Volume 5, Issue 5, MAY 2015 ISSN: 2277 128X

International Journal of Advanced Research in Computer Science and Software Engineering

www.ijarcsse.com

Research Paper

Available online at: www.ijarcsse.com

Overview of Framework for Spatially Scalable Video Coding

Snehal N Chaudhari
Communication Engineering,
MIT Aurangabad, Maharashtra, India

Prof. T. D. Shep
Electronics& Communication Engineering,
MIT Aurangabad, Maharashtra, India

Abstract: With the introduction of the H.264/AVC video coding standard, significant improvements have recently been demonstrated in video compression capability. The Joint Video Team of the ITU-T Video Coding Experts Group (VCEG) and the ISO/IEC Moving Picture Experts Group (MPEG) has now also standardized a scalable video coding (SVC) extension of the H.264/AVC standard. SVC enables the transmission and decoding of partial bit streams to provide video services with lower temporal or spatial resolutions or reduced fidelity while retaining a reconstruction quality that is high relative to the rate of the partial bit streams. Hence, SVC provides functionalities such as graceful degradation in loss transmission environments as well as bit rate, format, and power adaptation. These functionalities provide enhancements to transmission and storage applications. SVC has achieved significant improvements in coding efficiency with an increased degree of supported scalability relative to the scalable profiles of prior video coding standards. This paper provides an overview of the basic concepts for extending H.264/AVC towards SVC. Moreover, the basic tools for providing temporal, spatial, and quality scalability are described in detail and experimentally regarding their efficiency and complexity.

Keywords- SVC, H.264/AVC (Advance video coding), MPEG-4, VCEG

I. INTRODUCTION

Video coding today is used in a wide range of applications ranging from multimedia messaging, video telephony and videoconferencing over mobile TV, and wireless and Internet video streaming to standard- and high-definition TV broadcasting. In particular, the Internet and wireless networks gain more and more importance for video applications. Video transmission in such systems is exposed to variable transmission conditions, which can be dealt with using scalability features. Furthermore, video content is delivered to a variety of decoding devices with heterogeneous display and computational capabilities. In these heterogeneous environments, flexible adaptation of once-encoded content is desirable, at the same time enabling interoperability of encoder and decoder products from different manufacturers.

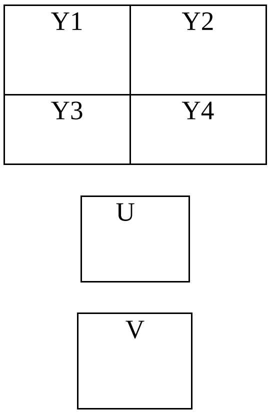
The objective in the development of SVC was to enable the encoding of a high-quality video bit stream that contains one or more subset bit streams that can themselves be decoded with a complexity and reconstruction quality similar to that achieved using the existing H.264/AVC design with the same quantity of data as that in the subset bit stream. Since the original H.264/AVC specification already includes the basic features necessary to enable scalability in terms of frame rate, the main objective was to add scalability in terms of picture size and reconstruction quality (fidelity). At the same time, the objective was to allow for straightforward and very-low-complexity manipulation and adaptation of scalable bit streams. Overall, the objective was to ensure the benefit of the scalable coding scheme compared to a simultaneous transmission of single-layer bit streams with different picture sizes and bit rates as required by an application, a method also referred to as simulcast. Scalable video coding has been an active research and standardization area for at least 20 years. The prior international video coding standards H.262 | MPEG-2 Video [3], H.263 [4], and MPEG-4 Visual [5] already include several tools by which the most important scalability modes can be supported. However, the scalable profiles of those standards have rarely been used. Reasons for that include the characteristics of traditional video transmission systems as well as the fact that the spatial and quality scalability features came along with a significant loss in coding efficiency as well as a large increase in decoder complexity as compared to the corresponding non-scalable profiles. It should be noted that two or more single-layer streams, i.e., non-scalable streams, can always be transmitted by the method of simulcast, which in principle provides similar functionalities as a scalable bit stream, although typically at the cost of a significant increase in bit rate. Moreover, the adaptation of a single stream can be achieved through trans-coding which is currently used in multipoint control units in video conferencing systems or for streaming services in 3G systems. Hence, a scalable video codec has to compete against these alternatives.

II. RELATED WORKS

International video coding standards such as H.261, MPEG-1, MPEG-2 Video, H.262, MPEG-4 Visual, and H.264/AVC have played an important role in the success of digital video applications. They provide interoperability among products from different manufacturers while allowing a high flexibility for implementations and optimizations in various application scenarios. The H.264/AVC specification represents the current state of- the-art in video coding. Compared to prior video coding standards, it significantly reduces the bit rate necessary to represent a given level of perceptual quality— a property also referred to as increase of the coding efficiency. The desire for SVC, which allows on the fly adaptation to certain application requirements such as display and processing capabilities of target devices, and varying transmission conditions, originates from the continuous evolution of receiving devices and the increasing usage of transmission systems that are characterized by a widely varying connection quality. Scalability has already been present in the video coding standards MPEG-2 Video, H.263, and MPEG-4 Visual in the form of scalable profiles. However, the provision of scalability in terms of picture size and reconstruction quality in these standards comes with a considerable growth in decoder complexity and a significant reduction in coding efficiency (i.e., bit rate increase for a given level of reconstruction quality) as compared to the corresponding non scalable profiles. These drawbacks, which reduced the success of the scalable profiles of the former specifications, are addressed by the new SVC amendment of the H.264/AVC standard.

A) H.261-A Video Coding Standard for ISDN Visual Telephony Applications -

The CCITT "Specialist Group on Coding for Visual Telephony" was given the task to standardize a video coding algorithm to support audiovisual Tele-services on ISDN, including videophone and videoconferencing applications. The resulting Recommendation H.261 video coding algorithm was designed and optimized for low target bit rate applications suitable for transmission of colour video over ISDN at *p* x 64 kb/s with low delay; here *p* specifies an integer number with values between 1 and 30 to allow transmission over more than one ISDN channel . Although the H.261 video coding standard was developed before the JPEG standard, both algorithms feature common elements. The H.261 standard specifies a Hybrid DCT/DPCM coding algorithm with motion compensation-which can be seen as a straightforward extension of the JPEG baseline algorithm towards inter frame coding. For an H.261 video coder the input video source consists of non interlaced colour frames of either CIF or quarter CIF (QCIF) format and a frame rate of 29.97 frames per second. Notice, that the CIF format specifies frames with 352 x 288 active luminance pixels *(Y)* and 176 x 144 pixels for each chrominance band *(U* or V), each pixel represented with fig 1. (a**)**. H.261 compatible decoders must be able to operate with QCIF frames-the CIF format is optional.



1.(a) With each macro-block (MB), information related to four luminance blocks *(Yl, Y2, Y3, Y4)* and two chrominance blocks *(U,V)* is coded. Each block contains 8 x 8 pixel

B) MPEG-1 - A Generic Standard for Coding of Moving Pictures and Associated Digital Storage Audio for Media at up to about ***1.5*** Mb/s

The video compression technique developed by MPEG- 1 covers many applications from interactive systems on CDROM to the delivery of video over telecommunications networks. Similar to JPEG, the MPEG-1 video coding standard is thought to be generic. To support the wide range of applications profiles a diversity of input parameters including flexible picture size and frame rate can be specified by the user. MPEG has recommended a constraint parameter set: Every MPEG-1 compatible decoder must be able to support at least video source parameters up to TV size: including a minimum number of 720 pixels per line, a minimum number of 576 lines per picture, a minimum frame rate of 30 frames/s and a minimum bit rate of 1.86 Mb/s. The standard video input consists of a non interlaced video picture format. It should be noted that by no means the application of MPEG-I is limited to this constrained parameter set. The MPEG-1 video algorithm has been developed with respect to the JPEG and H.261 activities. It was seek to retain a large degree of commonalty with the CCITT H.261 standard so that implementations supporting both standards were plausible. However, MPEG- 1 was primarily targeted for multimedia CD-ROM applications, requiring additional functionality supported by both encoder and decoder. Important features provided by MPEG- 1 include frame based random access of video, fast forwara7fast reverse (FF/FR) searches through compressed bit streams, reverse playback of video, and edit ability of the compressed bit stream.

C) MPEG-2 and ITU-T-H.262 Standards for Generic Coding of Moving Pictures and Associated Audio -

Worldwide MPEG-1 is developing into an important and successful video coding standard with an increasing number of products becoming available on the market. A key factor for this success is the generic structure of the standard supporting a broad range of applications and applications specific parameters. However, MPEG continued its standardization efforts in 1991 with a second phase (MPEG-2) to provide a video coding solution for applications not successfully covered or envisaged by the MPEG-1 standard. Specifically, MPEG-2 was given the charter to provide

video quality not lower than NTSC/PAL and up to CCIR 601 quality. Emerging applications, such as digital cable TV distribution, networked database services via ATM, digital VTR applications, and satellite and terrestrial digital broadcasting distribution, were seen to benefit from the increased quality expected to result from the new MPEG-2 standardization phase. Work was carried out in collaboration with the ITU-T SG 15 Experts Group for ATM Video Coding and in 1994 the MPEG-2 Draft International Standard was released. The specification of the standard is intended to be generic-hence the standard aims to facilitate the bit stream interchange among different applications, transmission and storage media. It is expected that the ITU-T Experts Group for ATM Video Coding will adapt the MPEG-2 International Standard, thus the ITU-T H.262 standard for ATM Video Coding will be identical **or** at least very similar to MPEG-2. Basically MPEG-2 can be seen as a superset of theMPEG-1 coding standard and was designed to be backward compatible to MPEG- l--every MPEG-2 compatible
decoder can decode a valid MPEG- 1 bit stream. Many video coding algorithms were integrated into a single syntax to meet the diverse applications requirements. New coding features were added by MPEG-2 to achieve sufficient functionality and quality, thus prediction modes were developed to support efficient coding of interlaced video. In addition scalable video coding extensions were introduced to provide additional functionality, such as embedded coding of digital TV and HDTV, and graceful quality degradation in the presence of transmission errors

D) MPEG-4 and ITU-TS Experts Groups for Coding of Video at Very Low Bit Rates -

Recent developments in telecommunications technologies and multimedia systems have prompted the demand for coding of audiovisual information at very low bit rates (5-64 kb/s) for storage and transmission. It is generally expected, that the delivery of video information over existing and future low-bandwidth communication networks will become increasingly important, such as audiovisual services operating over mobile radio networks as well as the PSTN. However, the success of these services in the market place will depend on the ability to encode video at very low bit rates with sufficient image quality. Existing video coding standards (e.g., H.261 or MPEG-1) have been optimized to achieve good video quality at bit rates higher than 64 kb/s. Accordingly the video quality provided by these algorithms is not sufficient for the applications envisaged at very low bit rates.

The ITU-TS Experts Group for Very Low Bit Rate Visual Telephony (ITU-TS is the former CCITT SG XV) has started activities in 1993 and has targeted its work into two areas: Near term work is directed towards a Rec.H.263 coding algorithm and a long term effort towards a H.263L coding scheme. It is expected that the first Rec.H.263 standards draft will be frozen in early 1995. To meet the rather short time schedule requirement the H.263 video coding algorithm will be an extension of Rec. H.261. However, to adapt H.261 for the videophone applications at very low bit rates between 9.6 kb/s . . . 28.8 kb/s, a number of significant changes will be required. Changes discussed at present include extended motion compensation accuracy and smaller motion vector search window sizes compared to H.261. The ITU work to develop Rec.H.263L is being accomplished in close collaboration with the IS0 MPEG-4 activity

E) H.264/AVC standard (video compression standard) –

Both the H.264/AVC standard and its SVC extension were developed by the Joint Video Team (JVT) consisting of experts from ITU-T's Video Coding Experts Group (VCEG) and ISO/IEC's Moving Pictures Experts Group (MPEG). VCEG is officially referred to as ITU-T SG16 and is a part of the Telecommunication Standardization Sector of the International Telecommunications Union (ITU-T). MPEG is officially referred to as ISO/IEC JTC1/ SC29/WG11, and it falls jointly under the International Organization for Standardization (ISO) and the International Electro technical Commission (IEC). In October 2003, MPEG issued a call for proposals on SVC technology. After the submitted proposals were analyzed regarding their potential for a successful future standard, an extension of H.264/AVC was chosen as a starting point. In January 2005, MPEG and VCEG agreed to jointly finalize the SVC project as an amendment of H.264/AVC within the JVT. The final draft of the SVC amendment was finalized in July 2007.

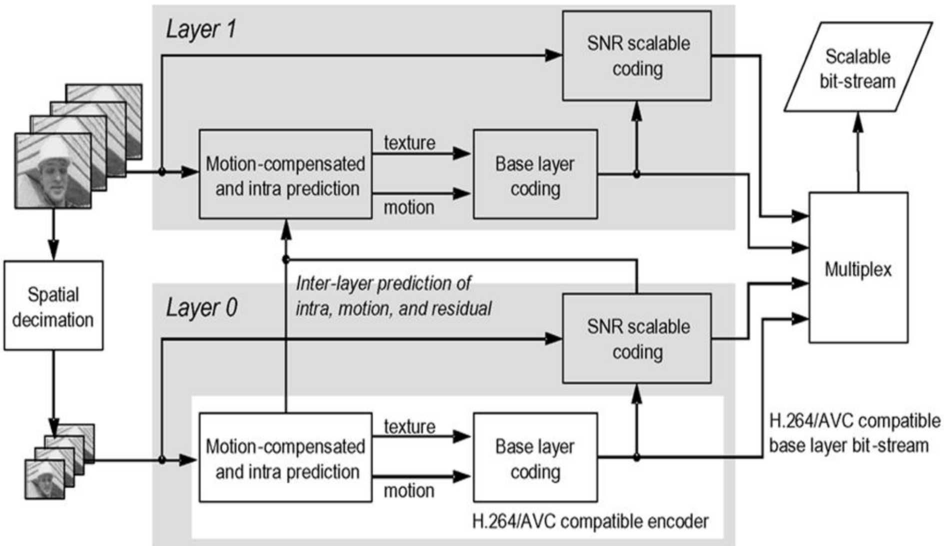


Fig.2(a). SVC encoder structure

SVC is specified as an Annex to H.264/AVC with a few backward-compatible changes of the base standard. As for all ITU-T and ISO/IEC video coding standards, only the bit stream syntax including certain constraints (e.g., for providing scalability features) and the decoding process are specified. The specification describes how encoded data that may represent a scalable sub stream are converted into a series of decoded pictures, and therefore it guarantees that compliant decoders produce identical outputs for the same bit stream. All other algorithms of a video transmission chain, including the encoding process, pre- and post processing, error concealment, the display process, and the processes for bit stream adaptation are outside the scope of the standard. Thereby, manufacturers are able to build optimized products while not interfering with the interoperability between different devices supporting SVC.

The SVC design shown at fig 2 (a) includes spatial, temporal, and fidelity scalability. Spatial scalability and temporal scalability describe cases in which subsets of the bit stream represent the source content with a reduced picture size (spatial resolution) and frame rate (temporal resolution), respectively. With fidelity scalability, the sub stream provides the same spatio-temporal resolution as the complete bit stream but with a lower fidelity. An SVC bit stream can provide a wide variety of combinations of these basic scalability types. Similar to MPEG-2 Video and MPEG-4 Visual, SVC supports spatial scalability with arbitrary resolution ratios. This means that the ratio of the picture sizes for the complete bit stream and the included sub streams are not restricted to a particular value. In addition, the pictures of bit stream subsets may contain additional parts beyond the borders of the pictures of the complete bit stream, or they may represent only a selected rectangular area of the pictures of the complete bit stream. The relation between the pictures of the complete bit stream and the pictures of the included sub streams may even be modified at any in time. The SVC design also supports a special mode of fidelity scalability, which allows a low-complexity rewriting of a fidelity scalable bit stream into a single layer H.264/AVC bit stream with identical output. Each SVC bit stream contains a subset bit stream—usually the one with the lowest spatial resolution and the lowest fidelity—that is compatible with a non scalable profile of H.264/AVC and can be decoded by legacy decoders. Since the support of fidelity and spatial scalability usually comes with an increase of the bit rate required for representing a given level of perceptual quality compared to the single-layer coding with H.264/AVC, the trade-off between the provided degree of scalability and coding efficiency can be adjusted according to the needs of an application.

III. CONCLUSION

This paper proposes an survey of scalable video coding technology with the goal of providing additional functionalities and improved coding efficiency is being investigated in the JVT. The current working fields are bit-depth scalability, chroma- format scalability, and fine-granular fidelity scalability with an improved drift control for low-delay coding.

REFERENCES

[1] H. Schwarz, D. Marpe, and T. Wiegand, “Overview of the scalable video coding extension of the H.264/AVC standard,” IEEE Trans. Circuits Syst. Video Technol., vol. 17, no. 9, pp. 1103–1120, Sep. 2007.

[2] C. A. Segall and G. J. Sullivan, “Spatial scalability within the H.264/AVC scalable video coding extension,” IEEE Trans. Circuits Syst., Video Technol., vol. 17, no. 9, pp. 1121–1135, Sep. 2007.

[3] C. A. Segall and A. Katsaggelos, “Resampling for spatial scalability,” in Proc. IEEE Int. Conf. Image Process. (ICIP*)*, Oct. 2006, pp. 181–184.

[4] K. Rose and S. L. Regunathan, “Toward optimality in scalable predictive coding,” IEEE Trans. Image Process., vol. 10, no. 7, pp. 965–976, Jul. 2001.

[5] J. Han, V. Melkote, and K. Rose, “Estimation-theoretic approach to delayed prediction in scalable video coding,” in Proc. 17th IEEE Int. Conf. Image Process. (ICIP*)*, Sep. 2010, pp. 1289–1292.

[6] J. Han, V. Melkote, and K. Rose, “A unified framework for spectral domain prediction and end-to-end distortion estimation in scalable video coding,” in *Proc*. 18th IEEE Int. Conf. Image Process. (ICIP), Sep. 2011, pp. 3278–3281.

[7] R. Zhang and M. L. Comer, “Efficient inter-layer motion compensation for spatially scalable video coding,” IEEE Trans. Circuits Syst., Video Technol., vol. 18, no. 10, pp. 1325–1334, Oct. 2008.

[8] R. Zhang and M. L. Comer, “Rate distortion analysis for spatially scalable video coding,” IEEE Trans. Image Process., vol. 19, no. 11, pp. 2947–2957, Nov. 2010.

