Multi-channel Audio Signal Compression and Quality Assessment for AV Communication

Jin-Woo Hong, Dae-Young Jang, and Seong-Han Kim

Realistic Telecom. Section, Electronics and Telecommunications Research Institute, Taegu, 305-600, Korea

Abstract: This paper describes a technique for multi-channel audio signal compression using two sets of MPEG audio, an implementation and the result on quality assessment of real time operating multi-channel audio codec, and a multiplexing method of synchronization for AV communication from the implementation of HDTV system using this codec. Subjective test for quality assessment was performed by Double blind and triple stimulus with hidden reference. All test items except one are awarded difference grade(diffgrade) better than -1.

1. INTRODUCTION

Audio signal compression technology currently stands out in bold relief on account of limitation of storage capacity and lack of transmission channel. Audio compression usually refers to the coding of high fidelity audio signal for consumer hi-fi, professional audio including motion picture and HDTV(High Definition Television) audio, and various multimedia systems. Audio coding has advanced rapidly in recent years spurred on by cost-effective digital technology and diverse commercial applications, especially by some international standards. Because standards are essential for compatibility of coding algorithm for various applications, standardization of audio compression technique has lately become a major activity of central importance to industry and government.

This paper focuses a special application on a multi-channel audio codec using two sets of MPEG stereo audio for signal compression and a practical configuration of target DSP system is adopted for real time operation with other software tips. The channel configuration of this codec on real time operation is composed of two front channels(front left, front right) and two rear channels(rear left, rear right).

2. THE IMPLEMENTATION OF CODEC AND AV MULTIPLEXING SCHEME

This codec was implemented as an encoder unit and a decoder unit by using universal DSP, TMS320C40. Total 4 channel audio inputs for multi-channel are encoded by two sets of MPEG stereo audio encoder. Encoder should provide the user interface that consists of several switches for parameter selection and LCD panel. Audio data for each channel is converted to parallel and latched by shift register with the channel synchronization signal. Then audio data are multiplexed and transmitted to FIFO. An interrupt signal is generated when almost full data are stored in FIFO. If DSP detects the interrupt signal, it downloads a frame of data from FIFO through DMA channel. The data in FIFO are stored in sequence of channel, and the channel identification flag is attached to the data to confirm the data for each channel. The decoder block diagram of this codec is shown in Figure 1.

![Decoder block diagram](image)

In case of 48kHz sampling and 384kbit/s bit rate, the input to output processing time for one frame of 1152 samples of PCM audio data must be less than 24ms. The MPEG audio algorithm applied to this codec, but, requires many computing times for the complicated calculations and the iterative processes, and then it is very difficult to operate software on real time, especially for encoder. Hence, special hardware configuration and software design scheme have to be considered because of this constraint. DSP must, that is, finish the assigned
processing itself for incoming frame within 24ms.

The program multiplexer performs the multiplexing through the two systems coding layers. At the most basic layer are packetized elementary streams (PESs) that carry ES of data for one application, video or audio, etc. In the TS layer, the PES packets are subdivided into payload of TS packets of 188 bytes. Synchronization of decoding and presentation process for audio and video at a receiver is a particularly important aspect of a real time program multiplexer. In order to solve this problem, MPEG system specifies a timing model in which the end-to-end delay through the entire codec system is constant. This is done by formatting transport stream with two types of timing elements: program clock reference (PCR) and presentation/decoding time stamp (PTS/DTS). PCR, which is a sampled value of the encoder’s 27 MHz system time clock (STC), is periodically added in the adaptation headers of selected TS packets. PCR serve as a reference for the system clock recovery at the decoder and establishes a common time base through the entire system. Synchronization between video and audio signals is accomplished by comparing both the audio presentation time stamps (PTS) and the video PTS with the STC.

3. QUALITY ASSESSMENT AND RESULT

As a subject, 17 listeners with normal hearing participated in this test. This test was done in the listening room according to the requirements of ITU-R. Test materials using in the test are collected in atmosphere and announcement of a ground, play of the orchestra, original sound track of movies and sound of musical instrument, etc. Figure 2 shows the test result with 95% confidence interval. Almost test materials except for material 4 are awarded mean difference grades better than -0.6, and three materials(material 4, 6, 10) are awarded diffgrades worse than -0.5. Three materials that achieved relatively bad diffgrades are identified as the material that easy to detect the back ground noise. That is material 4, pitch pipe sounds pure tone, therefore white noise can easily spoil the sound. Materials 6, 10 includes relatively long interval of silence, and in that interval, back ground noise is heard clearly.

![Figure 2: Subjective test results (95% confidence interval included).](image)

ACKNOWLEDGEMENTS

This work is supported by the Korea Telecom under the title of “A Development of HDTV transmission Technology.”

REFERENCES

1. ISO/IEC 11172-3 and 13818-3, International Standard, Coding of moving pictures and associated audio for digital storage media at up to about 1.5 Mbps