

Versatile Video Coding: a Next-generation Video Coding Standard

Seishi Takamura

Abstract

The scope of Subcommittee (SC) 29 of ISO/IEC (International Organization for Standardization/International Electrotechnical Commission) Joint Technical Committee 1 is coding of audio, picture, multimedia, and hypermedia information. Working Group (WG) 11, part of SC 29, is standardizing video coding, media transmission, streaming, audio coding, image/video retrieval, and genomic information coding. In April 2018, WG 11 initiated standardization of next-generation video coding called Versatile Video Coding (VVC), which is aimed at achieving higher compression, in conjunction with ITU-T (International Telecommunication Union - Telecommunication Standardization Sector) Study Group 19. This article introduces the background, target, and recent development status of VVC.

Keywords: ISO/IEC JTC 1/SC 29/WG 11, MPEG, Versatile Video Coding

1. Introduction

Subcommittee (SC) 29 (Coding of audio, picture, multimedia, and hypermedia information) is a subcommittee of the International Organization for Standardization/International Electrotechnical Commission (ISO/IEC), an international standardization body, Joint Technical Committee (JTC) 1, and it has been the hub of standardization activities related to the coding of multimedia information technology [1, 2]. There are two Working Groups (WGs), WG 1 and WG 11, under the auspices of SC 29. WG 11 is known as the Moving Picture Experts Group (MPEG), and it develops the standards for technologies such as video and audio coding, multiplexing/synchronizing audio and visual streams, high efficiency coding, three-dimensional video coding, and image/video retrieval. These standards are known as MPEG series standards.

The amount of Internet protocol (IP) traffic is increasing rapidly worldwide. It was reported that in 2017, the annual run rate for global IP traffic was 122 EB (exa (10^{18}) bytes) per month [3]. IP traffic is predicted to grow at a compound annual growth rate of 26% from 2017 to 2022, and to increase threefold (387 EB per month) by 2022. IP video traffic

accounted for 75% of all IP traffic in 2017, which is forecast to be 82% by 2022. Such video traffic content is already compressed (mainly by MPEG coding standards) down to a few hundredths of the original size. Thus, it is clear that not only video-based services but also network services as a whole would collapse without video coding standards. Also, with the trend for rapid growth of video traffic, more and more powerful compression techniques are necessary.

2. Latest video coding standard

Development of the most recent international video coding standard to date, H.265/MPEG-H HEVC (High Efficiency Video Coding), began in 2010 by the Joint Collaborative Team on Video Coding (JCT-VC) consisting of WG 11 and the International Telecommunication Union - Telecommunication Standardization Sector Study Group 16 (ITU-T SG16) Working Party 3, Question 6 (VCEG: Video Coding Experts Group). The initial version of the Final Draft International Standard (FDIS) was completed in January 2013. It is used in 4K/8K broadcasting and video transmission. More than 2 billion HEVC-compliant devices have been shipped worldwide.

Since the initial version of HEVC, various extensions and amendments have been published [4].

HEVC effectively achieves compression performance twice that or more than any of the previous standards. However, due to the strong demand for higher compression, in 2014 MPEG began working on achieving an even better video coding standard, aiming at a next-generation video coding standard named Future Video Coding (FVC) at that time.

At a meeting in Strasbourg, France, in October 2014, a brainstorming session was held inviting engineers from the communications, network service, and hardware fields as speakers. At that meeting, the requirements for FVC were discussed and defined as broadcasting, IP television, contribution quality video transmission, digital cinema, surveillance, mobile viewing, and virtual reality (VR), among others. In February 2015, JCT-VC began developing reference software for next-generation video coding called Key Technical Area (KTA).

At a meeting in Geneva, Switzerland, in October 2015, the Joint Video Exploration Team (JVET) between ISO/IEC and ITU-T was created, and this team subsequently inherited KTA and used it to develop reference software called Joint Exploration Model (JEM). MPEG issued the requirements for the functionalities and performance of FVC in June 2016 [5]. In January 2017, JEM was confirmed to have achieved a significant performance gain compared to the HEVC test model (HM) [6]. In April 2017, a Call for Evidence was issued [7].

In October 2017, a Call for Proposal (CfP) was issued [8]. At the same time, the team was renamed the Joint Video *Experts* Team (keeping the same abbreviation JVET). After the meeting, 26 CfP responses from 21 organizations were submitted. All the proposals were carefully compared via a subjective viewing test, and the results were compiled and reported at a meeting in San Diego, USA, in April 2018. The new standard is named ISO/IEC 23090 MPEG-I Part 3 Versatile Video Coding (VVC). The “I” in MPEG-I stands for Immersive Media. Based on the subjective results and proposed technological details, a technical description document (Working Draft 1: WD1) and reference software (VVC Test Model 1: VTM1) were created as a starting point.

3. Current standardization status of VVC and future development plan

VVC will support three types of video: Standard Dynamic Range (SDR), High Dynamic Range

(HDR), and 360° (VR-oriented, omnidirectional view). Various technologies are being proposed and evaluated in JVET’s exploration process. The target compression performance is a 30–50% bit-rate reduction compared to H.265/HEVC at the same subjective video quality.

JVET holds standardization meetings four times a year. At each meeting, proposed coding tools are intensively reviewed from subjective/objective aspects. Decisions are made on which tools will be adopted, and the WD is updated based on the consensus of attendees with various backgrounds. At the same time, implementation of adopted tools, encoding optimization (outside the scope of standardization) and performance enhancement, and bug fixes are carried out on the reference software VTM using an open/public repository.

Numerous contribution documents have been submitted so far, with 118 in April 2018 (beginning of standardization), 559 in July 2018, 690 in October 2018, 897 in January 2019, and 858 in March 2019. This is far more than the total submitted at the first five meetings on HEVC standardization.

After the March 2019 meeting, meetings will be held in Gothenburg, Sweden, (September 2019), Geneva (October 2019), Brussels, Belgium, (January 2020), Alpbach, Austria, (April 2020), and Geneva (June–July 2020), as well as a meeting in Rennes, France, (October 2020) to reach the FDIS stage.

The most recent reference software for VVC, VTM5.0, achieves a 33.14% bit-rate reduction (luminance (Y) Bjøntegaard Delta rate (BD-rate)), an encoding runtime of 6.71 times, and a decoding runtime of 1.03 times compared to HEVC reference software HM16.19, under the random access coding structure. More details are given in **Table 1**. Faster coding tools with more coding gain are demanded.

Some of the principal tools adopted in VVC so far are listed in **Table 2**. Among them, CST (chroma separate tree), CCLM (cross-component linear model), ALF (adaptive loop filter), AFF (affine motion compensation), MTS (multiple transform set), and DQ (dependent quantization) contribute to improving coding performance more than others. Many proposals, including those on neural network technologies, are intensively evaluated and selected at each meeting.

4. Future exploration of WG 11

This article has overviewed the background of VVC standardization, the latest trends, and future

Table 1. Coding performance of current VVC (VTM5.0) against HEVC (HM16.19).

Coding structure	Y BD-rate	Encoding runtime	Decoding runtime
All intra	−23.14%	22.46	1.04
Random access	−33.14%	6.71	1.03
Low delay B	−24.69%	4.82	0.89
Low delay P	−28.15%	4.39	0.92

Table 2. Principal tools adopted in HEVC and current VVC.

Category	H.265/HEVC	MPEG-I VVC (WD4) [9]
Partition structure	Maximum size of 64 x 64, quad-tree	Maximum size of 28 x 128, quad-/ternary-/binary-tree, CST (chroma separate tree)
Intra prediction	Direct current (DC) + Planar + 33 directional prediction, reference/boundary pixel filter	DC + Planar + 65 directional prediction + 28 wide angle + 3 CCLM (cross-component linear model) , PDPC (position dependent prediction combination), MLIP (multi-line intra prediction), CPR (current picture referencing)
Inter prediction	Skip, merge, weighted prediction	AFF (affine motion compensation) , CIIP (combined intra/inter prediction), triangular, BWA (bi-directional weighted average), decoder-side motion refinement (BDOF (bi-directional optical flow)), DMVR (decoder-side motion vector refinement), MV prediction (ATMVP (alternative temporal motion vector prediction), HMVP (history-based motion vector prediction)), AMVR (adaptive motion vector difference), PMC (pairwise merge candidate), MMVD (merge with motion vector difference)
Transform	Square transform (up to 32 x 32 size), fixed transform bases (discrete cosine transform (DCT)2, discrete sine transform (DST)7)	Square and rectangular transforms (up to 64 x 64 size), shape adaptive transform, MTS (multiple transform set) (DCT2, DST7, DCT8)
In-loop filter	Deblocking filter (DF), sample adaptive offset (SAO)	ALF (adaptive loop filter) , large-block adaptive DF, LADF (luma-adaptive DF), SAO
Quantization, entropy coding	Fixed quantization, position context	DQ (dependent quantization) , template context

plans. Other than VVC, current MPEG standardization efforts continue on video retrieval standard ISO/IEC 15938 MPEG-7 Part 15: Compact Descriptors for Video Analysis (CDVA), neural network coding, and genomic information coding standard ISO/IEC 23092 MPEG-G.

As for point cloud compression (PCC) standards, Video-based PCC (V-PCC) of MPEG-I Part 5 is scheduled to reach the FDIS stage in January 2020, and Geometry-based PCC (G-PCC) of MPEG-I Part 9 is scheduled to reach the FDIS stage in April 2020. Additionally, work is underway on high-density light field video coding, 6DoF (degree of freedom) video coding with which the user walks a few steps away from a central position in a scene, and 3DoF+ video coding, with which the user does not walk in the scene (e.g., sitting on a chair) [10].

For 3DoF+, a CfP was issued at the meeting in Marrakech, Morocco, in January 2019 [11], and it is

scheduled to reach the FDIS stage of MPEG-I Part 7 in July 2020. Moreover, standardization of Low Complexity Video Coding Enhancements, which enables two-layered spatial scalability [12], and Essential Video Coding (ISO/IEC 23094 MPEG-5 Part 1) [13] are under development. MPEG's coverage and activity is ever increasing.

References

- [1] ISO/IEC JTC 1/SC29, <https://www.itscj.ipsj.or.jp/sc29/>
- [2] S. Takamura, "Image and Video Coding Related Standardization Activities of ISO/IEC JTC 1/SC 29," NTT Technical Review, Vol. 13, No. 10, 2015.
<https://www.ntt-review.jp/archive/ntttechnical.php?contents=ntr201510gls.html>
- [3] Cisco Visual Networking Index: Forecast and Trends, 2017–2022 White Paper,
<https://www.cisco.com/c/en/us/solutions/collateral/service-provider/visual-networking-index-vni/white-paper-c11-741490.html>
- [4] JCT-VC, "Text of ISO/IEC FDIS 23008-2 (4th edition)," N18277, Jan. 2019.

- [5] ISO/IEC JTC 1/SC 29/WG 11, "Requirements for a Future Video Coding Standard v4," N16359, June 2016.
- [6] T. Suzuki, "BoG Report on Test Material," JVET-E0132, Jan. 2017.
- [7] JVET, "Joint Call for Evidence on Video Compression with Capability beyond HEVC," JVET-F1002, 6th JVET Meeting, Apr. 2017.
- [8] JVET, "Joint Call for Proposals on Video Compression with Capability beyond HEVC," JVET-H1002, 8th JVET Meeting, Oct. 2017.
- [9] B. Bross, J. Chen, and S. Liu, "Versatile Video Coding (Draft 4)," JVET-M1001, 13th JVET Meeting, Jan. 2019.
- [10] ISO/IEC JTC 1/SC 29/WG 11, "Summary on MPEG-I Visual Activities," N18166, Jan. 2019.
- [11] ISO/IEC JTC 1/SC 29/WG 11, "Call for Proposals on 3DoF+ Visual," N18145, Jan. 2019.
- [12] ISO/IEC JTC 1/SC 29/WG 11, "Call for Proposals for Low Complexity Video Coding Enhancements," N17944, Oct. 2018.
- [13] ISO/IEC JTC 1/SC 29/WG 11, "Working Draft 1 of Essential Video Coding," N18283, Jan. 2019.



Seishi Takamura

Senior Distinguished Engineer, Video Coding Group, NTT Media Intelligence Laboratories.

He received a B.E., M.E., and Ph.D. from the Department of Electronic Engineering, Faculty of Engineering, the University of Tokyo in 1991, 1993, and 1996. His current research interests include efficient video coding and ultrahigh-quality video processing. He has fulfilled various duties in the research and academic community in current and prior roles, including serving as Associate Editor of the Institute of Electrical and Electronics Engineers (IEEE) Transactions on Circuits and Systems for Video Technology (2006–2014), Editor-in-Chief of the Institute of Image Information and Television Engineers (ITE), Executive Committee Member of the IEEE Region 10 and Japan Council, and Director-General Affairs of ITE. He has also served as Chair of ISO/IEC JTC 1/SC 29 Japan National Body, Japan Head of Delegation of ISO/IEC JTC 1/SC 29, and as an International Steering Committee Member of the Picture Coding Symposium. From 2005 to 2006, he was a visiting scientist at Stanford University, CA, USA.

He has received 46 academic awards including ITE Niwa-Takayanagi Awards (Best Paper in 2002, Achievement in 2017), the Information Processing Society of Japan (IPSJ) Nagao Special Researcher Award in 2006, PCSJ (Picture Coding Symposium of Japan) Frontier Awards in 2004, 2008, 2015, and 2018, the ITE Fujio Frontier Award in 2014, and TAF (Telecommunications Advancement Foundation) Telecom System Technology Awards in 2004 and 2008, and in 2015 with highest honors, the Institute of Electronics, Information and Communication Engineers (IEICE) 100-Year Memorial Best Paper Award in 2017, and the Kenjiro Takayanagi Achievement Award in 2019.

He is a Fellow of IEEE, a senior member of IEICE and IPSJ, and a member of Japan Mensa, the Society for Information Display, the Asia-Pacific Signal and Information Processing Association, and ITE.