International Journal on New Computer Architectures and Their Applications (IJNCAA) 1(2): 316-324
The Society of Digital Information and Wireless Communications, 2011 (ISSN: 2220-9085)

A Scalable Block Compensation with Index Selection Mechanism to Achieve High Video Quality Retrieval

Ta-Te Lu[1], Tsung-Hsuan Tsai[2], Jia-Yuan Wu[1]
[1]Department of Computer Science & Information Engineering, Ching Yun University, Chung-Li, Taiwan.
[2]Department of Electric Engineering, Ching Yun University, Chung-Li, Taiwan.
{ttlu, m9852003}@cyu.edu.tw, tsunghsuantsai@gmail.com

ABSTRACT

Video quality retrieval is an important issue in reconstruction video sequences. The amounts of transmission bitstreams are very large to achieve high video quality reconstruction. Thus, high video quality retrieval is still a challenge problem. The aim of this paper is to reduce the transmission bitstreams and raise reconstruction video quality. In this paper, we propose a scalable block compensation with index selection mechanism using eigenvalue and local variance statistics for each group of pictures (GOP), which are regarded as the feature parameters for compensating motion blocks for each video frame from low-resolution to high-resolution sequences at the receiver ends. To reduce transmission overhead, the index selection technique is then applied to those selective compensation blocks to have variances larger than a threshold. The results of index numbers from index selection are regarded as reconstruction information. Simulation results show the percentage of non-compensation blocks is nearly 30% saving for Mobile, 36% saving for Foreman, and 48% saving for News. These results indicate that the proposed algorithm offers a more efficient way to decrease the percentage of compensation blocks about 30~48%. The proposed algorithm, on average, outperforms H.264/AVC about 2.8~9.8 dB from 100 kbps to 1 Mbps.

KEYWORDS

Scalable block compensation, eigenvalue, variance, index selection, H.264/AVC.

1 INTRODUCTION

The Internet is a communication infrastructure that interconnects a global community of end users and content servers. In recent years, multimedia video streaming distribution over the Internet has been an intensively studied topic [1]. However, the amount of video streaming distribution, coupled with limited network bandwidth, causes network delays in transmission. Thus, high video quality retrieval is still a challenging problem at the receiver ends. H.264 is the standard with the highest performance in video data compression [1]. Compared with MPEG-4 and H.263, H.264 nearly doubles the compression rate at the same video quality. Motion estimation and motion compensation methods in H.264 are used to reduce temporal redundancy. Loomans et al. [2]-[3] proposed highly parallel predictive search and enhanced predictive zonal search methods to improve prediction of motion estimation for scalable video coding. Paul [4] used phase correlation to select motion estimation and compensation modes directly. Xiao [5] based on the correlation between the multi-reference frame and the size of the motion vector, reduced the H.264/AVC encoder's computational complexity. Zhang [6] proposed an inter-layer motion compensation method to improve the visual quality of the reconstruction. Wang [7] combined the predictors of template matching and block motion compensation to save the bit-rate of motion compensation about 11%. Super-resolution (SR) gets a high-resolution (HR) frame from a series of low-resolution (LR) video frames at the receiver ends, which is one of methods to improve the visual quality of the reconstruction. SR reconstruction technology is applied to many applications, e.g. medical, remote sensing, video surveillance and video conference. The SR technique has the advantage of an existing low-resolution system that can be still used.

In recent years, SR researchers have focused on improving computational efficiency, fast reconstruction or robust SR algorithms [8]-[13]. Anantrasirichai and Canagarajah [8] used a weighting map to decrease the inaccuracy of motion estimation and a quantization noise model for low bit-rate in the super-resolution estimator. Hung and Siu [9] proposed a translational motion compensation model via frequency classification for video super-resolution systems. Banerjee [10] proposed a critical-based sampling function to apply higher sampling for high motion and lower sampling for edge regions. Zibetti and Mayer [11] proposed a class super-resolution algorithm to exploit the correlation among the frames of a video sequence. The weakness of this method is that the computational cost and the number of iterations are still high. Katartzis and Petrou [12] used the Bayesian estimation method for SR reconstruction, though the estimation algorithm was too complex for computing. Martins et al. [13] used Markov random fields for SR reconstruction and the iterated conditional models for computing the maximum a posteriori conditional probability. The weakness of this algorithm is blurring distortion. Overall, these SR methods still waste too much computing time for video reconstruction. Therefore, the aim of this paper is to reduce the transmission bitstreams and raise reconstruction video quality. To achieve this, we use eigenvalue to obtain the features in each Group of Pictures (GOP) and variance to estimate blocks' differences from high-resolution (HR) frames to low-resolution (LR) frames. Then, the proposed algorithm is applied for video compensation to achieve high visual quality retrieval.

The paper is organized as follows: Section 2 presents the block diagram of feature detection for the video encoder. Section 3 presents the proposed reconstruction structure. Simulation results are given in Section 4. Finally, conclusions are drawn in Section 5.

2 BLOCK DIAGRAM OF FEATURE DETECTION FOR VIDEO ENCODER

H.264 is the standard with the highest performance among H.263 and MPEG-4 in video data compression. It supplies several prediction modes in intra and inter frames. The original intra frame is divided into non-overlapping 16x16 macroblocks with each macroblock containing sixteen 4x4 blocks. For each macroblock, the one/sixteen prediction costs are estimated to obtain the minimum rate-distortion (R-D) cost from the left and up neighborhood macroblocks.

The block diagram of video encoder with the feature detection algorithm, scalable block compensation with index selection is illustrated in Fig. 1. The video coding basically includes three major parts – quantization, motion estimation, and motion compensation. First, DCT transform is used to reduce intra picture correlation. The DCT has the property of energy compaction, and most of the signal information concentrates in a few low-frequency components after DCT transform [14]. Second, motion estimation and motion compensation techniques are applied to reduce the temporal redundancy. The feature detection mechanism identifies each GOP's characteristic after motion compensation, and scalable block compensation is used to compensate values from high-resolution (HR) frames to low-resolution (LR) frames.

The index selection mechanism is applied to compensate for selective blocks that have variances larger than a threshold and to reduce the transmission amount. The results of index numbers from index selection are regarded as reconstruction information. The details of the features detections, scalable blocks compensation and index selections are described separately as follows.

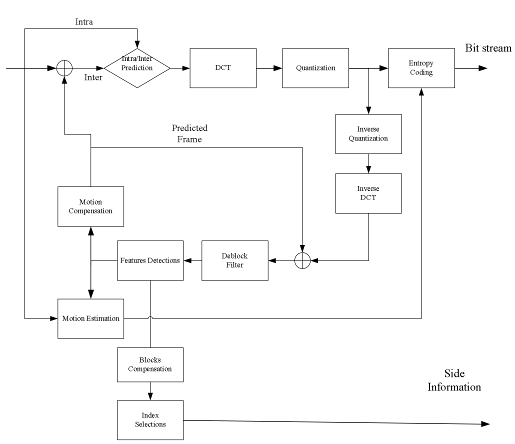


Fig. 1. The block diagram of video encoder with the feature detections.

2.1 FEATURE DETECTION

The feature detection mechanism identifies each GOP's characteristic after motion compensation. We denote video sequence with n GOPs, each GOP having m frames. Each frame has a matrix A with $N \times N$ coefficients, the autocorrelation matrix R_{NxN} is calculated from matrix A and A^T

$$R_{NxN} = E\left[AA^T\right], \quad (1)$$

where A^T with $N \times N$ coefficients, and then the eigenvalue $\lambda = \{\lambda_1, \lambda_2, \ldots \lambda_N\}$ is calculated from the characteristic equation as follows

$$R_{N \times N} = E[AA^T] = \lambda\ v, \quad (2)$$

$$\det(R - \lambda I) = 0, \quad (3)$$

where $v = \{v_1, v_2, \ldots, v_N\}$ represents eigenvectors of $E[AA^T]$ associated with λ.

Frame feature F_k in k-th frame can be calculated as

$$F_k = \frac{1}{N}\sum_{i=1}^{N}\lambda_i \quad (4)$$

, where F_k represents the mean of all eigenvalues in k-th frame. The total GOP features is denoted as $\{G_1, G_2, \ldots, G_n\}$, where G_j is the j-th GOP with m frames $\{F_1, F_2, \ldots, F_m\}$, which is calculated as

$$G_j = \frac{1}{m}\sum_{i=1}^{m} F_i \quad (5)$$

2.2 SCALABLE BLOCK COMPENSATION

To increase coding efficiency, the scalable block compensation method is applied to divide motion compensation blocks into compensation blocks and non-compensation blocks. Variances are used to obtain the features of each video frame from low-resolution (LR) to high-resolution (HR) sequences. The HR and LR video sequences have m frames. Each frame is partitioned into k non-overlapping macroblocks, each block having p pixels, and then the variance between HR and LR video frames is estimated to obtain the block compensation values as the feature in each video frame. The blocks' variances are denoted as $\sigma = \{\sigma_1^2, \sigma_2^2, \ldots, \sigma_k^2\}$, and then the j-th blocks' variance σ_j^2 is calculated as

$$\sigma_j^2 = \sum_{i=1}^{p} \left| (x_i - \eta)^2 - (y_i - \eta)^2 \right|, \qquad \textbf{(6)}$$

where x_i is the LR pixel value, y_i is the HR pixel value, $\eta = \frac{\eta_x + \eta_y}{2}$ is the average result of the LR mean η_x and the HR mean η_y. The blocks' variances are from $\{\sigma_{\min}^2, \ldots, \sigma_{\max}^2\}$; $\sigma_{\min}^2$ represents that the results $\{\sigma_1^2, \sigma_2^2, \ldots, \sigma_k^2\}$ are reordered coefficients having high correlation between HR blocks and LR blocks, $\sigma_{\max}^2$ represents that the correlation of the block coefficients between HR and LR is very low.

$$\begin{cases} f_b = 1, & \sigma_{\max}^2 - \sigma_{\min}^2 \geq T_0 \\ f_b = 0, & \sigma_{\max}^2 - \sigma_{\min}^2 < T_0 \end{cases} \qquad (7)$$

If the variance result larger than threshold T_0, then the block is set as a compensation block. Otherwise, the block is set as a non-compensation block. If the block is a non-compensation block, then the flag f_b is set to '0' for the non-compensation block. Otherwise, the flag f_b is set to'1' for the compensation block. To reduce transmission overhead, all pixels in compensation blocks indicate one index number in next section.

2.3 INDEX SELECTION

The transmission amount is very large for video sequences reconstruction. To reduce the transmission amounts, the index selection mechanism is applied to those pixels in the compensation blocks f_b=1. Otherwise, the pixels in the non-compensation blocks f_b=0 are set to zero. Each compensation block has p pixels, and then the differences between the pixels of the HR and LR video frames are estimated to obtain the blocks' distortion values.

$$D_{i,j}^l = HR_{i,j}^l - LR_{i,j}^l, \qquad \textbf{(8)}$$

where $l \in \{1, 2, \ldots, m\}$, $i \in \{1, 2, \ldots, k\}$, $j \in \{1, 2, \ldots, p\}$. The maximum estimation difference value $D_{\max}^l$ and the minimum estimation difference value $D_{\min}^l$ in the l-th frame are determined by (8). The range $[D_{\min}^l, D_{\max}^l]$ will be divided into t intervals and the interval size Δ is an integer,

$$D_i^l = D_{\min}^l + \Delta \times n_i, \qquad \textbf{(9)}$$

where $l \in \{1, 2, \ldots, m\}$, $n_i \in \{1, 2, \ldots, t\}$, $D_i^l \in \{D_{\min}^l + \Delta, D_{\min}^l + 2\Delta, \ldots, D_{\max}^l\}$.
The compensation value

$$D_{k_i, p_i}^l = D_{\min}^l + (n_i - \frac{1}{2}) \times \Delta \qquad \textbf{(10)}$$

for p_i-th coefficient in k_i-th compensation block, for which

$$D_{\min}^{l} + (n_i - 1)_{p_i} \times \Delta \le D_{k_i,p_i}^{l} < D_{\min}^{l} + (n_i)_{p_i} \times \Delta$$
.

3 Reconstruction process

The steps of reconstruction process are described as follows:

Step 1: LR video streaming and side information are received for reconstruction;
Step 2: Video decoder is applied to get the LR video sequence from LR video streaming;
Step 3: The same procedure is used as in section 2.1 to get the features of LR video sequence;
Step 4: The HR reconstruction video sequence is from block compensation. If the flag is set '1'for the k_i-th block in the l-th frame, the reconstruction p_i -th coefficient $\widetilde{R}_{ki,pi}^{l}$ is compensated in the block as

$$\widetilde{R}_{ki,pi}^{l} = LR_{ki,pi}^{l} + D_{\min}^{l} + (n_i - \frac{1}{2})_{p_i} \times \Delta. \quad \textbf{(11)}$$

Otherwise, the flag is set '0', then the compensation value can be ignored so that the reconstruction p_i -th coefficient $\widetilde{R}_{ki,pi}^{l}$ is $LR_{ki,pi}^{l}$.

4 SIMULATION RESULTS

Simulations are performed on various video-coding algorithms, including H.264/AVC and our proposed method at the same bit-rate. Class B and Class C sequences, e.g. Mobile, Foreman and Stefan, were sampled at frame rate 5 fps, GOP=10, encoded bit-rate from 100 kbps to 1 Mbps. All video sequences were in the QCIF (176x144) format. The experiments were conducted using Matlab with Intel CPU (1.66GHz), 1 GByte memory. The threshold T0 = 8 and $\Delta = \pm 10$ are chosen for the simulation test. Fig. 2, Fig. 3 and Fig. 4 show the reconstruction results of H.264/AVC and proposed algorithm for Mobile, Foreman and Stefan sequences from 100 kbps to 1Mbps, respectively. In the condition of the same bit rate including the overhead, on average, the proposed algorithm outperformed H.264/AVC by 9.8 dB for Mobile, 2.8 dB for Foreman, and 8.26 dB for Stefan, respectively. Fig. 5, Fig. 6 and Fig. 7 show the 64-th and 153-th reconstruction results for Mobile, the 190-th and 288-th reconstruction results for Foreman, and the 74-th and 144-th reconstruction results for Stefan, respectively. Both the perceptual quality and the peak signal-to-noise ratio (PSNR) of the proposed algorithm perform better than H.264/AVC. These results reveal that the proposed algorithm offers a more efficient way to decrease the percentage of compensation blocks about 30~48%. The proposed algorithm, on average, outperforms H.264/AVC about 2.8~9.8 dB at the receiver end.

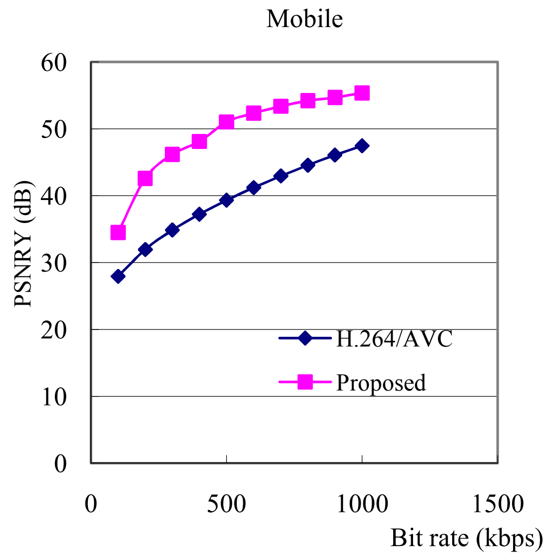


Fig. 2 PSNR results of H.264/AVC and the proposed algorithm for Mobile sequence from 100 kbps to 1 Mbps.

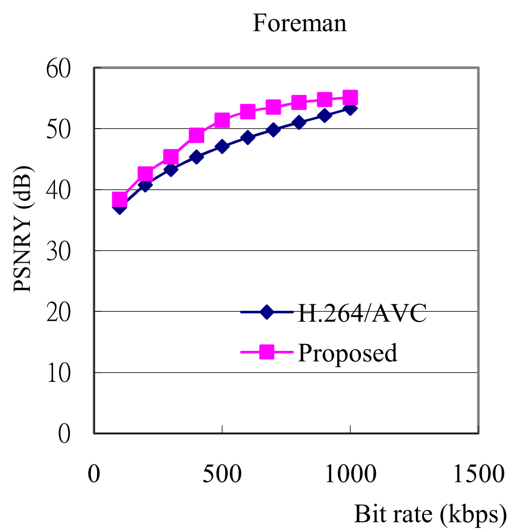


Fig. 3 PSNR results of H.264/AVC and the proposed algorithm for Foreman sequence from 100 kbps to 1 Mbps.

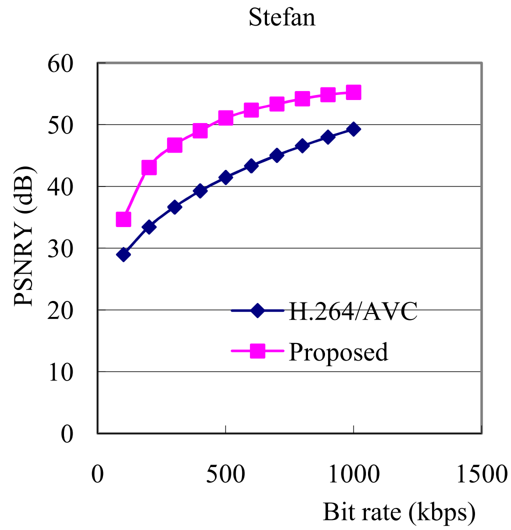
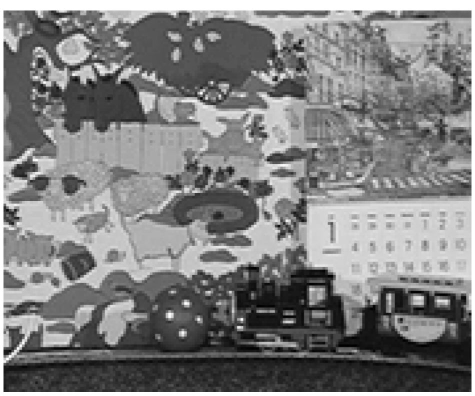


Fig. 4 PSNR results of H.264/AVC and the proposed algorithm for Stefan sequence from 100 kbps to 1 Mbps.

(a)

(b)

(c)



(d)

Fig. 5 Reconstruction video frames for Mobile at 30 kbps using H.264/AVC (a) 64-th, PSNR=17.96 dB (c) 153-th, PSNR=22.48 dB and the proposed algorithm (b) 64-th, PSNR=25.19 dB (d) 153-th, PSNR=43.91 dB

(a)

(b)

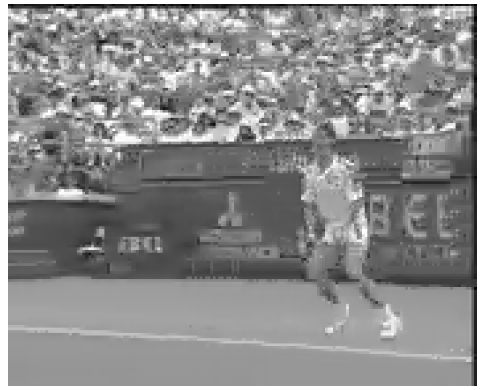
(c)

(d)

Fig. 6 Reconstruction video frames for Foreman at 30 kbps using H.264/AVC (a) 190-th, PSNR=30.46 dB (c) 288-th, PSNR=28.46 dB and the proposed algorithm (b) 190-th, PSNR=43.88 dB (d) 288-th, PSNR=42.40 dB

(a)

(b)



(c)

(d)

Fig. 7. Reconstruction video frames for Stefan at 30 kbps using H.264/AVC (a) 74-th, PSNR=24.03 dB (c) 144-th, PSNR=23.75 dB and the proposed algorithm (b) 74-th, PSNR=39.35 dB (d) 144-th, PSNR=40.08 dB

5 CONCLUSIONS

This paper presents a scalable block compensation with index selection mechanism that is appropriate for lower resolution compensation to achieve high visual quality at the receiver end. This approach is particularly suitable for coding sequences at very low-bit rates. Simulation results show that the proposed method decreases the percentage of compensation blocks about 30~48% and enhances video quality of H.264/AVC about 2.8~9.8 dB at the receiver end. Furthermore, this block compensation mechanism can be applied to other video coding schemes.

ACKNOWLEDGMENTS. The authors would like to thank the National Science Council of the Republic of China, Taiwan for financially supporting this research under Contract No. NSC 99-2221-E-231-031.

6 REFERENCES

1. Wiegand, T. G., Sullian, J. G., Bjontegaard, Luthra, A.: Overview of the H.264 Video Coding Standard. IEEE Trans. Circuits Syst. Video Technol., vol. 13, pp. 560-576, July 2003.
2. Loomans, Marijn J. H., Koeleman, Cornelis J., With, Peter H. N. de: Highly-parallelized Motion Estimation for Scalable Video Coding. In: Proc. 2009 IEEE International Conf. Image, pp. 1577 -1580, Nov. 2009.
3. Loomans, Marijn J. H., Koeleman, Cornelis J., With, Peter H. N. de: Enhanced Prediction for Motion Estimation in Scalable Video Coding. In: Proc. 2010 IEEE International Conf. Image, pp. 1301 -1304, Nov. 2010.
4. Peter, M., Lin, W., Lau, C. T., Lee, B. H.: Direct Intermode Selection for H.264 Video Coding Using Phase Correlation. IEEE Trans. Image Processing, vol. 20, pp. 461-473, Feb. 2011.
5. Xiao, M., Cheng, Y.: A Fast Multi-reference Frame Selection Algorithm for H.264/AVC. In. Proc. 2011 IEEE International Conf. Computer Science and Automation Engineering, pp. 615-619, July 2011.
6. Zhang, R., Comer, M. L.: Efficient Inter-Layer Motion Compensation for Spatially Scalable Video Coding. IEEE Trans. Circuits and Systems for Video Technology, vol. 18, pp. 1325-1334, Oct. 2008.
7. Wang, R., Huo, L., Ma, S., Gao W.: Combing Template Matching and Block Motion Compensation for Video Coding. In. Proc. 2010 International Symposium on Intelligent Signal Processing and Communication Systems, pp. 216 – 221, Dec. 2010.
8. Anantrasirichai, N., Canagarajah, C. N.: Spatiotemporal Super-Resolution for Lowbitrate H.264 Video. In: Proc. 2010 IEEE International Conf. Image, pp. 2809 -2812, Sept. 2010.
9. Hung, K. W., Siu, W. C.: New Motion Compensation Model via Frequency Classification for Fast Video Super-resolution. In: Proc. 2009 IEEE International Conf. Image, pp. 1193 – 1197, Nov. 2009.
10. Banerjee, S.: Low-Power Content-Based Video Acquisition for Super-resolution Enhancement. IEEE Trans. on Multimedia, vol. 11, pp. 455-464, April 2009.
11. Zibetti, M. V., Mayer, J.: A Robust and Computationally Efficient Simultaneous Super-Resolution Scheme for Image Sequences. IEEE Trans. Circuits Syst. Video Technol., vol. 17, pp. 1288-1300, Oct. 2007.
12. Katartzis, A., Petrou, M.: Robust Bayesian Estimation and Normalized Convolution for Super-resolution Image Reconstruction. In: Proc. 2007 IEEE International Conf. Computer Vision and Pattern Recognition, pp.1 – 7, June 2007.
13. Martins, A. L. D. M., Homem, R. P., Mascarenhas, N. D. A.: Super-Resolution Image Reconstruction using the ICM Algorithm. In: Proc. 2007 IEEE International Conf. Image, vol. 4, pp. 205 -208, Oct. 2007.
14. Rao, KR, Yip, P.: Discrete cosine transform: algorithms, advantages, applications, Advantage, Applications. New York: Academic, Boston, 1990.