Comments from Fraunhofer IIS on the Draft version of the MDA and IDA specification for the 2nd generation Digital Terrestrial Television Broadcasting System (DVB-T2) - Integrated Receiver Decoder T2

Submitter: **Fraunhofer Institute for Integrated Circuits IIS**

Address: **Am Wolfsmantel 33**
91058 Erlangen
Germany

Author: **Stefan Meltzer**
E-Mail: **stefan.meltzer@iis-extern.fraunhofer.de**
Date: **August 10, 2012**

Please direct all communication to the author.

Fraunhofer
IIS

Table of Contents

Summary of Major Points

Fraunhofer IIS main interest is on the audio section of the specification. Therefore our comments will focus on the audio specific topics.

In our understanding, the specification shall allow a cost effective and spectrum efficient implementation of the DVB-T2 system and the new receivers. Based on this assumption we agree with the selection of MPEG-4 HE-AAC v2 as the main audio codec, since it provides the most spectrum efficient solution. On the receiver side we would like to propose the following changes:

- Definition of a base line receiver with only stereo output and an optional extension with multichannel capabilities
- Implementation of receiver mix mode for the audio description for the advanced audio coding schemes MPEG-4 HE-AAC and Enhanced AC-3
- Reference to the upcoming version of the ETSI Specification 101 154 – Specification for the use of Video and Audio Coding in Broadcasting Applications based on MPEG-2 Transport Stream

The first proposed change will lead to a cost reduction for the majority of receivers since most of the receivers will reproduce the audio only over a stereo system. The cost reduction is realized by lower patent licensing cost and a reduced testing effort.

The implementation of a receiver mixed mode for audio description reduces the amount of transmitted data significantly, especially in case of multichannel audio transmissions. This feature is implemented for the advanced audio codec systems mentioned in the specification, namely MPEG-4 HE-AAC and Enhanced AC-3. It is already in use in various systems worldwide. Therefore no additional cost or delay in implementation is to be expected from the use of this mode.

The upcoming version of the ETSI Specification 101 154 contains a number of important clarifications for the use of MPEG-4 HE-AAC. It is currently under review at ETSI and according to the ETSI work schedule the publication is expected on October 9th 2012.

More details on the proposed changes can be found in the following sections. Beside the points mentioned in this section, we will also mention in a separate section some minor issues like typos and little inconsistencies in the draft specification.

Statement of Interest

Fraunhofer IIS is working for more than 20 years in the field of audio coding. The name of Fraunhofer IIS is strongly connected with the invention of MPEG Layer-3, better known as mp3. Fraunhofer IIS is also one of the main contributors of MPEG AAC and its variant HE-AAC. Since many years, Fraunhofer IIS is also a member of DVB. Within DVB Fraunhofer IIS is contributing to the Audio Video Coding groups of the commercial and technical module. Fraunhofer IIS is also working closely with key players in the broadcast industry on both, encoder and decoder side, to enable the use of MPEG-4 AAC and HE-AAC in broadcast systems.

Fraunhofer IIS' interest is to use its expertise in audio coding for broadcasting applications to support IDA and MDA in defining the audio section of the DVB-T2 Integrated Receiver Specification

Comments

Basic requirements for audio coding

Fraunhofer IIS would recommend dividing the section describing the audio decoding into a basic requirement part and an optional multichannel extension. This would allow manufacturers to offer cheaper receiver with stereo only output. Since the vast majority of consumers will use the build in speakers of the TV set or stereo audio systems, it is justified to define a stereo only receiver without the necessary additional cost in licensing and testing required for the multichannel support. However, we would strongly recommend specifying for MPEG-4 HE-AAC v2 the use of level 4 instead of level 3. Level 3 is a very uncommon and seldom used level. It offers only stereo decoding capabilities. MPEG-4 HE-AAC v2 level 4 offers multichannel decoding capabilities. In combination with a mandatory downmix to stereo, the license fee for the use of MPEG-4 HE-AAC v2 level 4 is identical to the license fee for level 3 or level 2. (For details on the MPEG-4 AAC /HE-AAC licensing see: http://www.vialicensing.com/licensing/aac-overview.aspx) Specifying a MPEG-4 HE-AAC v2 level 4 decoder for the basic stereo receiver offers the opportunity to transmit only one multichannel MPEG-4 HE-AAC audio bit stream to stereo and multichannel receivers instead of a simulcast of a stereo and a multichannel bit stream. Therefore this combination allows a more efficient use of the spectrum. On the implementation side, all chipsets used today to build STBs and TV sets with MPEG-4 AVC video decoding capabilities are also able to decode MPEG-4 HE-AAC v2 level 4 including downmix.

Fraunhofer IIS would also recommend allowing for MPEG-4 HE-AAC v2 a sampling rate of up to 48 kHz instead of limiting it to exactly 48 kHz. The lower sampling rates allow especially in combination with the version 2 of HE-AAC to transmit audio a very low data rates with still a good audio quality. This is of interest for radio programs or mobile applications transmitted alongside of the TV signal over the DVB-T2 network. Since the required computational complexity depends on the highest required sampling frequency, adding lower sampling frequencies doesn't add any cost to the chipset implementation and is already supported by the chipsets.

For the receiver output we would recommend to add for the stereo receiver the pass through of HE-AAC bit stream for optical/coaxial output as well as for the HDMI output. The pass through would allow to make use of a multichannel signal with an external decoder e.g. an AV Receiver. Adding the pass through capabilities does not increase any license fee.

For the multichannel receivers we would also recommend adding the bit stream pass through capabilities for MPEG-4 HE-AAC and Enhanced AC-3 for the optical/coaxial output.

For consistency, it is necessary to add the transcoding of MPEG-4 HE-AAC multichannel audio into DTS also to the transcoding sub section. The transcoding of MPEG-4 HE-AAC multichannel into DTS offers an alternative to the transcoding into AC-3.

Based on the above comments, Fraunhofer IIS would propose the following alternative text for the audio decoding section: (All changes are to the original text are highlighted. Additions are marked in red)

4.3 Audio Decoding

4.3.1 Basic Audio Decoding Requirements

4.3.1.1 MPEG-1 Layer II Audio

4.3.1.1.1 The IRD-T2 shall comply with the implementation guidelines outlined in § 6.1 of ETSI TS 101 154 [1] for decoding MPEG-1 Layer II audio, based on ISO/IEC 13813-3 [7] in single (mono), dual, joint stereo and stereo modes with bit rates between 64 kbit/s and 256 kbit/s, and sampling rates of 32 kHz, 44.1 kHz and 48 kHz.

4.3.1.1.2 The IRD-T2 shall provide convenient user control for appropriate audio output format switching between different audio modes.

4.3.1.2 MPEG-4 HE AAC Version 2 Audio (up to Level ~~3~~4)

The IRD-T2 shall comply with the implementation guidelines outlined in § 6.4 and Annex C5 of ETSI TS 101 154 [1] for decoding MPEG-4 HE AAC version 2 audio up to Level ~~3~~ 4 and sampling rates of up to 48 kHz, based on ISO/IEC 14496-3 [8]. The IRD-T2 shall apply bit-stream metadata parameters and downmix multichannel input configurations to stereo PCM.

4.3.1.3 Decoding & Presentation Options for 2 Channels of Decoder Output

Codec	Analogue Output / Speaker (IDTV)	Optical / Coaxial (SPDIF)	HDMI*
MPEG-1 Layer II	Mono / Stereo	PCM stereo	PCM stereo
MPEG-4 HE AAC	Mono / Stereo	PCM stereo Pass through of HE-AAC bit stream	PCM stereo Pass through of HE-AAC bit stream

Note: * Only applicable to IRD-T2 with HDMI output

Table 4: Presentation Options for 2 Channels of Decoder Output

4.3.~~4~~2 Optional Multi-Channel Audio

4.3.~~4~~2.1 Format

The IRD-T2 shall identify, accept and decode input bit-streams in the following formats:

a) Enhanced AC-3 as specified in ETSI TS 102 366 [9]; and

b) MPEG-4 HE AAC version 2 Level 4 as defined in ISO/IEC 14496-3 [8]

4.3.42.2 Metadata

The IRD-T2 shall apply bit-stream metadata parameters and downmix multichannel input configurations to stereo PCM for Enhanced AC-3 and MPEG-4 HE AAC in accordance with guidelines given in ETSI TS 101 154 [1], and as specified in ETSI TS 102 366 [9] and ISO/IEC 14496-3 [8] respectively.

4.3.42.3 Pass-through

The IRD-T2 shall pass through the native input bit-stream over the HDMI (and ARC on compatible HDMI inputs) output as well as the optical/coaxial output.

4.3.42.4 Trans-coding

The IRD-T2 shall transcode audio and metadata from E-AC-3 ~~and MPEG-4 HE AAC~~ input bit-streams to AC-3 output bit-streams at a data rate of 640kbps. The IRD-T2 shall transcode audio and metadata from input bit-streams to an AC-3 output bit-streams at a data rate of 640kbps or alternatively to a DTS output bit stream at a data rate of 1.5 Mbps. The number of channels on the output AC-3 or DTS bit-stream shall be equal to ~~or greater than~~ the number of channels contained within the input bit-stream. The AC-3 or DTS bit-stream shall be provided over the following outputs:

a) S/PDIF; and

b) HDMI (and ARC on compatible HDMI inputs)

4.3.42.5 Decoding & Presentation Options for Multi-Channel Decoder

Codec	Analogue Output / Speaker (IDTV)	Optical / Coaxial (SPDIF)	HDMI*
Enhanced AC-3 multi-channel	Down-Mixed Stereo	Trans-code to AC-3 bit-stream and pass through Pass through of E AC-3 bit-stream PCM stereo	Pass through of E AC-3 bit-stream Trans-code to AC-3 bit-stream and pass through PCM stereo and multi-channel
MPEG-4 HE-AAC multi-channel	Down-Mixed Stereo	Trans-code to AC-3 or DTS bit-stream and pass through Pass through of HE-AAC bit-stream PCM stereo	Pass through of HE-AAC bit-stream Trans-code to AC-3 or DTS bit-stream and pass through PCM stereo and multi-channel

Note: * Only applicable to IRD-T2 with HDMI output

Table 5: Presentation Options for Multi-Channel Decoder

4.3.~~5~~3 Audio handling when changing service or audio format

The IRD-T2 should gracefully handle change of service or audio format at the audio outputs without significant disturbances to the end user.

4.3.~~6~~4 Lip-Sync

For all supported formats of audio, the maximum timing misalignment between audio and video in reference with Program Clock Reference (PCR) - time stamp carried by Video shall be confined within ±20 ms.

4.3.~~7~~5 Loudness Matching

The IRD-T2 shall apply format dependent attenuation to decoded stereo PCM audio, in order to achieve loudness alignment between different input formats.

Audio Description

The audio description service is a service for visually impaired people and consists of an additional audio track, which describes the scene on the TV screen. The additional audio track is mixed with the standard audio before the audio is played back over the speaker. This mixing can either happen at the broadcaster side or at the receiver side. In case of the mixing at the broadcaster side, two complete audio tracks need to be transmitted, one with and one without audio description mixed in. In case of mixing on the receiver side, the receiver receives the standard audio and in addition an audio description track. Based on the receiver settings, the receiver can mix the audio description track with the standard audio. The mixing process can be steered by metadata transmitted alongside with the audio description track. Since the audio description track is in general a mono track, the receiver mix mode is more efficient than the broadcast mix mode. Due to the limited processing capabilities in early DVB receivers, most system nowadays still use the broadcast mix mode. Newer systems based on DVB-T2 and modern audio codecs make use of the receiver mix mode. A prominent example is the Freeview HD service of the BBC. Therefore Fraunhofer IIS would recommend specifying the receiver mix mode in combination of the use of MPEG-4 HE-AAC or Enhanced AC-3 as transmission codec.

The corresponding section of the specification could be changed as follows:

1.4 This Specification also requires that the IRD-T2 be capable of decoding Singapore's FTA DVB-T and DVB-T2 broadcast of television, radio and enhanced services (§ 4). This shall include the following capabilities:

a) Subtitling (where available and selected by viewer);

b) Audio Description in Broadcast-mix mode;in combination with MPEG-4 HE-AAC and Enhanced AC-3 Audio Description in Receiver-mix mode according to Annex E of ETSI 101 154 [1]

c) Electronic Program Guide;

d) Teletext;

e) Parental Lock Feature; and

f) Multiple Audio Selection

Latest Revision of ETSI 101 154 Specification

The latest revision of the ETSi 101 154 specification is currently under review at ETSI. This new revisions includes a major rework of the metadata section for MPEG-4 HE-AAC. The changes include clarifications, explanations and removal of ambiguities. The current schedule for the work item within ETSI states as publication date October 9th 2012 (see: http://webapp.etsi.org/WorkProgram/Report_Schedule.asp?WKI_ID=39522). The document is already publicly available as Blue Book on DVB web site (http://dvb.org/technology/standards/a157_DVB-AVC-MPEG2.pdf). Therefore Fraunhofer IIS strongly recommends to already reference to this new version of ETSI 101 154.

The reference would be changed as follows:

[1] ETSI TS 101 154 v1.1~~01~~.1 (~~2011-06~~ 2012-10) Digital Video Broadcasting (DVB); Specification for the use of Video and Audio Coding in Broadcasting Applications based on the MPEG-2 Transport Stream

Interfaces and Connectors

There is are inconsistencies between the requirements defined in section 4 " Audio Decoding" and section 8.9 "Digital Audio Data Stream Output". To avoid this inconsistencies, Fraunhofer IIS proposes to reference to the section 4 instead of listing the capabilities again in section 8.9.

The new proposed text is as follows:

8.9 Digital Audio Data Stream Output (Optional)

It is optional for the IRD-T2 to provide an S/PDIF digital audio output – electrical (coaxial) or optical (TOSLINK). The capabilities of this interface are defined in section 4 of this document. ~~This digital interface may carry PCM stereo audio and/or AC3 coded audio and/or MPEG1-Layer II audio streams to an external decoder multichannel sound system.~~

List of Abbreviations

We found one typo in the list and one missing abbreviation.

Typo:

HE-AAC High Efficiency Advanced Audio Coding

Missing:

DTS Digital Theater System

Conclusion

The audio section of the draft specification is already an excellent starting point. The guiding theme for Fraunhofer IIS' comments was to provide a cost effective solution for the users and an improved spectral efficiency as well as keep the solution future proof. The recommended changes will improve the specification in these points.

About Fraunhofer IIS

Fraunhofer IIS, based in Erlangen, Germany, has been working in compressed audio and digital broadcasting technology for more than 20 years and remains a leading innovator of technologies for cutting-edge multimedia systems. Fraunhofer IIS is the main inventor of mp3 and universally credited with the co-development of AAC (Advanced Audio Coding) as well as technologies for the media world of tomorrow, including MPEG Surround and data services like Journaline. In addition Fraunhofer IIS is active in the area of standardization, overall broadcast system design, receiver core development, and OEM broadcast server equipment. The technologies developed at Fraunhofer IIS have established themselves globally in satellite-based and terrestrial broadcasting systems, such as Digital Radio Mondiale DRM, DAB Digital Radio, Digital Video Broadcasting DVB, WorldSpace and Sirius XM Radio.

Through the course of more than two decades, Fraunhofer IIS has licensed its audio codec software and application-specific customizations to at least 1,000 companies. Fraunhofer estimates that it has enabled more than 5 billion commercial products worldwide using its mp3, AAC and other media technologies.

The Fraunhofer IIS organization is part of Fraunhofer-Gesellschaft, based in Munich, Germany. Fraunhofer-Gesellschaft is Europe's largest applied research organization and is partly funded by the German government. With nearly 20,000 employees worldwide, Fraunhofer-Gesellschaft is composed of 60 Institutes conducting research in a broad range of research areas. For more information, contact visit www.iis.fraunhofer.de/amm.