Audio Compliance with ISDB-T

Need and Implementation Challenges

Jun-2015

Abstract

This paper discusses the necessity of audio compliance with ISDB-T standard and how it is implemented.

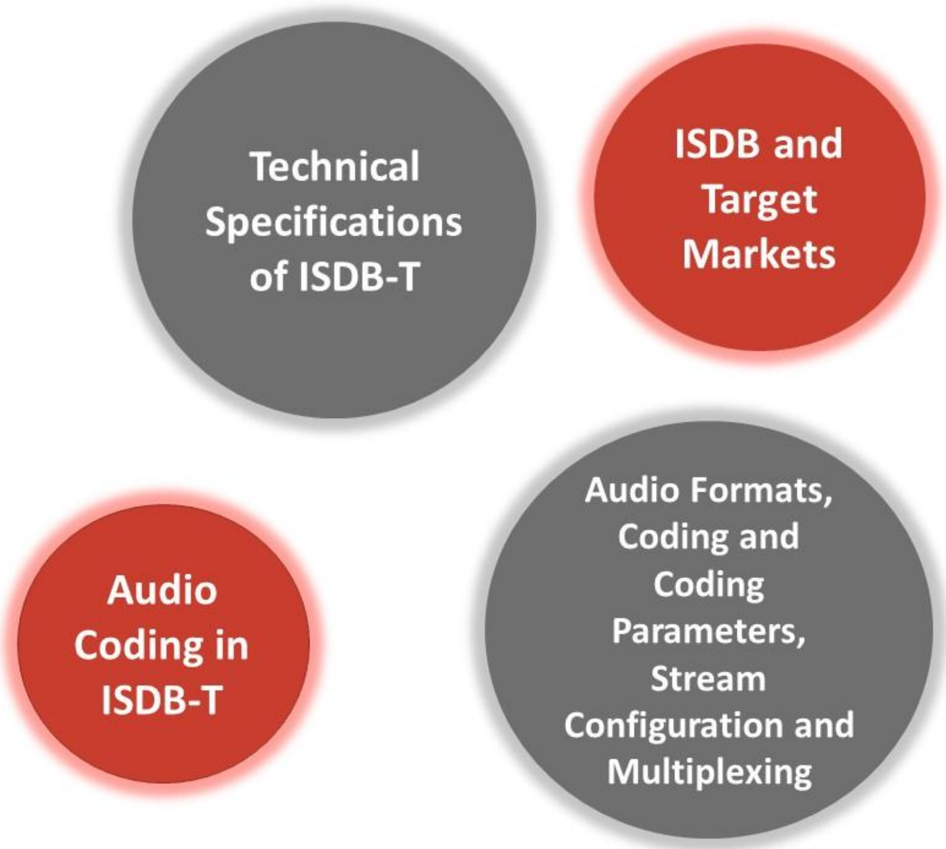

Figure 1 - Overview

Contents

Figures

Tables

Introduction

The **Integrated Services Digital Broadcasting** (ISDB) is a Japanese standard for digital television and digital radio used by Japan's radio and television networks. ISDB replaced the previously used MUSE Hi-vision analogue HDTV system.

ISDB is maintained by the Japanese organization ARIB (Association of Radio Industries and Businesses). The core standards of ISDB include the following:

- ISDB-T (terrestrial)
- ISDB-S (satellite television)
- ISDB-C (cable)
- 2.6 GHz band mobile broadcasting

All of these core standards are based on MPEG-2 or MPEG-4 standard for multiplexing in a transport stream (TS) structure and video and audio coding (MPEG-2 or H.264), and are capable of high definition television (HDTV) and standard definition television. ISDB-T and ISDB-Tsb are for mobile reception in TV bands. 1seg is the name of an ISDB-T service for reception on cell phones, laptop computers and vehicles.

ISDB-T International, a derivative of ISDB-T (Terrestrial) was developed by the Brazilian government, and is now being adopted in many South American countries.

The aim of this paper is to bring out the distinctions in ISDB-T, the Japanese terrestrial standard for digital television and radio, which are required for a regular MPEG-2 AAC (Low Complexity) Encoder to be ISDB-T compliant.

Target Markets

The ISDB-T recommendation was adopted in Japan, in December 2003, for commercial transmissions, and it currently comprises a market of about 100 million TV sets. Audio compliance with ISDB-T, standardized by ARIB is necessary for meeting the requirements of the Japanese digital television and radio broadcast formats. Apart from Japan, other Asian countries, such as Maldives, Philippines and Thailand are also adopting ISDB-T and are in the process of assessing trial broadcast runs.

Brazil, to move away from its current usage of an analogue TV system (PAL-M) to Digital Television format, has chosen ISDB-T as a base, calling it ISDB-Tb or internally SBTVD (Sistema Brasileiro de Televisão Digital - Terrestre). This was launched commercially in Brazil, in December 2007.

ISDB-Tb differs from the original ISDB-T standard by using H.264/MPEG-4 AVC as a video compression standard (ISDB-T uses H.262/MPEG-2 Part 2), a presentation rate of 30 frames per second in even portable devices (ISDB-T, One seg, uses 15 frame/s for portable devices) and powerful interaction using middleware Ginga, composed by Ginga-NCL and Ginga-J modules (while ISDB-T uses BML).

In January 2009, the Brazilian-Japanese study group for digital television finished and published a specification document, merging the Japanese ISDB-T with Brazilian SBTVD, resulting in a specification called ISDB-T International. ISDB-T International is the system that has been proposed by Japan and Brazil for other countries in South America and around the world.

Several countries in South America have already adopted ISDB-T International, whose audio encoding format is fundamentally based on ISDB-T, and have started broadcasting in digital format. These countries include Brazil, Uruguay, Peru, Argentina, Chile, Venezuela, Ecuador, Costa Rica, Paraguay and Bolivia. Countries such as Nicaragua and Guatemala are in the pre-implementation stage, while Belize is currently assessing this digital platform.

Technical Specifications of ISDB-T

Category	Sub-Category	Details
Transmission channel coding	Modulation	64QAM-OFDM, 16QAM-OFDM, QPSK-OFDM, DQPSK-OFDM (Hierarchical transmission)
	Error correction coding	Inner coding, Convolution 7/8,5/6,3/4,2/3,1/2 Outer coding: RS(204,188)
	Guard interval	1/32,1/16,1/8,1/4
	Interleaving	Time, Frequency, bit, byte
	Frequency domain multiplexing	BST-OFDM (Segmented structure OFDM)
Conditional Access		Multi-2
Data broadcasting		ARIB STD-B24 (BML, ECMA script)
Service information		ARIB STD-B10
Multiplexing		MPEG-2 Systems
Audio coding		**MPEG-2 Audio (AAC)**
Video coding		MPEG-2 Video
		MPEG-4 AVC /H.264*

Table 1 – Technical Specifications of ISDB-T

* H.264 Baseline profile is used in one segment (1seg) broadcasting for portables and Mobile phone.

* H.264 High profile is used in ISDB-Tb to high definition broadcasts.

With ISDB-T, one or more transport stream (TS) inputs, defined in ’MPEG-2 Systems‘, are re-multiplexed to create a single TS. This TS is then subjected to multiple channel-coding steps in accordance with the intentions of the service, and is finally sent as a single OFDM signal. ISDB-T also offers time interleaving to provide powerful channel coding for mobile-reception in which variations in field strength are inevitable.

The transmission spectrum of television broadcasting consists of 13 successive OFDM blocks (or ’OFDM segments‘), each bandwidth of which is equal to one fourteenth of a television-broadcasting channel. An OFDM-segment carrier configuration that allows connection of multiple segments makes it possible to provide a transmission bandwidth appropriate in terms of units of segment width for the target media, while at the same time enabling use of the same receiver for both ISDB-T and ISDB-Tsb.

Audio Coding in ISDB-T

The audio coding technique employed in ISDB-T conforms to the MPEG2 Advanced Audio Coding (AAC) standard ISO/IEC 13818-7 [2]. According to this ordinance, the input audio signal must comply with the following:

- The sampling frequency for audio signals shall be 32 kHz, 44.1 kHz, or 48 kHz.
- To configure stereophonic signals (consisting of two or more audio signals to achieve a three-dimensional reproduction of sound), the sampling timing for all signals shall be the same.
- The number of quantization bits for the input signal shall be 16 or more.
- The maximum number of audio input channels shall be five, in addition to the channel used to enhance low frequencies.

The specification also states that audio signals shall be coded by a combination of time-frequency transform coding (in which the input signal is transformed into frequency components by modified discrete cosine transform, and in which the amount of information is reduced by using the decrease in energy deviation of frequency components) and psychoacoustic weighted bit assignment (in which codes are weighted to minimize signal degradation in the frequency range that is readily perceived by humans). This can be shown in the form of a high level block diagram (Figure 2).

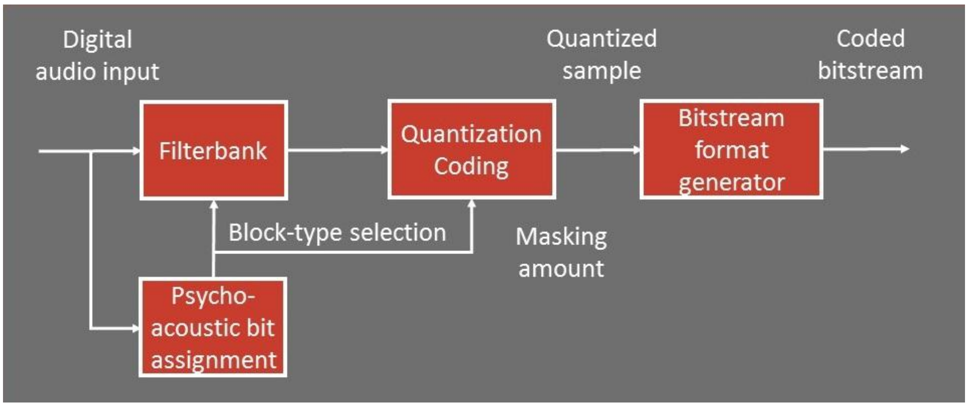

Figure 2 – Audio Compression and Transmission Procedures

The audio bit-stream configuration is depicted in Figure 3.

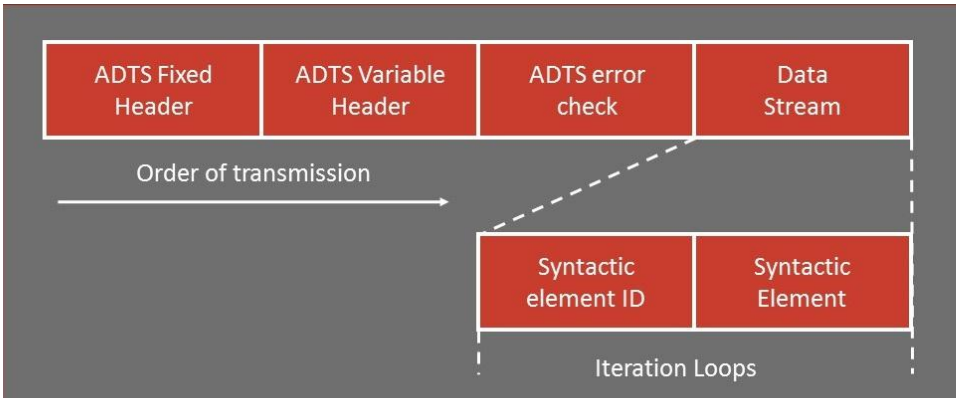

Figure 3 – Bit Stream Configuration

The ordinance states:

- The ADTS fixed header consists of synchronization and audio coded information defined in ISO/IEC 13818-7.
- The ADTS variable header consists of audio coded information defined in ISO/IEC 13818-7.
- ADTS error check consists of error detection information.
- The data stream consists of audio data coded according to ISO/IEC 13818-7.
- The syntactic element ID indicates the type of syntactic element that follows this ID or end of the data stream.
- The syntactic element consists of various components of audio data coded according to ISO/IEC 13818-7. It is iterated the number of times specified in the ADTS variable header.

There are certain restrictions on the audio coding parameters that need to be compliant with the ISDB-T specification, which are not restricted as such in the MPEG2 standards. The following section makes a mention of these operational restrictions.

Input Audio Format

The input audio format for digital broadcasting is subject to the following restrictions. These restrictions are imposed because the MPEG-2 AAC Standard, adopted as the audio coding system for terrestrial digital television, BS digital, broadband CS digital, and terrestrial digital audio broadcasting, provides no provisions that clearly specify audio modes. To ensure compliance with the provision regarding narrowband CS digital broadcasting, the possible audio modes are specified by ISDB-T. Furthermore, this list is trimmed to allow the audio modes to be advantageous, according to real world needs, and these are listed as the recommended audio modes in Table 2 .

Parameter		Restriction
Audio mode	Possible audio mode	Mono, stereo, multichannel stereo (3/0, 2/1, 3/1, 2/2, 3/2, 3/2+LFE), 2-audio signals (Dual mono), multi-audio (3 or more audio signals) and combinations of the above
	Recommended audio mode	Mono, stereo, multichannel stereo (3/1, 3/2, 3/2+LFE), 2-audio signals (dual mono)
Emphasis		None

Table 2 – Restriction on Input Audio Format

Audio Coding System

The audio coding system for digital broadcasting, stipulated by ISDB-T, is the MPEG-2 AAC audio. This section further mentions the restrictions on this MPEG-2 standard, as needed by ISDB-T. In the ISDB-T standard specification [1], the reasons for choosing the Low Complexity (LC) profile of MPEG-2 AAC are mentioned in detail. Even though LC and SSR (Scalable Sampling Rate) profiles meet the quality criteria required by ITU-R, the SSR profile specific features are not effective for the broadcast requirement. It was also seen that the LC profile could improve audio quality as a result of optimization and technical advancements and the chip costs were significantly lower than the requirement for the MAIN profile.

Major Parameters

The major parameters of the audio coding system, as enforced by the ISDB-T standard, are described in this section.

Parameter	Restriction
Bitstream Format	AAC Audio Data Transport Stream (ADTS)
Profile	Low Complexity (LC) profile
Max. Number of coded channels	5.1 channels per ADTS
Max. bitrate	Compliant to MPEG-2 standard (ISO/IEC 13818-7)

Table 3 – Restriction on Major Parameters of Audio Coding System

MPEG-2 AAC ADTS Coding Parameters

Fixed header of ADTS – According to ISDB-T, the presence of the Cyclic Redundancy Check bits are mandatory, while in the MPEG-2 AAC standard specification, this is optional. This is required to improve the tolerance of ADTS.

Parameter	Restriction
Protection absent	'0' (CRC check always present)
Profile	'1' for Low Complexity (LC) profile
Sampling frequency index	Selected from 0x3 to 0x8 (48k, 44.1k, 32k, 24k, 22.05k, 16k)
Channel configuration	This has been explained in detail in Table 7 and Table 8

Table 4 – Restriction on parameters in Fixed Header of ADTS

Variable Header of ADTS–The ISDB-T standard does not permit variable rate encoding, and hence imposes a restriction on the buffer fullness parameter. It also does not permit more than one raw data block in a frame under a single ADTS header, since for broadcast applications, a header loss can have a seriously adverse impact. Furthermore, in the event of even a single error, the adversities are inordinate, as the single header represents CRCs from all the stream elements from all the raw data blocks.

Parameter	Restriction
ADTS buffer fullness	Use of 0x7FF (for variable rate) not permitted
Number of raw data blocks in frame	0 (no. of raw data blocks per frame = 1)

Table 5 – Restriction on parameters in Variable Header of ADTS

Raw Data Stream

Parameter	Restriction
Coding mode in a single ADTS and raw data block configuration (order of transmission)	This has been explained in detail in Table 8
Handling of coupling channel option	Use of Coupling Channel option is not permitted
Handling of Program Configuration Element (PCE)	This has been explained in detail in the next section
Handling of Fill Element (FIL)	This has been explained in detail in the next section

Table 6 – Restriction on Raw Data Stream

Audio stream configuration and Multiplexing

Depending on the potential need for simultaneous reproduction, it was determined by ISDB-T, whether to use a single ADTS or multiple ADTSs for different input audio modes.

Provisions regarding input audio mode and ADTS configuration and multiplexing are given in Table 7.

Input Audio Mode	ADTS Configuration and Multiplexing
Mono, stereo	Comprises one ADTS
Multichannel Stereo (3/0, 2/1, 3/1, 2/2, 3/2, 3/2+LFE)	Comprises one ADTS
2 Audio signals (Dual mono)	Comprises one ADTS
Multiple audio signals other than dual mono (2/0 + 2/0)	Comprises the same number of ADTSs as that of audio streams (languages) and is multiplexed with the MPEG-2 systems layer.

Table 7 – ADTS Configuration and Multiplexing for different Input Audio Modes

For example, for 2-audio transmission, dual mono mode with one ADTS is used when simultaneous reproduction is requested. However, dual mono mode with two ADTSs can be used when simultaneous reproduction is not requested.

Detailed provisions regarding coding mode in a single ADTS and order of transmission in ADTS configuration are mentioned below. These are as stipulated by the MPEG-2 AAC standard:

Coding Mode	Channel configuration in ADTS fixed header	Syntactic Elements (SE) in order of transmission	Default Element to Speaker Mapping
Mono (1/0)	1	<SCE1><END>*	SCE1 = C**
Stereo (2/0)	2	<CPE1><END>*	CPE1 = L and R**
3/0	3	<SCE1><CPE1><END>	SCE1 = C, CPE1 = L and R
3/1	4	<SCE1><CPE1><SCE2><END>	SCE1 = C, CPE1 = L and R, SCE2 = MS**
3/2	5	<SCE1><CPE1><CPE2><END>	SCE1 = C, CPE1 = L and R, CPE2 = LS and RS**
3/2 + LFE	6	<SCE1><CPE1><CPE2><LFE><END>*	SCE1 = C, CPE1 = L and R, CPE2 = LS and RS, LFE = LFE**

Table 8 - Channel Configuration and Syntactic Elements in single ADTS configuration

* SCE – Single Channel Element

CPE – Channel Pair Element

LFE – LFE Channel Element

END – End/Terminator Element

** L: Left front speaker

R: Right front speaker

C: Center front speaker

LFE: Low frequency emphasis

LS: Left surround speaker

RS: Right surround speaker

MS: Monophonic surround speaker

To transmit channel configuration and down mix coefficients, a Program Configuration Element (PCE) may be used. This PCE should necessarily be consistent with the ADTS header. When the channel_configuration bit in ADTS header is 0, it is possible to accurately represent the intended state of reproduction by decoding PCE. The detailed provisions regarding Program Configuration Element transmission are provided below:

a) PCE should be transmitted when switching between audio modes 2/1, 2/2, 1/0+1/0 for which the channel configuration parameter in the ADTS fixed header is 0. This is during continuous service using the same service ID. The PCE parameter value has to match that included in the ADTS header.
b) When downmix coefficient is transmitted, PCE should be transmitted at an interval of less than 550 ms for that purpose. However, this applies only when channel_configuration = 5 or 6. When performing this operation, PCE shall always be transmitted during the period in which channel_configuration = 5 or 6, is in continuous service.
c) While PCE may be included in every ADTS frame, any modification of parameters other than changes made (for example) to channels and downmix coefficients is prohibited.
d) The following operational provisions are established for bits comprising PCE.
    a. The same value shall be assigned to Sampling_frequency_index and Profile as the header.
    b. Number of side channel elements shall be 0. Therefore the following flags do not exist:
        i. side_element_is_cpe
        ii. side_element_tag_select
    c. No specific provisions are established for num_assoc_data_elements. Note that <DSE> is treated as an option for broadcasts.
    d. num_valid_cc_element shall be 0. Therefore the following flags do not exist:
        i. cc_element_is_ind_sw
        ii. valid_cc_element_tag_select
    e. Mono_mixdown_present shall be 0. Therefore, mono_mixdown_element_number does not exist.
    f. Stereo_mixdown_present shall be 0. Therefore, stereo_mixdown_element_number does not exist.
    g. Comment_field_bytes shall be treated according to the AAC standard. Its content is meaningless as far as the system is concerned. It is treated as an option (for example) for bitstream control.

The detailed provisions regarding configuration of Fill Element (FIL) are provided below:

When the value of coding parameter sampling_frequency_index in the ADTS Fixed Header is in the range of 0x6 to 0x8 (24k, 22.05k, 16 kHz), EXT_SBR_DATA (‘1101’) and EXT_SBR_DATA_CRC (‘1110’) can be used in Fill Element (FIL).

(Note) For BS / broadband CS digital broadcasting, the value of sampling_frequency_index does not fall within the range of 0x6 to 0x8, therefore, EXT_SBR_DATA (‘1101’) and EXT_SBR_DATA_CRC (‘1110’) are not used.

Compatibility with 2-channel stereo-capable receiver

There is a strong likelihood that not only terrestrial digital television broadcasting and BS/broadband CS digital broadcasting receivers capable of reproducing multi-channel stereo, but also even those receivers capable of reproducing two-channel stereo will be commercially available. Hence, compatibility with 2-channel stereo-capable receiver should be accounted for when multi-channel stereo service is provided.

This section defines the conditions and lists considerations in relation to compatibility with 2-channel stereo-capable receiver when multichannel stereo service is provided.

1. Two-channel stereo simulcasting is not obligatory when multichannel stereo service is provided. Basically, 2-channel stereo-capable receiver shall handle the service by downmixing.
2. It shall be possible to transmit downmix coefficient using PCE according to the AAC Standard when 5-channel stereo (3/2) and 5.1-channel stereo (3/2+LFE) services are provided. This ensures that the transmitting side needs to transmit just a single stream, which improves efficiency in bitrate.
3. It shall be possible to provide 2-channel stereo simulcasting service at the request of broadcasting stations. In this case, two streams should be treated as different ADTSs, multiplexed, and stream-controlled by the systems layer.

Processing procedure of CRC bits for MPEG2 AAC ADTS

The MPEG-2 AAC Standard includes the following as CRC error detection data processing procedure (as described in ISO/IEC 11172-3 [4]). This is mentioned as adts_error_check in the standard. ISDB-T follows the same processing procedure.

The following bits are input to the CRC algorithm in order of appearance:

- All bits of the headers
- First 192 bits of any
  - Single_channel_element (SCE)
  - Channel_pair_element (CPE)
  - Coupling_channel_element (CCE)
  - Low Frequency enhancement channel (LFE)
- In addition, the first 128 bits of the second individual_channel_stream in the channel_pair_element shall be protected.
- All information in any program configuration element or data element shall be protected.

For any element where the specified protection length of 128 or 192 bits exceeds its actual length, the element is zero padded to the specified protection length for CRC calculation.

Ittiam Solutions

For complete compliance with the digital format ISDB-T being adopted in Japan, Philippines, Maldives and Thailand, and its derivative, ISDB-T International being adopted widely in South America, it is necessary, but not sufficient to have an audio encoding system that conforms to the requirements of the MPEG-2 AAC Standard [2]. It is also essential for this MPEG-2 AAC Encoding scheme to comply with the restrictions specified by the ISDB-T standard [1].

The AAC LC (Low Complexity) Encoder available at Ittiam Systems for broadcasters has been tested and is in use by Japanese broadcasters, thus certifying acquiescence with the ARIB standard, ISDB-T. The encoder is also integrated into Ittiam's Broadcast Media SDKs.

Conclusion

Audio compliance with ISDB-T, standardized by ARIB, is necessary to meet the requirements of the Japanese digital television and radio broadcast formats. However, since there are restrictions on the regular MPEG-2 AAC Encoding standards, which have been explained in detail in this document, for complete compliance with the ISDB-T standard, it is not sufficient, to have an audio encoding system according to the requirements of the MPEG-2 AAC standard.

References

1. ARIB STD-B32: Video Coding, Audio Coding, and Multiplexing Specifications for Digital Broadcasting (Version 2.1)
2. ISO/IEC 13818-7: 2003 Information technology – Generic coding of moving pictures and associated audio information: Advanced Audio Coding
3. ISO/IEC 13818-7: 2003/AMD 1: 2004 Information technology – Generic coding of moving pictures and associated audio information: Advanced Audio Coding AMENDMENT 1: Embedding of bandwidth extension
4. ISO/IEC 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

About Ittiam

Ittiam Systems is a global technology company with deep R&D driven solutions for media creation, management and consumption, providing advanced media codecs, software development platforms, systems and workflows for embedded and online applications. Its solutions are at the heart of tens of millions of lifestyle products that drive mobility, content access, networking and sharing.

For more information about Ittiam products, please visit www.ittiam.com or contact us at mkt@ittiam.com.

Disclaimer

This white paper is for informational purposes only. Ittiam makes no warranties, express, implied or statutory, as to the information in this document. The information contained in this document represents the current view of Ittiam Systems on the issues discussed as of the date of publication. It should not be interpreted to be a commitment on the part of Ittiam, and Ittiam cannot guarantee the accuracy of any information presented after the date of publication.

Complying with all applicable copyright laws is the responsibility of the user. Without limiting the rights under copyright, no part of this document may be reproduced, stored in or introduced into a retrieval system or transmitted in any form or by any means (electronic, mechanical, photocopying, recording, or otherwise), or for any purpose, without the express written permission of Ittiam Systems. Ittiam Systems may have patents, patent applications, trademarks, copyrights, or other intellectual property rights covering subject matter in this document. Except as expressly provided in any written license agreement from Ittiam Systems, the furnishing of this document does not give you any license to these patents, trademarks, copyrights, or other intellectual property.