|  |  |
| --- | --- |
| **Joint Collaborative Team on Video Coding (JCT-VC)**  **of ITU-T SG16 WP3 and ISO/IEC JTC1/SC29/WG11**  2nd Meeting: Geneva, CH, 21-28 July, 2010 | Document: JCTVC-B072 |

|  |  |  |  |
| --- | --- | --- | --- |
| *Title:* | **MB Mode with Joint Application of Template and Block Motion**  **Compensations** | | |
| *Status:* | Input Document to JCT-VC | | |
| *Purpose:* | Proposal | | |
| *Author(s) or Contact(s):* | Yi-Wen Chen, Chung-Hao Wu, Chung-Lin Lee, Tse-Wei Wang and Wen-Hsiao Peng 1001 Ta-Hsueh Rd., HsinChu 30010, Taiwan | Tel: Email: | +886-955-028-902 pawn@mail.si2lab.org |
| *Source:* | National Chiao Tung University (NCTU) /ITRI International | | |

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

# Abstract

This contribution introduces a MB mode, termed the P+ mode, to improve the prediction efficiency of P-MBs by a joint application of template and block motion compensations. The motion vector (MV) found by minimizing template matching error is viewed as an additional free MV, which can contribute to estimating pixel intensities in a P-MB. As in POBMC [1], the predictors derived from the template and block MVs are linearly combined based on a distance weighting criterion. In particular, given that the template MV tends to minimize the prediction error in the upper left quarter of a MB, the block MV search criterion is changed so that the resulting MV can contribute more to minimizing the error in the remaining part. When applied to 16x16 P-MBs, the proposed scheme is observed to have an average BD-Rate saving of 2.42% and a corresponding BD-PSNR gain of 0.08dB. It achieves the highest rate saving (4.21%) in S03 (Kimono, Class B) and the lowest (0.4%) in S10 (PartyScene, Class C).

# Introduction

POBMC was introduced in JCTVC-A123 [1] to extend the notion of OBMC in H.263 to accommodate variable block-size motion partitions in AVC. It solves for the optimal weights associated with different MVs as functions of the distances between the predicted pixel and its nearby block centers where these MVs are located. This far-reaching generalization provides a generic reconstruction framework, allowing MVs associated with multiple motion partitions of arbitrary shape to be optimally combined for motion compensation. Some examples are given in Fig. 1, where the block MVs are approximated as the motion of the block centroids.

|  |  |  |  |
| --- | --- | --- | --- |
|  |  |  |  |
| 1. Quad-tree | 1. TMP-like | 1. Geometry | 1. Asymmetric |

Figure 1. POBMC combined with various motion partitions

Among the various applications of POBMC in Fig. 1, the combination of template matching prediction (TMP) and block motion-compensated prediction is of particular interest. Similar to bi-prediction, it carries out (pixel-adaptive) weighted prediction for the target block based on two MVs: one obtained from template matching and the other from block motion estimation. Because the TMP MV is computed on the decoder side, this bi-prediction scheme need only transmit one block MV, making it particularly attractive. Unlike the conventional use of TMP, where it is intended for decoder side motion inference, the present invention aims to utilize TMP as a tool for improving prediction efficiency. Since we have less control over the selection of the TMP MV, a question that naturally arises is, how can we optimize the search of the block MV so that both MVs can be used to their best advantage for minimizing prediction error?

To gain some insights into the answer, Fig. 2 plots the mean squared prediction error surface with the single use of TMP MV for motion compensation of the target block. It is seen that this MV tends to minimize the prediction error in the upper left quarter, a result that is intuitively agreeable since it approximates the true motion associated with the template centroid. This observation implies that the block MV should be managed to contribute more to minimizing the error in the remaining part, especially in the bottom right region. We thus propose in this contribution a new search criterion for the block MV to achieve this objective.

In the interest of time, this new prediction mode is currently applied to 16x16 P-MBs only as an enhancement and is thus termed the P+ mode. Its notion however can readily be extended to any MB or sub-MB partition with single- or multi-hypothesis prediction. Moreover, our framework allows the TMP MV to be replaced with a MV inferred by other means.

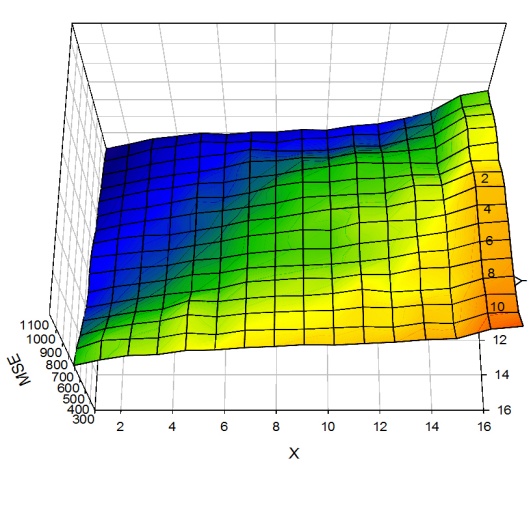


Figure 2. Mean squared prediction error surface of TMP using the sequence “Football”.

# P+ Prediction Mode

## Theoretical Analysis

The usual search criterion of a block MV is to minimize squared prediction error over the pixels in the prediction block. However, with the new prediction mode, we wish to arrive at a block MV that when applied together with the TMP MV for POBMC, results in minimum prediction error. To proceed analytically, we approximate the TMP MV as the motion of the template centroid and assume the block MV to be found corresponds to the true motion of some unknown pixel . The problem to find the block MV can now be cast as the search for that unknown pixel location that minimizes the sum of mean squared prediction error (SMSE) over the entire block :

(1)

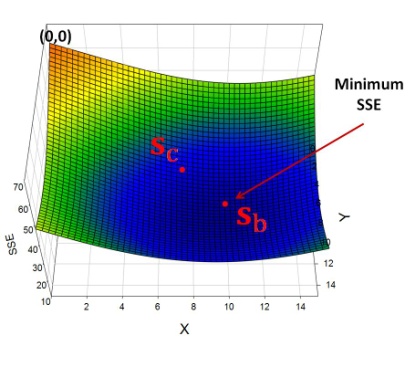
where represents the true motion of pixel **,** and the weighting factor is computed by POBMC with denoting the L₂ distance (measured in units of pixels) between and .

Proceeding in much the same way as in [1], we obtain

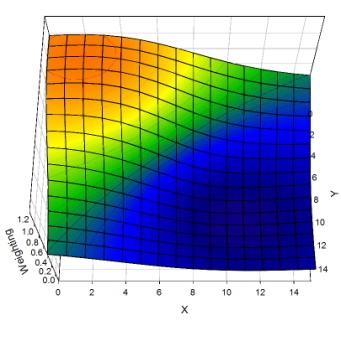
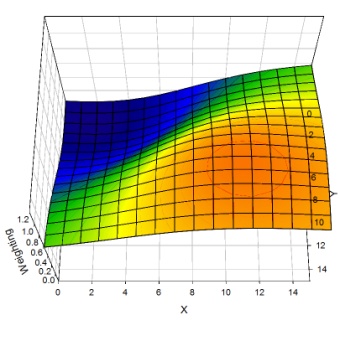
(2)

As an illustration of the relationship between the MSE and the location of pixel , consider the example in Fig. 3(a), where the target block is a 16x16 P-MB and the template size is 4. Fig. 3(b) plots the SMSE as a function of the location of pixel according to (2). From the figure, the minimum SMSE is achieved when pixel is located at around (10, 10), which is obviously not the block center . As expected, the optimal block MV, approximated by the true motion of pixel , has a higher correlation with the true motion of pixels in the bottom right quarter. This result also agrees with our previous conjecture.

Given the optimal location of pixel , Fig. 3(c) and Fig. 3(d) further show the window functions , of the template and block MVs, respectively. Interestingly, their distributions resemble a special case of geometry partition, where the two MVs are located on the diagonal.



1. (b)

(c) (d)

Figure 3. 16x16 P+ mode: (a) an example with template centroid located at (1.94, 1.94), (b) SMSE surface over different pixel , the x and y axis represent the horizontal and vertical indices of pixel . Given optimum sampling point **,** weighting windows of template motion and block motion are depicted in (c) and (d), respectively.

## Motion Search Criterion

The preceding result can readily be applied to design a block MV search criterion specifically optimized for the new prediction mode. We circumvent the difficulties of having to find the true motion of pixel by noting the one-to-one relationship between and the location of . If we substitute for in (1), the SMSE will be minimum only when . As a result, a block MV is likely to be if it minimizes the sum of squared predication error over block :

, (3)

We thus propose using (3) as the new block MV search criterion.

# Experimental Results and Discussions

The P+ mode has been integrated into KTA2.4r1 reference software [4] without enabling RDOQ and KTA tools. A flag is signaled in the bitstream for each macroblock to adaptively switch on/off P+ mode. The results are obtained based on the encoding of 100 frames from all CfP sequences [6] except for class D. Some other encoder settings are described in Table 1. The anchor is KTA2.4r1 reference software with RDOQ and KTA tools turning off.

Compared to the anchor, Table 2 below summarizes the results in terms of BD-Rate savings and BD-PSNR gain using the Bjontegaard tool [5]. The minimum BD-Rate saving is 0.4% (PartyScene, Class C), the maximum BD-Rate saving is 4.21% (Kimono, Class B). Average 2.42% BD-Rate reduction and 0.08 BD-PSNR gain are achieved over all testing sequences. Because skip mode is the dominant mode, especially in low-bitrate case (QP=37), it limits the coding gain of P+ mode. Figure 4 (a) and (b) show the distributions of coded modes including P+ mode in different QPs, we can see that as QP increases, the increasing skip mode suppresses the enabling of P+ mode and thus reduces the coding gains.

1. (b)

Figure 4. Mode distribution of coding S03 at (a) QP=22, (b) QP=37.

**Table 1. Encoding Configurations**

|  |  |
| --- | --- |
| Reference Frame | 4 |
| CABAC | On |
| 8x8 Transform | On |
| Deblocking | On |
| RDO | On |
| MV Search Range | 128 |
| Motion Search | EPZS |
| Sub-pel MC | On |
| Quantization Parameter (I/P) | 22/23, 27/28, 32/33, 37/38 |
| Adaptive Rounding | off |
| Prediction Structure | IPPP… |
| Block Partitioning | 16x16 to 4x4 |

**Table 2. Comparison of coding efficiency for CfP sequences**

|  |  |  |  |
| --- | --- | --- | --- |
|  | Sequences | BD-PSNR  Gain | BD-Rate  Saving |
| Class A | S01 | 0.082 | -2.24% |
|  | S02 | 0.070 | -1.51% |
|  | Average | 0.076 | -1.88% |
| Class B | S03 | 0.158 | -4.21% |
|  | S04 | 0.033 | -0.91% |
|  | S05 | 0.047 | -1.89% |
|  | S06 | 0.083 | -3.12% |
|  | S07 | 0.094 | -3.49% |
|  | Average | 0.083 | -2.72% |
| Class C | S08 | 0.102 | -2.64% |
|  | S09 | 0.097 | -2.20% |
|  | S10 | 0.018 | -0.40% |
|  | S11 | 0.046 | -1.04% |
|  | Average | 0.066 | -1.57% |
| Class E | S16 | 0.140 | -3.76% |
|  | S17 | 0.074 | -1.87% |
|  | S18 | 0.059 | -2.01% |
|  | Average | 0.091 | -2.55% |
| Overall Average | | 0.079 | -2.42% |

# Conclusions

In this contribution, a new prediction mode with joint application of TMP and BMC is proposed. In the interest of time, the proposed scheme was only integrated into 16x16 P-MBs, but even so, we already observed an average BD-Rate saving of 2.42% and a BD-PSNR gain of 0.08dB. In some test sequences, the gain can be as high as 3-4%. This work is still in its early stage. We believe there is still plenty of room for further improvement.

# References

1. Y. W. Chen, T. W. Wang, C. H. Chan, C. L. Lee, C. H. Wu, Y.­ C. Tseng, W. H. Peng, C. J. Tsai, and H. M. Hang, “Description of video coding technology proposal by NCTU,” JCTVC-A123, Dresden, Germany, April 2010.
2. Y. W. Chen, T. W. Wang, Y. C. Tseng, W. H. Peng, and S. Y. Lee, “A Parametric Window Design for OBMC with Variable Block Size Motion Estimates,” *IEEE Workshop on Multimedia Signal Processing*, 2009.
3. T. W. Wang, Y. W. Chen, and W. H. Peng, "Analysis of Template Matching Prediction and Its Application to Parametric Overlapped Block Motion Compensation," IEEE Int’l Symposium on Circuits and Systems, 2010.
4. KTA Reference Software, <http://iphome.hhi.de/suehring/tml/kta/>.
5. G. Bjøntegaard, “Improvements of the BD-PSNR Model,” ITU-T SG16 Q.6 Document, VCEG-AI11, Berlin, July 2008.
6. ITU-T Q6/16 and ISO/IEC JTC1/SC29/WG11, “Joint Call for Proposals on Video Compression Technology,” MPEG document w11113, 91st MPEG meeting, Kyoto, Japan, Jan. 2010

# Patent rights declaration(s)

(NOTE – Activities in the JCT-VC and contributions to the JCT-VC are subject to the common patent policy for ITU-T/ITU-R/ISO/IEC. A statement of that policy can be found at  
<http://www.itu.int/ITU-T/dbase/patent/patent-policy.html>, with further information available at <http://www.itu.int/ITU-T/ipr/index.html> and in [the ISO/IEC Directives](http://isotc.iso.org/livelink/livelink?func=ll&objId=4230455&objAction=browse&sort=subtype). The form to be used for the formal reporting of patent rights to ITU-T/ITU-R/ISO/IEC can be found at <http://www.itu.int/ITU-T/ipr/index.html>. Contributions to the JCT-VC proposing normative technical content shall contain a non-binding informal notice of whether the submitter may have patent rights that would be necessary for implementation of the resulting standard. The provided informal notice shall indicate the category of anticipated licensing terms according to the ITU-T/ITU-R/ISO/IEC patent statement and licensing declaration form. This obligation to provide an informal notice is supplemental to, and does not replace, any existing obligations of parties with technology included in a final or draft standard to submit formal IPR declarations to ITU-T/ITU-R/ISO/IEC. An example of an informal IPR notification statement for a contribution is provided below.)

**NCTU may have IPR relating to the technology described in this contribution and, conditioned on reciprocity, is prepared to grant licenses under reasonable and non-discriminatory terms as necessary for implementation of the resulting ITU-T Recommendation | ISO/IEC International Standard (per box 2 of the ITU-T/ITU-R/ISO/IEC patent statement and licensing declaration form).**