

## Research Article

# Quality Scalability Compression on Single-Loop Solution in HEVC

**Mengmeng Zhang, Hongyun Lu, and Zhi Liu**

*North China University of Technology, No. 5 Jin Yuanzhaung Road, Shijingshan District, Beijing 100144, China*

Correspondence should be addressed to Mengmeng Zhang; [zhang\\_mengmengbbb@163.com](mailto:zhang_mengmengbbb@163.com)

Received 2 April 2014; Accepted 28 May 2014; Published 24 June 2014

Academic Editor: Guoqiang Zhang

Copyright © 2014 Mengmeng Zhang et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

This paper proposes a quality scalable extension design for the upcoming high efficiency video coding (HEVC) standard. In the proposed design, the single-loop decoder solution is extended into the proposed scalable scenario. A novel interlayer intra/interprediction is added to reduce the amount of bits representation by exploiting the correlation between coding layers. The experimental results indicate that the average Bjøntegaard delta rate decrease of 20.50% can be gained compared with the simulcast encoding. The proposed technique achieved 47.98% Bjøntegaard delta rate reduction compared with the scalable video coding extension of the H.264/AVC. Consequently, significant rate savings confirm that the proposed method achieves better performance.

## 1. Introduction

In January 2013, joint collaborative team on video coding has finalized a final draft about the next generation video standard, that is, high efficiency video coding (HEVC) [1]. Scalable high efficiency video coding (SHVC) is being formulated as the extension of HEVC. Given the special performance of HEVC in delivering the target resolution and frame rates [2], SHVC will be also applied to a variety of consumer domains. Scalable video coding is a promising alternative similar to multiple description coding for robust transmission of information over unpredictable networks [3]. Scalable video coding is applied to cope with the heterogeneity of networks and devices used in the video service environment. With the rapid evolution in theory and techniques, the technology of scalable video coding may be applied to mobile terminal in the future. Therefore, the investigation of a scalable video coding scenario is important and necessary.

The scalable video coding has been investigated for more than twenty years. Many international video coding standards have been presented, and the scalable extension of H.264/AVC has been a particular new design. The scalable extension of H.264/AVC standard (referred to hereafter as SVC) achieved that goal by employing multilayer coding along with adaptive interlayer prediction and hierarchical temporal reference [4]. Generally, scalable video coding is

a highly attractive solution to the problems caused by the characteristics of modern video transmission systems. The scalable video coding method can be used to achieve the adaptation of a bitrate with features such as temporal and spatial scalabilities [5]. The quality scalability could be treated as a special case of spatial scalability with the same resolution in different layers [6]. The next development direction focuses on the combination of low bitrate and low complexity for SHVC. Multiloop solution for quality scalability based on all base layer (BL) modes in HEVC has been proposed in [7] in order to sufficiently reduce the bitrate. As motion compensation remains the most time-consuming module, accounting for half of the time on average [8], the multiloop solution increases complexity in the decoder. The single-loop solution will be in general less complex for the decoder and adopted in the proposal.

In this paper, the single-loop solution will be used to be the framework for the scalable video coding scenario. The single-loop solution will be applied to the interlayer interprediction including the intermerge prediction and intermotion prediction. When the base layer is coding in intramode, interlayer intraprediction is proposed to remove redundancy between the layers. In the scalable scenario, the interlayer prediction quadtree is derived from the corresponding base layer quadtree.

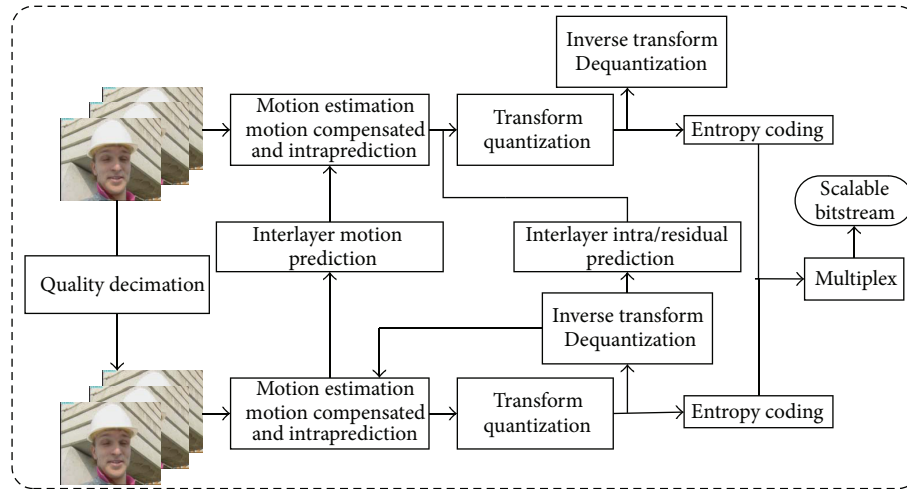


FIGURE 1: Framework of the H.264/SVC encoding.

The remainder of this paper is organized as follows. Section 2 introduces the background on SVC and HEVC. In Section 3, the proposed scalable video coding is introduced briefly in this paper. Finally, Sections 4 and 5 give the results and conclusion.

## 2. Background

**2.1. Scalable Extension of the H.264/AVC Standard.** With the introduction of the H.264/AVC video coding standard, the joint video team of the ITU-T VCEG and the ISO/IEC MPEG has also developed a scalable video coding extension of the H.264/AVC standard. As SVC is formulated as an extension of the H.264/AVC, its certain technical methods are inherited from the H.264/AVC. Similar to the splitting technique of H.264/AVC standard, the  $16 \times 16$  biggest block, using of the idea of ABT (adaptive block-size transform), and hybrid coding structure also appear in scalable video coding extension of the H.264/AVC standard [5]. In addition to these, single-loop structure is employed to improve compression efficiency in the scalable extension of H.264/AVC standard. SVC is investigated to remove the interlayer redundancy by the following three interlayer prediction methods. The three interlayer prediction methods are interlayer intraprediction, interlayer motion prediction, and interlayer residual prediction [5]. The framework of the H.264/SVC encoding is displayed as shown in Figure 1. The prediction mechanisms have been integrated in the coding of the enhancement layers (EL) to eliminate the redundancy between several layers. Besides, many transformations performed in the image domain will lead to some increase in bitrate in the single-loop coding structure.

**2.2. High Efficiency Video Coding.** HEVC is the latest video coding standard developed by a joint effort between ISO/IEC and ITU-T and succeeding H.264/AVC. The emerging high efficiency video coding standard aims to provide a doubling in coding efficiency with respect to the H.264/AVC high profile, delivering the same video quality at half the bitrate [9].

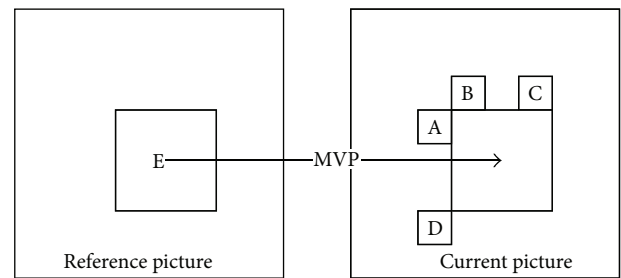


FIGURE 2: Motion merge candidates in interprediction.

This design still follows a traditional hybrid coding approach, such as interprediction based on the motion compensated, interprediction residuals of the two-dimensional transform, SAO, entropy coding, and quadtree. The quadtree data structure is commonly used in HEVC to decompose a frame into separate spatial regions to adaptively identify the type of quantizer used in various regions of a frame.

In terms of the interprediction, an intermode coding unit (CU) can be encoded with one of the following coding modes: they are `MODE_SKIP` and `MODE_INTER` [10]. There are two prediction techniques existing in `MODE_INTER` mode. They are the intermerge prediction and intermotion prediction. Intermerge prediction is to predict the current picture by the neighbouring prediction unit (PU). As shown in Figure 2, A~E is neighbouring PUs from the base layer, respectively. A~D express the spatial neighbouring PUs and E shows the temporal neighbouring PU. Intermotion prediction needs to transmit the motion vector or the motion vector difference (MVD) to the decoder. These main approaches are applied to the composing technology of the proposed scalable video coding in this paper.

## 3. Proposed Scalable Video Coding

The single-loop solution will be adopted as the framework in the proposed scalable video coding. In the single-loop

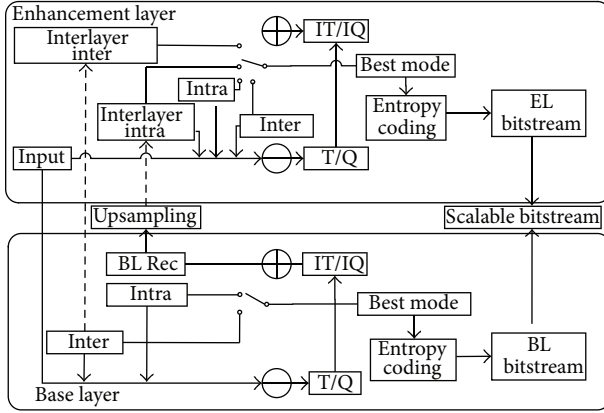


FIGURE 3: High-level block diagram of the encoder.

scalable compression, only the target layer needs to be reconstructed. The deblock filter (DF) and motion compensation of inter-CUs are skipped in base layer decoding to reduce the decoder complexity. For the intracoded CUs, the collocated base layer CUs which need to be reconstructed and upsampled are used as the prediction for the enhancement layer CUs. For the intercoded CUs, the information of neighbouring PU from enhancement layer will be obtained from the decoding base layer data. As enhancement layer and base layer are equipped with the same texture content, the motion parameters of the PUs from base layer can be reused for enhancement layer. In Figure 3, the dotted lines indicate the interlayer intraprediction and interlayer interprediction.

**3.1. Interlayer Intraprediction.** Interlayer intraprediction is proposed to remove redundancy because of the characteristic of the intracoding. The lower layer needs to be reconstructed entirely before the enhancement layer can be predicted. A CU level mode flag will be used to indicate the usage for the current enhancement layer CU. The transform and quantization processes of interlayer intraprediction predict that CUs are the same as an intrapredicted CU on the quantization parameter (QP) of the EL, in which discrete sine transform and discrete cosine transform are applied to the different types of TUs. The residual data, mode flag, and size of CU are the additional transmission signals in interlayer intraprediction. When the frame is coded in I-slice, the quadtree algorithm mechanism will also be used to the scalable video coding scenario similar to the HEVC. The split method of the quadtree is chosen by the rate-distortion optimization (RDO) by Lagrangian equation (1) just as that in HEVC [11]:

$$J(M, Q) = D(M, Q) + \lambda_{\text{MODE}} * R(M, Q). \quad (1)$$

$M \in \{\text{INTRA}, \text{INTER}, \text{SKIP}\}$  indicates a mode chosen for a particular block.  $Q$  is the selected quantizer step size.  $D$  is the distortion that is obtained by coding the considered block,  $R$  is the number of bits associated with choosing  $M$  and  $Q$ , and  $\lambda$  is the Lagrangian multiplier that is derived in the foundation of the used quantization parameter. According to the rate-distortion optimization method, the enhancement

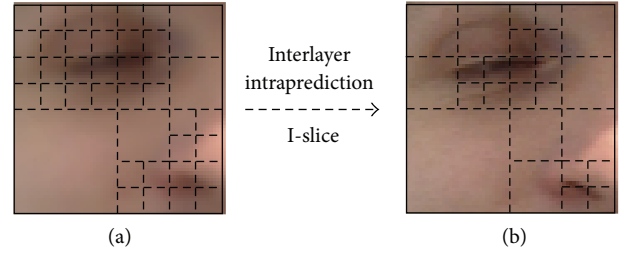


FIGURE 4: Possible adaptive divided CU, (a) reconstructed CU from BL, (b) possible PU partition from EL.

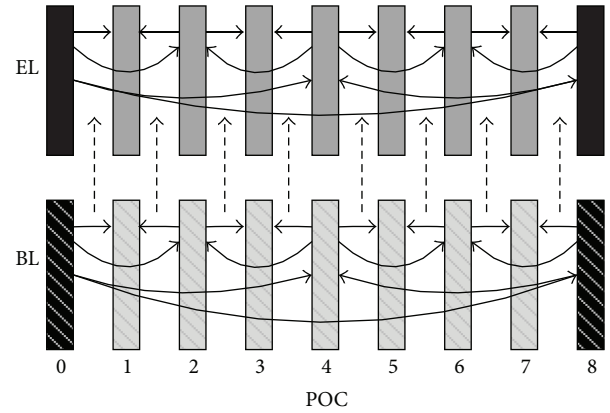


FIGURE 5: Prediction of motion parameters for EL.

layer image will be split on its texture features adaptively. The relatively smooth zone in the enhancement layer will be coded in a large block; otherwise, the block is to be split into more small blocks. The possible adaptive divided block can be represented by the dotted lines in Figure 4.

**3.2. Interlayer Interprediction.** The decoder motion parameter from base layer is utilized directly in interlayer interprediction to support the single-loop decoder solution. Simultaneously, the enhancement layer prediction quadtree is derived from the corresponding base layer quadtree for interlayer interprediction. The interlayer interprediction is presented in Figure 5. The black parts denote the key frame and the dotted lines indicate the interlayer motion parameters inheriting.

Next, the interlayer interprediction will be presented in two prediction situations: intermerge prediction and intermotion prediction. If the intermerge prediction is used in base layer blocks, the motion merge indexes from the base layer spatial or temporally neighbouring intercoded PU will be inherited and can be inferred as the ones for the current enhancement layer PU. The enhancement layer motion parameters will be gained according to the inherited motion merge indexes. The inherited motion merge indexes need not be transmitted to the decoder again in the EL. If the intermerge prediction is discarded, the intermotion prediction will be applied instead in base layer prediction. For the intermotion prediction, a lot of motion vectors need to be

TABLE 1: Performance of scalable HEVC compared to single layer, simulcast, and SVC.

Video sequences	BD rate performance			
	Single layer	Simulcast	SVC	$\Delta QP$
ParkScene	18.10%	-22.60%	-58.40%	6
BasketballDrive	19.20%	-19.90%	-46.90%	6
BasketballDrill	17.50%	-22.30%	-53.20%	6
PartyScene	21.10%	-20.80%	-50.60%	6
BasketballPass	23.70%	-18.50%	-49.20%	6
RaceHorses	23.60%	-23.80%	-43.10%	6
ChinaSpeed	18.50%	-19.70%	-39.20%	6
SlideShow	17.20%	-16.40%	-43.20%	6
Average	17.55%	-20.50%	-47.98%	—

searched and estimated to compensate the prediction image. The enhancement layer PU can be predicted by the base layer motion search block. Similarly, we still choose the inherited base layer motion parameters, thus attaining the motion vector, the reference image index, and prediction direction index for the interprediction. The prediction image will be obtained with the above motion parameters based on the reference pictures of the EL. After the prediction image in the EL is derived from the motion compensation in the above two situations, residual data will be transformed, quantized, and transmitted. The MODE\_SKIP mode is also used in interlayer interprediction. In MODE\_SKIP, any subpartitioning will not be allowed to continue to split except for the  $SIZE\_2N \times 2N$  and the default of its transform quantized coefficients is 0. After the transform and quantization in the residual pixel block, if all the blocks of transform coefficients are less than 1, the skip flag will be initialized to 1. The residual block is set to all-zero block and the process of the DCT/IDCT, Q/IQ and the reconstruction will be terminated, thus reducing the computational complexity of encoding and decoding system.

Above all, the motion parameter of the base layer will be applied in the decoding side and only the residual data is transmitted to the decoding side in enhancement layer prediction. So the inherited motion or merge parameters of the base layer blocks and the enhancement layer residual data can be used to reconstruct the enhancement layer.

#### 4. Experimental Results

The scalable video coding theme is implemented based on HM 8.0, which corresponds to the HEVC text specification draft 8 [12]. The common test configurations are defined in L1009 [13]. Two layers (one base layer and one enhancement layer) are evaluated in random access configuration. The GOP size of hierarchical-B coding (or random access, RA) is 8. The BL and EL are encoded in different QPs with a spatial ratio of 1:1. The QPs of base and enhancement texture views are  $Q_1$  and  $Q_2$ . The common conditions specify the biggish  $Q_1$  (28, 33, 38, and 43), and the gap of QPs between two layers is 6. A comparison is made with single layer, SVC of H.264/AVC standard, and simulcast encoding that means multistream switching. The single layer version had same QPs as the EL of the scalable variation. For simulcast, base

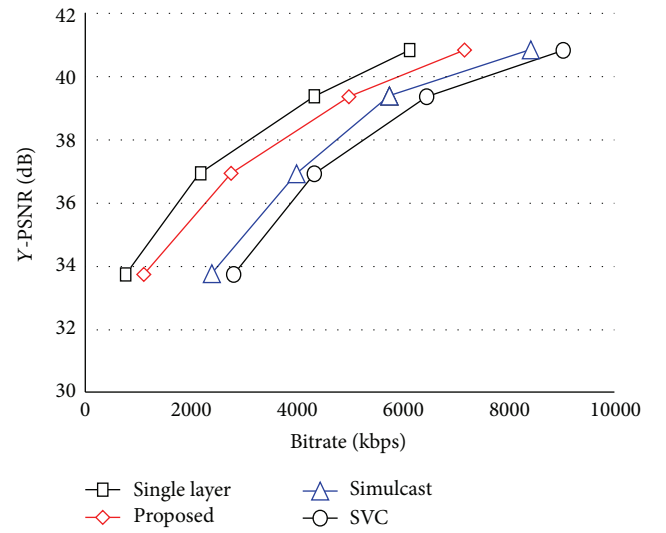


FIGURE 6: PSNR-bitrate among scalable HEVC, single layer, simulcast, and SVC of RaceHorse (832 × 480\_30).

layer and enhancement layer are encoded with quantization parameters  $Q_1$  and  $Q_2$ , respectively, and bitstream is saved on the server. Besides, the bitrate contains all the layers, and the PSNR is the highest enhancement layer with  $Q_2$ .

In Table 1, eight sequences are mainly to be tested, and the PSNR-bitrate graph of RaceHorse and ParkScene is shown in Figures 6 and 7. The results indicate a 20.50% Y-BD-rate decrease on average compared with simulcast. Besides, this proposed theme brings 47.98% Y-BD-rate saving compared with the SVC. As shown in Figures 6 and 7, the performance of the two sequences is presented obviously in different video resolution. The size of GOP and gap of QP are important factors leading to a different result of the sequences. The experimental results show that the coding performance is greatly improved in the scalable scenario in HEVC.

#### 5. Conclusion

The paper presents a scalable video coding theme on the emerging HEVC. The proposed scalable video coding supports two interlayer prediction methods by enabling a single-loop solution. The interlayer intraprediction method exploits



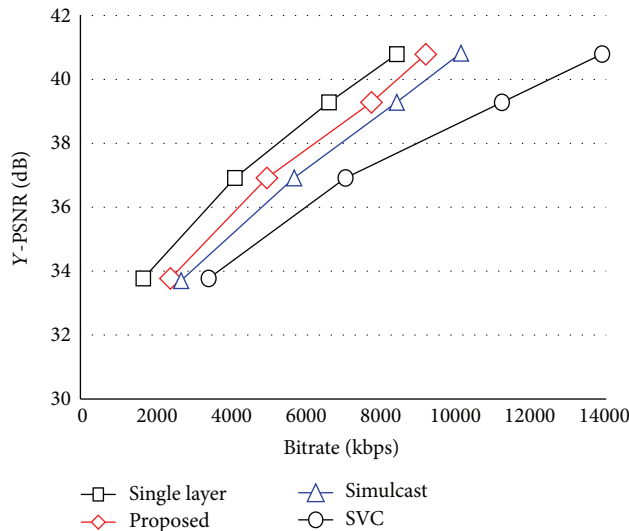


FIGURE 7: PSNR-bitrate among scalable HEVC, single layer, simulcast, and SVC of ParkScene (1920 × 1080\_24).

the interlayer correlation for intracoded CUs. The interlayer interprediction achieves significant bitrate decrease and reduces the complexity for intercoded CUs in the decoder. Experimental results demonstrate the effectiveness of our proposed scenario. Improvement of the coding performance of scalable video coding theme can be explored in the future. More interlayer prediction tools will be proposed to further exploit the temporal-spatial correlation and reduce the overhead bits of enhancement layers.

## Conflict of Interests

The authors declare that there is no conflict of interests regarding the publication of this paper.

## Acknowledgments

This work is supported by the Natural National Science Foundation of China (nos. 61370111, 61103113, and 61272051) and Beijing Municipal Education Commission General Program (KM201310009004).

## References

- [1] B. Bross, "High Efficiency Video Coding (HEVC) text specification draft 10 (for FDIS & Consent)," Document JCTVC-L1003, JCT-VC, Geneva, Switzerland, 2013.
- [2] G. J. Sullivan, J. Ohm, W.-J. Han, and T. Wiegand, "Overview of the high efficiency video coding (HEVC) standard," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 22, no. 12, pp. 1649–1668, 2012.
- [3] H. Bai, C. Zhu, and Y. Zhao, "Optimized multiple description lattice vector quantization for wavelet image coding," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 17, no. 7, pp. 912–917, 2007.

- [4] W.-H. Peng, J. K. Zao, H.-T. Huang, T.-W. Wang, and L.-S. Huang, "A rate-distortion optimization model for SVC inter-layer encoding and bitstream extraction," *Journal of Visual Communication and Image Representation*, vol. 19, no. 8, pp. 543–557, 2008.
- [5] H. Schwarz, D. Marpe, and T. Wiegand, "Overview of the scalable video coding extension of the H.264/AVC standard," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 17, no. 9, pp. 1103–1120, 2007.
- [6] Z. Shi, X. Sun, and F. Wu, "Spatially scalable video coding for HEVC," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 22, no. 12, pp. 1813–1826, 2012.
- [7] G. van Wallendael, J. de Cock, R. van de Walle, and M. Mrak, "Multi-loop quality scalability based on high efficiency video coding," in *Proceedings of the 29th Picture Coding Symposium (PCS '12)*, pp. 445–448, Krakow, Poland, May 2012.
- [8] F. Bossen, B. Bross, K. Suhring, and D. Flynn, "HEVC complexity and implementation analysis," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 22, no. 12, pp. 1685–1696, 2012.
- [9] J. Ohm, G. J. Sullivan, H. Schwarz, T. K. Tan, and T. Wiegand, "Comparison of the coding efficiency of video coding standards-including high efficiency video coding (HEVC)," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 22, no. 12, pp. 1669–1684, 2012.
- [10] I.-K. Kim, K. McCann, K. Sugimoto, B. Bross, and W.-J. Han, "High efficiency video coding (HEVC) test model 10 encoder description," document JCTVC-M1002, JCT-VC, Incheon, South Korea, 18–26 April 2013.
- [11] G. J. Sullivan and T. Wiegand, "Rate-distortion optimization for: video compression," *IEEE Signal Processing Magazine*, vol. 15, no. 6, pp. 74–90, 1998.
- [12] B. Bross, "High efficiency video coding (HEVC) text specification draft 8," document JCTVC-J1003, The Joint Collaborative Team on Video Coding (JCT-VC), Stockholm, Sweden, 2012.
- [13] X. Li, J. Boyce, P. Onno, and Y. Ye, "Common Test Conditions and Software Reference Configurations for the Scalable Test Model," in *Proceedings of the 12th Joint Collaborative Team on Video Coding Meeting (JCTVC '13)*, JCTVC-L1009, Geneva, Switzerland, January 2013.

