

Review of AVS Audio Coding Standard

ZHANG Tao, ZHANG Caixia, and ZHAO Xin

(School of Electronic Information Engineering, Tianjin University, Tianjin 300072, China)

Abstract

Audio Video Coding Standard (AVS) is a second-generation source coding standard and the first standard for audio and video coding in China with independent intellectual property rights. Its performance has reached the international standard. Its coding efficiency is 2 to 3 times greater than that of MPEG-2. This technical solution is more simple, and it can greatly save channel resource. After more than ten years' development, AVS has achieved great success. The latest version of the AVS audio coding standard is ongoing and mainly aims at the increasing demand for low bitrate and high quality audio services. The paper reviews the history and recent development of AVS audio coding standard in terms of basic features, key techniques and performance. Finally, the future development of AVS audio coding standard is discussed.

Keywords

Audio Video Coding Standard (AVS); audio coding; AVS1 audio; AVS2 audio

1 Introduction

The Audio Video Coding Standard (AVS) workgroup of China was approved by the Science and Technology Department under the former Ministry of Industry and Information Technology of People's Republic of China in June, 2002 [1]. The goal of the AVS workgroup is to establish the generic technical standards for high-quality compression, decompression, processing and representation of digital audio and video. The AVS workgroup also aims to provide the digital audio-video equipment and systems with high-efficient and economical coding/decoding technologies [2]. The formal name of AVS is "Information Technology-Advanced Audio and Video Coding", including four main technical standards: system, video, audio, digital rights management and the supporting standards, such as conformance testing. The members of AVS workgroup are domestic and international institu-

tions and enterprises that focus on the research of digital audio and video coding technology and the development of related products.

Since 2002, the AVS audio subgroup has drafted a series of audio coding standards, including AVS1-P3, AVS1-P10 and AVS-Lossless. Since 2009, the AVS audio subgroup started drafting the next generation audio coding standards. To identify the difference between these two series, the former serial is called AVS1 audio coding standards and the later one is called AVS2 audio coding standards. The AVS1 audio coding standards has been finished and widely used in various applications so far. The AVS2 audio coding standards is still under development and will be released soon.

This paper reviews AVS audio coding standards. It is organized as follows. In section 2, the series standards of AVS1 audio including AVS1-P3, AVS1-P10 and AVS-Lossless will be introduced. AVS2 audio coding scheme are presented in section 3. At last, the conclusion is given in section 4.

2 The Development of AVS1 Audio Coding Standards

The AVS audio subgroup started drafting the first generation AVS audio standard in 2003. The prime goal of the AVS audio subgroup is to establish an advanced audio codec standard with general performance equivalent or superior to MPEG AAC, on the premise of developing our own intellect property [2]. The first generation of AVS audio codec standard includes three parts: Information Technology—Advanced Audio Video Coding Part 3: Audio (AVS1-P3), Information Technology—Advanced Audio Video Coding Part 10: Mobile Voice and Audio (AVS1-P10 or AVS1-P10 audio), and AVS Lossless Audio Coding Standard (AVS-LS). **Tables 1** and **2** show the development of AVS1 audio coding standards [3].

2.1 AVS1-P3

After three years of effort, in 2005 the AVS working group finished the first AVS audio standard. The audio codec supports the scalable audio coding and is applied in mass information industries such as digital broadcasting of high resolution, intense laser digital storage media, wireless wideband multimedia communication, wideband stream media on the Internet and other related applications [1].

The AVS1-P3 encoder supports mono, dual and multichannel PCM audio signal. One frame audio signal includes 1024 samples. It is separated into 16 blocks, every 128-point block with 50% overlap is hanning windowed. The transform length is determined by the long/short window switching module: 2048 for long and 256 for short in order to accommodate both stationary and transient signals. The sampling rate ranges from 8 kHz to 96 kHz for the input signal. The output bitrate ranges from 16 kbps to 96 kbps per channel. The transparent audio quality could be provided at 64 kbps per channel. The com-

▼Table 1. History of AVS1 audio

Project	Project No. (national standard)	Working Draft (WD)	Committee Draft (CD)	Final Committee Draft (FCD)	Publicity	Final Draft (FD)	Final Draft for Standard (FDS)	Standard (GB)
AVS1-P3 (Dual Track)	20051305-T-339	2004.12	2005.3	2005.12		2006.4	2009.3	
AVS1-P3 (5.1)	20051305-T-339	2005.9	2005.12	2005.12		2006.4	2009.3	
AVS1-P3 (lossless)		2010.6	2010.7	2011.9				
AVS1-P10 (Mobile speech and audio coding)	20080548-T-469	2007.3	2008.9	2009.12	2010.6	2010.9	2011.6	2013.12

▼Table 2. Review of AVS1 audio

Time	Standard	Application and major coding tools
Dec. 2003	AVS1-P3	High quality application; multi-resolution analysis linear prediction, vector quantization, fine granularity scalable coding (FGSC)
Aug. 2005	AVS1-P10	Wireless network and mobile equipment; sampling rate conversion filter, ACELP/TVC mixed encoding, ISF vector quantization, Mono-signal High-Band Encoding (BWE), parametric stereo coding
Dec. 2009	AVS-Lossless	Reconstruction of original digital audio, data archive; lifting wavelet analysis, linear prediction

pression ratio is 10–16 [2].

2.1.1 Basic Encoding Process

Fig. 1 shows the framework of the AVS1-P3 audio codec [2]. Audio input PCM signals are analyzed by a psychoacoustic model. Then long/short window switch module determines the length of the analysis block depending on the transients. The signals are transformed to frequency domain by integer Modified Discrete Cosine Transform (intMDCT). For stereo signals, Square Polar Stereo Coding (SPSC) may be applied to the encoder if there are strong correlations between the channel pair. After that, frequency domain signals undergo nonlinear quantization. Context-Dependent Bitplane Coding (CBC) is used for entropy coding of quantized spectrum data. Finally the coded bits are written to the output bitstream based on the format defined in AVS1-P3 standard.

2.1.2 Key Technologies in AVS1-P3

The structure of the encoder in the AVS1-P3 standard is similar to that in Advanced Audio Coding - Low Complexity (AAC-LC). It has higher coding efficiency and better sound quality as a result of the use of new coding technologies, such as multi-resolution analysis, linear prediction in frequency domain, vector quantization and fine granularity scalable coding (FGSC).

Window switching is applied to reduce pre-echoes. A two-stage window switch decision is recommended in the AVS audio coding standard, which is called Energy and Unpredictability Measure Based Window Switching Decision (ENUPM - WSD) [2]. In the first stage, one frame audio signal is separated into 16 blocks, and then the energy variation of every sub-block is analyzed. If the maximum energy variation of a sub-block meets a given condition, the second stage based on unpredictability measurement in the frequency domain is ap-

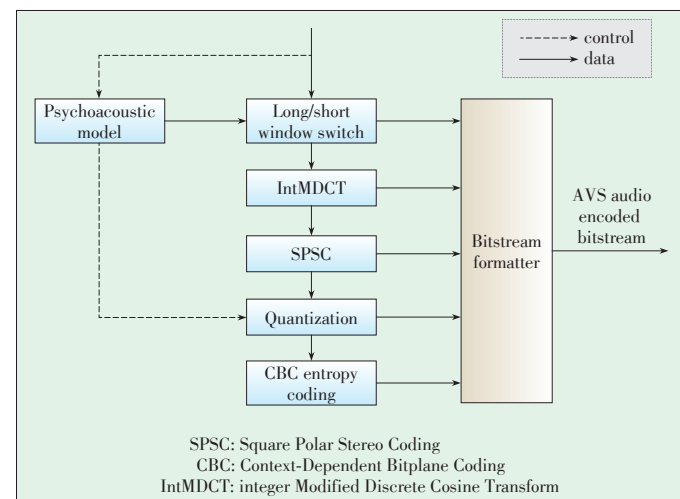
plied. Otherwise, it judges the window type by analyzing signal characteristics in the time domain and frequency domain. ENUPM - WSD has the merits of low complexity and high accuracy.

Considering the extension of lossless compression for the future, AVS audio subgroup adopted Integer MDCT as the time-frequency mapping module instead of traditional DCT [2]. IntMDCT can be used in lossless audio coding or combined perceptual and lossless audio coding. The advantages of MDCT, such as a good spectral representation of the audio signal, critical sampling and overlapping of

blocks are all reserved.

SPSC is applied as an efficient stereo coding scheme in AVS. Compared with Mid/Side stereo coding in AAC, its coding efficiency is higher. When SPSC is applied, one channel transmits the bigger value of the channel pair, and the other transmits the difference. In terms of quantization noise, the final decoded audio noise is smaller in SPSC than in M/S. Because in SPSC, noise superposition happens at only one channel at decoder end, while in Mid/Side, noise disperses to both channels.

In entropy coding, AVS adopts Context-dependent Bit-plane Coding (CBC). It is more efficient compared with Huffman coding. CBC entropy coding technology gets an improvement of 6% of bitrate in comparison with Huffman coding at 64 kbps/channel. The most striking characteristic of CBC is its Fine



▲Figure 1. AVS1-P3 audio encoder block diagram.

Review of AVS Audio Coding Standard

ZHANG Tao, ZHANG Caixia, and ZHAO Xin

Grain Scalability (FGS). CBC coded bitstream are evenly layered (16 to 96 layers, each layer 1 kbps), and as it changes from higher layer to lower layer, the audio quality downgrades from high to low, but still audible [1].

2.1.3. Performance

Figs 2, 3 and 4 show the informal subjective listening tests between AVS1-P3 and other three most popular audio compression formats, MP3 (lame 3.96), AAC (FAAC 1.24) and WMA (WMA 10) [1]. The tests use ITU-T P.800/P.830 test basic model. Four sequences, es02, sc02, si02, and sm02 were used in this test. These are speech, complex mixage audio, single instrument sound and simple mixage audio, respectively. Bitstreams were coded at 128 kbps.

The following conclusion could be obtained from Figs 2–4: at 128 kbps, AVS1-P3 is superior to MP3 (lame 3.93), the same as AAC (FAAC 1.24), but slightly worse than WMA

(WMA 10.0) [1].

2.2 AVS1-P10

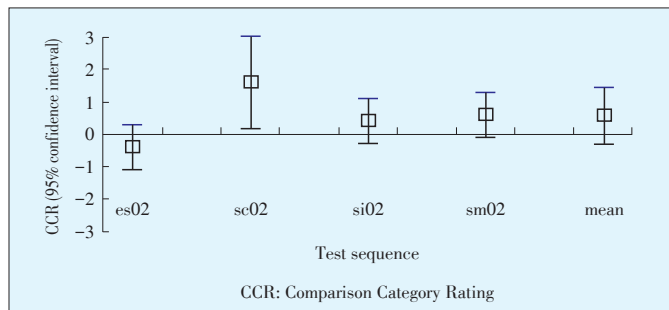
With the development of third-generation mobile communication, many challenges have arisen. There is a growing demand for low bitrate and high fidelity quality audio codec. At present, there have been many international audio standards for mobile applications such as G.711 series standard (ITU-T) and AMR series standard (3GPP).

In order to provide a mobile audio standard with independent intellectual property rights for the quickly developing mobile communication system, the AVS Audio Subgroup started drafting the AVS1-P10 in August 2005. The AVS1-P10 Final Committee Draft (FCD) was completed in December 2009. It is approved as the national standard in December 2013 and has wide applications in 3G communication, wireless broadband multimedia communications, broadband Internet streaming service and more. The advantages of AVS1-P10 include high efficiency, flexible compression quality, low complexity and strong error prevention mechanism [4].

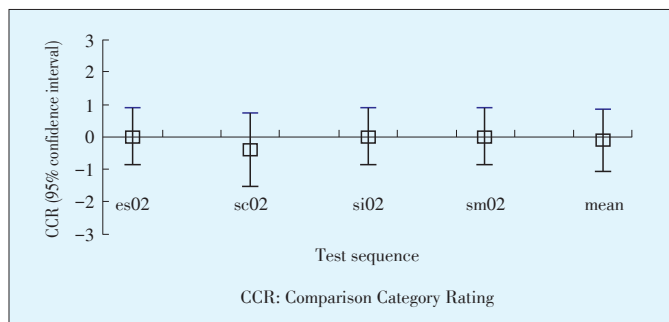
The encoder supports mono and stereo Pulse Code Modulation (PCM) signals with sampling rate of 8 kHz, 16 kHz, 24 kHz, 32 kHz, 48 kHz, 11 kHz, 22 kHz and 44.1 kHz. The output bitrate for mono ranges from 10.4 kbit/s to 24.0 kbit/s, and for stereo, ranges from 12.4 kbit/s to 32.0 kbit/s.

2.2.1 Basic Encoding Process

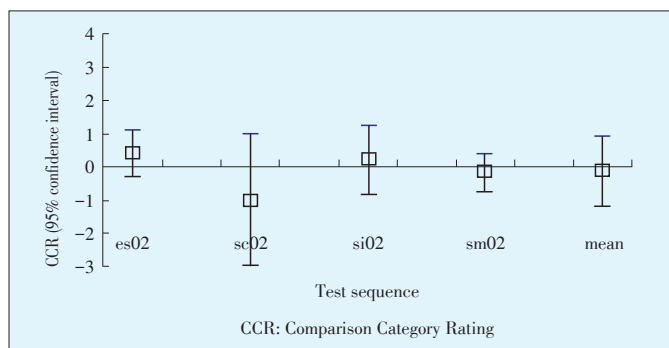
AVS1-P10 adopts the basic framework of AMR-WB+ (Fig. 5). It firstly convert the sampling frequency of the input signal into an internal sampling frequency F_s . For mono mode, the low frequency (LF) signal adopts Algebraic Code Excited Linear Prediction/Transform Vector Coding (ACELP/TVC) codec mode, while the high-frequency signal is encoded using a bandwidth extension (BWE) approach. For stereo mode, the same band decomposition as in the mono case is used. The HF part of the left channel and right channel is encoded by using parametric BWE on the two stereo channels. The LF part of the left channel and right channel is down mixed to main channel and



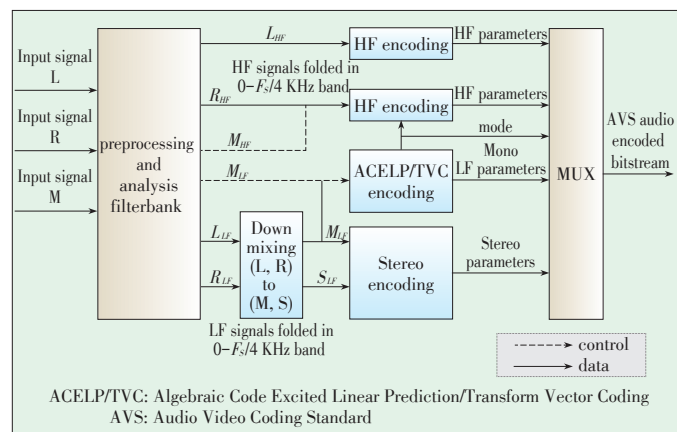
▲ Figure 2. AVS1-P3 128 kbps vs. MP3 128 kbps.



▲ Figure 3. AVS1-P3 128 kbps vs. FAAC 128 kbps.



▲ Figure 4. AVS1-P3 128 kbps vs WMA 128 kbps.



▲ Figure 5. AVS1-P10 encoder block diagram.

side channel (M/S). The main channel is encoded by ACELP/TVC module. The stereo encoding module processes the M/S channel and produces the stereo parameters [5].

2.2.2 Key Technologies in AVS1-P10

1) ACELP/TVC Mixed Encoding Module

The core codec module in AVS1-P10 is ACELP/TVC mixed encoding module. AVS1-P10 codec integrates ACELP coding and the Transform Vector Coding (TVC) into a mixed orthogonal encoder. It can choose the best encoding mode between two coding modes according to the signal type. ACELP mode is based on time-domain linear prediction, so it is suitable for encoding speech signals and transient signals [6]. On the other hand, TVC mode is based on transform domain coding, so it is suitable for encoding music signals. Thus it can encode a variety of complex audio signals.

Several coding methods, such as ACELP256, TVC256, TVC512, and TVC1024, can be applied to one superframe. There are 26 mode combinations of ACELP/TVC for each superframe [7]. The mode can be selected by adopting the closed-loop search algorithm or the open-loop search algorithm. The latter is relatively simple, but the mode selected may be not optimum.

2) High-Band Encoding

AVS1-P10 adopts BWE approach to code HF signal with the frequency components above $F/4$ kHz of the input signal. In BWE, energy information is sent to the decoder in the form of spectral envelop and gain [5]. However, the fine structure of the signal is extrapolated at the decoder from the decoded excitation signal in the LF signal. Besides, in order to keep the continuity of the signal spectrum at the $F/4$, the HF gain needs to be adjusted according to the correlation between the HF and LF gain in each frame.

3) Stereo Signal Encoding

In order to support stereo signal encoding, AVS1-P10 adopts a high effective configurable parametric stereo coding scheme in the frequency domain. This scheme provides a flexible and extensible codec structure with coding efficiency, which is similar to that of AMR-WB+. Because it avoids the resampling in the time domain, it reduces the complexity of encoder and decoder. The ability to flexibly configure the low frequency bandwidth determined by the coding bit rate is also available, which makes it a high-effective stereo coding approach [6].

2.2.3 Complexity Analysis

In order to analyze the complexity of the AVS1-P10 Codec, the AVS1-P10 Fixed-point Codec is developed, and the functions of complexity analysis are integrated into it. The Weighted Million Operation per Second (WMOPS) method approved by ITU is used to analyze the complexity of the AVS1-P10 Codec. The basic operation's weighting is given in **Table 3** [8]. The complexity of the AVS1-P10 Codec is shown in **Tables 4**

and **5** [8].

2.2.4 Performance

In order to evaluate the performance of AVS1-P10 codec, objective test - perceptual evaluation of speech quality (PESQ) and subjective test-complementary metal-oxide-semiconductor (CMOS) methods are applied to the 20 test sequences selected from AVS workgroup at the bit rate of 12 kbps and 24 kbps, including pure speech, pure music, noisy speech, speech over music and speech between music. Each of these contained two mono audio and two stereo audio.

1) Objective Performance Evaluation of AVS1-P10 Audio

PESQ result between AVS1-P10 (AVS1-P10) and AMR -

▼ **Table 3. Operation's weighting**

Operation	Weight
16bit add	1
16bit sub	1
16bit abs	1
16bit shift	1
16bit multiply	1
16bit divide	18
32bit add	2
32bit sub	2
32bit abs	3
32bit shift	2

▼ **Table 4. AVS1-P10 encoder complexity**

Condition	Command	Complexity (WMOPS)
12 kbps, mono	-rate 12 -mono	Average=56.318 Worst=58.009
24 kbps, mono	-rate 24 -mono	Average=79.998 Worst=80.055
12.4 kbps, stereo	-rate 12.4	Average=72.389 Worst=73.118
24 kbps, stereo	-rate 24	Average=83.138 Worst=83.183
WMOPS: Weighted Million Operation per Second		

▼ **Table 5. AVS1-P10 decoder complexity**

Condition	Command	Complexity (WMOPS)
12 kbps, mono	-mono	Average=9.316 Worst=9.896
24 kbps, mono	-mono	Average=13.368 Worst=13.981
12.4 kbps, stereo	None	Average=16.996 Worst=17.603
24 kbps, stereo	None	Average=18.698 Worst=19.103
WMOPS: Weighted Million Operation per Second		

Review of AVS Audio Coding Standard

ZHANG Tao, ZHANG Caixia, and ZHAO Xin

WB+ are given respectively in **Figs 6** and **7** [8].

2) Subjective Performance Evaluation of AVS1-P10 Audio

CMOS result between AVS1-P10 and AMR-WB+ are given respectively in **Figs 8** and **9** [8].

It can be concluded from these figures that the performance of AVS1-P10 is no worse than AMR-WB+ on average.

2.3 AVS Lossless Audio Coding

Nowadays, the ever-decreasing price of storage and ever-increasing Internet bandwidth make storing and distributing lossless coded audio possible. Lossless audio coding technology is designed to perfectly reconstruct original digital audio. After decoding, its output sequence is same as the original audio sequence. There are many lossless audio coding standards, such

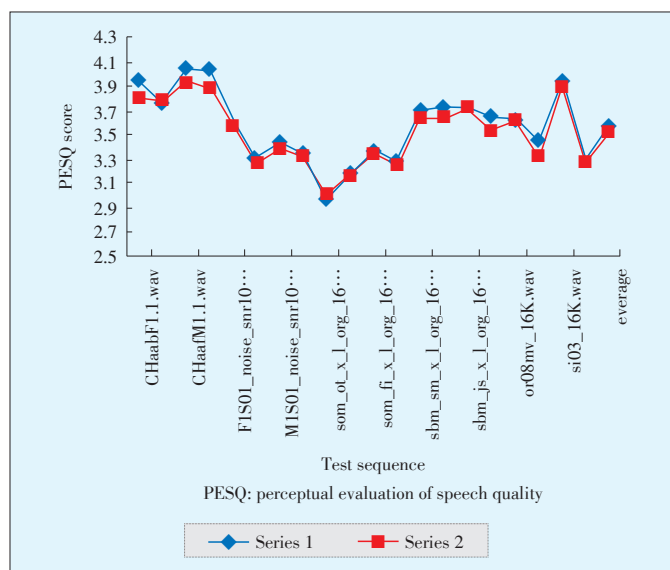
as MPEG-4 Audio Lossless Coding (MPEG-4 ALS), MPEG-4 Scalable to lossless standard (MPEG-4 SLS), Free Lossless Audio Codec (FLAC) and so on [9]. In order to make a lossless audio coding standard with independent intellectual property rights, AVS audio subgroup started to draft lossless audio coding standard in early 2010, and the FCD and reference software were finished in September 2010 [10].

2.3.1 Basic Encoding Process

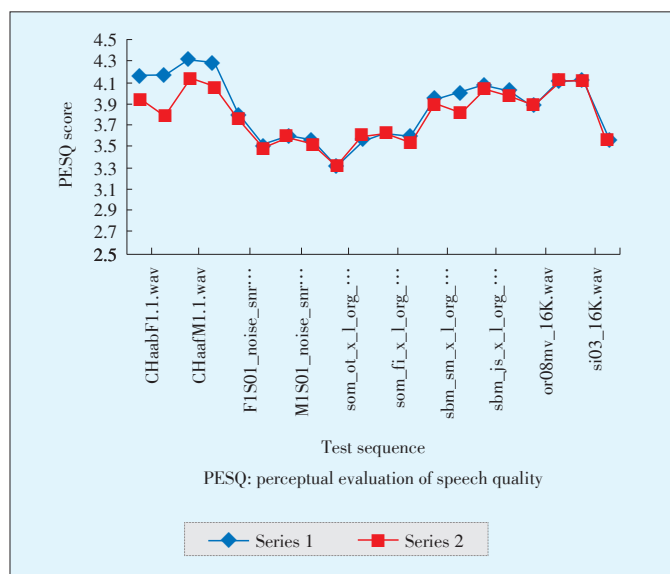
Fig. 10 shows the the AVS lossless encoder [9]. The input audio signal is decorrelated to remove the correlation between different channels. Then, the de-correlated signal is filtered into a high-frequency band and a low-frequency band by a lifting the wavelet filter. After that, subband signals go through the linear predictor to generate prediction residual signals. Then, the residual signal is sent to the preprocessor, where it is normalized. Finally, the normalized signal goes through the entropy codec and a bitstream is generated.

2.3.2 Key Technologies in AVS-Lossless

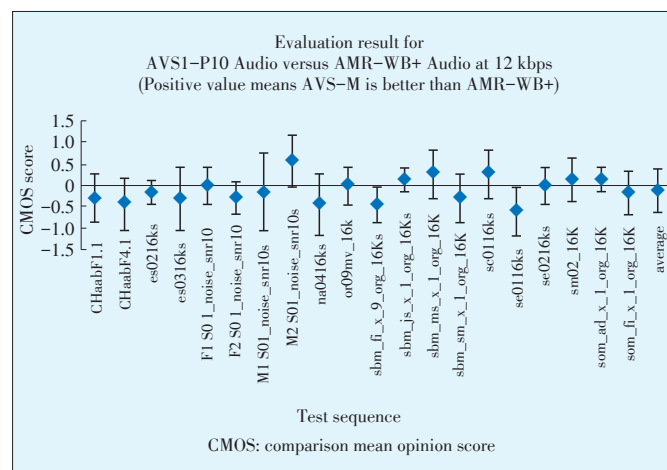
1) De-correlation



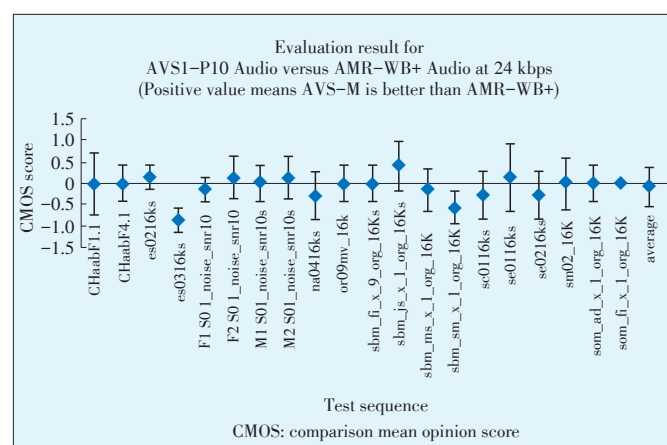
▲ Figure 6. PESQ test result at 12 kbps.



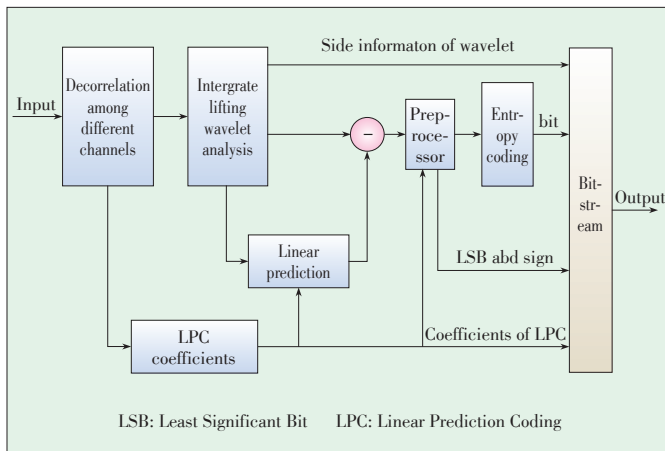
▲ Figure 7. PESQ test result at 24 kbps.



▲ Figure 8. CMOS test result at 12 kbps.



▲ Figure 9. CMOS test result at 24 kbps.



▲ Figure 10. Framework of AVS lossless audio encoder.

In an AVS lossless audio encoder, a sum and subtraction coding method is used to de-correlate multichannel signals sounds.

2) Integer Lifting Wavelet Transform

The coder of the lifting wavelet transform has two steps. First, an input signal is processed by the lifting wavelet transform to generate a high-frequency signal and a low-frequency signal. Then, the two subband signals are processed by the LPC algorithm. In the first step, the integer lifting wavelet is used to avoid the rounding error [9].

3) Entropy Coding

Before entropy coding, the residual signal is normalized because there is often a few large-valued samples in the first part of residual signal. The residual signal is coded according to the probability distribution. The entropy coder first segregates the input sequence, calculates the mean of each subsequence, and then quantifies the mean value. The MSB of mean index and residual signal is coded according to the probability table, which is generated by the probability model [9]. Gaussian-like probability distribution is used to fit the probability distribution of predicted residuals.

2.3.3 Performance

Table 6 shows the average compression rate of AVS-LS and

▼ Table 6. Average compress rate of AVS1-lossless coder (%)

Coder	32 kHz/16 bit	44.1 kHz/16 bit	96 kHz/22 bit	192 kHz/22 bit
AVS-LS	51.44	47.98	46.37	35.22
Monkey EH	50.77	47.02	47.19	35.57
Monkey normal	52.23	48.32	49.08	36.31
TAK	52.29	47.95	47.97	35.01
MPEG4 ALS (RICE)	52.49	48.6	47.89	35.91
MPEG4 ALS (BGMC)	51.95	48.09	47.57	35.54
FLAC	53.31	49.38	51.68	40.33
WavPack	54.39	50.64	51.38	46.03

other popular lossless coding methods, including Monkey's Audio (extra high/normal), TAK (normal), ALS RM21 (RICE/BGMC 1024 sample), FLAC (normal), and WavPack (default) [9]. The table shows that the average compression rates of AVS-LS are equivalent or superior to others.

3 Progress of AVS2 Audio Coding Standard

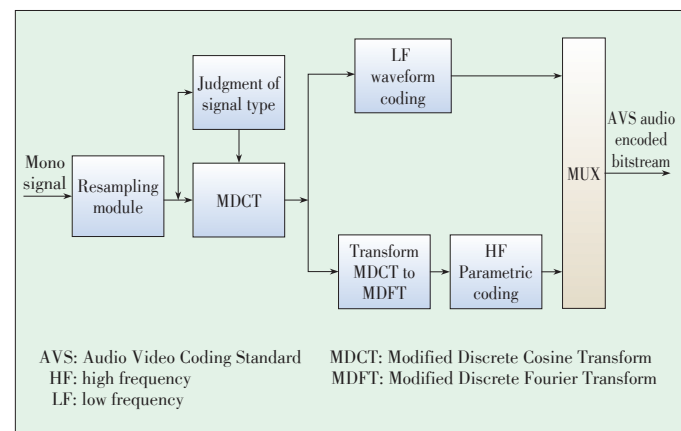
At the end of 2008, the AVS workgroup declared the project plan about the second generation of AVS standard, called "Information Technology - New Multimedia Coding Technology". AVS2 is an ongoing standardization effort for efficient coding performance compared to the previous standard (AVS1). It includes two compatible compression schemes, lossless audio coding and lossy audio coding. The later completely contains the first generation of AVS lossless audio coding. If the bit-stream is decoded completely the audio signals could be restored nearly transparent. Now the workgroup focuses more on surround sound encoding for super-high-definition television.

AVS1-P3 has good quality under high and medium bitrate but is not suitable for low bit rate coding. To solve this problem, the AVS Audio Subgroup has set up the project of AVS2-P3 in July 2013. The following three requirements are formally proposed for the AVS2 audio coding standard in the AVS M2845 in September 2011:

- 1) It is suitable for high, medium and low bit rate coding;
- 2) It can achieve hierarchical coding, or it can achieve the coarse scalable coding ability to meet different network application and the incremental download ability;
- 3) It has high scalability, low complexity and low decoding delay.

AVS2 audio called for proposals in December 2011. A hybrid audio coding technology of waveform coding and parameter coding proposed in the AVSM2977. AVSM2983 is adopted as the basic structure and reference model output (RM0) of the AVS2 audio coding (Fig. 11 [11]).

This scheme is based on the waveform coding in AVS1-P3 and uses frequency band extension technology, parameter ste-



▲ Figure 11. AVS2 encoder block diagram.

Review of AVS Audio Coding Standard

ZHANG Tao, ZHANG Caixia, and ZHAO Xin

reo coding technique and surround sound coding technique. The goal is to guarantee the quality of the medium and high bit rate coding and, at the same time, greatly increase the quality of the low bit rate audio coding. AVS2 codec is an initial reference release, so it will be improved for audio quality at the final release after more optimization.

At present, the AVS workgroup is collecting and evaluating the parametric stereo coding technology. In order to realize the low bit rate stereo and surround sound audio coding, AVS make full use of the inter-aural phase difference and intensity time difference (IPD/ITD), inter-aural intensity differences (IID), the binaural correlation, the binaural masking effect and other psychological mechanism.

4 Conclusions

AVS standardization is an indigenous innovation of China and has finished the transition from the development of technology to a maturing industry chain. The highly efficient and simple implementation of the AVS standard ensures its successful applications [12]. It fills up the current empty space in the development of our information industry and greatly promotes the development of audio-video industry in our nation. AVS2 audio coding standardization is ongoing. Although it is far from being a fully-fledged standard, it is moving forward steadily. With the continuous development and improvement of AVS standardization, it will serve better for our audio-video industry.

References

- [1] R. Hu, S. Chen, H. Ai, and N. Xiong, "AVS generic audio coding," in *Sixth International Conference on Parallel and Distributed Computing, Applications and Technologies*, Dalian, China, Dec. 2005, pp.679–683. doi: 10.1109/PD-CAT.2005.97.
- [2] R. Hu, Y. Zhang, and H. Ai, "Digital audio compression technology and AVS audio standard research," in *Intelligent Signal Processing and Communication Systems*, Hong Kong, China, Dec. 2005, pp. 757–759. doi:10.1109/

ISPACS.2005.1595520.

- [3] AVS Working Group. (2015, November 11). The Development of AVS [Online]. Available: <http://www.avs.org.cn/avsStandard.asp>
- [4] R. Hu and Y. Zhang, "Research and application on AVS-P10 audio standard," *Audio Engineering*, vol. 31, no. 7, pp. 65–70, Aug. 2007. doi: 10.3969/j.issn.1002-8684.2007.07.023.
- [5] T. Zhang, C.-T. Liu, and H.-J. Quan, "AVS1-M audio: algorithm and implementation," *EURASIP Journal on Advances in Signal Processing*, vol. 2011, no.4, pp. 307–323, Jan. 2011. doi:10.1155/2011/567304.
- [6] L. Jiang, R. Hu, X. Wang, et al., "AVS2 speech and audio coding scheme for high quality at low bitrates," in *IEEE International Conference on Multimedia and Expo Workshops (ICMEW)*, Chengdu, China, Jul. 2014, pp.1–6. doi: 10.1109/ICMEW.2014.6890693.
- [7] H. Zhang, "Research of AVS-P10 audio codec and implementation on DSP," M. S. Thesis, School of Electronic Information Engineering, Tianjin University, Tianjin, China, 2008.
- [8] W. Zhang, T. Zhang, L. Zhao, et al., "Performance analysis and evaluation of AVS-M audio coding," in *2010 International Conference on Audio Language and Image Processing (ICALIP)*, Shanghai, China, Nov. 2010, pp. 31–36. doi: 10.1109/ICALIP.2010.5685024.
- [9] W. He, Y. Gao, and T. Qu, "Introduction to AVS audio lossless coding/decoding standard," *IEEE COMSOC MMT E-Letter*, vol. 7, no. 2, pp. 21–29, Feb. 2012.
- [10] H. Huang, H. Shu, R. Yu, et al., "Introduction to AVS lossless audio coding standard: entropy coding," *Audio Engineering*, vol. 34, no. 12, pp. 69–71, 2010.
- [11] X. Pan, "Progress review and planning of AVS2 audio coding technology," in *Academic Forum and Academic Exchange Conference of Audio Engineering*, Ningbo, China, 2014.
- [12] W. Gao, Q. Wang, and S. Ma, "Digital audio video coding standard of AVS," *ZTE Communications*, vol. 12, no. 3, pp. 6–13, Aug. 2006.

Manuscript received: 2015-11-15

Biographies

ZHANG Tao (Zhangtao@tju.edu.cn) received his master's degree and PhD from Tianjin University, China in 2001 and 2004. He is an associate professor in the Texas Instruments DSP Collaboration Lab, Tianjin University. His principal interests are in acoustic signal processing, auditory model, and speech enhancement.

ZHANG Caixia (zhangcaixia@tju.edu.cn) is currently pursuing the master's degree at the School of Electronic Information Engineering, Tianjin University, China. Her research interests include audio coding and audio information hiding.

ZHAO Xin (zhaoxin_tju@126.com) is currently pursuing the master's degree at the School of Electronic Information Engineering, Tianjin University, China. His research interests include audio processing and DSP application.