Heiko Schwarz
and Mathias Wien

The Scalable Video Coding Extension of the H.264/AVC Standard

The Scalable Video Coding extension (SVC) of the H.264/MPEG-4 Advanced Video Coding (AVC) standard (H.264/AVC) is the latest amendment for this successful specification. SVC allows partial transmission and decoding of a bit stream. The resulting (decoded) video has lower temporal or spatial resolution or reduced fidelity while retaining a reconstruction quality that is close to that achieved using the existing single-layer H.264/AVC design with the same quantity of data as in the partial bit stream. SVC provides network-friendly scalability at a bit stream level with a moderate increase in decoder complexity relative to single-layer H.264/AVC. Furthermore, it provides the functionality of lossless rewriting of fidelity-scalable SVC bit streams to single-layer H.264/AVC bit streams. The SVC extension of H.264/AVC is suitable for video conferencing as well as for mobile to high-definition broadcast and professional editing applications.

The H.264/AVC standard was presented in this column in the March 2007 issue of *IEEE Signal Processing Magazine* (pp. 148–153). The current article focuses on the SVC extension in terms of technology, performance, and targeted application scenarios.

BACKGROUND

International video coding standards such as H.261, MPEG-1, MPEG-2 Video, H.263, 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 nonscalable 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.

[THIS ARTICLE FOCUSES ON THE SVC EXTENSION IN TERMS OF TECHNOLOGY, PERFORMANCE, AND TARGETED APPLICATION SCENARIOS.]

MOTIVATION

Video coding today is used in a wide range of applications ranging from multimedia messaging, video telephony and video conferencing 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.

OBJECTIVES

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

Digital Object Identifier 10.1109/MSP.2007.914712

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.

ISSUING BODY AND SCHEDULE

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 Q.6 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 Electrotechnical 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.

TARGET APPLICATIONS

SVC extends the target applications of the H.264/AVC standard to enable video transmission with heterogeneous clients and multicast with clients of different capabilities (e.g., for video conferencing or broadcasting scenarios from mobile up to high-definition applications). For applications under variable channel conditions, it enables robust transmission with graceful degradation in the presence of errors and bit stream adaptation using media-aware network elements. It is therefore suitable for video surveillance, erosion storage applications (where the reconstruction quality of the stored video may be readjusted according to age or storage space requirements), video streaming, and professional video editing and manipulation applications.

STRUCTURE OF THE STANDARD

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 substream 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 postprocessing, 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.

TECHNOLOGY

FUNCTIONALITIES

The SVC design 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 substream 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 substreams 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 substreams may even be modified at any point in time.

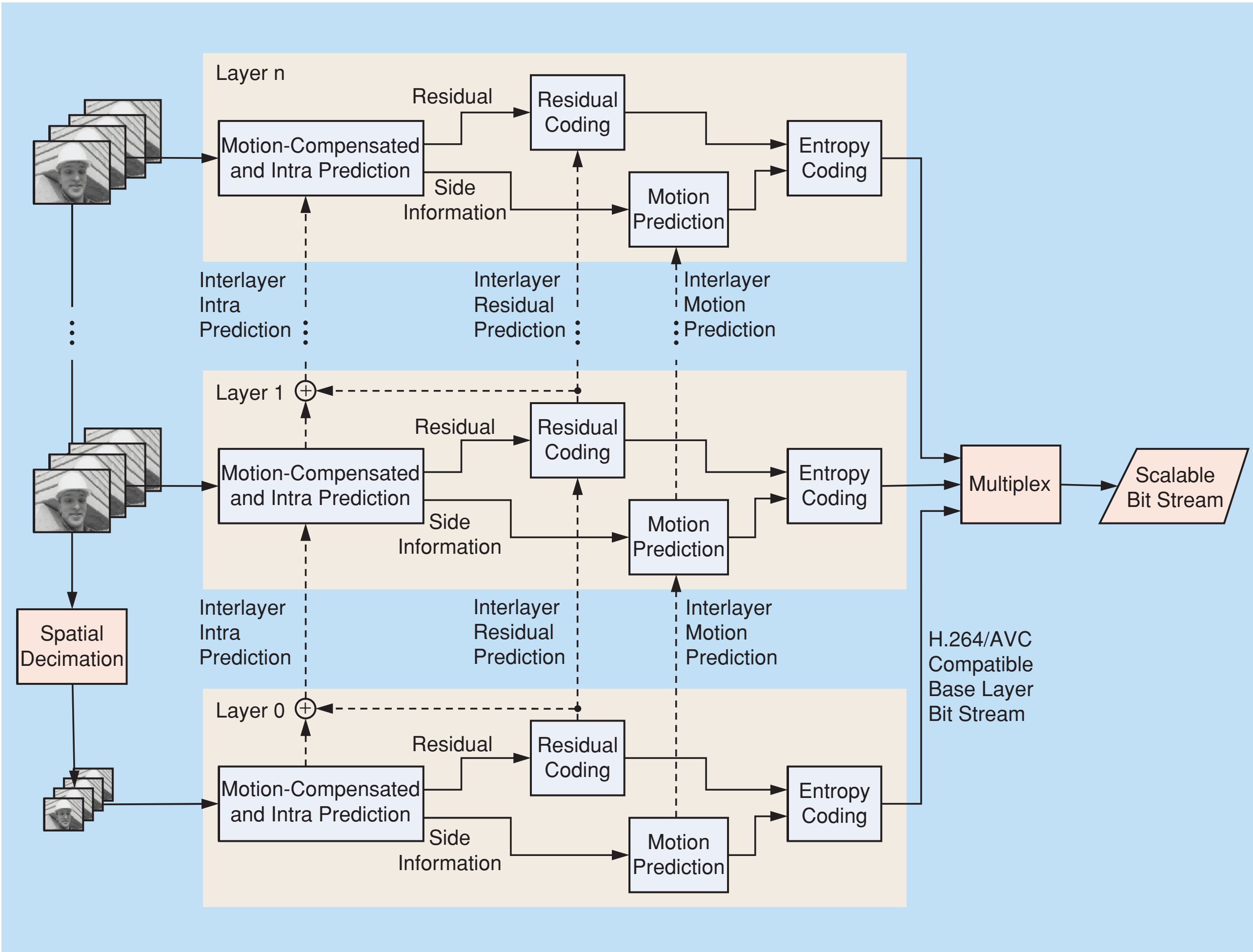
The SVC design also supports a special mode of fidelity scalability, which allows a low-complexity rewriting of a fidelity scalable bit stream into a single-layer H.264/AVC bit stream with identical output.

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 nonscalable 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 tradeoff between the provided degree of scalability and coding efficiency can be adjusted according to the needs of an application.

ARCHITECTURE

Similar to the underlying H.264/AVC standard, the SVC design includes a video coding layer (VCL) and a network abstraction layer (NAL). While the VCL represents the coded source content, the NAL formats the VCL representation and provides header information appropriate for conveyance by transport layers and storage media. The coded video data is organized into NAL units, which represent packets with an integer number of bytes. NAL units are classified into VCL NAL units, which contain coded video data, and non-VCL NAL units, which provide associated additional information. An access unit is a set of NAL units that results in exactly one decoded picture. A set of successive access units with certain properties constitutes a coded video sequence, which represents an independently decodable part of a bit stream.

Scalability is provided at the bit stream level. A bit stream with reduced spatio-temporal resolution and/or fidelity can be obtained by discarding NAL units or corresponding network packets



[FIG1] Simplified SVC encoder structure.

from a scalable bit stream. The NAL units that are required for decoding of a specific spatio-temporal resolution and bit rate are identified by syntax elements inside the NAL unit header or by a preceding so-called prefix NAL unit.

Although the standard actually specifies the decoding process, we focus on typical encoder techniques to explain the VCL design, since it is easier to understand. A simplified block diagram of an SVC encoder is illustrated in Figure 1. Each representation of the video source with a particular spatial resolution and fidelity that is included in an SVC bit stream is referred to as a layer (shaded area in Figure 1) and is characterized by a layer identifier. In each access unit, the layers are encoded in increasing order of their layer identifiers. For the coding of a layer, already transmitted data of another layer with a smaller layer identifier can be employed as described in the following paragraphs. The layer to predict from can be selected on an access unit basis and is referred to as the reference layer. The layer with a layer identifier equal to zero, which may only be present in some access units, is coded in conformance with one of the nonscalable H.264/AVC profiles and is referred to as the base layer. The layers that employ data of other layers for coding are referred to as enhancement layers. An enhancement layer is called a spatial enhancement layer when the spatial resolution changes relative to its reference layer, and it is called a fidelity enhancement layer when the spatial resolution is identical to that of its reference layer.

The number of layers present in an SVC bit stream is dependent on the needs of an application. SVC supports up to 128 layers in a bit stream. With the currently specified profiles, the maximum number of enhancement layers in a bit stream is limited to 47, and at most two of those can represent spatial enhancement layers.

The input pictures of each spatial or fidelity layer are split into macroblocks and slices. A macroblock represents a square area of 16×16 luma samples and, in the case of 4:2:0 chroma sampling format, 8×8 samples of the two chroma components. The macroblocks are organized in slices that can be parsed independently. For the purpose of intra prediction, motion-compensated prediction, and transform coding, a macroblock can be split into smaller partitions or blocks.

INTRALAYER CODING

Inside each layer, the SVC design basically follows the design of the underlying H.264/AVC standard for single-layer coding, which is described in the March 2007 issue of *IEEE Signal Processing Magazine* (pp. 148–153). The samples of each macroblock are either predicted by intrapicture or interpicture prediction. With intrapicture prediction, each sample of a block is predicted using spatially neighboring samples of previously coded blocks in the same picture. With interpicture prediction, the prediction signal of a partition is built by a spatially displaced region of a previously coded picture of the same layer. The residual representing the difference between the original and the prediction signal for a block is transformed using a decorrelating transform. The transform coefficients are scaled and quantized. The quantized transform coefficients are entropy coded with other information including the macroblock coding type, the quantization step size, and the intra prediction modes or the motion information consisting of identifiers specifying the employed reference pictures and corresponding displacement (or motion) vectors. The motion vector components are differentially coded using motion vectors of neighboring blocks as predictors. The decoded representation of the residual is obtained by inverse scaling and inverse transformation of the quantized transform coefficients. The obtained decoded residual is then added to the prediction signal, and the result is additionally processed by a deblocking filter before output and potential storage as a reference picture for interpicture coding of following pictures.

INTERLAYER CODING

In addition to the basic coding of H.264/AVC, SVC provides so-called interlayer prediction methods, which allow an exploitation of the statistical dependencies between different layers for improving the coding efficiency (reducing the bit rate) of enhancement layers. All interlayer prediction methods can be chosen on a macroblock or submacroblock basis, allowing an encoder to select the coding mode that gives the highest coding efficiency. For SVC enhancement layers, an additional macroblock coding mode is provided, in which the macroblock prediction signal is completely inferred from colocated blocks in the reference layer without transmitting any additional side information. When the colocated reference layer blocks are intracoded, the prediction signal is built by the potentially up-sampled (for spatial scalable coding) reconstructed intra signal of the reference layer—a prediction method also referred to as interlayer intra prediction. Otherwise, the enhancement layer macroblock is interpicture predicted as described above, but the macroblock partitioning—specifying the decomposition into smaller block with different motion parameters—and the associated motion parameters are completely derived from the colocated blocks in the reference layer. This is also referred to as interlayer motion prediction. For the conventional intercoded macroblock types of H.264/AVC, the (scaled) motion vectors of the reference layer blocks can also be used as replacement for the usual spatial motion vector predictor.

A further interlayer prediction method referred to as interlayer residual prediction targets a reduction of the bit rate required for transmitting the residual signal of intercoded macroblocks. With the usage of residual prediction, the (up-sampled) residual of the colocated reference layer blocks is subtracted from the enhancement layer residual (difference between the original and the interpicture prediction signal) and only the resulting difference, which often has a smaller energy then the original residual signal, is encoded using transform coding as described above. For fidelity enhancement layers, the interlayer intra and residual prediction are performed in the transform coefficient domain to avoid multiple inverse transform operations at the decoder side.

As an important feature of the SVC design, each spatial and fidelity enhancement layer can be decoded with a single motion compensation loop. For the employed reference layers, only the intracoded macroblocks and residual blocks that are used for interlayer prediction need to be reconstructed and the motion vectors need to be decoded. The computationally complex operations of motion-compensated prediction and deblocking (for intercoded macrobloccks) only need to be performed for the target layer to be displayed.

Similarly to H.264/AVC, temporal scalability can be achieved by partitioning the access units into a temporal base and one (or more) temporal enhancement layers and restricting the encoding structure for each access unit of a specific temporal layer so that only access units of the same or a coarser temporal layer are employed for interpicture prediction.

TOOLS

The SVC design includes all coding and error resilience tools of the underlying H.264/AVC standard. In addition, the following tools for scalable coding are included:

- interlayer intra prediction similar to MPEG-2 Video, H.263, and MPEG-4 Visual
- efficient new methods for interlayer prediction of motion and residual data
- low-complexity decoding with a single motion compensation loop
- resynchronization pictures for efficiently controlling the drift in fidelity scalable coding
- interlayer prediction in quantized transform coefficient domain for enabling low-complexity rewriting of fidelity scalable bit streams into single-layer H.264/AVC bit streams
- a system interface for easily identifying NAL units of a scalable bit stream that are required for decoding a representation with a specific spatio-temporal resolution and/or fidelity.

PROFILES AND LEVELS

Profiles and levels specify conformance points to facilitate interoperability between applications that have similar functional requirements. A profile defines a set of coding tools that can be used when generating a bit stream. A level specifies constraints on key parameters of the bit stream. All decoders conforming to a specific profile must support all included coding tools.

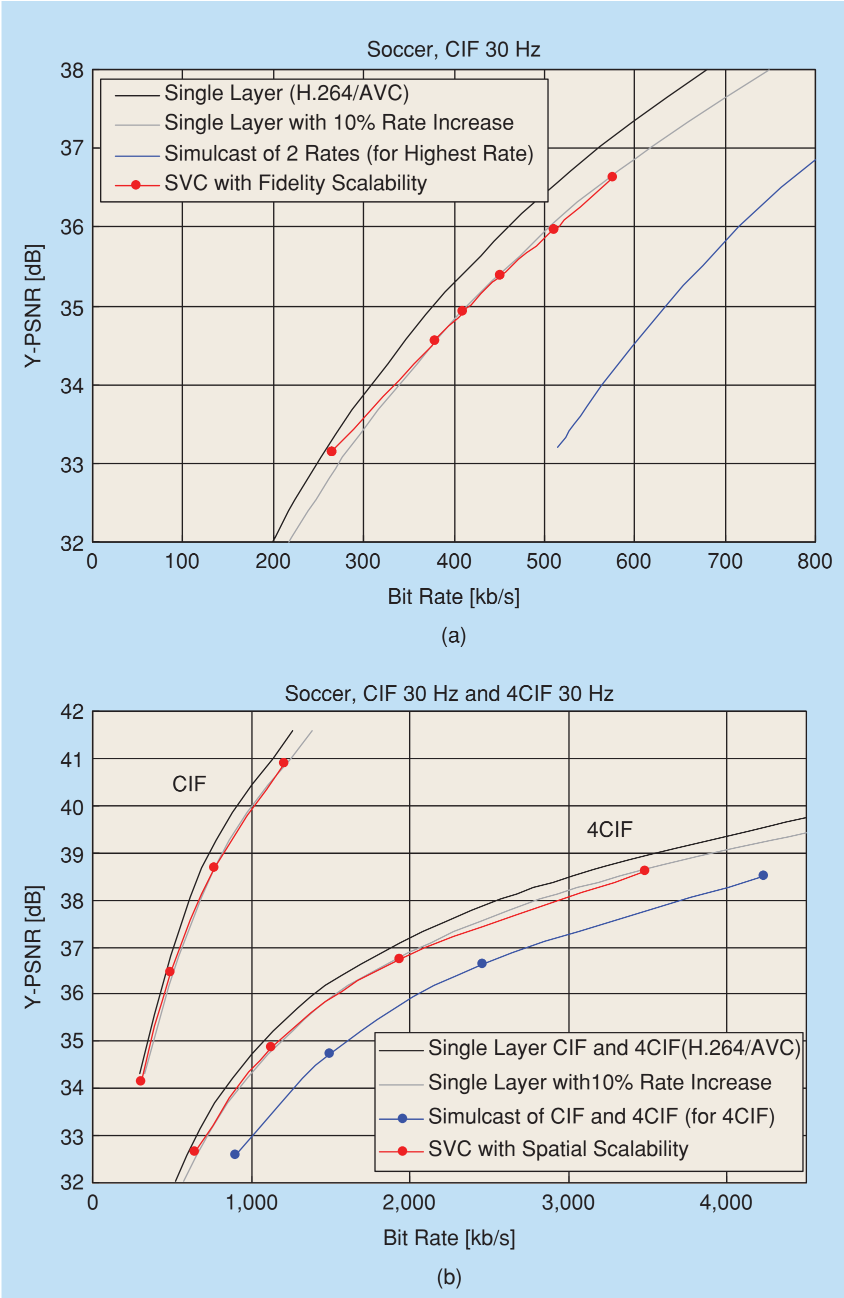
So far three profiles—the Scalable Baseline, the Scalable High, and the Scalable High Intra profile—have been defined for SVC. The Scalable Baseline profile is mainly targeted for conversational and surveillance applications requiring low decoding complexity. The Scalable High profile was designed for broadcast, streaming, and storage applications. The Scalable High Intra profile targets professional applications.

COMPARISON WITH OTHER STANDARDS

In contrast to the scalable profiles of prior video coding standards, SVC provides a mechanism to efficiently control the drift in packet-based fidelity scalable coding. With packet-based fidelity scalable coding, an arbitrary number of fidelity enhancement NAL units can be discarded (in decreasing order of the layer identifiers) in each access unit, and thus virtually any bit rate within the range given by the amount of available fidelity enhancements can be achieved. Drift is the effect of the prediction loops at the encoder and decoder not being synchronized. This effect occurs if fidelity enhancements that were employed for interpicture prediction in the encoder have been discarded from the bit stream and are therefore not available at the decoder.

In prior video standards, packet-based scalable coding was either not supported (for example, in MPEG-2 Video), or drift was completely avoided by always employing the layer with the lowest fidelity for interpicture prediction (for example in MPEG-4 Visual). The latter layer is always present when neglecting transmission errors. However, its usage for interpicture prediction dramatically decreases the coding efficiency of fidelity enhancement layers, since the portion of the bit rate that is spent for coding fidelity enhancement layer packets cannot be exploited for the coding of the enhancement layers of following pictures. By contrast, in SVC, the fidelity layer that is employed for interpicture prediction can be selected on a picture basis, enabling the necessary resynchronization between encoder and decoder as well as high coding efficiency by employing fidelity enhancement layers for interpicture prediction between the resynchronization pictures.

In comparison to nonscalable profiles of prior video coding standards, the single-layer H.264/AVC standard already provides a substantially increased coding efficiency. In comparison to the scalable profiles of prior video coding standards, SVC provides enhanced coding efficiency, increased flexibility, additional functionalities (such as low-cost



[FIG2] Rate-distortion comparison of SVC, single-layer H.264/AVC coding, and simulcast with H.264/AVC using the sequence Soccer: (a) SVC with fidelity scalability and (b) SVC with dyadic spatial scalability. Common intermediate format (CIF) and 4CIF denote sizes of pictures equal to 352×288 and 704×576 luma samples, respectively.

bit stream rewriting), and a reduced complexity overhead in comparison to the underlying single-layer coding scheme. With SVC, the bit rate increase necessary to provide the same reconstruction quality as a nonscalable H.264/AVC bit stream can be in the range of 10%, as shown in the next section. With scalable profiles of prior video coding standards, the corresponding bit rate increase in relation to the corresponding nonscalable profiles was significantly higher; often the coding efficiency for spatial and fidelity scalable coding was only slightly higher than that for simulcasting the spatial resolutions and fidelity representations required by an application.

PERFORMANCE

OBJECTIVE AND SUBJECTIVE QUALITY

For comparison, consider the example shown in Figure 2, where the coding efficiency of SVC providing fidelity scalability and spatial scalability is compared with that of single-layer H.264/AVC coding. As stated earlier, the coding efficiency is measured in terms of bit rate and visual quality, the latter of which is measured using the average luma peak signal-to-noise ratio (PSNR) of the video frames in our experiments. Temporal scalability with five levels is provided by all bit streams used for comparison and does not have any negative impact on the rate-distortion results.

In Figure 2(a), SVC fidelity scalability is compared to H.264/AVC single-layer coding. All SVC rate-distortion points are extracted from one single bit stream, while for single-layer coding,

SVC HARDWARE AND SOFTWARE PRODUCTS

Application Standards

- RTP payload format for SVC, current draft [Online]. Available: *http://www.ietf.org/internet-drafts/draft-ietf-avt-rtp-svc-01.txt*
- SVC file format, current draft: ISO/IEC JTC 1/SC 29/WG 11, "Study text of ISO/IEC 14496-15/PDAM2 (SVC file format)," document N9026, San Jose, CA, Apr. 2007.

Hardware Products

- Prototype demo by STMicroelectronics at the Consumer Electronics Show 2007 showing an implementation of SVC on existing ST products STi7109 and Nomadik 8810.

Software Products

- Videoconferencing system by Vidyo supporting resolutions from QCIF to HD and error resilience capabilities for up to 50% packet loss rates.

SVC RESOURCES

The Standard

- ITU-T and ISO/IEC JTC 1, "Advanced Video Coding for Generic Audiovisual Services," ITU-T Recommendation H.264 and ISO/IEC 14496-10 (MPEG-4 AVC), Version 8 (including the SVC extension): Consented in July 2007.

Tutorials

- *IEEE Trans. Circuits Syst. Video Technol. (Special Issue on Scalable Video Coding Standardization and Beyond)*, vol. 17, no. 9, Sept. 2007.

Overviews

- 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, Sept. 2007.
- T. Wiegand and G.J. Sullivan, "The H.264/AVC Video Coding Standard," *IEEE Signal Processing Mag.*, vol. 24, no. 2, pp. 148–153, 2007.

Test and Conformance Bit Streams

- Available for download from *http://ftp3.itu.int/av-arch/jvt-site*

Reference Software

- Available for download from *http://ftp3.itu.int/av-arch/jvt-site*

Discussion List

- JVT discussion list; enrollment through http://mailman.rwth-aachen.de/mailman/listinfo/jvt-experts
- SVC discussion list; enrollment through http://mailman.rwth-aachen.de/mailman/listinfo/jvt-svc

Resources for Further Development

- J.-R. Ohm, "Advances in scalable video coding," *Proc. IEEE*, vol. 93, no. 1, pp. 42–56, Jan. 2005.
- JVT documents [Online]. Available: http://ftp3.itu.int/av-arch/jvt-site

Applications

- A. Eleftheriadis, R. Civanlar, and O. Shapiro, "Multipoint videoconferencing with scalable video coding," in *J. Zhejiang Univ. – Science A*, vol. 7, no. 5, pp. 696–705, May 2006.
- T. Schierl, K. Gänger, T. Wiegand, and T. Stockhammer, "SVC-based multisource streaming for robust video transmission in mobile ad hoc networks," *IEEE Wireless Commun. Mag. (Special Issue on Multimedia in Wireless/Mobile Ad Hoc Networks)*, vol. 13, no. 5, pp. 96–103, Oct. 2006.

each rate-distortion point represents a separate nonscalable bit stream. The diagram additionally shows a rate-distortion curve representing the simulcast of the single-layer H.264/AVC bit stream with the fidelity of the SVC base layer and another single-layer H.264/AVC bit stream with the fidelity specified in the diagram.

In Figure 2(b), SVC spatial scalability is compared to H.264/AVC single-layer coding and simulcast of two spatial resolutions with H.264/AVC. Two spatial layers with a dyadic relation are provided. This means that the horizontal and vertical resolutions of the enhancement layer (4CIF) double the horizontal and vertical resolutions of the base layer (CIF). For each point on the 4CIF curve, the spatial scalable bit stream comprises the corresponding point on the CIF curve with about one-half to one-third of the overall bit rate.

The comparison shows that SVC can provide a suitable degree of scalability at the cost of approximately 10% bit rate increase compared to the bit rate of single-layer H.264/AVC coding. This bit rate increase usually depends on the degree of scalability, the bit rate range, and the spatial resolution of the included representations. The comparison also shows that SVC is clearly superior to simulcasting single-layer streams for different spatial resolutions or bit rates.

COMPLEXITY

Due to the substantially higher complexity of the underlying H.264/AVC standard relative to prior video coding standards, the computational resources necessary for decoding an SVC bit stream are usually higher than those for scalable profiles of older standards. However, due to the single-loop decoding feature of SVC, the required decoding complexity overhead for spatial and especially fidelity scalable coding relative to the underlying single-layer coding standard has been substantially reduced. For simple SVC configurations like fidelity scalability or dyadic spatial scalability, SVC decoders require 10–50% more computational resources than single-layer H.264/AVC decoders for the same target resolution and bit rate. The decoder complexity depends on the number of layers that are employed for interlayer prediction, the resolution ratios between the layers, and the bit rate. The minimization of the encoder complexity overhead for scalable coding has become an active research area in the video coding community.

FURTHER TECHNICAL DEVELOPMENTS

An extension of SVC with the goal of providing additional functionalities and improved coding efficiency is being investigated in the JVT. The current working items are bit-depth scalability, chroma format scalability, and fine-granular fidelity scalability with an improved drift control for low-delay coding.

RESOURCES

The main resources for SVC are the standard text itself, the standardized conformance bit streams, and the reference software. The standard is published by both ITU-T and ISO/IEC as so-called "twin texts" (i.e. they are technically aligned but published independently of each other). Other resources are listed in the "SVC Resources" sidebar.

PRODUCTS

The final draft of the SVC amendment was finalized in July 2007. Since SVC has not been released yet, products are in an early stage of deployment. The first product announcements and SVC implementations are listed in the "SVC Hardware and Software Products" sidebar.

AUTHORS

Heiko Schwarz (heiko.schwarz@hhi.fhg.de) is with the Fraunhofer Institute for Telecommunications, Heinrich Hertz Institute, Berlin, Germany. He is a coeditor of the SVC amendment for the H.264/AVC standard.

Mathias Wien (wien@ient.rwth-aachen.de) is with RWTH Aachen University, Germany. He is a coeditor of the SVC amendment for the H.264/AVC standard. [SP]

[dsp **APPLICATIONS**] continued from page 120

REFERENCES

[1] M.J. Wainwright and M.I. Jordan, "Graphical models, exponential families, and variational inference," Depart. Statistics, UC Berkeley, Tech. Rep. 649, Sept. 2003.

[2] J.S. Yedidia, W.T. Freeman, and Y. Weiss, "Understanding belief propagation and its generalizations," in *Exploring Artificial Intelligence in the New Millennium,* G. Lakemeyer and B. Nebel, Eds. San Mateo, CA: Morgan Kaufmann, 2002.

[3] B.J. Frey and N. Jojic, "A comparison of algorithms for inference and learning in probabilistic graphical models," *IEEE Trans. Pattern Anal. Machine Intell.*, vol. 27, no. 9, pp. 1392–1416, Sept. 2005.

[4] F.R. Kschischang, B.J. Frey, and H.-A. Loeliger, "Factor graphs and the sum–product algorithm," *IEEE Trans. Inform. Theory*, vol. 47, no. 2, pp. 498–519, Feb. 2001.

[5] D.J.C. MacKay, *Information Theory, Inference & Learning Algorithms*. Cambridge, U.K.: Cambridge Univ. Press, 2002.

[6] J. Pearl, *Probabilistic Reasoning in Intelligent Systems*. San Mateo, CA: Morgan Kaufman, 1988.

[7] P.F. Felzenszwalb and D.P. Huttenlocher, "Efficient belief propagation for early vision," *Int. J. Comput. Vis.*, vol. 70, no. 1, pp. 41–54, 2006.

[8] W.T. Freeman, E.C. Pasztor, and O.T. Carmichael, "Learning low–level vision," *Int. J. Comput. Vis.*, vol. 40, no. 1, pp. 25–47, 2000.

[9] D. Scharstein and R. Szeliski, "A taxonomy and evaluation of dense two-frame stereo correspondence algorithms," *Int. J. Comput. Vis.*, vol. 47, no. 1, pp. 7–42, 2002.

[10] J. Sun, N.-N. Zheng, and H.-Y. Shum, "Stereo matching using belief propagation," *IEEE Trans. Pattern Anal. Machine Intell.*, vol. 25, no. 7, pp. 787–800, July 2003.

[11] E.B. Sudderth, A.T. Ihler, W.T. Freeman, and A.S. Willsky, "Nonparametric belief propagation," in *Proc. IEEE Conf. on Computer Vision and Pattern Recognition*, 2003, vol. 1, pp. 605–612.

[12] J. Dauwels and H.-A. Loeliger, "Phase estimation by message passing," in *Proc. Int. Conf. Communications*, 2004, pp. 523–527. [SP]