International Telecommunication Union

# ITU-R

Radiocommunication Sector of ITU

# Recommendation ITU-R BS.1548-2

**(02/2006)**

# User requirements for audio coding systems for digital broadcasting

**BS Series**

**Broadcasting service (sound)**

International Telecommunication Union

## Foreword

The role of the Radiocommunication Sector is to ensure the rational, equitable, efficient and economical use of the radio-frequency spectrum by all radiocommunication services, including satellite services, and carry out studies without limit of frequency range on the basis of which Recommendations are adopted.

The regulatory and policy functions of the Radiocommunication Sector are performed by World and Regional Radiocommunication Conferences and Radiocommunication Assemblies supported by Study Groups.

## Policy on Intellectual Property Right (IPR)

ITU-R policy on IPR is described in the Common Patent Policy for ITU-T/ITU-R/ISO/IEC referenced in Annex 1 of Resolution ITU-R 1. Forms to be used for the submission of patent statements and licensing declarations by patent holders are available from http://www.itu.int/ITU-R/go/patents/en where the Guidelines for Implementation of the Common Patent Policy for ITU-T/ITU-R/ISO/IEC and the ITU-R patent information database can also be found.

### Series of ITU-R Recommendations

(Also available online at http://www.itu.int/publ/R-REC/en)

| Series | Title |
|---|---|
| **BO** | Satellite delivery |
| **BR** | Recording for production, archival and play-out; film for television |
| **BS** | **Broadcasting service (sound)** |
| **BT** | Broadcasting service (television) |
| **F** | Fixed service |
| **M** | Mobile, radiodetermination, amateur and related satellite services |
| **P** | Radiowave propagation |
| **RA** | Radio astronomy |
| **RS** | Remote sensing systems |
| **S** | Fixed-satellite service |
| **SA** | Space applications and meteorology |
| **SF** | Frequency sharing and coordination between fixed-satellite and fixed service systems |
| **SM** | Spectrum management |
| **SNG** | Satellite news gathering |
| **TF** | Time signals and frequency standards emissions |
| **V** | Vocabulary and related subjects |

***Note****: This ITU-R Recommendation was approved in English under the procedure detailed in Resolution ITU-R 1.*

*Electronic Publication*
Geneva, 2011

# RECOMMENDATION ITU-R BS. 1548-2*, **

## User requirements for audio coding systems for digital broadcasting

(Question ITU-R 19/6)

(2001-2002-2006)

### Scope

This Recommendation specifies the requirements relevant to the use of audio source coding systems in sound broadcasting, including television. The Recommendation covers the application of contribution and distribution, and emission.

The ITU Radiocommunication Assembly,

*considering*

a) that the multichannel sound system, with or without accompanying picture, is the subject of Recommendation ITU-R BS.775;

b) that audio coding for digital broadcasting is the subject of Recommendation ITU-R BS.1196;

c) that the coding systems recommended in Recommendation ITU-R BS.1196 offer monophonic, two-channel stereophonic and multichannel coding modes;

d) that the basic audio and stereo image quality required for sound systems for television and sound broadcasting is to be the highest possible, generally indistinguishable from the source material;

e) that required audio quality for some emission applications is to be equivalent to or better than good reception of FM or AM analogue broadcasting services;

f) that Recommendation ITU-R BS.1283 provides a guide to ITU-R Recommendations for subjective assessment of sound quality;

g) that interoperability and network operation involving programme connections such as contribution and distribution links should be carefully considered;

h) that interoperability with existing consumer multichannel audio equipment, such as matrix surround decoders and discrete multichannel decoders, should be carefully considered;

j) that, when introducing a multichannel sound system in an existing broadcasting service, compatibility with existing receivers to maintain the service must be considered;

---

* Radiocommunication Study Group 6 made editorial amendments to this Recommendation in November 2009 in accordance with Resolution ITU-R 1.

** Radiocommunication Study Group 6 made editorial amendments to this Recommendation in October 2010 in accordance with Resolution ITU-R 1.

k) that more generally, in view of the many applications of such systems, all technical, quality and operational requirements should be clearly specified;

l) that the performance of audio coding systems is widely dependent on the configuration under which the system is operated (bit rate, use of pre-matrixing, use of composite coding, etc.);

m) that several broadcast services already use or have specified the use of the systems recommended in Recommendations ITU-R BS.1196 ;

n) that, consequently, the broadcasters have a need of information necessary to set up all the available coding parameters of the recommended systems;

o) that the introduction of incompatible systems with similar performance characteristics is highly undesirable;

p) that those broadcasters which have not yet started services should be able to choose the system which is best suited to their application and which is the most cost-effective,

*recommends*

**1** that the audio coding systems for digital television and sound broadcasting for contribution and distribution applications should fulfil the requirements listed in Annex 1;

**2** that the audio coding systems for digital television and sound broadcasting for emission applications should fulfil the requirements listed in Annex 2;

**3** that the categories of audio quality listed in Annex 3 should govern the audio quality and applications in *recommends* 1 and 2.

NOTE 1 – Information about systems that have been shown to meet the quality and other requirements for contribution and distribution applications is included in Appendix 1 to Annex 1.

NOTE 2 – Information about systems that have been shown to meet the quality and other requirements for emission applications is included in Appendix 1 to Annex 2.

# Annex 1

# Requirements for contribution and distribution

The audio coding systems for digital television and sound broadcasting for both contribution and distribution applications should fulfil the requirements listed below.

## 1 Service requirements

### 1.1 Channel configurations

For audio services the following channel configurations should be supported according to the needs of applications (see Recommendation ITU-R BS.775 – Multichannel sterophonic sound system with and without accompanying picture):

| No. of channels | Channel configuration | Channel assignment |
|---|---|---|
| 1 channel | 1/0 | Mono |
| 2 channels | 2/0 | Left, right |
| 3 channels | 3/0<br>2/1 | Left, right, centre<br>Left, right/surround |
| 4 channels | 3/1<br>2/2 | Left, right, centre/surround<br>Left, right/surround left, surround right |
| 5 channels | 3/2 | Left, right, centre/surround left, surround right |

together with an optional low frequency effects (LFE) channel.

For contribution, in addition, it could be necessary to convey programmes produced in other formats than those listed above, e.g. 3/4, thus the coding system should allow for accommodation of additional high quality channels.

### 1.2 Flexible allocation of channels

A bit stream should provide identification data for signalling and controlling of sound configurations. It must be possible in the transmission system to switch dynamically among the channel configurations listed in § 1.1.

### 1.3 Ancillary data

The audio coding system should provide for the possibility of transmission of ancillary data. The ancillary data can convey various types of information, including dynamic range control, loudness control, user data, and any metadata required by the emission encoder that will encode the final audio for delivery to the consumer.

## 2 Performance requirements

### 2.1 Audio quality

#### 2.1.1 Basic audio quality

The quality of sound reproduced after a reference contribution/distribution cascade (five contribution codecs and three distribution codecs working consecutively) should be subjectively indistinguishable from the source for most types of audio programme material. Using the triple stimuli double blind with hidden reference test, described in Recommendation ITU-R BS.1116 – Methods for the subjective assessment of small impairments in audio systems including multichannel sound systems – this requires mean scores generally higher than 4.5 in the impairment 5-grade scale, for listeners at the reference listening position. The worst rated item should not be graded lower than 4 (Recommendation ITU-R BS.775).

NOTE 1 – The confidence interval (error bar) associated with the single mean score for a codec and item shows the range above and below the stated mean score in which the true score may fall, with some degree of certainty, usually 95%. The true score for a codec and item may be as poor as the lower limit of the confidence interval about the stated score. In order to make a meaningful evaluation of the expected performance of cascaded codecs, the confidence interval associated with the reported mean scores for the individual codecs must be approximately equal to or less than the difference between the scores being compared.

NOTE 2 – The contribution/distribution cascade, when placed in tandem with the emission codec, should not cause a significant reduction in quality compared to the basic audio quality of the emission codec. Precise specification requires further study.

NOTE 3 – The objective audio quality parameters for contribution/distribution can be incorporated later, conforming to Recommendation ITU-R BS.1387.

### 2.1.2 Quantization resolution

The required resolution should be at least 18 bits for distribution and 20 bits or greater is preferable for contribution.

### 2.1.3 Sampling frequency

In agreement with Recommendation ITU-R BS.646 – Source encoding for digital sound signals in broadcasting studios, the sampling frequency should be 48 kHz.

### 2.1.4 Bandwidth

Main audio channels: 20-20 000 Hz.

LFE channel: 15-120 Hz.

### 2.1.5 Emphasis

The audio coding system should be emphasis free.

### 2.1.6 Tandem capability

The tandem capability required depends on the application according to the following table:

| | |
|---|---|
| Distribution | 3 codecs in cascade |
| Contribution | 5 codecs in cascade |

These figures have been taken from previous experiments done to evaluate two-channel sound broadcasting systems (see Recommendation ITU-R BS.1196 – Audio coding for digital broadcasting) and may not be representative of the practical radio and television broadcasting operational situations. More information is required to specify this aspect better.

### 2.1.7 Post-processing capability

The post-processing capability required is strongly dependent on the application. For distribution crossfades can be applied together with dynamic range control.

## 2.2 Coding delay

Coding delay for all channels in a programme must be identical. The coding delay should be as low as possible, considering the coding performance (i.e. amount of bit rate reduction) required. In case of television sound, the delay of audio must be matched with the delay of video. It is desirable that the audio coder produces encoded audio frames (access units) that correspond exactly to the time period of the matching video frame.

### 2.3 Error resilience

A mechanism must be provided in the audio bit stream to allow the decoder to identify residual channel errors and to adopt proper concealment methods.

### 2.4 Recovery time

The recovery time should be as low as possible. In case of audio access unit (AAU) applied, the recovery time should be within a few AAU, and preferably the audio should resume upon receipt of the first error free AAU.

## 3 Functional and operational requirements

### 3.1 Bit rate and coding scheme

For distribution and contribution links, Recommendation ITU-R BS.1196 recommends the International Organization for Standardization/International Electrotechnical Commission (ISO/IEC) IS 11172-3 Layer II at a bit rate of 180 kbit/s per channel or above. For several reasons the system may be applied at a different bit rate or other systems may be employed.

These reasons may include the following:

– additional coding margin to support signal processing that may be inserted between coding generations (this was not tested or verified in the development of Recommendation ITU-R BS.1196);

– to obtain a lower bit rate in the distribution and contribution link;

– to obtain a higher quality;

– suitability of synchronization and switching with accompanying video signals.

### 3.2 Composite coding

Two-channel or multichannel programme material often contains some inter-channel statistical correlation. Composite coding can be an effective way to reduce the inter-channel irrelevance or redundancy, thus increasing the coding efficiency. Some coding systems use perceptual criteria to eliminate part of the inter-channel irrelevance by joining together two or more channels in frequency regions where the ability of the human ear to discriminate the direction of the source is poor. The disadvantage of this technique is that it is not possible to correctly reposition the sound information generally in the original channels at a later stage. For contribution and many distribution applications such composite coding schemes should not be used.

# Appendix 1
# to Annex 1

## Information about coding systems that have been demonstrated to meet quality, and other, user requirements for contribution and distribution

Table 1 lists, in the left hand column, the requirements specified in Annex 1. Right hand column show the ability of a specific codec to meet these requirements. It is anticipated that future revisions to this Recommendation will contain additional information about additional codecs.

TABLE 1

| List of requirements from Annex 1 | Codec: Dolby E [ref. 1] |
|---|---|
| 1.1 Channel configurations | Fulfilled [ref. 1, p. 6] |
| 1.2 Flexible channel allocation | Fulfilled [ref. 1, p. 15] |
| 1.3 Ancillary data | Fulfilled [ref. 1, p. 14] |
| 2.1.1 Basic audio quality | Fulfilled [ref. 2] |
| 2.1.2 Quantization | Fulfilled [ref. 1, p. 5] |
| 2.1.3 Sampling frequency | Fulfilled [ref. 1, p. 5] |
| 2.1.4 Bandwidth | Fulfilled [ref. 1, p. 9] |
| 2.1.5 Emphasis | Fulfilled [ref. 1] |
| 2.1.6 Tandem capability | Fulfilled [ref. 2] |
| 2.1.7 Post processing | Not demonstrated |
| 2.2 Coding delay | Fulfilled(1) [ref. 1, p. 7] |
| 2.3 Error resilience | Fulfilled [ref. 1, p. 15] |
| 2.4 Recovery time | Fulfilled [ref. 1, p. 15] |
| 3.1 Bit rate and coding | Fulfilled(2) [ref. 1, p. 6] |
| 3.2 Composite coding | Fulfilled [ref. 1] |

(1) To facilitate operation with television sound, the encode or decode delay is identical to a corresponding video frame rate (1/24, 1/25, 1/30 s). Access units correspond to video frames.

(2) The bit rate/channel is 250 kbit/s in order to obtain the advantages indicated in the first, third, and fourth bullets under § 3.1.

## References

[1] FIELDER, L. D., LYMAN, S. B., VERNON, S. and TODD, C. C. [September 1999] Professional audio coder optimized for use with video. 107th AES Convention, New York, NY, United States of America.

[2] GRANT, D., DAVIDSON, G. and FIELDER, L. [21-24 September 2001] Subjective evaluation of an audio distribution coding system. 111th AES Convention, New York, NY, United States of America.

# Annex 2

# Requirements for emission

The audio coding systems for digital television and sound broadcasting for emission applications should fulfil the requirements listed below.

## 1 Service requirements

### 1.1 Channel configurations

For audio services the following channel configurations should be supported according to the needs of applications (see Recommendation ITU-R BS.775):

| No. of channels | Channel configuration | Channel assignment |
|---|---|---|
| 1 channel | 1/0 | Mono |
| 2 channels | 2/0 | Left, right |
| 3 channels | 3/0<br>2/1 | Left, right, centre<br>Left, right/surround |
| 4 channels | 3/1<br>2/2 | Left, right, centre/surround<br>Left, right/surround left, surround right |
| 5 channels | 3/2 | Left, right, centre/surround left, surround right |

together with an optional low frequency effects (LFE) channel.

### 1.2 Audio services

Together with a main audio service, the following associated audio services can be provided according to the needs of applications:

– a multilingual service – consisting of one or more independent channels used to distribute a programme with commentary in one or more languages,

– audio services for the hearing and visually impaired – the service for the visually impaired usually contains a vocal description of the picture content while the service for the hearing impaired would contain the clean dialogue without, or with a lower level of, music and special effects to improve the intelligibility of the speech,

– ancillary data – to convey various types of information including: dynamic range control, loudness control and user data (Recommendation ITU-R BS.775).

The various services can be grouped as:

– *Main service* (every channel of a main service is assigned to the same programme, including the optional LFE channel).
– *Extended service(s)*, which could be:
  – *Independent services* (for additional programmes which are independent of the main service programme, such as commentary, or other services containing two or more channels; channel configurations can be chosen according to the table in § 1.1).
  – *Alternative services* (for programmes which are intended to replace one or more of the main service channels, such as multilingual, hearing impaired).
  – *Additional services* (containing channels to be added to channels of the main service, such as commentary, or additional channels for enhanced sound systems as 3D TV).

As any transmission system should include a system layer able to perform multiplexing operations, it is not required that all the audio services listed above be conveyed by a single bit stream.

### 1.3 Flexible allocation of channels

A bit stream should provide identification data for signalling and controlling of the sound configurations. The transmission system must provide the ability to switch dynamically among any of the channel configurations listed in § 1.1.

### 1.4 Ancillary data

The audio coding system should provide for the possibility of transmission of ancillary data. The ancillary data can convey various types of information, including dynamic range control, loudness control and user data.

## 2 Performance requirements

### 2.1 Audio quality

Two categories of audio quality are assumed for emission applications as shown in Annex 3. These are high-quality ("CD quality") emission and intermediate quality emission.

Audio quality is characterized by several parameters, particularly audio coding methods, sampling rates and bit rates. Required bit rates to satisfy the required audio quality are dominated by audio coding methods and sampling rates.

#### 2.1.1 Basic audio quality

##### 2.1.1.1 High-quality emission

The broadcaster typically has the ability to trade off audio quality against the bit rate applied to audio. Ideally, the quality of the sound reproduced after decoding will be subjectively similar to the original signal for most types of audio programme material. Using the triple stimuli double blind with hidden reference test, described in Recommendation ITU-R BS.1116, this requires mean values consistently higher than 4 on the Recommendation ITU-R BS.1116 impairment 5-grade scale at the reference listening position. In practice, commercial requirements sometimes lead to operation with bit rates lower than that necessary to achieve this level of quality. However, the system should offer the broadcaster the option to operate at this level of quality.

NOTE 1 – The objective audio quality parameters for contribution/distribution can be incorporated later, conforming to Recommendation ITU-R BS.1387.

#### 2.1.1.2 Intermediate quality emission

In some emission applications, audio quality below "CD quality" but equivalent to or better than good reception of FM or AM analogue broadcasting services may be required. Using the MUSHRA method described in Recommendation ITU-R BS.1534, the mean score corresponding to "excellent" or "good" grade may be required. Low-pass filtered versions of unprocessed audio signals used as anchors in the test may also be used, as these represent the audio quality of the existing analogue sound broadcasting systems.

### 2.1.2 Stereo image quality

In the case of two-channel stereophonic or multichannel configurations, the quality of the sound image of source material should be preserved. For the configurations which include a centre channel (3/0, 3/1, 3/2) the directional stability of the frontal sound image should be maintained within reasonable limits over a listening area larger than that provided by conventional two-channel stereophony. For the configurations including surround (2/1, 2/2, 3/1, 3/2) the sensation of spatial reality (ambience) should be significantly enhanced over that provided by conventional two-channel stereophony (Recommendation ITU-R BS.775).

### 2.1.3 Quantization resolution

The required resolution should be at least 16 bits.

### 2.1.4 Sampling frequency

#### 2.1.4.1 High-quality emission

In agreement with Recommendation ITU-R BS.646, the sampling frequency should be 48 kHz.

#### 2.1.4.2 Intermediate quality emission

The use of lower sampling frequencies than 48 kHz should be permitted when "CD quality" is not required. In agreement with Recommendation ITU-R BS.1196, the sampling frequency should be either 32 kHz or 48 kHz. Considering further that perceived audio quality for very low bit rates is improved by the use of a reduced sampling frequency and that MPEG-2 audio allows the use of lower sampling frequencies, namely, half sampling frequencies (16, 22.05 and 24 kHz) and quarter sampling frequencies (8, 11.025 and 12 kHz), lower sampling frequencies may be appropriate for intermediate quality emission.

### 2.1.5 Bandwidth

#### 2.1.5.1 High-quality emission

Main audio channels: 20-20 000 Hz.

LFE channel: 15-120 Hz.

#### 2.1.5.2 Intermediate quality emission

The bandwidth depends on the sampling frequency.

### 2.1.6 Emphasis

The audio coding system should not employ emphasis.

### 2.1.7 Post-processing capability

The post-processing capability required is strongly dependent on the application. For emission links, it can be restricted to equalization and dynamic range adjustment (e.g. to match the dynamic range of the programme material to that of the listening environment).

## 2.2 Coding delay

Coding delay for all channels in a programme must be identical. In case of television sound, the delay of audio must be matched with the delay of video.

## 2.3 Error resilience

A mechanism must be provided in the audio bit stream to allow the decoder to identify residual channel errors and to adopt proper concealment methods.

## 2.4 Recovery time

The recovery time should be as low as possible. For systems that provide Audio Access Units (AAUs), the recovery time should be within a few AAU, and ideally within a single AAU.

# 3 Functional and operational requirements for mutichannel systems

## 3.1 Compatibility with mono/stereo systems

### 3.1.1 Downward compatibility (Recommendation ITU-R BS.775)

A multichannel bit stream format must be decodable by classes of decoders of varying complexity. It must be possible in the decoder to arrange a presentation with a number of channels lower than the number of transmitted channels, according to the user reproduction capabilities, without impairment other than the loss of the stereo or multichannel localization effect.

Two methods have been identified which provide downward compatibility with low receiver complexity. The first requires the use of the matrix process. A low-cost receiver then only requires the A- and B-channels as in the case of the 2/0 system, i.e. a system which does not use a backwards compatibility matrix. The second method is applicable to the discrete 3/2 delivery system. The delivered signals are digitally combined using the equations, which enable the required number of signals to be provided. In the case of low bit rate source coded signals, the downward mixing of the 3/2 signals may be performed prior to the synthesis portion of the decoding process (where the bulk of the complexity lies).

### 3.1.2 Backward compatibility

This requirement applies in situations where an existing mono/stereo application must be upgraded to multichannel sound while services to existing receivers must be maintained. In systems that already employ mono or stereo, backward compatibility for multichannel low bit rate coding means that a decoder should properly decode basic stereo information, constituted by an appropriate down mix of the audio information from all source channels. To fulfil this requirement either the simulcast method or the matrixing method may be applied.

*Simulcast method*

One method is to continue providing the existing mono/stereo service and to add the new 3/2 channel service. This approach is referred to as a simulcasting operation. The advantage of this approach is that the existing mono/stereo service could be discontinued at some point in the future, and the 2/0 and 3/2 programme mixes may be independently optimized.

*Matrixing method*

Another method is the use of compatibility matrices in order to produce the wanted number of audio channels by a linear combination of the signals conveyed in the emission channels. The matrix equations may be used to provide compatibility with existing receivers. In this case, the existing left and right emission channels are used to convey the compatible A and B matrix signals. Additional emission channels are used to convey the T, $Q_1$, and $Q_2$ matrix signals. The advantage of this approach may be that less additional data capacity is required to add the new service.

### 3.1.3 Forward compatibility

For applications where the new multichannel system must coexist with the mono/stereo system, it may be required that decoders are able to decode the mono/stereo audio bit stream.

## 3.2 Bit rate

Recommendation ITU-R BS.1196 specifies required bit rates for a stereo signal for high-quality emission application. Thus two and a half times the bit rate (i.e. 5/2 × 144 kbit/s through 5/2 × 256 kbit/s) can be considered an upper limit for the five-channel main service in case that backward compatibility (see § 3.1.2) is not required. As the composite coding techniques would provide an additional coding gain, obvious reduction of bit rates should be achieved by new multichannel coding systems for the audio quality defined in § 2.1.

## 3.3 Decoder complexity

The decoder for the audio programme should be of not unduly high complexity so that the decoder cost may be kept low. In the case where a smaller number of channels, *M*, is to be reproduced from an audio programme containing *N* channels, the decoder complexity should be smaller than the complexity of the complete *N* channel decoder.

# Appendix 1
# to Annex 2

## Information about coding systems that have been demonstrated to meet quality, and other, user requirements for emission

Table 2 and Table 3 list, in the left hand column, the requirements for high-quality emission and intermediate quality emission, respectively, specified in Annex 2. Other columns (of which four exist at this time) show the ability of specific codecs to meet these requirements. It is anticipated that future revisions to this Recommendation will contain additional information about additional codecs.

TABLE 2

**High-quality emission**

| List of requirements from Annex 2 | AAC LC profile | AC-3 | MPEG-2 Layer II |
|---|---|---|---|
| 1.1 Channel configurations | Fulfilled | Fulfilled | Fulfilled |
| 1.2 Audio services | Fulfilled | Fulfilled | Fulfilled |
| 1.3 Flexible allocation of channels | Fulfilled | Fulfilled | Fulfilled |
| 1.4 Ancillary data | Fulfilled | Fulfilled | Fulfilled |
| 2.1.1 Basic audio quality | Fulfilled at 144 kbit/s per 2 channels [1] | Fulfilled at 192 kbit/s per 2 channels [1] | Fulfilled at 256 kbit/s per 2 channels [1] |
| 2.1.2 Stereo image quality | Fulfilled | Fulfilled | Fulfilled |
| 2.1.3 Quantization resolution | Fulfilled | Fulfilled | Fulfilled |
| 2.1.4 Sampling frequency | Fulfilled | Fulfilled | Fulfilled |
| 2.1.5 Bandwidth | Fulfilled | Fulfilled | Fulfilled |
| 2.1.6 Emphasis | Fulfilled | Fulfilled | Fulfilled |
| 2.1.7 Post processing | Not demonstrated | Not demonstrated | Not demonstrated |
| 2.2 Coding delay | Fulfilled(1) | Fulfilled(1) | Fulfilled(1) |
| 2.3 Error resilience | Fulfilled | Fulfilled | Fulfilled(2) |
| 2.4 Recovery time | Fulfilled | Fulfilled | Fulfilled |
| 3.1.1 Downward compatibility | Fulfilled | Fulfilled | Fulfilled |
| 3.1.2 Backward compatibility | Fulfilled by simulcast method | Fulfilled by simulcast method | Fulfilled by matrixing method |
| 3.1.3 Forward compatibility | Fulfilled by dual decoders | Fulfilled by dual decoders | Fulfilled |
| 3.2 Bit rate | Fulfilled | Fulfilled | Fulfilled |
| 3.3 Decoder complexity | Fulfilled | Fulfilled | Fulfilled |

(1) The inherent coding delay is sufficiently low that applications may readily match the video and audio delays.

(2) Some error resilience is provided in the Layer II elementary stream and additional resilience is typically provided by the application.

TABLE 3

**Intermediate quality emission**

| List of requirements from Annex 2 | HE-AAC | | | HE-AAC v2 |
|---|---|---|---|---|
| 1.1 Channel configurations | Fulfilled | Fulfilled | Fulfilled | Fulfilled |
| 1.2 Audio services | Fulfilled | Fulfilled | Fulfilled | Fulfilled |
| 1.3 Flexible allocation of channels | Fulfilled | Fulfilled | Fulfilled | Fulfilled |
| 1.4 Ancillary data | Fulfilled | Fulfilled | Fulfilled | Fulfilled |
| 2.1.1 Basic audio quality | Fulfilled (Excellent) at 48 kbit/s per 2 channels [2], [4] | Fulfilled (Good) at 32 kbit/s per 2 channels [2], [4] | Fulfilled (Good) at 24 kbit/s per 1 channel [3] | Fulfilled (Good) at 24 kbit/s per 2 channels [2] |
| 2.1.2 Stereo image quality | Fulfilled | Fulfilled | Fulfilled | Fulfilled |
| 2.1.3 Quantization resolution | Fulfilled | Fulfilled | Fulfilled | Fulfilled |
| 2.1.4 Sampling frequency | Fulfilled | Fulfilled | Fulfilled | Fulfilled |
| 2.1.5 Bandwidth | N/A | N/A | N/A | N/A |
| 2.1.6 Emphasis | Fulfilled | Fulfilled | Fulfilled | Fulfilled |
| 2.1.7 Post processing | Not demonstrated | Not demonstrated | Not demonstrated | Not demonstrated |
| 2.2 Coding delay | Fulfilled(1) | Fulfilled(1) | Fulfilled(1) | Fulfilled(1) |
| 2.3 Error resilience | Fulfilled | Fulfilled | Fulfilled | Fulfilled |
| 2.4 Recovery time | Fulfilled | Fulfilled | Fulfilled | Fulfilled |
| 3.1.1 Downward compatibility | Fulfilled | Fulfilled | Fulfilled | Fulfilled |
| 3.1.2 Backward compatibility | Fulfilled by simulcast method | Fulfilled by simulcast method | Fulfilled by simulcast method | Fulfilled by simulcast method |
| 3.1.3 Forward compatibility | Fulfilled by dual decoders | Fulfilled by dual decoders | Fulfilled by dual decoders | Fulfilled by dual decoders |
| 3.2 Bit rate | Fulfilled | Fulfilled | Fulfilled | Fulfilled |
| 3.3 Decoder complexity | Fulfilled | Fulfilled | Fulfilled | Fulfilled |

N/A: Not applicable.

(1) The inherent coding delay is sufficiently low that applications may readily match the video and audio delays.

## References

[1] GRANT D., DAVIDSON, G. and FIELDER, L. [21-24 September 2001] Subjective evaluation of an audio distribution coding system. 111th AES Convention, New York, NY, United States of America.

[2] ISO/IEC JTC 1/SC 29/WG 11 N6009 [October, 2003] Report on the Verification Tests of MPEG-4 High Efficiency AAC.

[3] ISO/IEC JTC 1/SC 29/WG 11 N7137 [April, 2005] Listening test report on MPEG-4 High Efficiency AAC v2.

[4] KOMORI, T, SUGIMOTO, T. and KUROZUMI, K. [2005] AAC + SBR Audio coding quality used for the mobile digital terrestrial broadcasting. Proc. Spring meeting of the Acoustical Society of Japan.

# Annex 3

## Categories of audio quality for broadcasting applications

The following three categories of audio quality are assumed for broadcasting applications.

| **Category** | **Audio quality** | **Application** |
|---|---|---|
| (1) | Very high quality, with sufficient quality margin to allow cascade (concatenation) and post-processing | Contribution, distribution, production, and post-production |
| (2) | Subjectively transparent quality, sufficient for the highest quality broadcasting | High-quality ("CD quality") emission |
| (3) | Equivalent to or better than good FM service quality, or equivalent to or better than good AM service quality | Intermediate quality emission |