Scalable ROI Algorithm for H.264/SVC-Based Video Streaming

Jung-Hwan Lee and Chuck Yoo
Department of Computer Science and Engineering, Korea University, South Korea

***Abstract*- H.264/SVC has been widely adopted to smart CE devices for network-adaptive video transmission. However, current H.264/SVC does not support sufficiently high subjective QoS (Quality of Service) because its scalability methods provide only coarse-grained network adaptivity. To address this problem, we propose Scalable ROI (SROI) algorithm that applies ROI (Region of Interests) concept to SVC (Scalable Video Codec). While existing methods use fixed size of enhancement layers, our method utilizes only ROI area as an enhancement layer so that it can improve subjective QoS with low transmission overhead. Through some experiments, we show that our method can increase PSNR (Peak Signal-to-Noise Ratio) by 4~5 dB in ROI areas.**

I. INTRODUCTION

In the wireless network, network bandwidth and error rate are frequently changed because of user's mobility and dynamic channel condition. This can be a critical problem to video streaming applications because video data is generally very sensitive to delay and error. In this reason, video codecs need to be more network-friendly and should have adaptive bit-rate functions [1]. To address these issues, ITU-T and ISO/IEC MPEG (Moving Picture Experts Group) have published H.264/SVC (Scalable Video Codec) draft. The main feature of H.264/SVC is to provide bandwidth-optimized transmission for video streaming by observing current network condition [2]. That is, encoded video data can be selectively transmitted according to the type of the contents and network condition using bitstream extractor. To support scalable quality, H.264/SVC uses CGS (Coarse Grain Scalability), MGS (Medium Grain Scalability) and FGS (Fine Grain Scalability) methods, which generate one base layer and several enhancement layers. Then each enhancement layer can be transmitted additionally as current available bandwidth increases. In other words, using those methods, we can obtain better objective quality, which is expressed by SNR (Signal Noise Rate), when the network condition is enough good to accommodate the enhancement layers. However, the problem is that existing methods have too high computing complexity to provide fine grained scalability. With coarse-grained scalability, the video quality is fluctuated easily by network dynamics, so it would be difficult to obtain high video quality stably.

In this paper, we propose Scalable ROI (SROI) algorithm, which can support fine-grained scalability with low computing complexity so that we can achieve better video quality. Note that most users are more sensitive to subjective quality rather than objective quality, and most video frames contain interesting area that is related to subjective quality. The interesting area is so-called ROI (Region of Interests), which has been already used in surveillance systems and medical image processing areas [3].

To increase subjective quality, ROIs should have higher quality than the rest of ROIs. By ROI coding, we can provide different levels of quality within one frame and extract higher level of quality for ROI regions. In this manner, the size of enhancement layers can be reduced leading to fine-grained scalability. As a result, a client would feel better subjective quality even though the network condition is unstable.

This work was supported by the Seoul R&BD Program(WR080951)
This work was supported by the National Research Foundation of Korea(NRF) grant funded by the Korea government(MEST) (No. 2010-0029180)

II. SCALABLE ROI ALGORITHM

In this section, we explain our Scalable ROI algorithm in detail. First, ROI area should be determined. Since ROI detection algorithm has been already proposed in the literature [3], this paper simply assumes that the center of a frame has higher priority than the marginal area. Then, our algorithm generates slice groups according to the ROI priority as shown in Fig. 1 (left up). ROI-based encoding can be realized using FMO (Flexible Macroblock Ordering) Type 2. It is used to mark rectangular areas (i.e. ROI) in each frame [4] [5].

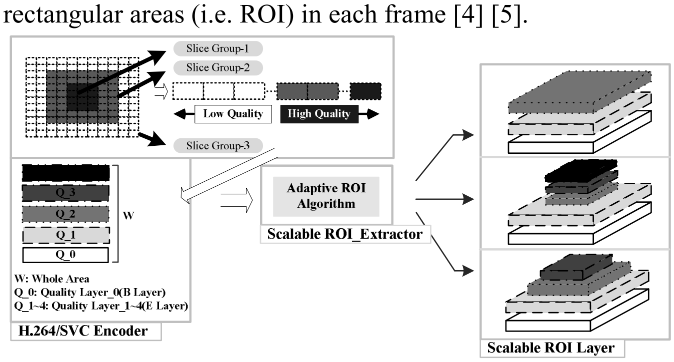

Fig. 1 Scalable ROI structure

Second, Scalable ROI_Extractor periodically analyzes current network condition using the interval between acknowledgements (ACKs). And then, it calculates available network bandwidth, quality layer (only for enhancement layer), quality coefficient, and ROI area.

Finally, the enhancement layers are generated according to the available network bandwidth. Our SROI algorithm has three methods for making Scalable ROI Layer as shown in Fig. 1 (right). The first method (*Sr*0) is to use no ROI coding for the case of high network bandwidth. The second method (*Sr*1) stacks the enhancement layers only for ROI regions. The third method (*Sr*2) uses wider ROI regions, but has smaller number of enhancement layers than the second (*Sr*1).

To select the best method, our SROI algorithm performs the following analysis using layer information and parameters described in Table 1.

$$Bw(x) \ge \sum_{n=0}^{m} BR\left(Wp_{id(n)}(C)\right) + \sum_{l=m+1}^{j} BR\left(Sr_{id(l)}(c)\right) \quad (1)$$

$$if\left[f\left(Wp_{\mathrm{id}}(m)\right) = f\left(\sum_{l=m+1}^{j} Ext_{\mathrm{id}}(l)\right)\right]$$

$$Sr1_{\mathrm{id}}(c) = \sum_{l}^{j}\left[{f\left(Ext_{\mathrm{id}}(l)\right)}/{j+1}\right] \quad (2)$$

$$if\left[f\left({Wp_{\mathrm{id}}(m)}/{2}\right) \le f\left(\sum_{l=m+1}^{j} Ext_{\mathrm{id}}(l)\right)\right]$$

$$Sr2_{\mathrm{id}}(c) = \sum_{l}^{j} MAX\left[\left[{f\left(\mathrm{Ext}_{\mathrm{id}}(l)\right)}/{j}\right], \left[{f\left(\mathrm{Ext}_{\mathrm{id}}(l+1)\right)}/{j}\right]\right] \quad (3)$$

$$if\left[f\left({Wp_{\mathrm{id}}(m)}/{2}\right) > f\left(\sum_{l=m+1}^{j} Ext_{\mathrm{id}}(l)\right)\right]$$

$$Sr0_{\mathrm{id}}(c) = none \quad (4)$$

The formula (1) means that the bitrates of non-ROI layer and ROI layer are not above the available network bandwidth. The *C* coefficient implies objective quality in each layer, and it has the maximum value when the first method (*Sr*0) is used. The *c* coefficient means the level of MGS quality for the second or third method. Finally, the *c* coefficient is determined by (2), (3) and (4).

TABLE I
UNITS FOR SCALABLE ROI ALGORITHM PROPERTIES

Symbol	Quantity
Bw()	Available network bandwidth
BR()	Bit-rate
Wp()	Whole picture of quality parameter
Sr()	ROI picture of quality parameter
f()	Count of quality flag
Ext()	Extractor of quality scalability
m	Number of non-ROI layers
j	Number of ROI layers
id	Content ID

III. EXPERIMENT

We use JSVM 9.13 for the encoding and decoding processor [6], and implement our Scalable ROI algorithm into the bitstream extractor. H.264/SVC video data is extracted with scalable quality based on the bitrates using our bitstream extractor. Then, the bitstream extractor determines which SROI method should be used among the three methods (*Sr*0~*Sr*2) explained in the previous section. Fig.2 shows the standard crew content's (CIF size) PSNR (Peak Signal Noise Rate) using the JSVM 9.13 encoder/decoder with MGS method. And then, the bitstream extractor extracts the encoded data at 900Kbps of bitrate. In this experiment, our SROI algorithm selects the third method (*Sr*2) for Scalable ROI layer. Fig. 2 shows PSNR comparison between traditional and proposed methods, and we can confirm that ROI areas have higher PSNR than non-ROI areas for all frames in the case of the proposed method.

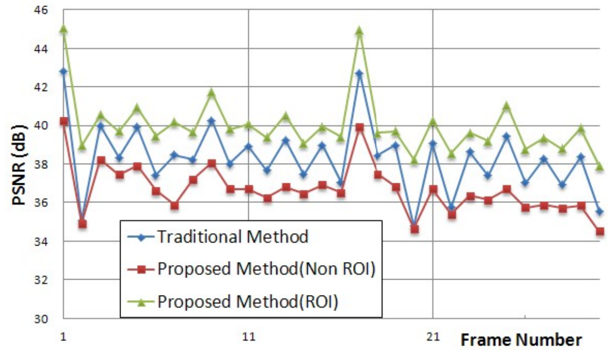

Fig. 2 PSNR comparison between traditional and proposed methods

(a) Traditional method (b) Proposed method

Fig. 3 Comparison of decoded images

The Fig. 3(b) shows more enhanced quality for ROI area (white dotted box) than Fig. 3(a) by our SROI algorithm. On the contrary, the top right area around the man's hand of Fig. 2(b) is fainter than that of Fig. 3(a). This is because the purpose of *Sr*2 method is to enhance subjective quality with small number of enhancement layers.

IV. CONCLUSION

This paper proposes Scalable ROI algorithm considering users' region of interests in various areas on frames. The traditional coarse-grained method cannot provide high video quality when network condition is unstable. On the other hand, our algorithm can support high subjective quality with fine-grained scalability by applying ROI method to scalable video codec (H.264/SVC). The proposed algorithm can be practically utilized in surveillance camera systems because surveillance camera videos usually have less motion and obvious ROI.

V. REFERENCE

[1] T. Schierl, T. Stockhammer, and T. Wiegand, "Mobile Video Transmission Using Scalable Video Coding", *IEEE Transactions on Circuits and Systems for Video Technology*, Vol. 17, No. 9, pp. 1204-1217, Sep. 2007.

[2] H. Sun, A. Vetro, and J. Xin, "An overview of scalable video steaming" *Wireless Communications & Mobile Computing,*Vol. 7, No. 2, pp. 159-172, Feb. 2007.

[3] Q. Chen, G. Zhai, X. Yang, and W. Zhang, "Application of Scalable Visual Sensitivity Profile in Image and Video Coding", in *Proc. IEEE International Symposium on Circuits and Systems*, May 2008.

[4] S. Benierbah and M. Khamadja, "A New Technique for Quality Scalable Video Coding with H.264", *IEEE Transactions on Circuits and Systems for Video Technology*, Vol.15, No. 11, pp. 1332–1340, Nov. 2005.

[5] M. A. A. Dewan, M. J. Hossain, and O. Chae, "Segmentation of Moving Object for Content Based Applications", in *Proc. International Conference on Consumer Electronics (ICCE)*, Jan. 2009.

[6] SVC Reference Software, [Online]. Available: CVS server(host: garcon.ient.rwth-aachen.de)