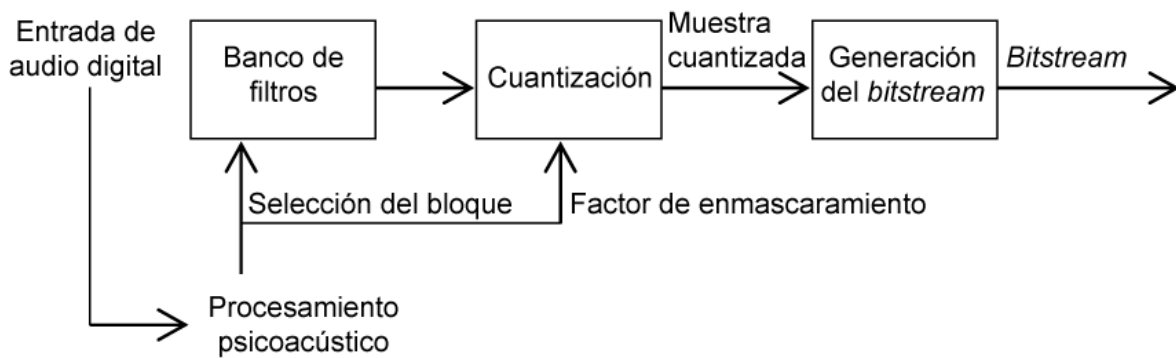


Figura 1 — Procedimientos de transmisión y codificación de audio

El banco de filtros debe convertir obligatoriamente la entrada de señal de audio digital del dominio del tiempo para el dominio de la frecuencia. A partir de ahí, el banco de filtros aplica la transformada discreta de los cosenos modificada y funciones de visualización por ventanas a bloques de la señal de entrada, de acuerdo con características fisiológicas audibles.

Los procesos psicoacústicos deben calcular obligatoriamente la cantidad de enmascaramiento (límite para el discernimiento entre una señal específica de audio de otras señales) y alimentan el banco de filtros con bloques de la señal de entrada.

Las muestras se deben cuantiar obligatoriamente después del procesamiento por el banco de filtros, con base en el factor de enmascaramiento calculado por los procesos de psicoacústica. Del mismo modo, no se deberá exceder el número total de bits utilizado por cada bloque.

El flujo de bits se debe configurar obligatoriamente conforme ISO/IEC 14496-3.

8.2 Perfiles y niveles

La codificación de audio debe ser compatible obligatoriamente con la ISO/IEC 14496-3. Los siguientes perfiles y niveles del estándar MPEG-4 AAC se deben permitir obligatoriamente:

- a) LC (*low complexity*), perfil básico del estándar AAC; niveles L2 y L4;
- b) HE (*high efficiency*), perfil avanzado de alta eficiencia, combinando el perfil LC con el uso de la herramienta SBR (*spectral band replication*) para la versión 1 de este perfil, niveles L2 y L4;
- c) HE combinado con la herramienta PS (*parametric stereo*) para la versión 2 de este perfil; nivel L2.

El perfil y el nivel del codificador MPEG-4 AAC se deben señalar obligatoriamente conforme ABNT NBR 15603-2.

8.3 Capa de transporte y multiplexación

La codificación y el empaquetado (*framing*) intermedio del audio deben ser compatibles obligatoriamente con LATM/LOAS, de conformidad con la ISO/IEC 14496-3. El *elementary stream* obligatoriamente debe ser primero encapsulado en el formato de transporte LATM y debe utilizar obligatoriamente el elemento de multiplexación *AudioMuxElement()*.

La capa de sincronización del transporte de audio (LOAS) debe utilizar el formato de transmisión *AudioSyncStream()* conforme ISO/IEC 14496-3.

El audio MPEG-4 transportado en el flujo de transporte MPEG-2 (TS), utilizando la sintaxis de transporte LATM/LOAS debe ser identificado obligatoriamente por *stream_type* 0x11 de acuerdo con el *stream_type assignments* en la ISO/IEC 13818-1:2007.

Para decodificación de audio el receptor debe identificar obligatoriamente el tipo, el perfil y el nivel transmitidos y ser capaz de extraer los *payloads* de los objetos de audio. Para transmisión de audio MPEG-4 sobre flujo de transporte MPEG-2, es obligatorio el uso de señalización explícita de SBR sin alineamiento de PES.

Los receptores deben ser capaces obligatoriamente de procesar la herramienta SBR. La señalización de la presencia de SBR debe usar obligatoriamente el mecanismo de señalización explícito *non-backward compatible*, de acuerdo con la ISO/IEC 14496-3.

La Tabla 2 describe los campos de la sintaxis LATM/LOAS de transporte en *StreamMuxConfig* que deben estar obligatoriamente formateados para la identificación y recuperación de los *payloads* de audio, de acuerdo con la ISO/IEC 14496-3.

Tabla 2 — Principales parámetros LATM a ser configurados

Parámetro LATM	Descripción de uso
<i>audioMuxVersion</i>	Debe asumir obligatoriamente el valor "0"
<i>allStreamsSameTimeFraming</i>	Debe asumir obligatoriamente el valor "1"
<i>numSubFrames</i>	Debe asumir obligatoriamente el valor "0" indicando un único <i>PayloadMux()</i> (<i>access unit</i>) presente en el <i>AudioMuxElement()</i>
<i>numProgram</i>	Debe asumir obligatoriamente el valor "0" indicando un programa por multiplex LATM
<i>numLayer</i>	Debe asumir obligatoriamente el valor "0" indicando una sola capa
<i>frameLengthType</i>	Debe asumir obligatoriamente el valor "0" indicando que el tamaño de cuadro del <i>payload</i> puede tener extensión variable. La extensión del <i>payload</i> en bytes se especifica directamente en <i>PayloadLengthInfo()</i> con palabras de 8 bits

9 Restricciones de los parámetros de codificación de audio

9.1 Restricciones de los parámetros de codificación de audio para servicio *full-seg*

9.1.1 Modos de codificación de audio

El modo de codificación determina el número de canales disponible en el servicio de audio. Los modos de codificación de audio para transmisión digital deben cumplir obligatoriamente las restricciones descritas en la Tabla 3.

Tabla 3 — Restricciones de modos de codificación de audio

Parámetro	Restricción
Modos de audio permitidos	Monoaural (1/0), estéreo (2/0 y 2/0 + LFE) ^a , estéreo multicanal (3/0, 2/1, 3/1, 2/2, 3/2, 3/2+LFE) ^a , dos señales de audio independientes (monoaural dual), multi-audio (tres o más señales de audio) y combinaciones de éstos
Modos de audio recomendados	Estéreo (2/0), multicanal (3/2+LFE)
<i>Downmix</i>	Para la configuración 5.0 y 5.1, se debe utilizar obligatoriamente el esquema conforme Tabla 1. En las demás configuraciones multicanal, otros esquemas de <i>downmix</i> pueden ser usados por el receptor, desde que mantengan la integridad del audio y el nivel de inteligibilidad. El esquema de <i>downmix</i> estéreo para mono no está cubierto en esta Norma, aunque el <i>clipping</i> se debe evitar obligatoriamente

^a Número de canales para altavoces delanteros/traseros.

EJEMPLO 3/1 = 3 delanteros + 1 trasero; 3/2 = 5.0 = 3 canales delanteros y 2 traseros.

El decodificador debe ser capaz obligatoriamente de procesar cualquiera de los modos recomendados de audio.

La configuración de los canales según el modo de operación y su orden de transmisión dentro del *payload* deben estar de acuerdo obligatoriamente con la Tabla 4.

Tabla 4 — Configuración de canales y modos estándar recomendados en el MPEG-4 AAC

Modo	Configuración de canales	Orden de transmisión del SE ^a	Elemento estándar para mapeo de altavoces ^b
Monoaural (1/0)	1	<SCE1><TERM>	SCE1 = C
Estéreo (2/0)	2	<CPE1><TERM>	CPE1 = L e R
3/0	3	<SCE1><CPE1><TERM>	SCE1 = C, CPE1 = L e R
3/1	4	<SCE1><CPE1><SCE2><TERM>	SCE1 = C, CPE1 = L e R, SCE2 = MS
Multicanal 5.0 (3/2)	5	<SCE1><CPE1><CPE2><TERM>	SCE1 = C, CPE1 = L e R, CPE2 = LS e RS
Multicanal 5,1 (3/2 + LFE)	6	<SCE1><CPE1><CPE2><LFE><TERM>	SCE1 = C, CPE1 = L e R, CPE2 = LS y RS, LFE = LFE

^a Abreviaturas relacionadas a elemento sintáctico (SE): SCE – *single channel element*, CPE – *channel pair element*, LFE – *LFE channel element*, TERM – *terminator*.

^b Abreviaturas relacionadas al arreglo de altavoces: L – altavoz frontal izquierdo / R – altavoz frontal derecho / C – altavoz frontal central / LFE – énfasis de baja frecuencia / LS – altavoz *surround* trasero izquierdo / Rs – altavoz *surround* trasero derecho / MS – altavoz *surround monoaural*.

En el caso de la transmisión de dos señales de audio independientes (monoaural dual o 1/0+1/0), se recomienda el orden de transmisión del SE de la siguiente forma: <SCE1><SCE2><TERM>, siendo SCE1 el primer (principal) canal y el SCE2 el segundo canal en el programa.

Si la configuración utilizada no se encuentra entre las mencionadas en la Tabla 4, se debe reproducir obligatoriamente usando una configuración con el mismo número de canales y con la respectiva señalización.

9.1.2 Principales parámetros

Los principales parámetros del sistema de codificación de audio deben cumplir obligatoriamente la Tabla 5.

Tabla 5 — Principales parámetros del sistema de codificación de audio – Servicio *full-seg*

Parámetro	Restricción
Mecanismos de transporte permitidos	LATM/LOAS (conforme ISO/IEC 14496-3)
Números de canales recomendados	Mono (1.0), 2 canales (estéreo o 2.0), o multicanal (5.1)
Perfiles y niveles permitidos	<i>Low complexity AAC</i> : nivel 2 (LC-AAC@L2) para dos canales <i>Low complexity AAC</i> : nivel 4 (LC-AAC@L4) para multicanal <i>High Efficiency (HE)</i> : nivel 2 (HE-AAC v1@L2) para dos canales <i>High Efficiency (HE)</i> : nivel 4 (HE-AAC v1@L4) para multicanal
Tasa máxima de bits permitida	Conforme ISO/IEC 14496-3
Muestras por cuadro	frameLengthFlag en GASpecificConfig() debe tener valor 0, indicando que la extensión del cuadro debe ser de 1024 muestras para AAC y 2048 al usar SBR. 960 muestras para AAC (o 1 920 al usar SBR) no pueden ser utilizadas

Para transmisión en alta fidelidad se recomienda utilizar el perfil/nivel AAC@L4 para modo multicanal y el perfil/nivel AAC@L2 para modo estéreo. En la transmisión de audio estéreo, no se debe usar el nivel 4 (L4).

Las señales se pueden codificar en cualquier tasa soportada en el perfil y nivel seleccionado. Al mismo tiempo la señal multicanal puede emplear cualquier frecuencia de muestreo del perfil.

Se pueden emplear las herramientas de control de rango dinámico del MPEG-4 AAC.

9.1.3 Restricciones operativas referentes a la compatibilidad con receptor estéreo

Cuando el servicio multicanal está disponible:

- a) debe ocurrir obligatoriamente la transmisión de por lo menos un programa en dos canales (2/0 o estéreo) o de un programa multicanal (3/2);
- b) la transmisión simultánea en dos canales no es obligatoria cuando el servicio multicanal 3/2 está disponible. Básicamente, el receptor de dos canales (estéreo) debe ser capaz obligatoriamente de procesar la señal a través de *downmixing*;
- c) el receptor debe ser capaz obligatoriamente de interpretar el coeficiente de *downmix* utilizando PCE de acuerdo con el estándar AAC (ver Tabla 1) cuando los servicios de cinco canales estéreo (3/2) y 5.1 canales estéreo (3/2 + LFE) estén disponibles.

9.2 Restricciones de los parámetros de codificación de audio para servicio *one-seg*

9.2.1 Modos de codificación de audio

El modo de codificación determina el número de canales disponible en el servicio de audio. Los modos de codificación de audio para transmisión digital deben cumplir obligatoriamente las restricciones descritas en la Tabla 6.

Tabla 6 — Restricciones de modos de codificación de audio – Servicio *one-seg*

Parámetro	Restricción
Modos de audio permitidos	Monoaural (1/0), estéreo (2/0)

El decodificador de audio debe ser capaz obligatoriamente de procesar cualquiera de los modos recomendados de audio.

La configuración de los canales según el modo de operación, y su orden de transmisión dentro del payload deben obligatoriamente estar de acuerdo con la Tabla 7, adoptada como estándar en la ISO/IEC 14496-3.

Tabla 7 — Configuración de canales y modos estándar en el MPEG-4 AAC

Modo	Configuración de canales	Orden de transmisión del SE ^a	Elemento estándar para mapeo de altavoces ^b
Monoaural (1/0)	1	<SCE1><TERM>	SCE1 = C
Estéreo (2/0)	2	<CPE1><TERM>	CPE1 = L y R

^a Abreviaturas relacionadas a elemento sintáctico (SE): SCE – *single channel element*, CPE – *channel pair element*, LFE – *LFE channel element*, TERM – *terminator*.

^b Abreviaturas relacionadas al arreglo de altavoces: L – altavoz frontal izquierdo / R – altavoz frontal derecho / C – altavoz frontal central.

9.2.2 Principales parámetros

Los principales parámetros de codificación de audio para dispositivos portátiles deben cumplir obligatoriamente la Tabla 8.

Tabla 8 — Principales parámetros de codificación de audio para servicios *one-seg*

Parámetro	Restricción
Mecanismos de transporte permitidos	LATM/LOAS, conforme ISO/IEC 14496-3
Perfiles y niveles permitidos	<i>High efficiency</i> (HE): nivel 2 (HE-AAC v2@L2)
Número máximo de canales codificados	2 canales por flujo de bits (estéreo o 2 canales monoaurales)
Tasa máxima de bits	Conforme ISO/IEC 14496-3

La versión 2 del MPEG-4 AAC-HE debe ser adoptada obligatoriamente para transmisión para dispositivos portátiles y también es obligatoria para dispositivos fijos y móviles, si éstos van a recuperar el servicio *one-seg*.

Las señales se pueden codificar en cualquier tasa de bits y con cualquier tasa de muestreo soportado en el perfil y nivel de la Tabla 8.

Para uso de la extensión PS, el decodificador de audio debe ser capaz obligatoriamente de procesar el elemento sintáctico *sbr_extension()*, cuyo *bs_extension_id* debe ser obligatoriamente igual a EXTENSION_ID_PS, de acuerdo con la ISO/IEC 14496-3 (señalización implícita del PS).

Bibliografía

- [1] ETSI TS 101 154:2007, *Digital Video Broadcasting (DVB); Implementation guidelines for the use of video and audio coding in broadcasting applications based on the MPEG-2 transport stream*
- [2] ARI B STD-B32-2:2006, *Video coding audio coding and multiplexing specifications for digital broadcasting – Part 2: Audio signal and coding systems*