Filtering Scheme for ROI Coding by Dynamic Range Compression and Updating Source Picture Filter

Kazuhito Sakomizu^{a, b, *}, Takashi Nishi^a, Takao Onoye^b

^aOki Electric Industry Co., Ltd., 2-6-8 Bingomachi Chuo-ku, Osaka, 541-0051, Japan

^bOsaka University, 1-1 Yamadaoka, Suita 565-0871, Japan

*Corresponding Author: sakomizu688@oki.com

Abstract

The region of interest (ROI) coding, which compresses background region more strongly than foreground region, is valuable technology in restricted communication environments such as limited bandwidth and/or storage capacity. This paper proposes a filtering scheme which reduces bit amount by filtering background region in order to support efficient ROI coding on ordinary devices without equipping ROI coding functionality. In particular, dynamic range compression (DRC) filter and updating source picture (USP) filter are added to a low-pass filter, which has been used for conventional filtering scheme. The DRC filter compresses dynamic range of all pixels by down-scaling. The USP filter measures distance between current and previous picture by block basis, and then overwrites the block of the current picture when the distance is smaller than predetermined threshold. Experimental results show that the proposed filtering scheme reduces bit-rate by half compared to the conventional low-pass filter as PSNR-Y of background region is about 30 dB. Compared to uniform quality coding, the proposed filtering scheme reduces bit-rate by 62 %. These results promise to cut down on cost of communication and storage.

Keywords: region of interest, filtering scheme, video coding.

1. Introduction

Region of interest (ROI) coding is technology to ensure quality in ROI on limited bandwidth by allocating more bit amount to ROI than background region. In case of a part of video surveillance system, it is sufficient if a

security officer can identify what persons in field angle are who⁽¹⁾. In this case, face region must be done finely to identify the individual adequately, while in other region rough quality which is enough to understand monitoring location is allowed. In this situation, ROI coding can reduce bit amount in background region, and then bit-rate reduction is realized⁽²⁾. Recently computer vision technique such as face detection has been developed to a practical level⁽³⁻⁸⁾, and commercialization is also in progress⁽⁹⁾. We expect that combination of face detection and ROI coding resolves cost problem of video surveillance systems.

The simplest method for ROI coding is QP control scheme, which uses built-in function for quality control^(10,11). For example, the H.264 / MPEG4 AVC⁽¹²⁾, which is mainstream coding standard in video surveillance systems, has syntax to control quality by a macro block (MB). This scheme has the advantage that coding efficiency of background region is high, while there is restriction that we have to select video coding standard and implementation which have built-in function to control quality of a MB^(13,14). In video surveillance systems, hardware accelerator is used for video encoding because video encoding is commonly known as heavy process. However that encoder necessarily does not support the built-in function. In addition, this scheme is also a disadvantage that range of trade-off between quality and bit-rate is limited within used video coding standard. In case of H.264 / MPEG4 AVC, range of the quality parameter is defined as value from 0 to 51.

Another method for ROI coding is filtering scheme which filters background region of a picture to reduces bit amount before encoding⁽¹³⁻¹⁵⁾. For instance, a low-pass filter such as the Gaussian filter can reduce bit amount in background region. Generically video encoder compresses bit amount by removing spatial and temporal redundancy,

and therefore the amount of bits in region where redundancy is increased by the low-pass filter is expected to be reduced. An advantage of filtering scheme is that it can also be used in the situation where using video coding implementation does not support ROI coding, because filters are independent from the codec. Moreover this scheme is also an advantage that range of trade-off between quality and bit-rate is not limited. This paper considers these features to be important, and discusses the filtering system in the following.

Filtering scheme has the advantages as described above, while coding efficiency is low compared to the QP control scheme. This is because discussion in terms of filtering design is not enough since the H.264 / MPEG4 AVC supports syntax for quality control of a MB. In case of a low-pass filter, high-frequency components are suppressed, and the amount of bits is reduced. However the QP control scheme also varies bit amount of background region by controlling degree of entropy reduction in quantization process and degree of skip mode selected in prediction process. Since measures for these differences are not sufficient, coding efficiency of filtering scheme in background region is low.

This paper proposes a new filtering scheme that are consisted of dynamic range compression (DRC) filter and updating source picture (USP) filter in addition to a conventional low-pass filter. The DRC filter compresses dynamic range of all pixels by down-scaling. The filter realizes the function which differentiates the amount of bits by changing dynamic range, similar to entropy, of the source signal in QP control scheme. The USP filter measures distance between current and previous picture by block basis, and then overwrites the block of the current picture to induce skip mode when that distance is smaller than predetermined threshold. This filter realizes the function by which skip mode is likely to be selected in background region. The proposed filtering scheme reduces bit amount of background region efficiently by combination of these filters. If filtering scheme can sufficiently reduce the amount of bits, it is expected to dramatically increase the applicable range of ROI coding.

This paper is organized as follow. Section 2 describes a proposed system for ROI coding using the proposed filtering scheme. Section 3 explains detail of proposed filters such as the DRC filter and the USP filter. This section is main contribution of this paper. Section 4 presents performance of the proposed filtering scheme. Finally, Sec. 5 concludes this paper.

Fig. 1. Overall of the proposed system for ROI coding.

Fig. 2. Normalization of ROI.

2. The Proposed System for ROI Coding

The proposed system filters background region in a source picture based on ROI information detected by face detection engine, and encodes the filtered picture. Figure 1 is overall image of the proposed system. Detail is shown in following paragraphs.

Face detection engine searches face region from the source picture, and output rectangular coordinates of the face region. Several face detection methods has been proposed as previously described. In this paper, face detection method using the Haar-like feature^(4,5) is used for the face detection engine.

The face detection engine outputs rectangular coordinates of face areas by a pixel. In contrast, the encoder divides a source picture into MBs, and encodes the MBs. In this paper, the system normalize the coordinates of the ROI from the face detection engine, and makes ROI be divided by a MB by the ROI normalization in Fig. 1. Figure 2 shows an image of normalization process. That process is formulated below.

Fig. 3. Block diagram of pre-filter.

$$X_{L,T} = x_{L,T}/N, \quad (1)$$

$$Y_{L,T} = y_{L,T}/N, \quad (2)$$

$$X_{R,B} = (x_{R,B} + 15)/N - 1, \quad (3)$$

$$Y_{R,B} = (y_{R,B} + 15)/N - 1, \quad (4)$$

where $x_{L,T}$ and $y_{L,T}$ are coordinates of top left pixel of ROI. $x_{R,B}$ and $y_{R,B}$ are those of bottom right pixel. $X_{L,T}$ and $Y_{L,T}$ are those of top left pixel of normalized ROI. $X_{R,B}$ and $Y_{R,B}$ are those of bottom right pixel. N is size of MB. In this paper, $N = 16$ because H.264 / MPEG4 AVC is oriented.

The normalized coordinates are transmitted to the pre-filter. The pre-filter reduces the amount of bits of region belonging to the MB determined as background. Detail of information about the pre-filter is described in following section.

Finally a filtered picture of which the amount of bits in background region is reduced is inputted to the H.264 / MPEG4 AVC encoder. In this paper, the baseline profile of open-source software x264 is used as the H.264 / MPEG4 AVC encoder.

In the decoder system, the H.264 / MPEG4 AVC decoder receives and decodes the transmitted stream. Then the decoded picture is transmitted to the post-filter. In the proposed system, it is necessary to decompress dynamic range to the original scale at the decoding side, since the pre-filter generates the picture of which dynamic range in a background region is compressed as described above. The post-filter is a module for this function. Output of the post-filter is that of the proposed decoder system.

3. Pre-filter and Post-filter

This section describes detail of the pre-filter and post-filter those are the core of the proposed filtering

Fig. 4. Flow chart of the DRC filter.

scheme. First, a structure of pre-filter is shown in Fig 3. The pre-filter is consisted of a low-pass filter, the DRC filter and the USP filter. Purpose and behavior of each module is described in following subsections.

3.1 Low-pass Filter

Low-pass filter has been conventionally used in filtering scheme^(13,15). This filter reduces the amount of bits in background region by suppressing the high-frequency components. In this paper the general Gaussian filter is used as this filter. Definition of the used Gaussian filter is following.

$$G(x, y) = \frac{1}{2\pi\sigma^2} e^{-\frac{x^2+y^2}{2\sigma^2}}, \quad (5)$$

$$\sigma = 0.3 \times \left(\frac{k}{2} - 1\right) + 0.8, \quad (6)$$

where is the variance parameter of the Gaussian filter, and the value is obtained from the kernel size k . In the proposed filtering scheme, k is assumed to be a 3, 7, 15 or 31.

3.2 The DRC Filter

The QP control scheme differentiates the amount of bits by changing dynamic range, similar to entropy, of the source signal. The module to realize the similar function in filtering scheme is the DRC filter. The DRC filter compresses dynamic range by performing the down-scaling processing for all pixels in pixel domain. By compressing dynamic range, entropy of picture signal becomes low, and therefore the amount of bits is reduced.

Figure 4 shows behavior of the DRC filter. First subtraction 128 from the pixel value p_i is performed to make center value of scaling process 128. Then the value down-scaled by using a predetermined scaling factor Re ,

Fig. 5. Flow chart of the USP filter.

Fig. 6. Flow chart of the post-filter.

and rounded to an integer value. Finally compressed value p_o is produced by adding 128 to the integer value. By the way, by using this filter, background region becomes dim since dynamic range of that region is narrow. When viewing the decoded video, to solve this problem is a post-filter to be described later.

3.3 The USP Filter

In the QP control scheme, skip mode is likely to be selected in background region. This influences on the differentiation of the amount of bits. The module to realize the similar function in filtering scheme is the USP filter. The USP filter measures distance between current and previous picture by block basis, and then overwrites the block of the current picture to induce skip mode when that distance is smaller than predetermined threshold, Figure 5 shows behavior of the USP filter. In this filter, it is determined whether the overwriting process is performed or not by threshold processing of following.

$$\text{MAD}(B_t(m, n), B_{t-1}(m, n)) < R_e \times T, \quad (7)$$

where MAD is a function to calculate the mean absolute difference (MAD) of the two blocks. t is picture index of the block, and $m; n$ represents position of the block in the picture. $B_t(m, n)$ represents 4×4 block to be filtered, and $B_{t-1}(m, n)$ represents the co-location block of $B_t(m, n)$ in the filtered picture transmitted to the encoder just before time. T is the parameter which is used to calculate the threshold. In the proposed filtering scheme, dynamic range of background region is compressed, and the higher scaling factor R_e becomes, the lower the difference of $B_t(m, n)$ and $B_{t-1}(m, n)$ are. Therefore multiplying R_e to threshold is performed to determine whether the overwriting process is performed or not in a stable manner regardless of R_e .

If the condition (7) is met, update processing $B_t(m, n)$ is applied.

$$B_t(m, n) = B_{t-1}(m, n). \quad (8)$$

By copying the filter block $B_{t-1}(m, n)$, the USP filter encourages that $B_t(m, n)$ is encoded by skip mode in the H.264 / MPEG4 AVC encoder.

3.4 The post-Filter

In the proposed system, it is necessary to decompress dynamic range to the original scale at the decoding side, since the pre-filter generates the picture of which dynamic range in a background region is compressed as described above. The post-filter is a module for this function. The post-filter performs inverse operation of the DRC filter to return the compressed pixel value the original scale. Without this filter, it is possible to view a decoded picture, while color of background region becomes dim. Figure 6 shows behavior of the post-filter. q_i and q_o represents pixel value. The basic behavior is the same as the DRC filter. The difference is that the scaling factor is reciprocal of R_e , $R_d = 1/R_e$, and that cropping process is added to the end. In the cropping process, the negative pixel values are shifted to 0, and 256 or more pixel values are shifted to 255.

4. Simulation

This section evaluates reduction efficiency of bit-rate of the proposed filtering scheme by simulation. Evaluation is performed by comparing the proposed filtering scheme with a conventional filtering scheme and QP control scheme. In this paper, the Gaussian filter is used the conventional filtering scheme.

4.1 Evaluation Method

To evaluate the reduction efficiency of bit-rate by ROI encoding, the graph of which horizontal axis is reduction rate, and vertical axis is quality of decoded pictures in background region is drawn. In this paper, reduction rate Q is presented by following equation.

$$Q = \frac{F_B - F_R}{F_B}, \quad (9)$$

where F_B is a reference bit-rate calculated by uniform quality coding. F_R is a bit-rate of ROI coding. As the quality of decoded pictures, the PSNR-Y of background region is used. This is because this paper assumes that quality of ROI is same in the uniform quality coding.

4.2 Parameter Configuration

Video used to draw the graph is Bowling created by FUJITSU LABORATORIES LTD. We select video which is taken by the fixed camera since we assume a surveillance system. Resolution of this video is 352 times 288, and that frame rate is 30 fps.

For QP, the value of ROI is fixed as PSNR-Y of ROI becomes about 39 dB. This condition is derived from that PSNR-Y of ROI becomes about 39 dB when the video is encoded with uniform quality in QP = 26. In the QP control scheme, QP of background region is varied from 27 to 51 to differentiate the amount of bits of ROI and background region. In filtering schemes, QP of both ROI and background region is fixed by 26, and the filters differentiate the amount of bits.

As described above, the proposed filtering scheme is formed by combining three kinds of filters. Some of these filters, it is assumed that valuable tool set changes with depending on target bit-rate. Therefore whether the each filter is used or not is also treated as a parameter. Other parameters of the proposed filtering scheme are kernel size of the Gaussian filter k , the scaling factor of the DRC filter R_e , and the parameters of the USP filter T . In this paper, the graph is drawn by selecting efficient point in data generated from performing compression and decompression by a number of combinations of above parameters.

In case of the conventional filtering scheme, the graph is drawn by varying the kernel size in $k = 3, 15, 31, 63, 127, 255$.

4.3 Simulation Results

Figure 7 shows a simulation result of reduction efficiency and PSNR-Y. When compared to case of the

Fig. 7. Relationship between reduction efficiency and PSNR-Y of background region.

proposed filtering scheme and the conventional scheme, at all the points the proposed filtering scheme more efficiently reduces the amount of bits than the conventional scheme does. For example, when PSNR-Y is approximately 30 dB, reduction rate of the conventional scheme is approximately 25 %, while that of the proposed filtering scheme is about 62 %. On the basis of bit-rate of the conventional scheme, the proposed filtering scheme can reduce bit-rate to half roughly while maintaining quality.

On the other hand, comparison of the proposed filtering scheme and the QP control scheme shows that the proposed filtering scheme is slightly inferior than the QP control scheme is. When PSNR-Y is approximately 30 dB, the reduction rate of the QP control scheme is about 69 %, and difference between two approaches is about 7 %.

Figure 8, 9 10, 11 show comparison of decoded pictures when performing the uniform quality encoding Fig. 8, the QP control scheme Fig 9, the conventional filtering scheme Fig. 10 and the proposed filtering scheme Fig. 11. PSNR-Y of each decoded pictures are approximately 30 dB. Comparing ROI, each ROI codings can achieve fine quality enough to identify the individual. In background region, tendency of the degradation between each scheme is different although the same quality is achieved in terms of PSNR-Y. For the QP control scheme, since the amount of bits is reduced by coarse quantization, block noise is conspicuous. On the other hand, the conventional filtering scheme reduces the amount of bits just by the low-pass filter, and therefore block noise is not observed while blur is observed over the picture instead. In case of the proposed

Fig. 8. A decoded picture of uniform quality coding.

Fig. 10. A decoded picture of conventional filtering scheme.

Fig. 9. A decoded picture of QP control scheme.

Fig. 11. A decoded picture of proposed filtering scheme.

filtering scheme, tendency is similar to the QP control scheme. The QP control scheme is better than the proposed filtering scheme with respect to reproducibility of color, and the proposed filtering scheme is fine with respect to reproducibility of detail such as edges.

5. conclusion

In this paper, a new filtering scheme is proposed to reduce the amount of bits in background region efficiently. The proposed filtering scheme is designed at the thought of the way that the QP control scheme reduces the amount of bits. It is consisted of the DRC filter and the USP filter in addition to the conventional low-pass filter. The simulation results show that the proposed filtering scheme can reduces bit-rate to half roughly while maintaining quality when

PSNR-Y is approximately 30 dB. Compared to the uniform quality coding, 62 % of bit-rate can be reduced, and we expect a significant reduction in storage cost and communication cost of a wireless environment.

References

- (1) S. Daly, K. Matthews, and J. Ribas-Corbera : "Face-based visually-optimized image sequence coding", in Proceedings 1998 International Conference on Image Processing, Vol. 3, pp. 443-447, 1998.
- (2) T. Hata, N. Kuwahara, T. Nozawa, D. Schwenke, and A. Vetro : "Surveillance System with Object-Aware Video Transcoder", in 2005 IEEE 7th Workshop on Multimedia Signal Processing, pp. 1-4, 2005.
- (3) M. Turk and A. Pentland : "Face recognition using eigenfaces", in Proceedings. 1991 IEEE Computer

- Society Conference on Computer Vision and Pattern Recognition, pp. 586-591, 1991.
- (4) P. Viola and M. Jones : "Rapid object detection using a boosted cascade of simple features", in Proceedings of the 2001 IEEE Computer Society Conference on Computer Vision and Pattern Recognition. CVPR 2001, Vol. 1 , pp. I-511-I-518, 2001.
 - (5) R. Lienhart and J. Maydt : "An extended set of Haar-like features for rapid object detection", in Proceedings. International Conference on Image Processing, Vol. 1, pp. I-900-I-903, 2002.
 - (6) Rein-Lien Hsu, M. Abdel-Mottaleb, and A. Jain : "Face detection in color images", IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol. 24, No. 5, pp. 696-706, 2002.
 - (7) Yao-Xin Lv, Zhi-Qiang Liu, and Xiang-Hua Zhu : "Real-time face detection based on skin-color model and morphology filters", in Proceedings of the 2003 International Conference on Machine Learning and Cybernetics, Vol. 5, pp. 3203-3207, 2003.
 - (8) Mei-Juan Chen, Ming-Chieh Chi, Ching-Ting Hsu, and Jeng-Wei Chen, : "ROI video coding based on H.263+ with robust skin-color detection technique", IEEE Transactions on Consumer Electronics, Vol. 49, No. 3, pp. 724-730, 2003.
 - (9) C. Zhang and Z. Zhang, : "A Survey of Recent Advances in Face Detection," Microsoft Research, p. 17, 2010.
 - (10) Y. Liu, Z. Li, Y. Soh, and M. Loke : "Conversational Video Communication of H.264/AVC with Region-of-Interest Concern," in 2006 International Conference on Image Processing, pp. 3129-3132, 2006.
 - (11) Yu Sun, I. Ahmad, Dongdong Li, and Ya-Qin Zhang : "Region-based rate control and bit allocation for wireless video transmission", IEEE Transactions on Multimedia, Vol. 8, No. 1, pp. 1-10, 2006.
 - (12) T. Wiegand, G. Sullivan, G. Bjontegaard, and A. Luthra : "Overview of the h. 264/avc video coding standard", Circuits and Systems for Video Technology, IEEE Transactions on, Vol. 13, No. 7, pp. 560-576, 2003.
 - (13) L. Karlsson and M. Sjostrom : "Improved ROI video coding using variable Gaussian pre-filters and variance in intensity", in IEEE International Conference on Image Processing 2005, pp. II-313, 2005.
 - (14) Y. Zheng, X. Tian, and Y. Chen : "Adaptive Frequency Coefficient Suppression for ROI-Based H.264/AVC Video Coding", in 2008 IEEE International Conference on Networking, Sensing and Control, pp. 714-718, 2008.
 - (15) D. Grois and O. Hadar : "Complexity-aware adaptive bit-rate control with dynamic ROI pre-processing for scalable video coding", in 2011 IEEE International Conference on Multimedia and Expo, pp. 1-4, 2011.