

Quality scalable coding of selected region

Wook-Joong Kim, Jong-Won Yi, Seong-Dae Kim

Dept. of Electrical Engineering, Korea Advanced Institute of Science and Technology,
373-1 Kusong-Dong, Yusong-Gu, Taejeon, 305-701, Korea

Tel: +82-42-869-5430; fax: +82-42-869-3410

e-mail: wjk@sdvision.kaist.ac.kr

ABSTRACT

If a region is semantically more important than others, it is appropriate that a image compression scheme is capable of handling the regional semantic difference because the information loss of the interested region is more severe. We propose the quality scalable coding with its model by introducing the quality scale parameter. It is more extended and generalized image compression philosophy than a conventional coding. As an implementation of the proposed quality scalable coding, a H.263 based scheme is presented. This scheme can control the temporal and spatial quality efficiently, and improve the reconstructed image quality of the interested region.

1 INTRODUCTION

Generally, visual information coding systems are designed to minimize overall distortions of the input image without regarding its semantic importance. However, there are certain cases when a region of the input image is more meaningful than others such as a person who is speaking in the videoconference, a player in sports games, or an enemy's vehicle in tactical scenes. Therefore, because the information loss of this region (foreground region) is more severe than other region (background region), a visual communication system should be capable of handling the difference of regional characteristics.

In order to meet this requirement, it is appropriate to discriminate the amount of allocating bits for each region. As the classical rate-distortion theory represents, higher reconstructed quality for the foreground region can be obtained by allocating more bits than the background region. Some techniques[1]–[7] are proposed to handle the bit quantity of the foreground region. Lee[1] proposed 'selective coding' for tactical images, Labit[2] proposed 'region of interest (ROI) based encoding' and presented simple mathematical model of ROI coding, and Eleftheriadis[3] proposed H.261 based coding scheme. Object scalability which is a functionality of MPEG-4 requirements and JPEG-extension[8] can be also considered as the method for this purpose. However, these methods only concerned presenting method-

ologies of system implementation. Unifying mathematical model or definitive principle is not presented, so there is no basis how one method is compared with another, and no rule how much amount of the foreground quality can be improved.

In this paper, we define the foreground quality variable coding as 'quality scalable coding of selected region (QSC)'. To resolve the ambiguities which previous works posses, we present the mathematical model of the QSC using the 'quality scalable parameter (QSP)' which is normalized value between 0 and 1. And as an example of the QSC implementation, a H.263 compatible coding scheme is presented.

2 THE MODELING OF THE QSC

Assume an image or sequence of images are given, and the foreground region and the background are predefined. Let R_f , R_b are the allocated bits for foreground region and background region respectively. The coding without considering the difference of information for each region can be represented as follows.

$$\begin{aligned} &\text{Maximize } Q(\text{foreground} + \text{background}) \\ &\text{subject to } R_f + R_b \leq R_{\text{budget}} \end{aligned} \quad (1)$$

where, R_{budget} is the bit budget and $Q(\cdot)$ represents the reconstructed quality of a region. Let Q'_f be the foreground quality and R'_f , R'_b be the allocated bits for the foreground and background region respectively when the Eq.(1) is satisfied. Consider a case when a foreground quality is lower than Q'_f despite the sum of allocated bits are equal to the sum of R'_f and R'_b . It means the allocated bits for the foreground region is less than R'_f , and the allocated bits for the background region is higher than R'_b . In this case, the background region is treated as more important neglecting the assumption that the information loss of the foreground region is more severe than the background region. Therefore, we conclude that Q'_f is the reasonable minimum quality for the foreground region.

Next, consider a different case when the foreground region is maximally emphasized. This can be repre-

sented as Eq.(2).

$$\begin{aligned} & \text{Maximize } Q(\text{foreground}) \\ & \text{subject to } R_f + R_b \leq R_{\text{budget}} \end{aligned} \quad (2)$$

In constart to the Eq.(1), the foreground quality, Q_f'' , which satisfies the Eq.(2) is the possible maximum value satisfying the bit budget. Therefore, from Eq.(1) and Eq.(2), it is possible to consider that we can control the foreground quality according to the relative importance compared to background region within the range from Q_f' to Q_f'' – this is the goal of the QSC.

Introducing the quality scale parameter (QSP), the QSC is represented as follows: the QSC is controlling the foreground quality as

$$Q(\text{foreground}) \propto QSP \quad (0 \leq QSP \leq 1) \quad (3)$$

If $QSP = 0$, it means the foreground and background region are evenly treated as the Eq.(1). And if $QSP = 1$, it means the foreground region is maximally emphasized as the Eq.(2).

3 H.263 BASED IMPLEMENTATION OF THE QSC

In this section, as an example of the QSC, the H.263 based moving picture coding scheme is presented. We assume that the foreground region is segmented by a segmentation algorithm such as [1, 3, 6, 9, 10]. And it is also assumed that the decoder is only capable for the bit stream which is generated by the default H.263 coder. Because H.263 is a block based coding system, the macroblock which is in the foreground region or on the boundary of two regions is regarded as the foreground MB (MB^f), and remaining macroblocks are considered as the background MB (MB^b).

Because an image sequence is a 3-dimensional signal, we consider the QSC in two ways: the temporal quality scaling which handles the frame rate and the spatial quality scaling which manages the foreground PSNR.

3.1 Temporal quality scaling

In low bit rate coding situations, it is necessary to make a compromise between the reconstructed image quality and the frame rate. As the frame rate increases, we can improve the subjective quality of an image sequence because we can obtain smooth motion of objects in the sequence – that is, decrease of the motion jerkiness. On the contrary, due to the limitation of transmission rate, the degradation of the reconstructed image quality is unavoidable.

Meanwhile, as one way for achieving the higher frame rate, instead of inserting the basic picture (BP) which is coded without considering the difference of the regional importance like plain pictures in conventional video coding systems, consider that the foreground emphasized picture is inserted for the skipped frame. The foreground emphasized picture (EP) represents the frame

that the most of the background information is intentionally disregarded by assigning the most coarse quantizer (i.e., QP=31) to MB^b while the QP of MB^f is determined by the rate controller. Surely, the degradation of reconstructed quality is also unavoidable in this case. However, because we already suppressed the unimportant information in the background region, we can get the effect of the frame rate increment with lower degradation of the reconstructed foreground quality than the BP inserted case. This is the goal of the temporal quality scalable coding (TQSC). Fig. 1 shows the conceptual diagram of the TQSC.

We formulate the TQSC as follows: let FR_{max} be the maximally possible input frame rate and FR_{min} be the minimum frame rate of the BP that should be guaranteed. The TQSC is controlling the frame rate (FR_{TQSC}) by inserting the EP as Eq.(4) by the temporal quality scaling parameter ($QSP_{temporal}$, $0 \leq QSP_{temporal} \leq 1$).

$$FR_{TQSC} = (FR_{max} - FR_{min}) \cdot QSP_{temporal} + FR_{min} \quad (4)$$

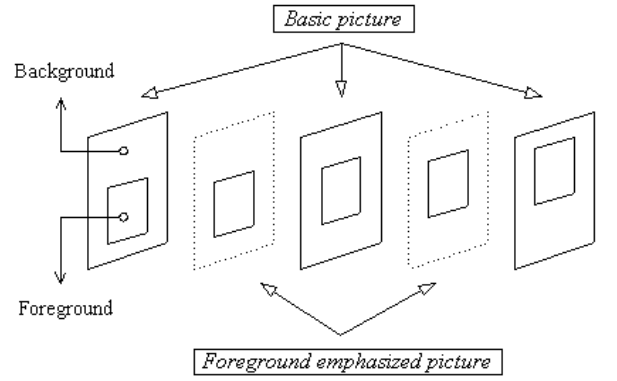


Figure 1: The conceptual diagram of the TQSC

3.2 Spatial quality scaling

The goal of the spatial quality scalable coding (SQSC) is improving the foreground quality – in this paper, we consider the PSNR as the spatial quality measure. This can be accomplished by saving bits from the background region and adding them to the foreground region.

Remaining the foreground allocated bits by the rate controller, we apply following methods to decrease the background bits – that is, the following methods are only applied to MB^b in the BP. In controlling the foreground quality, the spatial quality scaling parameter ($QSP_{spatial}$, $0 \leq QSP_{spatial} \leq 1$) is used as a scaling index.

3.2.1 The change of quantization parameter

One easy way of reducing bits is increasing the quantization parameter. We change the QP for MB^b as Eq.(5).

$$QP_{MB^b} = \text{Int}[(31 - QP_{rate}) \cdot QSP_{spatial} + QP_{rate}] \quad (5)$$

where $Int[\cdot]$ is rounding function to the nearest integer and QP_{rate} is the quantization parameter determined by the rate controller.

3.2.2 Residual error suppression

In moving picture coding systems, the residual error is coded to compensate the motion prediction error. In [11, 12], it is proposed that the compression efficiency can be enhanced by transmitting only the relevant part of the error without causing the noticeable degradation of the subjective quality. We use this technique to suppress the background bits. First, residual errors below a threshold are set to zero, and remaining errors are low pass filtered. Because the magnitude of the threshold decided the reduction rate of residual errors, we determine it as

$$Int[QSP_{spatial} \times QP_{MB}] \quad (6)$$

And the low pass filter of its coefficients

$$\left[\frac{1}{12}, \frac{1}{6}, \frac{1}{2}, \frac{1}{6}, \frac{1}{12} \right]$$

is used.

4 SIMULATION RESULTS

In simulations, we test the proposed scheme to the QCIF ‘Stefan’ sequence. As a segmentation result of the foreground and background region, the two-level alpha channel image is used. Fig. 2 shows the test image and its mask image – the region of a tennis player is considered as the foreground region. To make the reconstructed frame rate equals to the maximum input frame rate, the last decoded frame is repeated for the skipped frame. And, the transmission bit rate is fixed at 64Kbits/sec for all simulations.

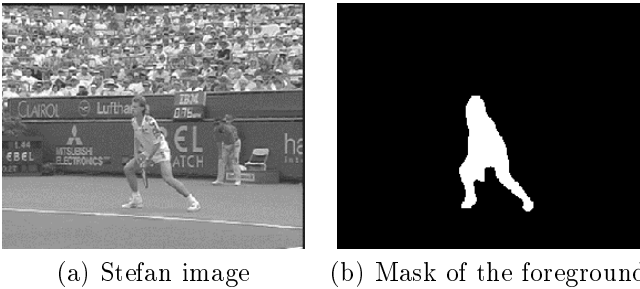


Figure 2: The test image and the foreground make

First, the TQSC is examined. Fig. 3 shows the foreground PSNR for some $QSP_{temporal}$ values when $FR_{min} = 8.33$ (frames/sec) and $FR_{max} = 30$ (frames/sec). For $QSP_{temporal} = 0.0$, the foreground PSNR varies abruptly because of the frame repetition for skipped frames. As $QSP_{temporal}$ increases, though there is small degradation of the foreground PSNR about 1 - 2 dB in average, the PSNR variation among

frames becomes small. This exhibits that the foreground region is reconstructed with smooth motion, and the motion related artifacts (e.g. motion jerkiness) are reduced.

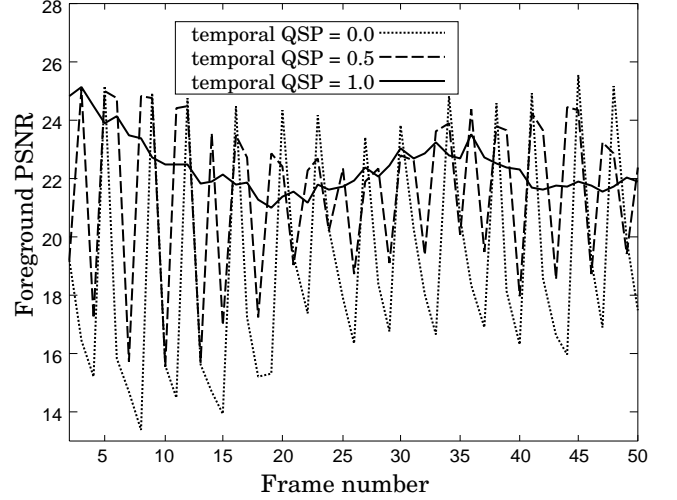


Figure 3: The results of the TQSC

Next, the SQSC is examined with the frame rate of 8.33 frames/sec. Fig. 4 shows the foreground PSNR obtained by the SQSC varying $QSP_{spatial}$. For this figure, the PSNR of the transmitted frame is only shown. As the $QSP_{spatial}$ increases, the foreground PSNR is also enhanced.

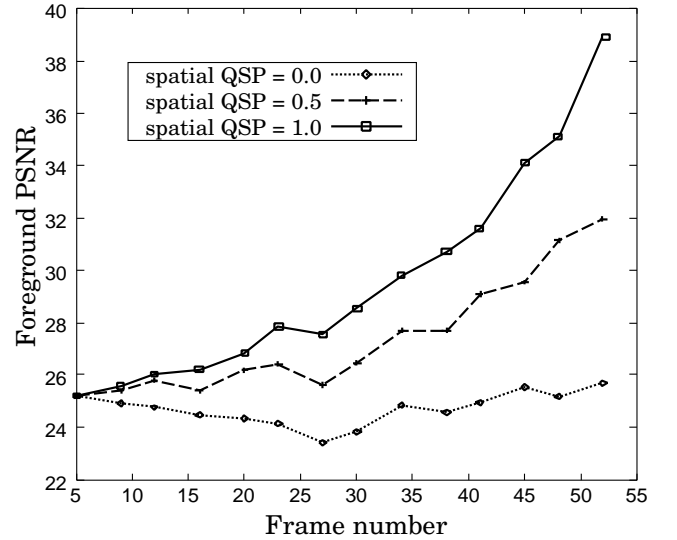


Figure 4: The comparison results of the SQSC

Finally, the temporal and spatial quality scaling are combined simultaneously. For the simulation we set FR_{max} to 30 frames/sec and FR_{min} to 8.33 frames/sec. Fig. 5 shows the results. As the frame number increase, the foreground PSNR enhancement and the reduction of quality variation compared with H.263 coding situation

are both achieved by adjusting the scale parameters.

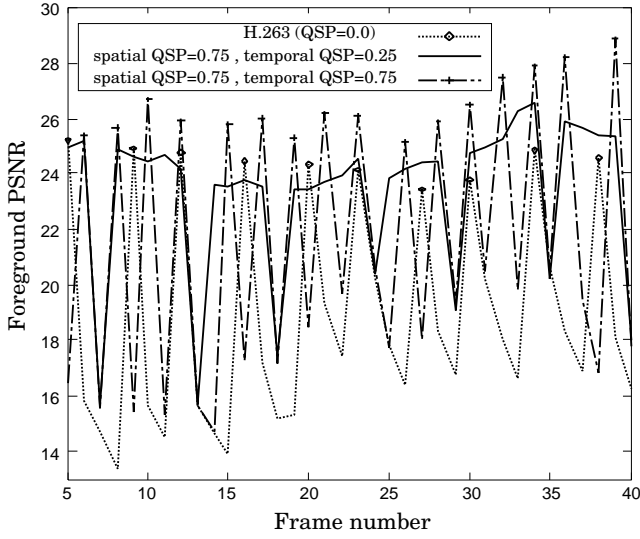


Figure 5: The results of the temporal and spatial quality scalable coding

Fig. 6 shows a comparison of the reconstructed images by the proposed QSC scheme and H.263 method. Although the background region of (b) has more degradation than H.263 case (a), we obtain the higher quality foreground region.

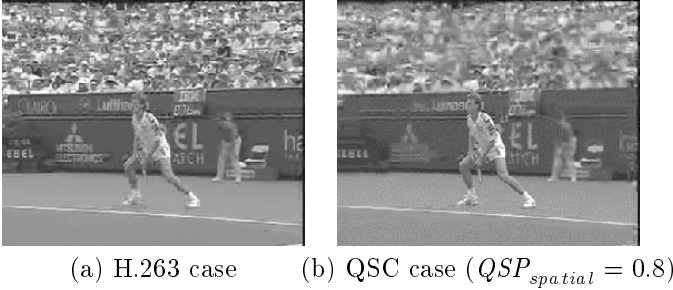


Figure 6: The comparison of a reconstructed image

5 CONCLUSION

In this paper, the quality scalable coding of selected region is proposed. We presented the model of the QSC by introducing the quality scale parameter. This model is more extended coding concept than conventional video coding, and unifies the coding concepts which are proposed in [1]–[7]. As an example of the QSC, the H.263 based scheme is presented. The simulation results show that the proposed system can efficiently control the foreground quality by changing the spatial and temporal quality scale parameters.

References

- [1] J.B. Lee, S.D. Kim, “A selective coding method based on global/local motion information”, *Signal Processing*. Vol. 55, 1996, pp339-350.
- [2] C. Labit, J.P. Leduc, “Very low bit rate(VLBR) coding schemes: a new algorithmic challenge?”, *Proc. of VCIP*, 1994, pp25-38.
- [3] A. Eleftheriadis, A. Jacquin, “Automatic face location detection for model - assisted rate control in H.261-compatible coding of video”, *Signal Processing : Image Communication*, Vol. 7, 1995, pp435-455.
- [4] O. Egger et. al, “Region based coding scheme with scalability features”, *Proc. of EUSIPCO*, 1996, pp747-750.
- [5] A. Susanto et. al., “Layered coding of check image using foreground and background segmentation”, *Proc. of VCIP*, 1996, pp1040-1049.
- [6] C.S. Kim, W.J. Kim, S.D. Kim, “A novel selective coding scheme based on layered structure”, *Proc. of ICIP*, 1996, pp355-358.
- [7] F. Muller, K. Illgner, “Hybrid video compression considering regions of interest”, *Proc. of VCIP*, 1997, pp666-675.
- [8] ISO/IEC 10918-3-ITU-T Recommendation T.84, “Information technology – Digital compression and coding of continuous-tone still image : Extension”, Nov. 1994.
- [9] J.B. Lee, S.D. Kim, “Target region extraction and image coding based on motion information”, *Proc. of VCIP*, 1995, pp705-716.
- [10] J.G. Choi, S.D. Kim, “Multi-stage segmentation of optical flow field”, *Signal Processing*, Vol. 54, 1996, pp109-118.
- [11] Mark. R.B. et. al., “Low bit rate coding using robust motion vector regeneration in the decoder”, *IEEE trans. on Image Processing*, Vol. 3, No. 5, Sep. 1994, pp652-665.
- [12] B. Hurtgen, M. Gilge, W. Guse, “Coding of moving video at 1Mbit/s - movie on CD”, *Proc. of VCIP*, 1990, pp1092-1103.