2010 International Symposium on Intelligent Signal Processing and Communication Systems (ISPACS 2010) December 6-8, 2010

COMBINING TEMPLATE MATCHING AND BLOCK MOTION COMPENSATION FOR VIDEO CODING

Ronggang Wang[1], Longshe Huo[2], Siwei Ma[3], Wen Gao[3]

School of Computer & Information Engineering, Peking University Shenzhen Graduate School[1]
China Unicom Research Institute[2]
School of Electronics Engineering and Computer Science, Peking University[3]

ABSTRACT

A new motion compensation mode of TM_BMC is proposed by this paper to improve video coding efficiency. In TM_BMC, the inter-prediction efficiency is improved by combining the predictor of template matching and predictor of traditional block motion compensation, the coding cost for motion vector of traditional block motion compensation is reduced by utilizing motion vector derived by template matching as motion vector predictor. Experiment results on H.264 KTA platform show that average 7.07% and up to 11.26% bit-rate can be saved by locally adaptive motion compensation mode between the traditional motion compensation mode of BMC and the proposed mode of TM_BMC.

***Index Terms*—** H.264/AVC, template matching, multihypothesis prediction, motion vector prediction

1. INTRODUCTION

H.264/AVC is state of the art video coding standard, great technique progresses over previous standards lie in motion compensation improvement, such as multiple reference frames, variable block size motion compensation, fractional pixel motion compensation and in-loop filtering etc. Recently, ITU-T VCEG is exploiting new technique evidences which can further improve the coding performance of H.264/AVC, and sets up a new experiment platform named as H.264 KTA platform [1]. Some new techniques related to motion compensation have been adopted, such as adaptive interpolation filter [2], 1/8 pixel motion compensation [3], and motion vector competition [4] etc.

Template matching is applied in inter-frame coding in [5]. The motion vector of the target block is estimated at both encoder and decoder side by template matching process, so there's no need for coding motion vectors. Similarly, a decoder-side motion vector derivation (DMVD) scheme is proposed to H.264 KTA platform [6].

The inter-prediction scheme by averaging multiple block motion compensation predictors (Multihypothesis Prediction) to improve the prediction efficiency of single predictor is proposed in [7]. The similar idea is applied in template matching prediction. Template matching prediction is improved by averaging multiple template matching predictors both for intra-prediction and inter-prediction in [8] and [9]. However, to achieve a reasonable performance gain, up to 16 template matching predictors [9] have to be estimated both at encoder side and decoder side, which involves heavy computing and memory access load for codec.

A new inter-prediction scheme based on template matching is proposed in this paper. The traditional motion compensation prediction is enhanced by adding a new motion compensation mode of TM_BMC. The predictor of TM_BMC is derived by combining predictor of template matching and predictor of traditional block motion compensation. The selection between the traditional block motion compensation BMC mode and the new TM_BMC mode is based on Rate Distortion Optimization criteria. The motion vector derived by template matching is utilized as motion vector predictor of block motion compensation in TM_BMC mode, to save the coding cost of motion vectors. Experiment results on H.264 KTA platform show great coding efficiency improvement by adding the proposed TM_BMC mode as a new optional motion compensation mode.

The rest of this paper is organized as follows. The proposed scheme by adding the proposed TM_BMC mode as a new optional motion compensation mode is described in detail in section 2. The experiment results by integrating the proposed scheme into H.264 platform are showed in section 3, followed by conclusions and future works in section 4.

2. PROPOSED METHOD

The traditional video encoder architecture of H.264/AVC is showed in Fig. 1. Generally a macroblock is a basic processing unit for encoding process. For an inter-coded macroblock, the motion vectors of the macroblock's partitions according to its partition mode are derived by the module of Block Motion Estimation. Then, the macroblock's coding mode with the least coding cost is selected by the module of Mode Selection, the motion parameters of the

This work is supported by grants from National Basic Research Program of China under contract No. 2009CB320907.

978-1-4244-7371-7/10/$26.00 ©2010 IEEE

selected mode are fed to the module of Motion Compensation to get the motion compensated predictor of target macroblock. The performance of the traditional motion compensation is constrained by the accuracy of single motion vector. It is well known that by averaging multiple predictors, the performance of motion compensation can be greatly improved. Multihypothisis prediction [7] and OBMC [11, 12] are examples of this kind of technique. In this paper, template matching technique is exploited to get an additional predictor, and then combine it with the traditional block motion compensation predictor to improve the motion compensation performance. A great advantage of the proposed method is that the additional predictor is involved without any coding cost for motion vectors. We named this motion compensation mode as TM_BMC mode.

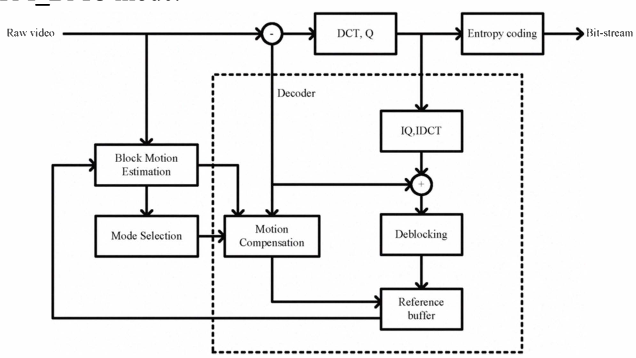

Fig.1. Traditional video encoder architecture of H.264/AVC

In this paper, TM_BMC mode is added as a new optional motion compensation mode upon the traditional motion compensation. The new proposed encoder architecture is showed in Fig.2. The grey modules of Block Motion Estimation and Motion Compensation in Fig.1 are changed, and a new module of Template Matching is added by this proposed method.

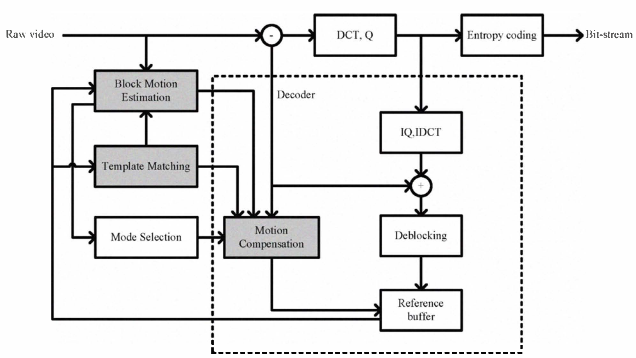

Fig.2. The proposed encoder architecture

For each macroblock partition with TM_BMC mode, to get its motion vector, a template matching process is first performed on the target macroblock partition. The detail template matching process is showed in Fig.3. A template region is defined as the reconstructed pixels extended to the left and top of the target partition. We search the predictor of the template region in the reference pictures, and calculate template matching cost of each candidate predictor. The template matching cost is defined as the difference between the template region and its predictor. The motion vector (and reference picture index) of the predictor with the least template matching cost is set as the motion vector (and reference picture index) of the template matching. We displace the target partition in its reference picture by the motion vector of its template matching, and get one predictor for the target partition. We call this predictor as the template matching predictor.

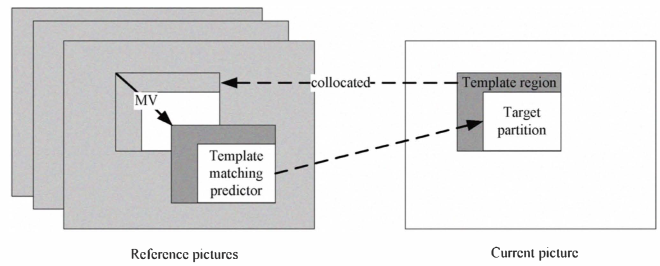

Fig.3. Template matching process illustration

Then, the traditional block motion estimation is performed based on the template matching predictor to get the motion vector of macroblock partition with TM_BMC mode. Each candidate motion vector is searched to get the block motion compensation predictor. The final motion compensation predictor of TM_BMC mode is calculated by combining template matching predictor and the block motion compensation predictor as following:

$$pred_{tm\_bmc} = \alpha * pred_{tm} + (1-\alpha) * pred_{bmc}, 0 < \alpha < 1 \quad (1)$$

Where $pred_{tm\_bmc}$ denotes motion compensation predictor of TM_BMC mode, $pred_{tm}$ denotes the template matching predictor, and $pred_{bmc}$ denotes the block motion compensation predictor, the combining parameter α controls the contribution of template matching predictor and block motion compensation to the final motion compensation predictor. In our current implementation we set α as 0.5. The candidate motion vector which is used to get the best $pred_{tm\_bmc}$ given $pred_{tm}$ and α is selected as the motion vector of macroblock partition with TM_BMC mode. Where, the best $pred_{tm\_bmc}$ is the predictor with the least differences from the raw target macroblock partition.

Up to now, there're two motion compensation modes for each macroblock partition. One is the traditional block motion compensation mode, its inter-prediction predictor is obtained by displacing in its reference picture by the block motion vector of this partition. The other is the TM_BMC mode, its inter-prediction predictor is calculated by combining template matching predictor and block motion compensation predictor. The motion compensation mode between the two modes is selected for each macroblock partition by Rate Distortion Optimization [10] criteria.

The motion cost of a macroblock partition with motion vector MV, reference index RF, and motion compensation mode MC_MODE is calculated, according to the following formula:

$$COST_m = D_m + \lambda_m * (MV_r + RF_r + MC\_MODE_r) \quad (2)$$

Where D_m represents the differences between motion compensated predictors and raw values of current macroblock partition, λ_m represents the trade-off between D_m and the bits for coding motion vector, reference picture index and motion compensation mode, MV_r represents the bits for coding motion vector, RF_r represents the bits for coding reference picture index, and MC_MODE_r represents the bits for coding motion compensation mode.

The variable D_m measures the overall degree of differences between motion compensated predictors and raw values of target partition based on sum of square differences (SSD). The SSD is calculated as the following form:

$$SSD = \sum_x \sum_y (s_{x,y} - pred_{x,y})^2 \quad (3)$$

Where:

x is the row index of sample in a macroblock partition,

y is the column index of sample in a macroblock partition,

$s_{x,y}$ is a raw pixel in the xth row and yth column of the macroblock partition in current frame,

$pred_{x,y}$ is a sample in the xth row and yth column of the maroblock partition's motion compensated predictor.

It will be appreciated that the lower is the difference indicated by SSD, the more efficient for motion compensated prediction.

The motion vector of macroblock partition with TM_BMC mode is coded as motion vector difference between the real motion vector and its predictor, the motion vector predictor is the one derived from template matching process other than the motion vector prediction result of H.264/AVC.

The motion compensation mode with the least motion cost calculated by formula (2) is selected for the target macroblock partition.

According to the motion information of each macroblock partitions obtained from Block Motion Estimation, a mode decision process is performed to select a macroblock coding mode by the module of Mode Selection. The optimal macroblock coding mode for the macroblock is the one that minimizes the Lagrange cost. The Lagrange cost is calculated according to the following formula:

$$COST = D + \lambda * (MB\_MODE_r + RES_r) \quad (4)$$

Where D represents distortion between the raw macroblock and the reconstructed macroblock under the target coding mode, λ represents a trade-off coefficient between distortion and the coded bits of the macroblock, MB_MODEr represents the bits needed for coding the macroblock coding mode, and RESr represents the bits needed for coding the residues under the target coding mode.

Once the coding mode of one macroblock is decided, then it is coded into bit-stream by entropy coding. The syntax structure on macroblock level of coded bit-stream according to the proposed scheme is showed in Fig.4. Compared with syntax structure on macroblock level of H.264/AVC, the syntax of motion compensation mode (MC mode) is added to the bit-stream, the MC mode is 1 bit long variable (either BMC or TM_BMC) can be coded by entropy coding scheme of either CAVLC or CABAC.

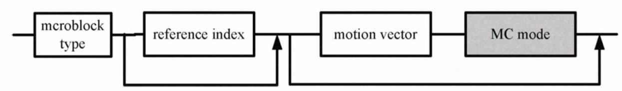

Fig.4. Macroblock syntax of the proposed scheme

At the decoder side, when a macroblock partition with motion compensation mode of TM_BMC mode is decoded, firstly the template matching process is performed to get the template matching predictor, and its associated motion vector is exploited as motion vector predictor to reconstruct the motion vector of block motion compensation, then the block motion compensation predictor is obtained by the reconstructed motion vector. The final motion compensation predictor is calculated by formula (1) to be used to reconstruct the target macroblock partition.

3. EXPERIMENT RESULTS

To verify the efficiency of the proposed scheme, a set of experiments is made on the platform of VCEG KTA (JM11.0KTA2.5) [1]. We use the anchor based on baseline profile for comparison and set QPs of I Slices as 22, 27, 32, and 37, QPs of P Slices as 23, 28, 33 and 38 respectively. The picture prediction structure is set as IPPP…, motion compensation block size is set as 16x16, the reference frame number is one, and the motion estimation is set up to 1/4 pixel accuracy. We compare the RD performance between the anchor of H.264/AVC and the proposed locally switched BMC and TM_BMC mode over six typical sequences: 3 CIF sequences of Foreman, Mobile and Tempete, and 3 720P sequences of City, Crew and Shuttle_start. The bit-rate savings by the proposed scheme over these sequences are showed in Fig.5. From which it shows that average 7.07% bit-rate is saved at the same quality compared with anchor by the proposed method. The largest bit-rate saving is on sequence of City, 11.26% bit-rate is saved. An example of the RD curve comparison between anchor and proposal over

City is showed in Fig.6. There's hardly performance improvement at low bit-rate operation point, since the performance improvement on PSNR is offset by bit-rate increase of coding motion compensation mode by the proposed method at low bit-rate.

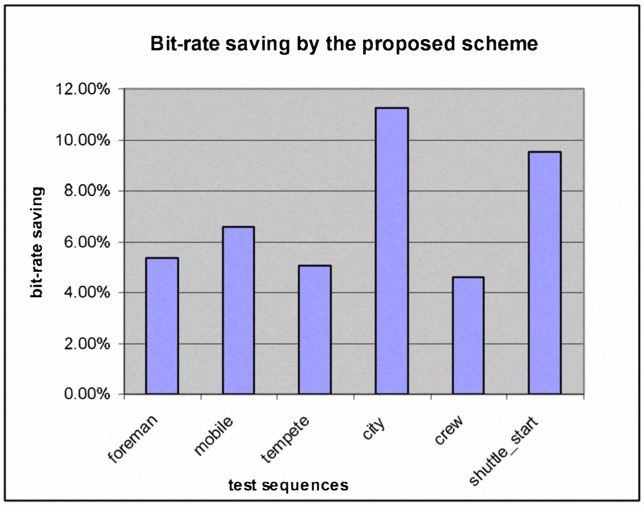

Fig.5. Performance gain by the proposed scheme

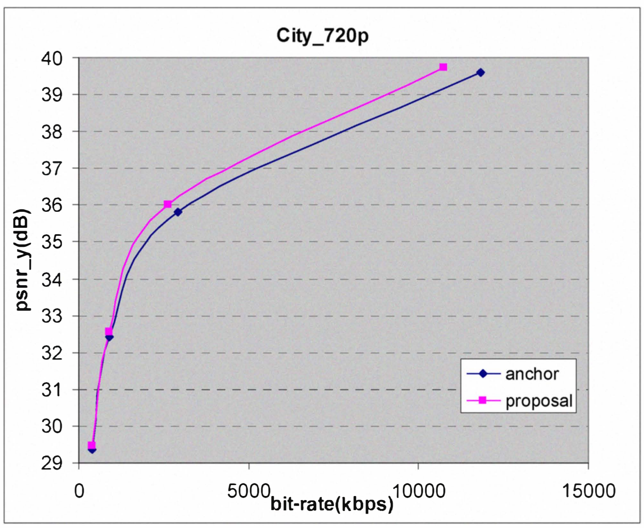

Fig.6. RD curve comparison between H.264/AVC and the proposed scheme

4. CONCLUSIONS AND FUTURE WORKS

A locally adaptive motion compensation mode between BMC and TM_BMC is proposed by this paper. When TM_BMC mode is selected as the motion compensation mode of target macroblock partition, the inter-prediction predictor of the target macroblock partition is calculated by combining template matching predictor and block motion compensation predictor, and the motion vector derived from template matching is set as the motion vector predictor of the motion vector of block motion compensation. The proposed method is integrated into H.264 KTA platform, the experiment results testified the effective of the proposed method, up to 11.26% bit-rate can be saved compared with H.264/AVC. The experiments on multiple reference frames and variable block size motion compensation (and compare with other schemes) will be performed in the next step.

5. REFERENCES

[1]Http://www.tnt.uni-hannover.de/~vatis/kta/.

[2]Y. Vatis, B. Edler, D. Thanh Nguyen, and J. Ostermann, "Two-dimensional non-separable Adaptive Wiener Interpolation Filter for H.264/AVC," *ITU-T SGI 6/Q.6 Doc.VCEG-Z17,* Busan, South Korea, April 2005.

[3]J. Ostermann, M. Narroschke, "Motion compensated prediction with 1/8-pel displacement vector resolution," *ITU-T SG16/Q.6 Doc. VCEG-AD09,* Hangzhou, China, October 2006.

[4]J. Jung, G. Laroche, "Competition-Based Scheme for Motion Vector Selection and Coding," *ITU-T SG16/Q.6 Doc. VCEG-AC06*, Klagenfurt, Austria, July 2006.

[5]K. Sugimoto et. al., "Inter Frame Coding with Template Matching Spatio-Temporal Prediction," *Proc. ICIP 2004,* Singapore, October 24-27, 2004.

[6]S. Kamp,M. Evertz, and M. Wien, " Decoder Side Motion Vector Derivation with Multiple Reference Pictures," *ITU-T SG16/Q.6 Doc.* VCEG-AH15, Antalya, Turkey, 12-13 January, 2008.

[7]M. Flierl et. al, "Rate-Constrained Multihypothesis Prediction for Motion-Compensated Video Compression," *IEEE Trans. CSVT, vol. 12, No. 11,* November 2002.

[8]T.K. Tan et. al, "Intra Prediction by Averaged Template Matching Predictors," *Proc. CCNC 2007,* Las Vegas, NV, USA, January 11-13 2007.

[9]Y. Suzuki, C.S. Boon, T.K. Tan, "Inter Frame Coding with Template Matching Averaging," *Proc. ICIP 2007,* San Antonio, TX, USA, Sept. 2007.

[10]Gary J. Sullivan and Thomas Wiegand, " Rate-distortion optimization for video compression," *IEEE Signal Processing Magazine*, pp. 74－90, Nov. 1998.

[11]H.Watanabe and S.Singhal,"Windowed motion compensation," in Proc. SPIE Conf. visual Commun, Image Proc., vol. 1605, pt, 2, pp.582-589, Nov.1991.

[12]M. T. Orchard and G. J. Sullivan, "Overlapped block motion compensation: An estimation-theoretic approach," IEEE Trans. Image Processing, vol. 3, pp. 693–699, Sept. 1994.