

Primeira edição
30.11.2007

Válida a partir de
01.12.2007

Versão corrigida
07.04.2008

Televisão digital terrestre — Codificação de vídeo, áudio e multiplexação
Parte 2: Codificação de áudio

Digital terrestrial television – Video coding, audio coding and multiplexing
Part 2: Audio coding

Palavras-chave: Televisão digital terrestre. Codificação de fonte. AAC. Nível e perfil.

Descriptors: Digital terrestrial television. Source coding. AAC. Level and profile.

ICS 33.160.01

ISBN 978-85-07-00560-5



ASSOCIAÇÃO
BRASILEIRA
DE NORMAS
TÉCNICAS

Número de referência
ABNT NBR 15602-2:2007
12 páginas

© ABNT 2007

Todos os direitos reservados. A menos que especificado de outro modo, nenhuma parte desta publicação pode ser reproduzida ou por qualquer meio, eletrônico ou mecânico, incluindo fotocópia e microfilme, sem permissão por escrito pela ABNT.

Sede da 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 no Brasil

Sumário

Página

Prefácio.....	iv
1 Escopo.....	1
2 Referências normativas.....	1
3 Termos e definições.....	1
4 Abreviaturas.....	2
5 Formato de entrada de áudio.....	3
5.1 Condições gerais.....	3
5.2 Principais parâmetros.....	3
5.2.1 Formatos.....	3
5.2.2 Interfaces.....	3
5.2.3 Níveis de sinal de áudio.....	4
5.2.4 Modos ou configurações multicanal.....	4
5.2.5 Metadados.....	4
6 Serviços de áudio e canais auxiliares.....	5
7 Sistema de codificação de áudio.....	6
8 Procedimentos para compressão e transmissão de áudio.....	6
8.1 Visão geral do padrão de codificação.....	6
8.2 Perfis e níveis.....	7
8.3 Camada de transporte e multiplexação.....	7
9 Restrições dos parâmetros de codificação de áudio.....	8
9.1 Restrições dos parâmetros de codificação de áudio para serviço <i>full-seg</i>	8
9.1.1 Modos de codificação de áudio.....	8
9.1.2 Principais parâmetros.....	9
9.1.3 Restrições operacionais a respeito da compatibilidade com receptor estéreo.....	10
9.2 Restrições dos parâmetros de codificação de áudio para serviço <i>one-seg</i>	10
9.2.1 Modos de codificação de áudio.....	10
9.2.2 Principais parâmetros.....	11
Bibliografia.....	12

Prefácio

A Associação Brasileira de Normas Técnicas (ABNT) é o Foro Nacional de Normalização. As Normas Brasileiras, cujo conteúdo é de responsabilidade dos Comitês Brasileiros (ABNT/CB), dos Organismos de Normalização Setorial (ABNT/ONS) e das Comissões de Estudo Especiais (ABNT/CEE), são elaboradas por Comissões de Estudo (CE), formadas por representantes dos setores envolvidos, delas fazendo parte: produtores, consumidores e neutros (universidades, laboratórios e outros).

Os Documentos Técnicos ABNT são elaborados conforme as regras da Diretivas ABNT, Parte 2.

A Associação Brasileira de Normas Técnicas (ABNT) chama atenção para a possibilidade de que alguns dos elementos deste documento podem ser objeto de direito de patente. A ABNT não deve ser considerada responsável pela identificação de quaisquer direitos de patentes.

A ABNT NBR 15602-2 foi elaborada pela Comissão de Estudo Especial de Televisão Digital (ABNT/CEE-00:001.85). O Projeto circulou em Consulta Nacional conforme Edital nº 07, de 29.06.2007 a 28.08.2007, com o número de Projeto 00:001.85-002/2.

Esta Norma é baseada nos trabalhos do Fórum do Sistema Brasileiro de Televisão Digital Terrestre, conforme estabelecido no Decreto Presidencial nº 5.820, de 29.06.2006.

A ABNT NBR 15602, sob o título geral "*Televisão digital terrestre – Codificação de vídeo, áudio e multiplexação*", tem previsão de conter as seguintes partes:

- Parte 1: Codificação de vídeo;
- Parte 2: Codificação de áudio;
- Parte 3: Sistemas de multiplexação de sinais.

Esta versão corrigida da ABNT NBR 15602-2:2007 incorpora a Errata 1 de 07.04.2008.

Televisão digital terrestre — Codificação de vídeo, áudio e multiplexação

Parte 2: Codificação de áudio

1 Escopo

Esta parte da ABNT NBR 15602 especifica os parâmetros para os sinais de áudio e o sistema de codificação e decodificação de som a ser utilizado no sistema brasileiro de televisão digital terrestre (SBTVD).

2 Referências normativas

Os documentos relacionados a seguir são indispensáveis à aplicação deste documento. Para referências datadas, aplicam-se somente as edições citadas. Para referências não datadas, aplicam-se as edições mais recentes do referido documento (incluindo emendas).

ABNT NBR 15602-3:2007, *Televisão digital terrestre – Codificação de áudio, vídeo e multiplexação – Parte 3: Sistemas de multiplexação de sinais*

ABNT NBR 15603-2:2007, *Televisão digital terrestre – Multiplexação e serviços de informação (SI) – Parte 2: Sintaxes e definições da informação 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: Áudio*

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

3 Termos e definições

Para os efeitos desta parte da ABNT NBR 15602, aplicam-se os seguintes termos e definições.

3.1

codificação

processo de transformação de sinais externos em bits que representem tais sinais

NOTA A codificação se dá, por exemplo, através de amostragem, e a informação obtida pode ainda ser compactada.

3.2

decodificação

processo responsável pela recuperação do sinal original através dos bits recebidos do codificador

NOTA A decodificação pode, eventualmente, realizar também a descompactação da informação recebida.

3.3

downmix

operação para transformar uma matriz de n canais e obter menos de n canais, usualmente realizada para a conversão de um programa multicanal para estéreo ou mono

3.4 LATM/LOAS

mecanismo de transporte definido no MPEG-4 que utiliza duas camadas, uma de multiplexação e outra de sincronização

NOTA A camada de multiplexação LATM gerencia a multiplexação de vários *payloads* de áudio (dados de áudio) e seus dados de configuração constantes nos elementos de *AudioSpecificConfig()*. A camada de sincronização LOAS especifica uma sintaxe para auto-sincronismo no fluxo de transporte de áudio MPEG-4.

3.5 nível

número máximo permitido de canais e de sua taxa de amostragem e indica a complexidade computacional para o decodificador

3.6 receptor *full-seg*

dispositivo capaz de decodificar informações de áudio, vídeo, dados etc., contidas na camada do fluxo de transporte de 13 segmentos, destinada ao serviço fixo (*indoor*) e móvel

NOTA A classificação *full-seg* é aplicada aos conversores digitais, também conhecido por *settop box*, e aos receptores de 13 segmentos integrados com tela de exibição, mas não exclusivos a estes. Este tipo de receptor é capaz de receber e decodificar sinais de televisão digital terrestre de alta definição e, a critério do fabricante, também receber e decodificar informações transportadas na camada "A" do fluxo de transporte, aplicada para os serviços direcionados aos receptores portáteis, definidos como *one-seg*.

3.7 receptor *one-seg*

dispositivo que decodifica exclusivamente informações de áudio, vídeo, dados etc., contidas na camada "A" locada no segmento central dos treze segmentos

NOTA A classificação *one-seg* é destinada aos receptores do tipo portátil, também conhecidos por "*handheld*", especialmente recomendados para telas de exibição de dimensões reduzidas, normalmente até 7 polegadas. Entre os produtos classificados como *one-seg*, estão os receptores integrados com telefone celular, PDA, *dongle* e televisores portáteis, os quais são energizados por uma bateria interna e, portanto, sem necessariamente demandar uma fonte externa de energia, bem como aqueles destinados a veículos automóveis. Este tipo de receptor é capaz de receber e decodificar apenas sinais de televisão digital terrestre transportado na camada "A" do fluxo de transporte e, conseqüentemente, apenas sinais de perfil básico, destinados aos dispositivos portáteis de recepção.

3.8 unidade de acesso de áudio

porção de áudio de um *elementary stream* que é acessível individualmente

NOTA Para os efeitos desta parte da ABNT NBR 15602, uma unidade de acesso de áudio equivale a um *rawdatablock()*.

4 Abreviaturas

Para os efeitos desta parte da ABNT NBR 15602, aplicam-se as seguintes abreviaturas:

AAC Codificação de Áudio Avançada (*Advanced Audio Coding*)

CPE *Channel Pair Element*

HDMI *High-definition Multimedia Interface*

LATM *Low Overhead AudioTransport Multiplex*

LFE *Low Frequency Enhancement*

LOAS *Low Overhead Audio Stream*

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	Fluxo de Transporte (<i>Transport Stream</i>)

5 Formato de entrada de áudio

5.1 Condições gerais

As condições gerais para o formato de entrada de áudio devem obrigatoriamente ser as seguintes:

- freqüência de amostragem do sinal de áudio: 32 kHz, 44,1 kHz ou 48 kHz;
- configuração de sinais estereofônicos e multicanal (ou seja, sinais consistindo em dois ou mais sinais de áudio para obter uma reprodução envolvente ou espacial do som); a taxa de amostragem para todos os sinais deve obrigatoriamente ser a mesma;
- quantização dos sinais de entrada deve empregar 16 bits ou 20 bits;
- um programa de áudio deve obrigatoriamente ter no mínimo um canal de áudio. O número máximo de canais no programa deve obrigatoriamente ser limitado ao número máximo de canais permitidos pela ISO/IEC 14496-3;
- é recomendado que os programas multicanal sejam preparados conforme a ITU Recommendation BS.775-1;
- os programas de áudio em modo multicanal compatíveis com os modos previstos na ITU Recommendation BS.775-1 devem obrigatoriamente estar em uma das configurações permitidas na Tabela 3;
- no caso de transmissão de somente um programa multicanal sem transmissão de um programa estéreo, o programa multicanal deve obrigatoriamente estar em modo 3/2 (5.0 ou 5.1, com ou sem adição do canal LFE de enriquecimento das baixas freqüências) para permitir o *downmix* para estéreo.

5.2 Principais parâmetros

5.2.1 Formatos

Devem obrigatoriamente ser admitidos fluxos de bits ou arquivos contendo áudio digital não comprimido em formato PCM, como WAVE ou AIFF, estéreo e multicanal.

5.2.2 Interfaces

Entre as interfaces (barramentos) de entrada/saída digital permitidos, devem obrigatoriamente estar AES3 (AES/EBU, contendo dois canais PCM por fluxo de bits), SDI, HD-SDI e HDMI.

5.2.3 Níveis de sinal de áudio

O nível de referência para a intensidade ou pressão sonora deve obrigatoriamente ser igual a 0 dB. A faixa dinâmica admissível de excursão deve obrigatoriamente ser limitada a + 20 dB (*headroom*) e - 70 dB com respeito à referência, correspondendo a uma faixa dinâmica típica de 90 dB. Convém que os níveis de áudio médio estejam a - 20 dBFS (0 dB), para possibilitar homogeneidade no volume entre canais distintos. O sinal deve acomodar picos de no mínimo 4 vezes sua potência média RMS.

5.2.4 Modos ou configurações multicanal

O modo de transmissão se refere à configuração multicanal utilizada, ao número de canais disponível no fluxo de bits e à forma de codificação desses canais.

O número de canais de áudio fonte deve obrigatoriamente ser no mínimo um para uma configuração básica, dois para transmissão padrão estéreo típico e cinco canais mais um canal de baixas frequências (LFE) para transmissão multicanal "5.1" padrão. Os sinais fontes devem obrigatoriamente ser pré-processados e/ou combinados previamente à entrada do codificador, para produzirem os canais de transmissão que devem obrigatoriamente estar presentes no fluxo de bits.

Uma mesma programação de áudio pode ser transmitida em mais de um modo, por exemplo, em estéreo (dois canais) mais modo multicanal 3/2 (5.1) simultaneamente, porém a transmissão simultânea não é obrigatória.

No caso da transmissão exclusiva em modo multicanal 3/2 (5.1), os receptores devem obrigatoriamente ser capazes de sintetizar o sinal estéreo por meio de conversão (*downmixing*), operações de replicação, *dematrixing*, combinação e processamento de sinal no âmbito funcional do sistema de reprodução de áudio do receptor. Os modos multicanal permitidos para codificação e transmissão devem obrigatoriamente atender ao descrito em 9.1.1.

5.2.5 Metadados

Dados auxiliares devem obrigatoriamente conter informações como descrições de conteúdo dos programas de áudio, parâmetros de configuração dos serviços de áudio e parâmetros dos sinais de áudio transmitidos no fluxo de bits.

Podem ser admitidos como tipos de dados auxiliares:

- a) descrição do conteúdo dos programas de áudio sendo transmitidos (por exemplo, classificação de programa sonoro, descrição dos objetos de áudio mixados no conteúdo, descrição do conteúdo do canal de áudio auxiliar etc.);
- b) modo multicanal;
- c) volume de referência para operações de equalização na reprodução no terminal de acesso.

Dados auxiliares e a descrição de conteúdo de programas de áudio podem ser classificados em dois níveis.

Um primeiro nível deve obrigatoriamente ser normativo. Esse nível deve obrigatoriamente afetar diretamente a operação do receptor (decodificação dos fluxos de bits) como, por exemplo, informação de quantidade e modo dos canais e perfil e nível de codificação extraídas diretamente das tabelas PSI. Os dados nesta categoria devem obrigatoriamente ser essenciais para a decodificação e reprodução correta do serviço de áudio no receptor.

Um segundo nível deve obrigatoriamente ser informativo. Esse nível não deve afetar a decodificação, mas sim trazer informações sobre os conteúdos dos programas de áudio associados a cada PID. Os dados nesta categoria devem obrigatoriamente ser usados para processamento de informação sobre os programas no receptor.

A Tabela 1 resume os tipos de dados auxiliares para áudio admissíveis no sistema.

Tabela 1 — Tipos de dados auxiliares para áudio

Parâmetros	Descrição e uso
<i>matrix_downmix_idx</i>	<p>Descrição: indicador do coeficiente a ser utilizado no <i>downmix</i> de multicanal para estéreo. Deve obrigatoriamente ser transmitido no fluxo de bits como metadado, conforme especificado em ISO/IEC 14496-3</p> <p>Uso: obrigatório quando um programa em modo multicanal for transmitido</p> <p>Quando o parâmetro <i>matrix_mixdown_idx_present</i> estiver ativo em “1” no respectivo PCE do programa (PID), deve-se obrigatoriamente utilizar o sistema de <i>downmix</i> descrito na ISO/IEC 14496-3:2005, SubSeção 4.5.1.2.2 e Tabela 4.70</p> <p>Quando o terminal receptor realizar operação de <i>downmix</i>, a saída estéreo analógica deve obrigatoriamente estar sempre ativa com este sinal</p>
<i>program_ref_level</i>	<p>Descrição: valor representativo da intensidade média do volume do programa de áudio de longo prazo para todos os canais combinados em relação à referência 0 dBFS. Representado em 128 níveis (7 bits), quantizado em passos de 0,25 dB, totalizando uma excursão de 32 dB em relação ao fundo de escala (0 dBFS)</p> <p>Este parâmetro deve conter obrigatoriamente uma descrição informativa do volume de referência adotado pela emissora (0 dB) em relação ao fundo de escala (0 dBFS), para efeito de normalização do diálogo e para tornar a troca de canais mais confortável para o usuário</p> <p>Uso: obrigatório. Recomenda-se utilizar <i>prog_ref_level</i> = 80 (0x50), que corresponde a um valor indicativo de – 20 dBFS como referência 0 dB, conforme a ISO/IEC 14496-3</p> <p>O parâmetro deve ser transmitido como junto à estrutura DRC, conforme ISO/IEC 14496-3:2005, Seção 4.5.2.7</p>
<i>Dynamic Range Control (DRC)</i>	<p>Descrição: o controle de faixa dinâmica é especialmente indicado às transmissões em modo multicanal e pode ser sinalizado na forma de metadados, conforme a ISO/IEC 14496-3:2005, Seção 4.5.2.7</p> <p>Uso: na codificação é opcional, contudo o decodificador deve obrigatoriamente suportar essa ferramenta. No caso de a informação de DRC não ser enviada pelo codificador, o decodificador não deve aplicar a ferramenta DRC</p>

6 Serviços de áudio e canais auxiliares

Serviços de áudio incluem a transmissão de programas de áudio adicionais ao programa principal e são obrigatoriamente considerados serviços opcionais, com exceção do serviço de descrição de áudio (DA), cuja transmissão é obrigatória conforme legislação vigente.

A transmissão destes serviços deve ser realizada através da alocação de canais de áudio auxiliares adicionais em programas de áudio (PID) distintos, ou no mesmo fluxo de bits de um mesmo PID, respeitando-se sempre o número máximo de canais permitidos no fluxo de bits pelo perfil/nível de codificação usado.

Canais adicionais ao programa principal podem ser utilizados para transmitir áudio em outros idiomas (como, por exemplo, serviço de programa de áudio secundário SAP), para transmitir serviços de descrição de áudio (DA), para transmitir programas adicionais ao programa principal e áudio secundário proveniente de outras tomadas de som (conteúdo adicional, como, por exemplo, efeitos).

Todos os canais adicionais referentes a serviços de áudio auxiliares devem ser obrigatória e apropriadamente sinalizados utilizando uma identificação válida de tipo de componente (*component_type*) no respectivo descritor de áudio (*audio_component_descriptor*) do programa.

Os canais auxiliares devem obrigatoriamente ser transmitidos em programas distintos (PID distintos), com a devida sinalização e identificação de seus canais, para serem selecionados, decodificados e reproduzidos juntamente com ou em substituição aos canais de áudio do programa principal.

O serviço de DA (descrição de áudio) consiste usualmente em um canal monaural de voz e fornece uma descrição da cena como um subcomponente associado a um serviço de televisão. Ele deve obrigatoriamente auxiliar no entendimento do entretenimento principal (mas não exclusivamente) para telespectadores com problemas visuais.

A transmissão de DA deve obrigatoriamente ser implementada utilizando ao menos um dos mecanismos abaixo:

- a) como canal auxiliar (monaural ou estéreo) contendo a descrição de áudio previamente mixada com um programa de áudio principal;
- b) como canal auxiliar contendo a descrição de áudio em separado, para posterior mixagem com o programa de áudio no receptor.

Em ambos os casos, convém sinalizar o serviço através do parâmetro *component_type* disponível no descritor "Audio component descriptor", conforme ABNT NBR 15603-2:2007, Tabela 28.

A habilidade de misturar um ou mais canais suplementares de descrição com o áudio do programa principal pode ter outras aplicações, incluindo comentários em outras línguas (multilíngual), interatividade e propósitos educacionais.

7 Sistema de codificação de áudio

Os sinais de áudio devem obrigatoriamente ser codificados por uma combinação de codificação transformada no tempo e em frequência. A transformada em frequência deve decompor o sinal de entrada em seus componentes de frequência empregando a transformada discreta do cosseno (DCT – *Discrete Cosine Transform*) quando a quantidade de informação é reduzida, diminuindo-se o desvio em frequência de cada componente.

Como ferramenta de compressão adicional empregam-se ponderações psico-acústicas do sinal, onde os códigos devem obrigatoriamente ser ponderados de forma a minimizar a degradação nas faixas de frequências perceptíveis pelo ouvido humano.

A compressão de áudio e os procedimentos de transmissão devem obrigatoriamente ser compatíveis com a ISO/IEC 14496-3.

O decodificador deve obrigatoriamente ser construído assumindo-se que qualquer estrutura válida da ISO/IEC 13818-1, incluindo descritores privados, pode ser incluída no fluxo de bits mesmo que no início das transmissões esses descritores não sejam utilizados. O decodificador de áudio deve obrigatoriamente desconsiderar estruturas "reservadas" ou aquelas que correspondem a funções não implementadas pelo receptor.

8 Procedimentos para compressão e transmissão de áudio

8.1 Visão geral do padrão de codificação

A Figura 1 mostra os procedimentos para compressão e transmissão de áudio.

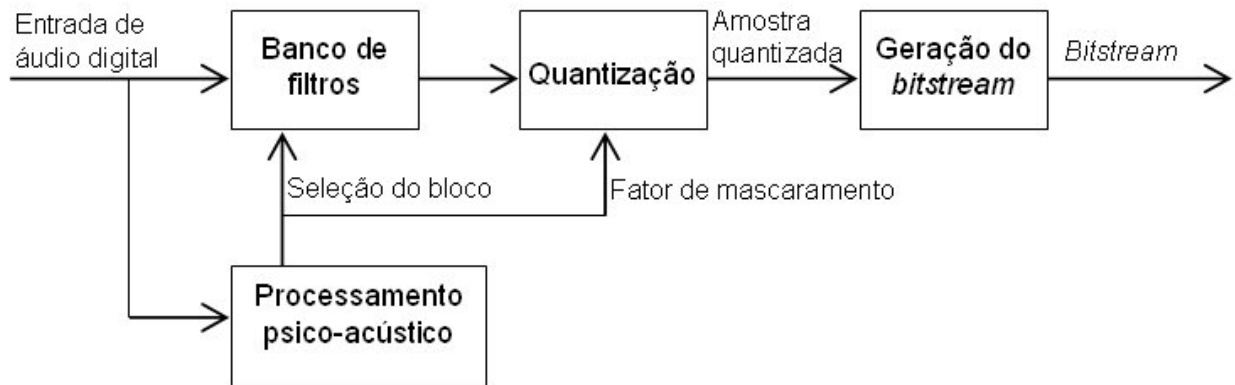


Figura 1 — Procedimentos de transmissão e codificação de áudio

O banco de filtros deve obrigatoriamente converter a entrada de sinal de áudio-digital do domínio do tempo para o domínio da frequência. A partir de então, o banco de filtros aplica a transformada discreta dos cossenos modificada e funções de janelamento a blocos do sinal de entrada, de acordo com características fisiológicas audíveis.

Os processos psico-acústicos devem obrigatoriamente calcular a quantidade de mascaramento (limite para o discernimento entre um sinal específico de áudio de outros sinais) e alimentam o banco de filtros com blocos do sinal de entrada.

As amostras devem obrigatoriamente ser quantizadas após o processamento pelo banco de filtros, com base no fator de mascaramento calculado pelos processos de psico-acústica. Deste modo, o número total de bits utilizado por cada bloco não deve ser excedido.

O fluxo de bits deve obrigatoriamente ser configurado conforme ISO/IEC 14496-3.

8.2 Perfis e níveis

A codificação de áudio deve obrigatoriamente ser compatível com a ISO/IEC 14496-3. Os seguintes perfis e níveis do padrão MPEG-4 AAC devem obrigatoriamente ser permitidos:

- LC (*low complexity*), perfil básico do padrão AAC; níveis L2 e L4;
- HE (*high efficiency*), perfil avançado de alta eficiência, combinando o perfil LC com o uso da ferramenta SBR (*spectral band replication*) para a versão 1 deste perfil, níveis L2 e L4;
- HE combinado à ferramenta PS (*parametric stereo*) para a versão 2 deste perfil; nível L2.

O perfil e o nível do codificador MPEG-4 AAC devem obrigatoriamente ser sinalizados conforme ABNT NBR 15602-3 e ABNT NBR 15603-2.

8.3 Camada de transporte e multiplexação

A codificação e o empacotamento (*framing*) intermediário do áudio devem obrigatoriamente ser compatíveis com LATM/LOAS, conforme a ISO/IEC 14496-3. O *elementary stream* deve obrigatoriamente ser primeiramente encapsulado no formato de transporte LATM e deve obrigatoriamente utilizar o elemento de multiplexação *AudioMuxElement()*.

A camada de sincronização do transporte de áudio (LOAS) deve utilizar o formato de transmissão *AudioSyncStream()* conforme em ISO/IEC 14496-3.

O áudio MPEG-4 transportado no fluxo de transporte MPEG-2 (TS), utilizando-se a sintaxe de transporte LATM/LOAS deve obrigatoriamente ser identificado por *stream_type* 0x11 de acordo com o *stream_type assignments* na ISO/IEC 13818-1:2007.

Para decodificação de áudio o receptor deve obrigatoriamente identificar o tipo, o perfil e o nível transmitidos e ser capaz de extrair os *payloads* dos objetos de áudio. Para transmissão de áudio MPEG-4 sobre fluxo de transporte MPEG-2, é obrigatório o uso de sinalização explícita de SBR sem alinhamento de PES.

Os receptores devem obrigatoriamente ser capazes de processar a ferramenta SBR. A sinalização da presença de SBR deve obrigatoriamente usar o mecanismo de sinalização explícito *non-backward compatible*, de acordo com a ISO/IEC 14496-3.

A Tabela 2 descreve os campos da sintaxe LATM/LOAS de transporte em *StreamMuxConfig* que devem obrigatoriamente estar formatados para a identificação e recuperação dos *payloads* de áudio, de acordo com a ISO/IEC 14496-3.

Tabela 2 — Principais parâmetros LATM a serem configurados

Parâmetro LATM	Descrição de uso
<i>audioMuxVersion</i>	Deve obrigatoriamente assumir o valor "0"
<i>allStreamsSameTimeFraming</i>	Deve obrigatoriamente assumir o valor "1"
<i>numSubFrames</i>	Deve obrigatoriamente assumir o valor "0" indicando um único <i>PayloadMux()</i> (<i>access unit</i>) presente no <i>AudioMuxElement()</i>
<i>numProgram</i>	Deve obrigatoriamente assumir o valor "0" indicando um programa por multiplex LATM
<i>numLayer</i>	Deve obrigatoriamente assumir o valor "0" indicando apenas uma camada
<i>frameLenghtType</i>	Deve obrigatoriamente assumir o valor "0" indicando que o tamanho de quadro do <i>payload</i> pode ter extensão variável. A extensão do <i>payload</i> em bytes é diretamente especificada em <i>PayloadLengthInfo()</i> com palavras de 8 bits

9 Restrições dos parâmetros de codificação de áudio

9.1 Restrições dos parâmetros de codificação de áudio para serviço *full-seg*

9.1.1 Modos de codificação de áudio

O modo de codificação determina o número de canais disponível no serviço de áudio. Os modos de codificação de áudio para transmissão digital devem obrigatoriamente atender às restrições descritas na Tabela 3.

Tabela 3 — Restrições de modos de codificação de áudio

Parâmetro	Restrição
Modos de áudio permitidos	Monaural (1/0), estéreo (2/0 e 2/0 + LFE) ^a , estéreo multicanal (3/0, 2/1, 3/1, 2/2, 3/2, 3/2+LFE) ^a , dois sinais de áudio independentes (monaural dual), multi-áudio (três ou mais sinais de áudio) e combinações destes
Modos de áudio recomendados	Estéreo (2/0), multicanal (3/2+LFE)
<i>Downmix</i>	Para a configuração 5.0 e 5.1, deve-se obrigatoriamente utilizar o esquema conforme Tabela 1. Nas demais configurações multicanal, outros esquemas de <i>downmix</i> podem ser usados pelo receptor, desde que mantenham a integridade do áudio e o nível de inteligibilidade. O esquema de <i>downmix</i> estéreo para mono não está coberto nesta Norma, porém o <i>clipping</i> deve obrigatoriamente ser evitado

^a Número de canais para alto-falantes dianteiros/traseiros.

EXEMPLO 3/1 = 3 dianteiros + 1 traseiro; 3/2 = 5.0 = 3 canais dianteiros e 2 traseiros.

O decodificador deve obrigatoriamente ser capaz de processar qualquer um dos modos recomendados de áudio.

A configuração dos canais segundo o modo de operação e sua ordem de transmissão dentro do *payload* devem obrigatoriamente estar de acordo com a Tabela 4.

Tabela 4 — Configuração de canais e modos-padrão recomendados no MPEG-4 AAC

Modo	Configuração de canais	Ordem de transmissão do SE ^a	Elemento-padrão para mapeamento de alto-falantes ^b
Monaural (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 e RS, LFE = LFE

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

^b Abreviaturas relacionadas ao arranjo de alto-falantes: L – alto-falante frontal esquerdo / R – alto-falante frontal direito / C – alto-falante frontal central / LFE – ênfase de baixa frequência / LS – alto-falante *surround* traseiro esquerdo / RS – alto-falante *surround* traseiro direito / MS – alto-falante *surround monaural*.

No caso da transmissão de dois sinais de áudio independentes (monaural dual ou 1/0+1/0), recomenda-se a ordem de transmissão do SE da seguinte forma: <SCE1><SCE2><TERM>, sendo SCE1 o primeiro (principal) canal e o SCE2 o segundo canal no programa.

Se a configuração utilizada não estiver entre as citadas na Tabela 4, deve obrigatoriamente ser reproduzida usando uma configuração com o mesmo número de canais e com a respectiva sinalização.

9.1.2 Principais parâmetros

Os principais parâmetros do sistema de codificação de áudio devem obrigatoriamente atender à Tabela 5.

Tabela 5 — Principais parâmetros do sistema de codificação de áudio – Serviço *full-seg*

Parâmetro	Restrição
Mecanismos de transporte permitidos	LATM/LOAS (conforme ISO/IEC 14496-3)
Números de canais recomendados	Mono (1.0), 2 canais (estéreo ou 2.0), ou multicanal (5.1)
Perfis e níveis permitidos	<i>Low complexity</i> AAC: nível 2 (LC-AAC@L2) para dois canais <i>Low complexity</i> AAC: nível 4 (LC-AAC@L4) para multicanal <i>High Efficiency</i> (HE): nível 2 (HE-AAC v1@L2) para dois canais <i>High Efficiency</i> (HE): nível 4 (HE-AAC v1@L4) para multicanal
Taxa máxima de bits permitida	Conforme ISO/IEC 14496-3
Amostras por quadro	frameLengthFlag em GASpecificConfig() deve ter valor 0, indicando que a extensão do quadro deve ser de 1024 amostras para AAC e 2048 quando usando SBR. 960 amostras para AAC (ou 1 920 quando usando SBR) não devem ser utilizadas

Para transmissão em alta fidelidade recomenda-se utilizar o perfil/nível AAC@L4 para modo multicanal e o perfil/nível AAC@L2 para modo estéreo. Na transmissão de áudio estéreo, não deve ser empregado o nível 4 (L4).

Os sinais podem ser codificados em qualquer taxa suportada no perfil e nível selecionado. Ao mesmo tempo o sinal multicanal pode empregar qualquer frequência de amostragem do perfil.

As ferramentas de controle de range dinâmico do MPEG-4 AAC podem ser empregadas.

9.1.3 Restrições operacionais a respeito da compatibilidade com receptor estéreo

Quando o serviço multicanal estiver disponível:

- a) deve obrigatoriamente ocorrer a transmissão de pelo menos um programa em dois canais (2/0 ou estéreo) ou de um programa multicanal (3/2);
- b) a transmissão simultânea em dois canais não é obrigatória quando o serviço multicanal 3/2 estiver disponível. Basicamente, o receptor de dois canais (estéreo) deve obrigatoriamente ser capaz de processar o sinal através de *downmixing*;
- c) o receptor deve obrigatoriamente ser capaz de interpretar o coeficiente de *downmix* utilizando PCE de acordo com o padrão AAC (ver Tabela 1) quando os serviços de cinco canais estéreo (3/2) e 5.1 canais estéreo (3/2 + LFE) estiverem disponíveis.

9.2 Restrições dos parâmetros de codificação de áudio para serviço *one-seg*

9.2.1 Modos de codificação de áudio

O modo de codificação determina o número de canais disponível no serviço de áudio. Os modos de codificação de áudio para transmissão digital devem obrigatoriamente atender às restrições descritas na Tabela 6.

Tabela 6 — Restrições de modos de codificação de áudio – Serviço *one-seg*

Parâmetro	Restrição
Modos de áudio permitidos	Monoaural (1/0), estéreo (2/0)

O decodificador de áudio deve obrigatoriamente ser capaz de processar qualquer um dos modos recomendados de áudio.

A configuração dos canais segundo o modo de operação, e sua ordem de transmissão dentro do *payload* devem obrigatoriamente estar de acordo com a Tabela 7, adotada como padrão na ISO/IEC 14496-3.

Tabela 7 — Configuração de canais e modos-padrão no MPEG-4 AAC

Modo	Configuração de canais	Ordem de transmissão do SE ^a	Elemento-padrão para mapeamento de alto-falantes ^b
Monoaural (1/0)	1	<SCE1><TERM>	SCE1 = C
Estéreo (2/0)	2	<CPE1><TERM>	CPE1 = L e R

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

^b Abreviaturas relacionadas ao arranjo de alto-falantes: L – alto-falante frontal esquerdo / R – alto-falante frontal direito / C – alto-falante frontal central.

9.2.2 Principais parâmetros

Os principais parâmetros de codificação de áudio para dispositivos portáteis devem obrigatoriamente atender à Tabela 8.

Tabela 8 — Principais parâmetros de codificação de áudio para serviços *one-seg*

Parâmetro	Restrição
Mecanismos de transporte permitidos	LATM/LOAS, conforme ISO/IEC 14496-3
Perfis e níveis permitidos	<i>High efficiency</i> (HE): nível 2 (HE-AAC v2@L2)
Número máximo de canais codificados	2 canais por fluxo de bits (estéreo ou 2 canais monaurais)
Taxa máxima de bits	Conforme ISO/IEC 14496-3

A versão 2 do MPEG-4 AAC-HE deve obrigatoriamente ser adotada para transmissão para dispositivos portáteis e também é obrigatória para dispositivos fixos e móveis, se estes forem recuperar o serviço *one-seg*.

Os sinais podem ser codificados em qualquer taxa de bits e com qualquer taxa de amostragem suportada no perfil e nível da Tabela 8.

Para uso da extensão PS, o decodificador de áudio deve obrigatoriamente ser capaz de processar o elemento sintático *sbr_extension()*, cujo *bs_extension_id* deve obrigatoriamente ser igual a EXTENSION_ID_PS, de acordo com a ISO/IEC 14496-3 (sinalização implícita do PS).

Bibliografia

- [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] ARIB STD-B32-2:2006, *Video coding audio coding and multiplexing specifications for digital broadcasting – Part 2: Audio signal and coding systems*