MISB MOTION IMAGERY STANDARDS BOARD	MISB EG 1001
Engineering Guideline 1001: Audio Encoding in MPEG-2 TS	**20 May 2010**

1 Scope

Many motion imagery systems incorporate sound features, and each year brings a greater diversity. These systems record audio information along with video and metadata, while still other systems may index, edit, or playback these data streams. There are a wide variety of audio formats in use. This document serves to tailor the use of external audio standards, for the purpose of archival storage of audio, particularly recorded speech, together with video in an MPEG-2 transport stream.

2 Normative References

[1] ISO 13818-1:2007, *Information technology - Generic coding of moving pictures and associated audio information – Part 1: Systems*, 16 October 2007

[2] ISO 11172-3:1993, *Information technology - Coding of moving pictures and associated audio for digital storage media at up to about 1.5 Mbit/s - Part 3: Audio*, 1 August 1993

[3] ISO 13818-3:1994, *Information technology - Generic coding of moving pictures and associated audio information: Audio*, 11 November 1994

[4] ISO 13818-7:2004, *Information technology - Generic coding of moving pictures and associated audio information - Part 7: Advanced Audio Coding (AAC)*, 15 October 2007

3 Informative References

[5] ITU-T Recommendation P.800, *Methods for subjective determination of transmission quality*, August 2006

[6] ITU-T Recommendation P.862, *Perceptual evaluation of speech quality (PESQ): An objective method for end-to-end speech quality assessment of narrow-band telephone networks and speech codecs*, February 2001

[7] ITU-T Recommendation P.862.1, *Mapping function for transforming P.862 raw result scores to MOS-LQO*, November 2003

4 Acronyms

AAC	Advanced Audio Coding
AAC-LC	Advanced Audio Coding Low Complexity

MPEG	Motion Picture Experts Group
PNS	Perceptual Noise Substitution
PS	Parametric Stereo
SBR	Spectral Band Replication

5 Introduction

There are many audio formats and codecs in use today. A subset of them has been standardized for use with MPEG-2 Transport Streams. Because of the importance of the MPEG-2 Transport Stream[1] as a standard container format and the associated audio format constraints, we will restrict our attention to this set of options.

In general, a standards document with a higher degree of constraint will tend to benefit interoperability. On the other hand, more constraints hinder the ability of implementers to innovate with new features in new products. This document delineates a list of accepted audio formats, which are in common enough use to promote interoperation among systems, while still allowing room for manufacturers to distinguish themselves through innovation.

Through technical exchange with members of our community, we have arrived at a list of formats, which offers a good balance between ease of use, breadth of use, cost, and state of the art technology. This list is summarized in Table 1.

6 Audio Formats

Encoder manufacturers may use any of the audio formats contained within Table 1.

Format Name	Defining Document
MPEG-1 Layer II	ISO 11172-3:1993 [2]
MPEG-2 Layer II	ISO 13838-3:1998 [3]
MPEG-2 AAC-LC	ISO 13818-7:2004 [4]

Table 1 – Audio Formats

6.1 MPEG-1 Layer II

This format provides support for legacy systems. Newer formats are expected to deliver higher performance at an equivalent data rate.

When using MPEG-1 Layer II audio, it is recommended to set protection_bit = 0 and include the crc_check in the audio header. See ISO 11172-3, sections 2.4.1 and 2.4.3 for details.

It is recommended to use constant bitrate audio streams when working with MPEG-1 Layer II.

6.2 MPEG-2 Layer II

This format is similar to MPEG-1 Layer II, with additional channels (surround sound support) and allowed data rates. This format also provides support for legacy systems. Newer formats are expected to deliver higher performance at an equivalent data rate.

For simplicity and to promote interoperability, the use of the ISO 13818-3 multilingual and/or multi-channel audio features are not recommended.

It is recommended to use constant bitrate audio streams when working with MPEG-2 Layer II.

6.3 MPEG-2 AAC-LC

MPEG-2 Part 7, Advanced Audio Coding (AAC) is the most advanced format of the types available here, and preferred for new systems.

When using MPEG-2 AAC-LC, it is recommended to set protection_absent = 0 in the adts_fixed_header, and include the corresponding error checking information. See ISO 13818-7, sections 6.2 and 8.1 for details.

7 Interoperability

The introduction of AAC allowed greater performance, and also broke from the traditional decoder models used in MPEG-1 and MPEG-2 Layers I & II. As a result, the AAC and Layer II formats are not interoperable. Table 2 summarizes the compatibility of various data formats with corresponding decoder types.

Table 2 - Interoperability Summary

		Decoder Type		
		MPEG-1 Layer II	MPEG-2 Layer II	MPEG-2 AAC-LC
Data Format	MPEG-1 Layer II	Ok	Ok	No
	MPEG-2 Layer II	Partial[1]	Ok	No
	MPEG-2 AAC-LC	No	No	Ok

1 An MPEG-1 decoder will be able to decode and play stereo sound from an MPEG-2 encoded data stream. If more MPEG-2 audio channels are present (e.g. center, left surround, and right surround) these would **not** be playable by an MPEG-1 decoder. At the transport stream level, MPEG-1 audio is encoded with stream_type 0x03 while MPEG-2 audio is encoded with stream_type 0x04.

It is noted that many systems offer the capability to decode several distinct audio formats. To guarantee interoperability, a decoder must be able to accept and decode each of the allowed audio data formats.

8 Data Rate

This section provides informative guidance for the selection of data rates when compressing audio data. It is not a normative part of this document.

Depending on the mode, MPEG-1 Layer II allows data rates at discrete intervals between 32 and 384 Kbits/sec. MPEG-2 Layer II includes the data rates of MPEG-1 Layer II, and also allows optional extension to rates as low as 8 Kbits/sec. Constant bitrates are recommended when using Layer II audio.

MPEG-2 AAC uses a 23-bit number to encode the data rate, which theoretically allows a rate between 1 bit/sec and approximately 16.8 Mb/sec. Most encoders will place more practical restrictions on this range–a range somewhat closer to that of the Layer II rates.

8.1 Speech Encoding

One application of MPEG-compatible audio in the NSG is the archival of recorded speech, either for *in situ* mission narration, or after-the-fact analysis. Recording fidelity is important to the value of an archive. At the same time, recording data rate impacts the storage requirements–and therefore cost. Speech has unique sonic characteristics. Industry groups have put a great deal of effort into the characterization of voice transmission systems, resulting in standard metrics such as ITU Recommendation P.862[6,7].

P.862 defines the Perceptual Evaluation of Speech Quality (PESQ) metric. PESQ is used to measure the quality of telephone systems in addition to speech codecs. This metric compares a transmitted (possibly degraded) speech recording to a reference. Distortions due to channel errors, low data rate, filtering, variable delays and other factors are measured and combined into a score. The PESQ-MOS (Mean Opinion Score) score is calibrated to fit data from multiple subjective surveys, where people are asked to rate quality on a five point scale (Excellent = 5, Good = 4, Fair = 3, Poor = 2, Bad = 1). The PESQ metric, which correlates well with experimental data, is useful as a predictor of quality as measured by the Mean Opinion Score.

When speech sample clips are compressed at a given data rate, then decompressed, the resulting sound file may be compared against the original to measure the loss of quality as a function of the compression. MPEG-2 Layer II and MPEG-2 AAC results are plotted in Figure 1 below:

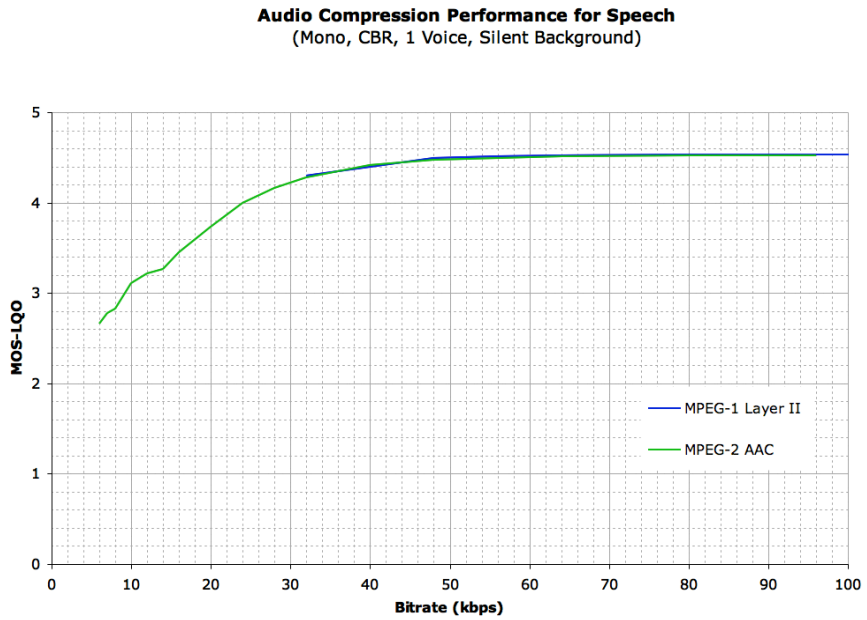

Figure 1 - Loss of Quality as a function of compression Data Rate for MPEG-2 Layer II and MPEG-2 AAC

As can be seen from Figure 1, compression with these two codecs has virtually no measurable effect on quality down through 48 Kb/sec. Even at 32 Kb/sec, the effect on the quality score is negligible. Below 32 Kb/sec, however, the sound quality falls off noticeably.

To allow efficient encoding while maintaining adequate quality, a minimum data rate of 32 Kb/sec is recommended with 48 Kb/sec or greater preferred.