OVERVIEW OF IEEE 1857 VIDEO CODING STANDARD

Siwei Ma, Shiqi Wang, Wen Gao
{swma,sqwang, wgao}@pku.edu.cn
Institute of Digital Media, Peking University

ABSTRACT

IEEE 1857 is a multi-part standard for multimedia data compression and transmission in a highly efficient way under constraints, which include limited complexity and bandwidth. IEEE 1857 video part is organized in an application-oriented way that suitable tools are chosen to get a high compression ratio. This paper will give an overview of its six profiles (Main, Broadcasting, Enhanced, Portable, Surveillance Baseline, and Surveillance) which are designed respectively for normal video, high-definition video broadcasting, digital cinema, video communication among portable devices and surveillance video. In the paper, the background of application requirement and the key coding tools of IEEE 1857 video coding standards are introduced, and the performance comparison between IEEE 1857 and the state-of-the art coding standards are provided.

***Index Terms*—**IEEE 1857, AVS, internet video, surveillance, background picture

1. INTRODUCTION

Multimedia data has become the majority of data flooding the Internet, especially for video data. According to Cisco VNI report, mobile video traffic exceeded 50% for the first time in 2011, and internet video traffic is predicted to be 86% of all consumer Internet traffic in 2016. The video contents created and consumed by people in modern life are increasing exponentially. However, video compression and communication efficiency only increases linearly from one generation to the next. It poses a great challenge to the development of video compression and communication technologies.

In the past, video compression standards are mainly developed by two world organizations ISO/IEC MPEG and ITU-T VCEG. So far, after the development of several decades, many coding standards have been established successfully and used widely in various applications, such as MPEG-1, MPEG-2, H.264/AVC [1]. And the state-of-the-art HEVC (High Efficiency Video Coding) standard has shown as large as 50% or more bits saving over H.264/AVC with the comparable visual quality [2].

IEEE 1857 originated from China Audio Video Coding Standard (AVS), which was developed by the China AVS working group [3], founded in June 2002. The role of the group is to establish general technical standards for the compression, decoding, processing, and the representation of digital audio-video, thereby enabling digital audio-video equipment and systems with high-efficiency and economical coding/decoding technologies. After more than ten years, AVS has established a series of standards, and AVS standards are becoming more and more internationalized. One part of AVS video coding standard has been accepted as an option by ITU-TFGIPTV for IPTV applications in 2007. Recently IEEE 1857 developed under the organization of IEEE will be released soon.

IEEE 1857 is a multi-part standard for multimedia data compression and communication in a highly efficient way under constraints that include limited complexity and bandwidth, in particular for the emerging applications like video surveillance and internet video. Actually, IEEE 1857 video coding standard can be viewed as an internationalized AVS1 video coding standard, which was the first stage of AVS standard work and has been finished (AVS2 is still under developing). This paper will give an overview of the background and technical features of the IEEE 1857 video coding standard. In the remained of the paper we will abbreviate IEEE 1857 video coding standard to IEEE 1857 for simplification.

The rest of the paper is organized as follows: Section 2 gives a brief overview to the framework of IEEE 1857. Section 3 gives an introduction to the key technical features of IEEE 1857. Section 4 provides the experimental results for the performance testing of IEEE 1857, and Section 5 concludes the paper.

2. AN OVERVIEW OF IEEE 1857

IEEE 1857 video part is organized in an application-oriented way that suitable tools are chosen to get a high compression ratio. There are six profiles (Main, Surveillance Baseline and Surveillance, Portable, Broadcasting, Enhanced) in IEEE 1857, which are designed respectively for normal video, surveillance video, video communication among portable devices, high-definition video broadcasting and digital cinema respectively (*Note: In the formal specification of IEEE 1857, a combination of "Group" and "Category" is used to denote "Profile", where "Category" is defined for 2D or stereo video applications. Here we still use "Profile" for easing the reader's understanding*).

Table 1. History of IEEE 1857/AVS Standard

Time	Profile (Group)	Applications and Major Coding Tools
Dec. 2003	Main	Broadcasting; 8x8 block based intra prediction, transform and deblock filter; variable block size motion compensation (16x16~8x8)
Jun. 2008	Surveillance Baseline	Video surveillance; Background-predictive picture for video coding, Adaptive Weighting Quantization (AWQ), core frame coding
Sep. 2008	Enhanced	Digital Cinema; Context Binary Arithmetic Coding (CBAC), Adaptive Weighting Quantization (AWQ)
Jul. 2009	Portable	Mobile video communication; 8x8/4x4 block transform
Jul. 2011	Survaillance	Video surveillance; Background picture model based coding
May 2012	Broadcasting	High definition TV broadcasting; CBAC, AWQ, enhanced field coding

Table 1 shows the development history of IEEE 1857/AVS standard. Main Profile is the first defined profile and mainly focuses on digital video applications like commercial broadcasting and storage media, including high-definition video applications. It has been approved as the national standard in 2006 in China. After that, an extended work of Main Profile, called Enhanced profile, was started to further improve the coding efficiency of Main Profile, which targets to fulfill the needs of multimedia entertainment, such as movie compression for high-density storage. Surveillance Baseline and Surveillance Profile focus on standardizing the solutions for the video surveillance applications, considering the characteristics of surveillance videos, i.e. the random noise, relatively lower encoding complexity affordable, and friendliness to events detection and searching required. Portable profile is defined in targeting to mobility video applications featured with lower resolution, low computational complexity and robust error resiliency ability to meet the wireless transporting environment. The latest developed Broadcasting Profile is also an improvement of Main Profile, which targets to high quality High Definition TV (HDTV) broadcasting, and it has been approved as industry standard by the State of China Broadcasting Film and Television Administration in Jul. 2012.

IEEE1857/AVS video encoder is developed based on the traditional transform/prediction hybrid framework. The coding framework has been widely used in the previous video coding standards. However, compared to the preceding coding standards, IEEE 1857/AVS standards are more applications oriented, which provide a good tradeoff between the performance and complexity for the applications, because all coding tools are selected by jointly considering the coding complexity and performance gain for the target applications. For example, in Main Profile, both the intra prediction and transform are 8x8 block based, because it was found that the blocks smaller than 8x8 do not have much performance improvement on high definition videos. On the contrary, smaller block size partition was used for mobile video for better prediction in Portable Profile. Moreover, for surveillance video, it was found that background picture prediction can improve the coding efficiency significantly. And background picture model based coding was proposed for surveillance video coding, as shown in Fig. 1. We will give a brief introduction to the background picture based coding technique later in the later sections.

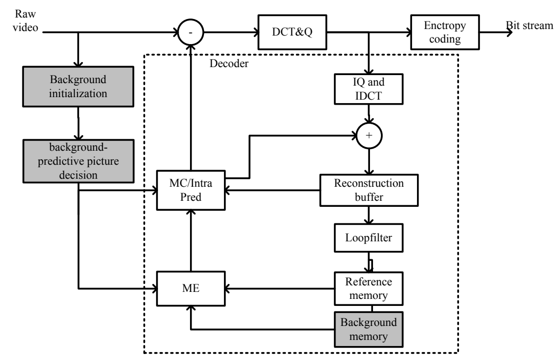

Fig.1. Background picture based surveillance video coding scheme [4].

3. KEY TECHNICAL FEATURES OF IEEE 1857

Due to the space limitation, we will give a brief introduction to the key technical features of IEEE 1857 and more details can be found in the previous overview papers [5][6] and book chapters [7]. The major coding tools in IEEE 1857 are summarized as follows:

- Intra prediction
- Variable Block-size Motion Compensation (VBMC)
- Multiple reference picture inter prediction
- Quarter/eighth sub-pixel motion interpolation
- Improved coding modes for B picture
- Integer transform and quantization
- Entropy coding
- Loop filter
- Background picture prediction (in Surveillance Baseline and Surveillance Profile)

In the remained of this section, we will introduce the above coding tools one by one. Moreover, as the Surveillance Profile is the latest developed profile and it achieves significant performance improvement for surveillance video, we will talk more about the Surveillance Profile in the paper.

Intra Prediction

Spatial domain intra prediction has been used in video coding standards since H.264/AVC. In IEEE 1857, the intra prediction is based on 8×8 luma blocks for Main, Enhanced, Broadcasting and Surveillance related Profiles. While adaptive 4×4 and 8×8 luma blocks based prediction is used for Portable Profile. The reason for that is mobile communications cover from low resolution coding to high resolution coding. On the contrary, other profiles mainly targeted to high resolution, such as Standard/High Definition (SD/HD) coding where the smaller block prediction is not as efficient as the bigger block prediction for coding performance improvement. Moreover, IEEE 1857 has made some delicate design for the intra coding modes, e.g. reducing the number of intra prediction modes for high resolution video coding to reduce the coding complexity.

Variable Block-size Motion Compensation (VBMC)

Variable Block-size Motion Compensation (VBMC) is very efficient for prediction accuracy and coding efficiency improvement; however it is also the most costly in video encoder due to the high cost of motion estimation and the rate-distortion optimization mode decision. In IEEE 1857, the minimum macroblock partition is 8×8 block, because the block size smaller than 8×8 does not produce much improvement, especially for high resolution coding.

Multiple Reference Picture Inter Prediction

Multiple reference picture coding can further improve coding efficiency compared to using one reference. In general, two or three reference pictures give almost the best performance and more reference pictures will not bring significant performance improvement but increase the complexity greatly. IEEE 1857 restricts the maximum number of reference picture to be 2. Actually, for IEEE 1857 setting the maximum number of reference to be 2 does not increase the reference buffer size relative to MPEG-1 and MPEG-2. In the previous video coding standards, such as MPEG-1 and MPEG-2, although P pictures use only one previous picture to predict the current picture, B pictures use one previous picture and one future picture as references, therefore the reference buffer size in a decoder has to be twice of the picture size. Compared with MPEG-1 and MPEG-2, IEEE 1857 can improve coding efficiency while using the same reference buffer size.

Quarter/Eighth Sub-pixel Interpolation

Fractional pixel motion accuracy has already been used since MPEG-1. Since MPEG-4 Advanced Simple Profile, quarter pixel interpolation has been developed. For fractional pixel interpolation, the interpolation filter has important effect on the coding efficiency. In H.264/AVC standard, a 6-tap filter is used for half pixel interpolation. But in IEEE 1857, a 4-tap filter is used for half pixel interpolation in all profiles. This is because the 6-tap filter does better for low resolution video, such as QCIF and CIF [7]. But a 4-tap filter can achieve similar performance with 6-tap filter while with much lower computing and memory access complexity. For Portable Profile, an eighth pel interpolation filter is added to improve the coding efficiency of low resolution video.

Improved Coding Modes for B Picture

The direct mode previously existing in H.263+ and MPEG-4 for global motion is improved in IEEE 1857. In H.264/AVC, it is classified as temporal and spatial direct mode according to the motion vector derivation scheme. The temporal direct mode and spatial direct mode are independent of each other. However, in IEEE 1857, the temporal direct mode and spatial direct mode are combined together. In the prediction process of direct mode, spatial prediction will be used when the co-located macroblock for temporal prediction is intra coded.

Moreover, in IEEE 1857, a new symmetrical mode is used for B picture coding to replace the traditional bi-direction coding. For the symmetric mode, only a forward motion vector is coded and the backward motion vector is derived from the forward motion vector. That is to say at most one motion vector is coded for a block in B picture in IEEE 1857.

Transform and Quantization

From H.264/AVC, the traditional float DCT is replaced with integer transform. In general, a larger size transform has better energy compaction property, while a smaller size transform has the advantages of reducing ringing artifacts at edges and discontinuities. In the development of H.264/AVC, adaptive block transform (ABT) was proposed and has been adopted into H.264/AVC High Profile. As larger transform size is more efficient for high resolution coding, an 8×8 integer transform is used in all profiles in IEEE 1857. For Portable Profile, an additional 4x4 transform is added for intra 4x4 block based prediction. Moreover, for Enhanced Profile and Broadcasting Profile, Adaptive Weighting Quantization (AWQ) technique was proposed to improve the visual quality with lower bit rate.

Entropy Coding

In H.264/AVC, two entropy coding schemes for transformed coefficients are included: CAVLC and CABAC. IEEE 1857 also has two entropy coding schemes to achieve different level compression efficiency with different coding complexity. One is Context-based 2D–VLC (C2DVLC) entropy coding and the other one is improved Context-based Adaptive Binary Arithmetic Coding (CBAC) [7]. Both two entropy coding methods are supported in all profiles except for Broadcasting Profile, where only CBAC is supported. The readers are referred to [9] for the details of CBAC.

Loop Filter

Block-based video coding often produces blocking artifacts especially at low bit rates. Similar as H.264/AVC, IEEE 1857 also adopts adaptive in-loop deblocking filter to improve the decoded visual quality, which handles 8×8 block boundaries. In IEEE 1857, the boundary may be between two 8×8 luma or chroma blocks except for picture or slice boundaries. For each edge, boundary strength (Bs) is derived based on the macroblock type, motion vector, and reference index and different filters will be selected as well.

Background picture prediction

In IEEE 1857 Surveillance Baseline Profile, background picture and background-predictive picture are defined to further exploit the temporal redundancy, suppress background noise and facilitate video event generation such as object segmentation and motion detection. The background picture is similar to the long term reference picture in H.264/AVC, however, in AVS it is explicitly coded as a special I-picture. The reconstructed background picture is stored in a separate background memory. The background-predictive picture is a special P-picture, which is coded by only being predictable from the reconstructed background picture.

In IEEE 1857 Surveillance Profile, the coding efficiency of background picture prediction coding was further improved by incorporating a background picture model. Firstly, the background frame is generated by the background modeling procedure and then encoded, where a small quantization parameter value is set to keep the quality of the background frame. Then, for the sequenced input frames, a difference frame is generated by the difference calculation module using a reconstructed background frame. Finally, the difference frame is coded by a 9-bit encoder. Both the coded background frames and the difference frames are written into the bit-stream. In decoding, the sequence is reconstructed through the difference compensation module using the decoded background and difference frames.

4. CODING PERFORMANCE OF IEEE 1857

The performance of IEEE 1857 has been well tested in the development of AVS standards, especially for Main Profile, where all the experimental results show that IEEE 1857/AVS achieve the similar performance as H.264/AVC with similar or lower complexity [7], e.g. less intra prediction modes and inter block partitions. For the recent developed profiles, the performance of Broadcasting Profile is close to H.264/AVC High Profile. But for Surveillance Profile, IEEE 1857 outperforms H.264/AVC significantly.

As shown in Fig. 2, JM-COMMON and JM-OPT are two different configurations for H.264/AVC reference software JM16.0, as tabulated in Table 2. JM-COMMON is a commonly used high-profile configuration, and JM-OPT is a specially configured encoder, which is more efficient for surveillance videos with QP=0 coded long term intra frame and GOP-Length=3000 (this configuration achieves very high performance according to experiments). *BgEncoder* is IEEE 1857 coder configured with Surveillance Profile. From the curves in Fig. 2, it can be seen that *BgEncoder* does much better than H.264/AVC even tested with near optimal configurations. Moreover, the complexity of *BgEncoder* is similar to JM encoder, or even less at low bitrate coding. So, we believe the background picture model based coding in IEEE 1857 would be very useful for surveillance video coding in the future. And we also believe for a specific video application or a specific kind of video data, the space for compression is still huge to be explored in the future. The background picture model based surveillance coding is a good example so far.

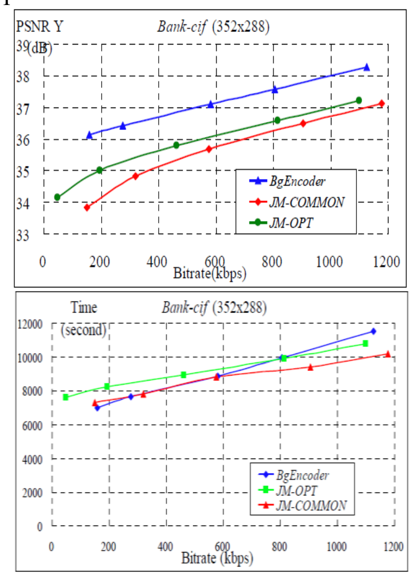

Fig. 2. Rate distortion performance (left) and the computation complexity (right) comparison of the *BgEncoder* and H.264/AVC under different configurations [8].

Table 2. Testing configurations for JM-COMMON and JM-OPT

Encoder	I slice QP	Long term reference frame	Open GOP	Reference reordering	Reference frame number	GOP length
JM-COMMON	/	Disable	Enable	Enable	5	30
JM-OPT	0	Enable	Disable	Disable	5+1 (long term)	3000

5. CONCLUSION

This paper gives an overview of IEEE 1857 from the development history to the major coding tools. Compared to the other video coding standards, IEEE 1857 is organized in an application-oriented way that suitable tools are chosen to get high compression efficiency. Especially for surveillance video coding, IEEE 1857 achieves significant coding bits saving with background picture model based coding technique. The future of IEEE 1857 is targeting on a higher compression ratio, as well as new features for emerging applications.

6. ACKNOWLEDGMENT

This research is supported by the National Science Foundation of China (61121002, 61103088), National High-tech R&D Program of China (863 Program, SS2012AA010805) and the 973 program (2009CB320900) of China, which are gratefully acknowledged.

7. REFERENCES

[1] T. Wiegand, G. Sullivan, G.Bjøntegaard, A. Luthra: Overview of the H.264/AVC Video Coding Standard. IEEE Transactions on Circuits and Systems for Video Technology, 13(7) 560-576 (2003).

[2] J. R. Ohm, G. J. Sullivan, F. Bossen, T. Wiegand, V. Baroncini, M. Wien, and J. Xu,"JCT-VC AHG report: HM subjective quality investigation (AHG22)", JCTVC-H0022, San José, CA, Feb., 2012.

[3] AVS Working Group Website: http://www.avs.org.cn.

[4] W. Gao, S. Ma, L. Zhang, L. Su, D. Zhao, "AVS Video Coding Standard," Intelligent Multimedia Communication: Techniques and Applications, ISBN 978-3-642-11685-8, vol.280/2010, pp.125-166, Jan. 2010.

[5] F. Liang, S. Ma, F. Wu, "Overview of AVS video standard," IEEE International Conference on Multimedia and Expo, pp.423-426, 2004.

[6] L. Yu, F. Yi, J. Dong, C. Zhang, "Overview of AVS-Video: Tools, Performance and Complexity," Visual Communication and Image Processing, 2005.

[7] S. Ma, L. Zhang, D. Zhao, W. Gao, "An Application-Oriented Video Coding Standard," Chapter 10, Multimedia Image and Video Processing, Second Edition, CRC Press 2012.

[8] X. Zhang, L. Liang, Q. Huang, Y. Liu, T. Huang, and W. Gao, "An efficient coding scheme for surveillance videos captured by stationary cameras," Proceedings of SPIE conference on Visual Communication and Image Processing, vol.7744, pp.77442A-1-10, 2010.

[9] L. Zhang, Q. Wang, N. Zhang, D. Zhao, X. Wu, W. Gao, "Context-based Entropy Coding in AVS Video Coding Standard," Signal Processing: Image Communication, vol.24, no.4, pp.263-276, Apr. 2009.