

Adaptation of Scalable Video Coding to Packet Loss and its Performance Analysis

Euy-Doc Jang*, Jae-Gon Kim*, Truong Cong Thang**, Jung Won Kang**

*Korea Aerospace University, 100 Hanggongdae gil, Hwaajeon-dong, Deogyang-gu, Goyang-city, Gyeonggi-do, 412-791, KOREA

**Electronics and Telecommunications Research Institute, 138 Gajeongno, Yuseong-gu, Daejeon, 305-700, KOREA
{jangeuydoc, jgkim}@kau.ac.kr, {tchang, jungwon}@etri.re.kr

Abstract— SVC (Scalable Video Coding) is a new video coding standard to provide convergence media service in heterogeneous environments with different networks and diverse terminals through spatial-temporal-quality combined flexible scalabilities. This paper presents the analysis of performance on packet loss in the delivery of SVC over IP networks and an efficient adaptation method to packet loss caused by buffer overflow. In particular, SVC with Medium Grained Scalability (MGS) as well as spatial and temporal scalabilities is addressed in the consideration of packet-based adaptation since finer adaptation is possible with a sufficient numbers of quality layers in MGS. The effect on spatio-temporal quality due to the packet loss of SVC with MGS is evaluated. In order to minimize quality degradation resulted by packet loss, the proposed adaptation of MGS based SVC first sets adaptation unit of Access Unit (AU) or Group of Picture (GOP) corresponding to allowed delay and then selectively discards packets in order of importance in terms of layer dependency. In the experiment, the effects of packet loss on quantitative spatio-temporal qualities are analyzed and the effectiveness of the proposed adaptation to packet loss is shown.

Keywords— SVC, MGS, Adaptation, Packet Loss

I. INTRODUCTION

In convergent media environments, users may access and interact with multimedia contents with different type of terminals such as PC, TV, PDA, or cellular phones over heterogeneous networks. In such environments, much more attention is being focused on video adaptation as a promising approach for keeping the Quality of Service (QoS) of video contents [1].

In general, scalable format is required to provide efficient video adaptation which is performed by very simple operation of truncation of parts of bitstream. Scalable video coding (SVC), which is recently developed by the Joint Video Team (JVT) from ISO/IEC MPEG and the ITU-T VCEG as an extension of the H.264/AVC, is a new video coding standard to provide efficient video adaptation [2], [3].

Scalable bitstreams of SVC are composed of a base layer that is compatible to H.264/AVC and one or more enhancement layers, which may enhance the spatio-temporal resolution and/or quality of the base layer. Based on such

scalable layered structure, SVC can be easily adapted to meet diverse constraints imposed by devices and connections by simple truncation.

In general, packet loss in IP networks is caused by transmission errors or buffer overflow due to traffic congestion. In this paper, we analyze the effect of packet loss on SVC in terms of spatio-temporal qualities. Based on the results, we propose an adaptation method of SVC to deal with packet loss caused by buffer overflow. For effective SVC transmission and adaptation, we assume that support of MANEs (Media Aware Network Elements), which could deal with SVC scalability structure [4], [5], is available. In other words, it is assumed that the processing of information of Network Abstraction Layer (NAL) header and flow-based buffer (queue buffer) control [6] are supported by MANEs.

Quite flexible scalability with the combination of spatial, temporal, and quality scalabilities is provided by SVC. In particular, SVC with MGS as well as spatial and temporal scalabilities is addressed in the consideration of packet-based adaptation since finer adaptation is possible with a sufficient numbers of quality layers in MGS. Many methods to improve the performance of SVC adaptation have been studied. Some approaches [7], [8] utilized interoperable metadata for flexible and optimal adaptation. Multilayer adaptation of MGS-based SVC by slightly modifying the syntax is proposed [8] and the SVC adaptation is adopted for selective discarding of packets in error prone IP networks [10].

In this paper, first the effect of random packet loss on the received quality of SVC with MGS is evaluated in terms of spatial quality and temporal resolution [10]. Based on the results, then we propose an adaptation method of SVC to reduce the quality degradation caused by packet loss. The adaptation method first set the adaptation unit as multiple times of AU or GOP according to allowed delay or buffering in a MANE, and then accommodating packet losses, which may occur by buffer overflow due to scheduling in a network node, by discarding the packets in order of their importance in terms of layer dependences among scalable layers.

The remainder of this paper is organized as follows. In Section II we briefly introduce the scalable structure of MGS. The proposed adaptation of SVC with MGS is presented in Section

III. Section IV presents the experiments to evaluate the effect of packet loss and the performance improvement of the proposed adaptation. Finally, conclusion is provided in Section V.

II. SCALABILITY: MGS

The SVC bitstream may be composed of multiple spatial, temporal, and SNR layers in a combined way. Typical SVC includes multiple spatial layers, each of which has a base quality layer with multiple enhanced quality layers. In addition, temporal scalability is provided based on hierarchical-B picture structure. Quality scalability is provided by either of coarse-grain quality scalability (CGS) layer or MGS.

In this paper, we mainly deal with MGS in which refinement DCT coefficients can be split into several fragments to cope with packet losses with sufficient number of quality layers [8]. In MGS, the coded data corresponding to a quantization step size can be fragmented into at most 15-layer. In other words, a base-layer with minimum quality is quantized with a larger step size and then at the quality enhancement layer, quantization error signal that is the difference between original image and the decoded image is quantized with a smaller step size. The set of MGS layers corresponding to a quantization step size is named as *MGS stack* [8], [9].

Illustration of an AU coded with MGS is shown in Fig. 1. This example of AU contains two spatial layers. Both of the low spatial layer and enhancement spatial layer are enhanced by two MGS stacks, each of which is composed 3 MGS layers.

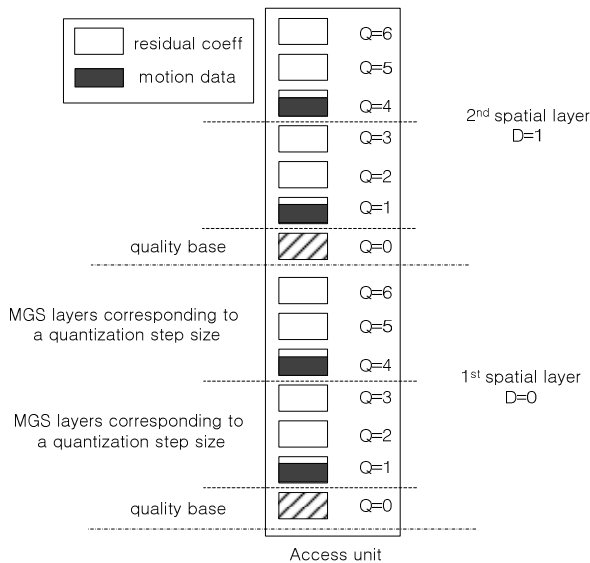


Fig. 1 MGS quality-layer structure in an AU.

III. EFFECT OF PACKET LOSS AND ADAPTATION

In this Section, assuming SVC with MGS, we evaluate the effect of packet losses on the quality of decoded SVC sequences. Based on the evaluation, we propose an

adaptation method to reduce the quality degradation caused by packet losses.

A. Effect of Packet Loss

In general, we suppose that the SVC bitstream is consists of spatial, temporal, and quality layers resulted in combined scalability, and each layer is composed as a NAL unit and then each NAL unit is payloaded into RTP packets [9]. In this case, when a packet is lost during delivering over IP networks, the quality of corrupted SVC sequences may be severely degraded due to error propagation due to layer dependent coding as well as the lost data itself.

The scalable layer of SVC is identified using a triple ID, consisting of the Dependency ID, Temporal ID, Quality ID, which is referred as triple (D,T,Q). For example, a base layer NAL unit of the lowest temporal resolution and SNR scalability is identified as (D,T,Q) = (0,0,0). As mentioned before, layers in SVC are highly interdependent on each other, which means that the loss of a NAL unit of a certain layer may cause a severe degradation of quality or even prevent the correct decoding of other layer due to error propagation. For instance, Fig. 2 shows an interdependent structure in a SVC bitstream encoded with three spatial layers, three temporal layers and MGS layers. In a GOP, if a NAL unit (1,2,0) is lost, then all NAL units that depend on it (i.e., NAL unit (1,2,1)) can't be decoded even if they are correctly received. In addition, other temporal and spatial layers also have interdependency, and then they are cause error propagation if the packet containing the reference data is lost.

In the case of NAL unit loss caused by packet loss, NAL units that depend on the lost NAL unit can't be decoded as well as the lost NAL unit itself. Therefore, it is necessary to modify the corrupted stream into a valid stream to be decoded correctly by discarding the dependent NAL unit additionally, prior to decoding [6].

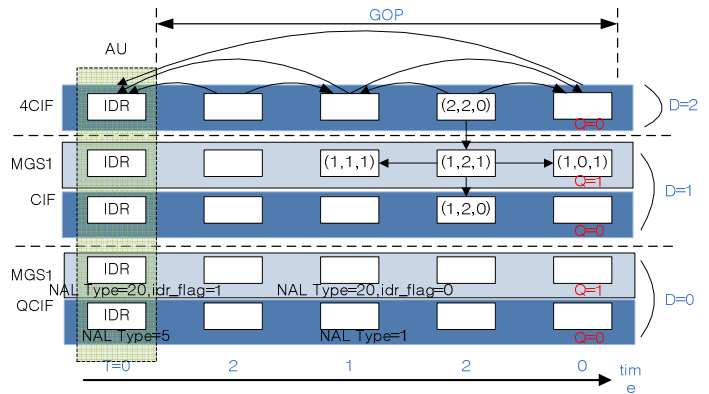


Fig. 2 Interdependency structure of a SVC.

B. Adaptation Method

As shown in the previous Section, the SVC bitstream could constitute an interdependent scalable structure, which in turn leads to poor quality when it is exposed to random packet loss. In order to alleviate this problem, in this Section proposes a

adaptation mechanism based on selective discarding instead of random loss, and evaluates reception quality using PSNR.

In the selective adaptation, first we set adaptation unit of AU or GOP corresponding to allowed delay. When a longer delay is allowed, the adaptation unit can be extended to multiples of AU or GOP. When a packet loss is occurred over an adaptation unit, a scheduler selectively discards a packet that is the least important in terms of layer dependency in a given adaptation unit.

In the experiments to show the impact of random packet loss on SVC bitstream, video sequences are encoded by using the SVC reference software JSVM (Joint Scalable Video Model) 9.14 [12]. Assumption that each scalable layer in the SVC bitstream is composed in a NAL unit and each NAL unit is packetized into a RTP packet. In addition, it is assumed that the base layer of each AU is not affected from packet loss. For this case we assume that applying unequal error protection (UEP) solutions to the base layer [13]. Each NAL unit embedding coded data of a scalable layer has different packet length. In the experiments, the internet packet loss model [14] is used to reflect packet loss in IP networks.

Table I presents the test condition of the experiment with the test sequences and coding parameters. The sequences are coded in SVC with two spatial layers, with five levels of temporal scalability and MGS quality scalability.

Table 1. SVC Coding Parameter

Coding rate	Base layer	254.1 kbps
	Enhancement layer	824.2 kbps
Spatial/Temporal resolution	Base layer	QCIF, 30 Hz
	Enhancement layer	CIF, 30 Hz
Coding option used	Fast search mode, MV search range: 32, CAVLC, loop filter, adaptive inter-layer prediction	
Test sequences	City, Crew (134 frames)	
MGS QP	Base layer: 38	
	Enhancement layer: 26 ~ 42	

C. Performance of MGS

We used SVC bitstream with MGS shown in Fig. 1. Both of the low spatial layer and enhancement spatial layer are enhanced by two MGS stacks, each of which is composed 3 MGS layers. In this experiment we evaluate the basis performance of the MGS in terms of rate-adaptation range and correspond to qualities. As shown in Fig. 3, City can be adapted over the range of 381 ~ 825 kbps with qualities of 30.0 ~ 36.1 dB in Fig. 3.

‘Standard’ and ‘QualityLayer’ denotes adaptation by using JSVM BitstreamExtractor tool and *priority_id* in NAL header, respectively [15]. Adaptation using priority-id gives better performance over the same rate-adaptation range.

It is also shown that there is a large gap in the comparison of coding efficiency between single layer and MGS over the same rate range. The case of MGS, different rates are obtained by adaptation of the same bitstream without individual encoding with each rate like in the case of single

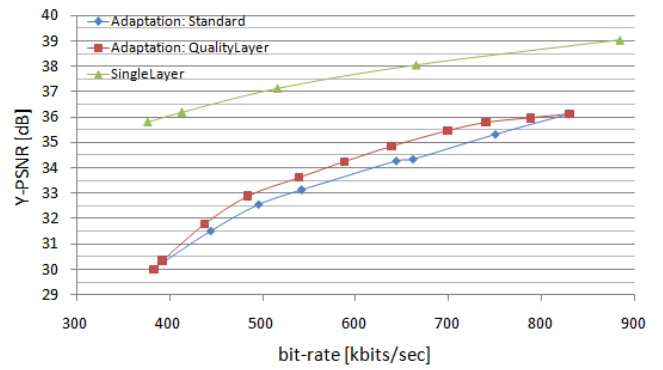


Fig. 3 Basic performances of MGS: Rate-adaptation ranges and resulting qualities.

layer. Therefore, it is clear “Single layer” is not appropriate in IP networks with time-varying available bandwidth.

D. Random Packet Loss

Fig. 4 presents the results of random packet loss in the City sequence. As shown, a small increase packet loss rate (PLR) corresponds to a significant increase of the amount of data that is received but cannot be decoded due to absence of lower NAL unit as a consequence of interlayer dependency. In this figure, the sum of data loss at the receiver includes such data to be discarded additionally due to layer dependency and the packet loss data.

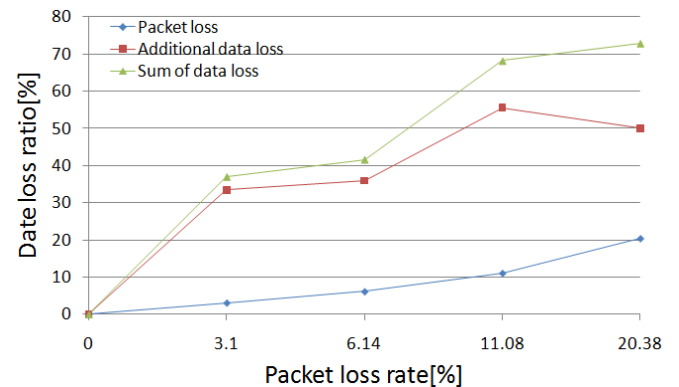


Fig. 4 Impact of random packet loss (Crew).

Random packet losses affect spatial and temporal quality which is measured by PSNR and the number of frames that may be correctly decoded, respectively. Fig. 4 shows the reduction of both qualities as the PLR increases for the Crew sequence. It can be noted that not only the PSNR value of decoded frames decreases, but also that there is a reduction in the number of frames that are decoded, which is represented in Fig. 4 by an absence of PSNR values for the associated frames. Fig. 5 presents the reduction of the mean frame rate of City sequences, showing a pronounce decrease in this value as the PLR incases, reaching nearly 14 fps for a PLR of 3.14%.

E. Adaptation Method Based on Selectively Discarding

Adaptation based on selective discarding considers two adaptation units of AU and GOP. In the case of AU-unit, when a packet loss is occurred due to buffer overflow, the packet with the lowest dependency in AU is discarded. In other word, discard the highest quality layer NAL unit in AU which is not used as a reference for any other layers. This adaptation mechanism needs delay of an AU. However, the AU-unit adaptation may not be effective preventing quality degradation since it does not consider the importance of different AUs in hierarchical B structure.

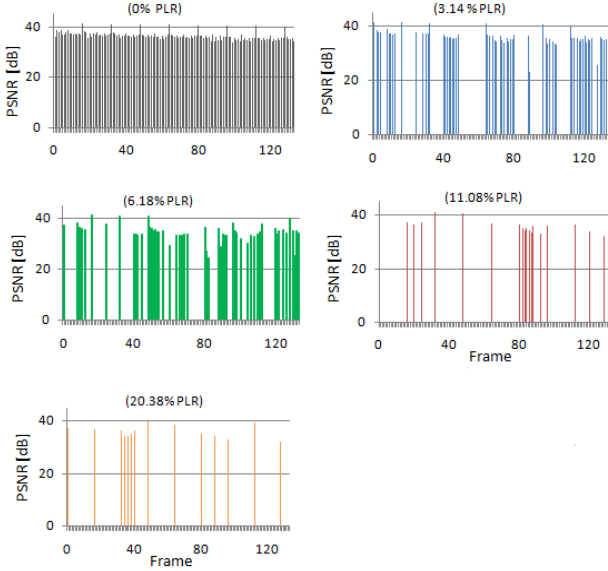


Fig. 4 Impact of random packet loss in term of PSNR of decoded frames (Crew, CIF).

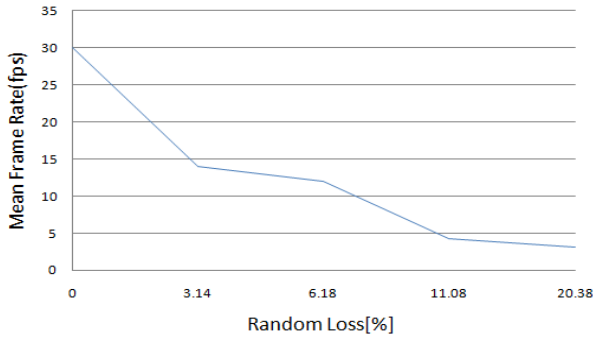


Fig. 5 Impact of random packet loss in term of mean frame rate of decoded frames (Crew)

As shown in Fig. 6, the quality is steeply degraded with increase of packet loss ratio since packet discarding is allowed even in the AU corresponding to lower temporal layer that would be referred to higher temporal layers in AU-unit adaptation

Adaptation unit is extended to GOP in the GOP-unit adaptation with the same discarding mechanism. In GOP-unit, the importance between AUs with different temporal layers is

also considered in the selective discarding. Therefore when a packet loss is occurred in GOP, the packet corresponding to the lowest interlayer dependency in quality layer as well as temporal layer and spatial layer is discarded. This mechanism prevents additional data loss results in keeping the quality as much as possible as shown in Fig. 6

At the experiment, first we consider temporal and quality layer packet dependency, and then to reflect the number of packet loss N_l , and determine the value of (D_d, T_d, Q_d) of the packet that should be discarded. N_T given by (1) presents the number of packet that is able to be discarded at arbitrary temporal layer T in GOP. $Q_{\max}$ is the value of maximum quality layer.

$$N_T = 2^{(T-1)} \times (Q_{\max} + 1) \quad (1)$$

As can be seen, the increases in PLR corresponds to increase of packet that be discarded in GOP. When increase in packet loss in GOP, maximum temporal layer $T_{\max}$ cannot accommodate packet loss. In this case, we consider packet of lower temporal layer $T_d = (T_{\max} - 1)$ to reflect packet loss. First decide temporal layer by using expression of (1), and then determine the quality layer that be able to reflect packet loss by using expression of (2).

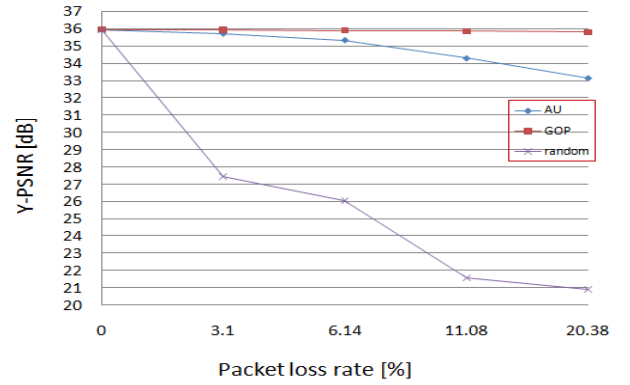


Fig. 6 Quality of selective discarding packet.

$$Q_d = Q_{\max} - \left\{ \frac{N_l}{2^{(T-1)}} - 1 \right\} \quad (2)$$

In case adaptation in GOP present the more performance in PSNR at same packet loss ratio compare with AU adaptation unit. Even though increases packet loss ratio, adaptation in GOP different from AU prevent radically quality degradation but need to long term delay of GOP.

Proposed adaptations of selective discard mechanism (see Fig. 6) as compared with random packet loss (see Fig. 4 and 5) efficiently cope with SVC quality degradation due to packet loss on IP network. In particular, adaptation unit of GOP provide efficiently adaptation mechanism in sudden increase of packet loss ratio.

IV. CONCLUSION

In this paper, the effect of packet losses on SVC is evaluated in terms of decoded spatial-temporal quality. Based on the evaluation, an adaption method of SVC to cope with packet losses is proposed and its performance is analyzed. MGS based SVC as well as spatial and temporal scalabilities is mainly addressed in the consideration of packet-based adaptation in IP network. The random packet loss in SVC lead to serious quality degradation, but the proposed adaptation is able to reduce the error propagation caused by packet loss by selectively discarding less important packets with the layer-dependency consideration. The effectiveness of the proposed adaptation is shown in the experiments with some SVC bitstreams with MGS. In particular, GOP-unit adaptation is effective in keeping quality even in sudden increase of packet loss ratio. The proposed method could be used for efficient QoS adaptation in SVC delivery over error prone IP networks.

ACKNOWLEDGMENT

This research was partly supported by the MKE under ITRC support program supervised by the IITA (IITA-2009-C1090-0903-0011) and GRRC program of Gyeonggi province [GRRC-KAU-2009-B01, Advanced Broadcasting Media Technology Research Center].

REFERENCES

- [1] S.-F Chang, A. Vetro, "Video adaptation: concepts, technologies, and open issues", *Proceedings of the IEEE*, vol. 93, pp.148 - 158, Jan. 2005.
- [2] T. Wiegand, G. J. Sullivan, J. Scharz, and M. Wien, "Joint Draft 11 of SVC Amendment," Joint Video Team, doc. JVT-X201, Geneva, Switzerland, July 2007.
- [3] H. Schwarz, D. Marpe, 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.
- [4] S. Wenger, Y.-K. Wang, and T. Schierl, "Transport and signaling of SVC in IP networks," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 17, no. 9, pp. 1164-1173, Sep. 20007.
- [5] Y.-K. Wang, M. M. Hannuksela, S. Pateux, A. Eleftheriadis, and S. Wenger, "System and transport interface of SVC," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 17, no. 9, pp. 1149-1163, Sep. 20007.
- [6] J. M. Monterio, C. T. Calafate, and M. S. Nunes, "Evaluation of the H.264 scalable video coding in error prone IP networks", *IEEE Trans. Broadcasting*, vol. 54, no. 3, pp. 652-659, Sep. 2008.
- [7] T. C. Thang, Y.-S. Kim, Y.-M. Ro, J.-W. Kang, J.-G. Kim, "SVC Bitstream Adaptation in MPEG-21 Multimedia Framework," in *Proc. 15th Int. Packet Video Workshop (PV2006)*, vol. 7, no. 5, pp.764-772, May,2006
- [8] T. C. Thang, J. W. Kang, J.-J Yoo, J.-G. Kim, "Multilayer adaptation for MGS-based SVC bitstream," in *Proc. ACM Multimedia 2008*, Oct. 2008.
- [9] T. C. Thang, J.-G. Kim, J. W. Kang, J.-J Yoo, "SVC adaptation: Standard tools and supporting methods," *Signal Processing: Image Comm.*, vol. 24, no. 3, pp. 214-228, Oct. 2009.
- [10] E.-D. J, J.-G. Kim, T. C. Thang, and J. W. Kang, "Adaptation of SVC to Packet Loss and its Performance Analysis", *J. Broadcasting Engineering*, vol. 14, no. 6, pp. 796-806, Dec. 2009.
- [11] A. Eleftheriadis and S. Wenger, "System and transport Internet of SVC," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 17, no. 9, pp. 1149-1163, Sep. 2007.
- [12] J. Reichel, H. Schwarz, and M. Wien, Joint Scalable Video Model 9.14 , Joint Video Team, Doc. JVT-X202, Jul. 2008.
- [13] S. Deawon and C. W. Chen, "Scalable H.264/AVC video transmission over MIMO wireless systems with adaptive channel selection based on partial channel information," *IEEE Trans.Circuit Syst. Video Technol.*, vol. 17,no. 9, pp. 1218-1226, Sep. 2007.
- [14] Y. Guo, H. Q. Li and Y. K. Wang, "SVC/AVC loss simulator," ISO/IEC JTC 1/SC 29/WG 11, JVT-Q069, Oct. 2005.
- [15] L. Amonou, N. Cammas, S. Kervadec, and S. Pateux, "Optimized rate-distortion extraction with quality layer in the scalable extension of H.264/AVC," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 17, no 9, pp. 1186-1193, Sep. 2007.