Low-cost Multi-hypothesis Motion Compensation for Video Coding

Lei Chen[a], Shengfu Dong[a], Ronggang Wang[a], Zhenyu Wang[a], Siwei Ma[b], Wenmin Wang [a], Wen Gao[b]
[a] Peking University, Shenzhen Graduate School, Shenzhen, China; [b]Institute of Digital Media, Peking University, Beijing, China

ABSTRACT

In conventional motion compensation, prediction block is related only with one motion vector for P frame. Multi-hypothesis motion compensation (MHMC) is proposed to improve the prediction performance of conventional motion compensation. However, multiple motion vectors have to be searched and coded for MHMC. In this paper, we propose a new low-cost multi-hypothesis motion compensation (LMHMC) scheme. In LMHMC, a block can be predicted from multiple-hypothesis with only one motion vector to be searched and coded into bit-stream, other motion vectors are predicted from motion vectors of neighboring blocks, and so both the encoding complexity and bit-rate of MHMC can be saved by our proposed LMHMC. By adding LMHMC as an additional mode in MPEG internet video coding (IVC) platform, the B-D rate saving is up to 10%, and the average B-D rate saving is close to 5% in Low Delay configure. We also compare the performance between MHMC and LMHMC in IVC platform, the performance of MHMC is improved about 2% on average by LMHMC.

Keywords: motion compensation, multi-hypothesis, generalized B frame, MPEG IVC

1. INTRODUCTION

In conventional motion compensation for P frame, prediction block is related only with one motion vector. However, this kind of motion compensation does not make good use of spatial and temporal correlation between frames, thus causing the prediction block to be not accurate. In bidirectional motion compensation for B frame, prediction block is derived by two motion vectors, one forward and one backward, thus, making the prediction block more accurate. Generalized B frame [1] uses the similar idea to improve the prediction performance of P frame by multi-hypothesis motion compensation (MHMC) [2-3]. However, in MHMC, at least two motion vectors have to be searched and coded in bit stream, thus increases both the encoding complexity and bit-rate.

In this paper, we introduce a new motion compensation technology: Low-cost Multi-hypothesis Motion Compensation (LMHMC). This kind of motion compensation is different from conventional motion compensation as prediction block is not only related with the motion vector obtained by motion estimation, but also related with motion vectors of neighboring blocks so as to provide higher accuracy for prediction block. Besides, the encoding complexity and bit-rate do not increase since only one motion vector needs to be searched and transmitted in bit-stream.

The proposed method is implemented in the MPEG IVC reference software ITM4.0 [4] as an additional mode for P frame. The mode is selected by the RDO process. The modified ITM4.0 with our proposed LMHMC method saves up to 10% B-D rate and the average B-D rate saving is about 5%.

The rest of this paper is organized as follows. Section 2 introduces traditional motion compensation. Section 3 presents our proposed low-cost multi-hypothesis motion compensation in detail. Section 4 gives the experimental results. Finally, we conclude this paper in Section 5.

2. CONVENTIONAL MOTION COMPENSATION

In traditional motion compensation for P frame, prediction block is only relative to one motion vector searched through motion estimation process and only one motion vector is transmitted in bit stream. Rate-constrained motion estimation for a block S is performed by minimizing the Lagrangian cost function:

$$J(\lambda_{SAD}) = D_{SAD}(S) + \lambda_{SAD} \times R(MV - MV\,\mathrm{Pr}ed) \tag{1}$$

Visual Information Processing and Communication V, edited by Amir Said, Onur G. Guleryuz, Robert L. Stevenson, Proc. of SPIE-IS&T Electronic Imaging, SPIE Vol. 9029, 902903
 · doi: 10.1117/12.2039512

Here MV and MVPred denote the estimated motion vector and its corresponding prediction signal, respectively. MV is calculated using Eq.(1). R(MV - MVPred) denotes the number of bits to code the difference between MV and MVPred. The distortion term of Eq. (2) is further defined as:

$$D_{SAD}(S) = \sum_{(x,y)} | S(x,y) - (S_{ref}(x + MVX, y + MVY)) | \quad (2)$$

Here (x, y) is the coordinates of pixels in the current block. Sref represents the reference picture. MVX and MVY are the x and y component of MV, respectively.

After the motion estimation process, prediction block can be derived through the searched motion vector in corresponding reference picture.

In bidirectional motion compensation for B frame [5-6] or generalized B frame, prediction block is more accurate since it is derived by two motion vectors. The motion estimation process needs to be carried out twice so as to get two prediction blocks which increases the encoding complexity. The motion estimation process is the same as the motion estimation process for P frame. After motion estimation, each motion vector points to a separate prediction block, and the final prediction block P is calculated through Eq. (3) and shown in figure 1. As in bidirectional motion compensation or generalized B frame, two motion vectors must be transmitted in bit stream, thus induces the increasing of bit-rate.

$$P(x,y) = (P1(x,y) + P2(x,y))/2 \quad (3)$$

Here (x,y) is the coordinates of pixels in the current block

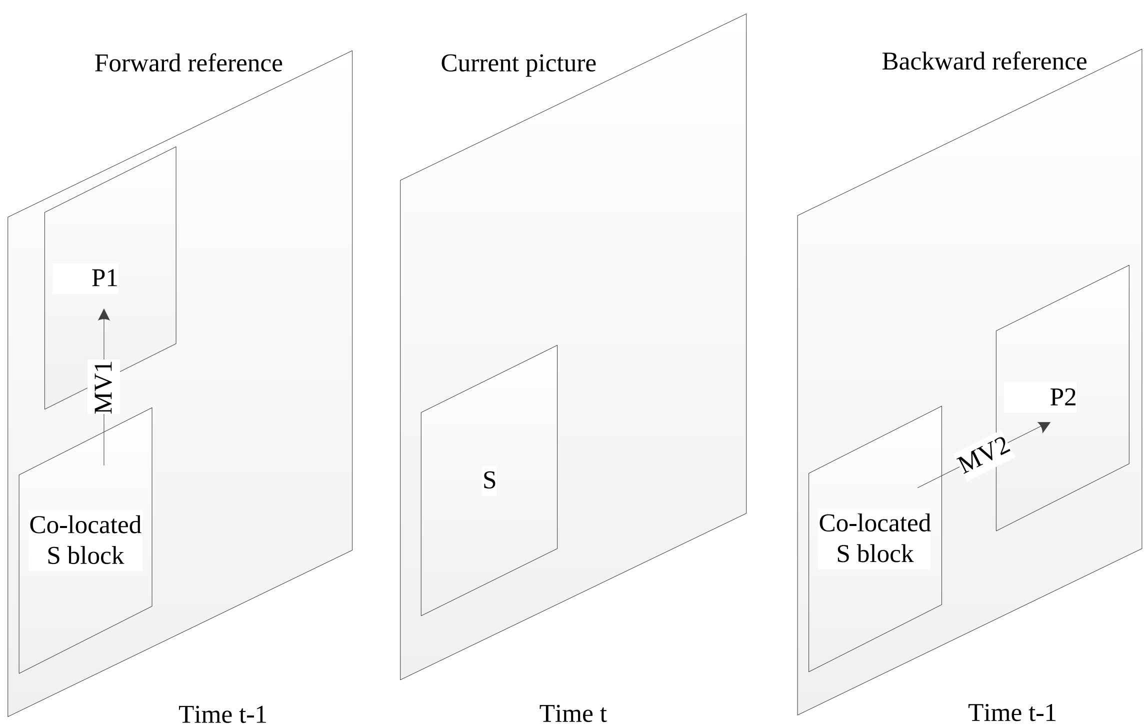

Figure 1. Motion compensation for B frame

3. LOW-COST MULTI-HYPOTHESIS MOTION COMPENSATION (LMHMC) SCHEME

The block diagram of our proposed LMHMC coding framework is depicted in Figure 2. There are two paths for getting the prediction block. Path 1 represents the conventional motion compensation that already exists in IVC platform. Path 2 is our proposed LMHMC. For an inter block, both conventional motion compensation and LMHMC are conducted and the best mode is selected by the RDO process similar to H.264 [5] and HEVC [7].

For our proposed LMHMC, prediction block is not only related with the motion vector searched through motion estimation process, but also related with motion vectors of neighboring blocks. The prediction block with LMHMC mode is derived through two motion vectors: MV1 and MV2. MV1 is derived by neighboring blocks. MV2 is derived through joint motion estimation which refers MV1. The two motion vectors point to prediction block P1 and P2. The final prediction block is the averaging of P1 and P2. In entropy coding process, only difference between MV2 and predictor of MV2 is coded and transmitted in bit-stream, MV1 can be deduced in both encoder and decoder and there is no need to transmit it.

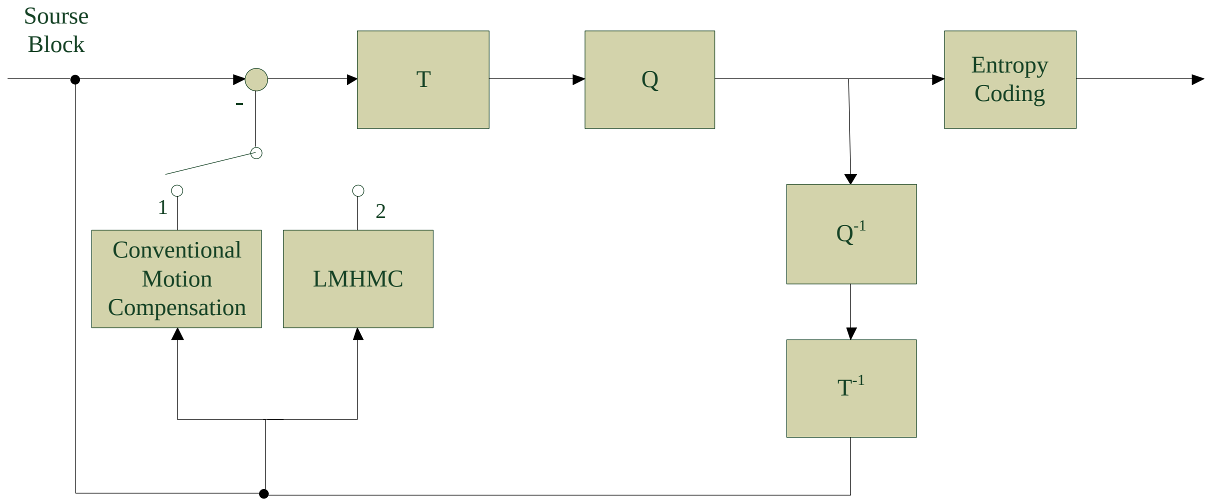

Figure 2. The proposed LMHMC coding framework

3.1 Derivation process of MV1

MV1 is derived by motion vector prediction [8], such as median prediction used in H.264/AVC. In MPEG IVC, four motion vectors of neighboring blocks (shown in figure 3) are used for motion vector prediction. In this paper, the derivation process of MV1 follows MPEG IVC motion vector prediction technology proposed in [9].

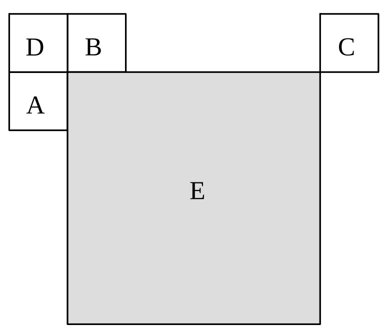

Figure 3. Motion vector prediction from neighboring blocks

3.2 Derivation process of MV2

MV2 is derived through joint motion estimation which is the same as the derivation process of backward motion vector in bi-directional motion compensation. The joint motion estimation can be easily implemented with the same motion estimation process as the conventional motion estimation does. Derivation process of MV2 for a coded block S is performed by minimizing the Lagrangian cost function:

$$J(\lambda_{SAD}, MV1) = D_{SAD}(S, MV1) + \lambda_{SAD} \times R(MV2 - MV1) \quad (4)$$

Here MV1 denotes the motion vector derived by motion vector prediction. MV2 represents the searched motion vector and is calculated using Eq.(4). R(MV2-MV1) denotes the number of bits to code the difference between MV2 and MV1. The distortion term is further defined as Eq.(5).

$$D_{SAD}(S, MV1) = \sum_{(x,y)} | S(x,y) - (S_{\text{ref}}(x + MV2\text{x}, y + MV2\text{y}) + S_{ref}(x + MV1\text{x}, y + MV1\text{y})) >> 1 | \quad (5)$$

Here, (x, y) is the coordinates of pixels in current block. Sref represents the reference picture. MV1x and MV1y are the x and y component of MV1, and MV2x and MV2y are the x and y component of MV2.

The difference between MV2 and MV1 is transmitted in bit-stream.

3.3 Low-cost multi-hypothesis motion estimation

After MV1 and MV2 are derived, corresponding prediction block P1 and P2 can be obtained in reference picture as shown in Figure 4. The final prediction block P of current coded block is calculated through Eq.(6).

$$P = (P1 + P2) / 2 \quad (6)$$

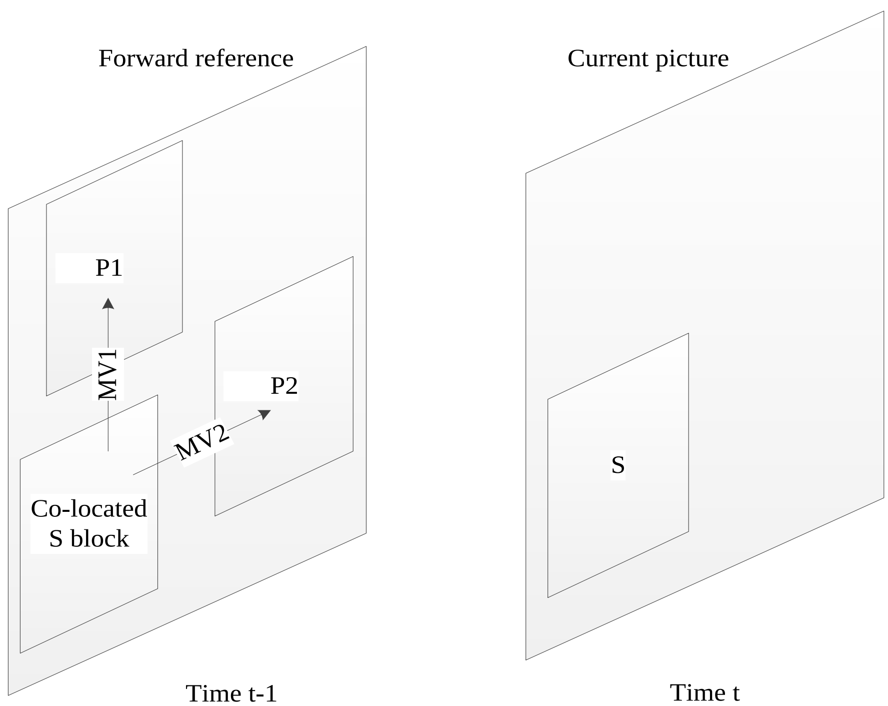

Figure 4. Prediction blocks of motion compensation

4. EXPERIMENTAL RESULT

We integrate the proposed LMHMC method into MPEG IVC reference software ITM4.0 [4] as an additional mode for inter frame. Besides, we also implemented MHMC in ITM4.0 as an additional mode for inter frame for comparison with LMHMC. For MHMC method, the prediction block is derived by two motion vectors as generalized B frame does, and the motion estimation process conducts twice. The sequences in common test condition of MPEG IVC are tested.

ITM4.0 is used as anchor for all experiments. The tested configure of Low Delay is used and only one reference picture is used.

Table 1 shows the performance comparison of anchor, MHMC, and LMHMC. As shown in table 1, LMHMC saves B-D rate in all sequences and the bit-rate saving is up to 10% for the WQVGA sequence BQSquare compared with anchor. The average BD rate saving is about 5%. For MHMC, the average BD rate saving is 2.6% and its performance is worse than LMHMC for all sequences. Figure 5 shows rate-distortion curve of RaceHorses sequence in class B.

Table 1. Performance comparison among MPEG IVC anchor, MHMC and LMHMC.

Resolution	Sequence	MHMC vs. anchor	LMHMC vs. Anchor
class 0 4K	Traffic	-1.3%	-2.6%
	PeopleOnStreet	-1.6%	-3.8%
class A 1080P	Kimono	2.3%	-1.9%
	ParkScene	-1.1%	-4.0%
class B WVGA	BasketballDrill	-1.7%	-2.6%
	BQMall	-3.2%	-6.0%
	PartyScene	-2.2%	-3.6%
	RaceHorses	-5.5%	-8.0%
class C WQVGA	BasketballPass	-2.7%	-5.1%
	BQSquare	-8.7%	-10.0%
	BlowingBubbles	-2.7%	-4.2%
	RaceHorses	-4.2%	-6.9%
class D 720P	FourPeople	-0.9%	-3.6%
	Johnny	-3.7%	-4.9%
	KristenAndSara	-1.1%	-3.0%
total		-2.6%	-4.7%

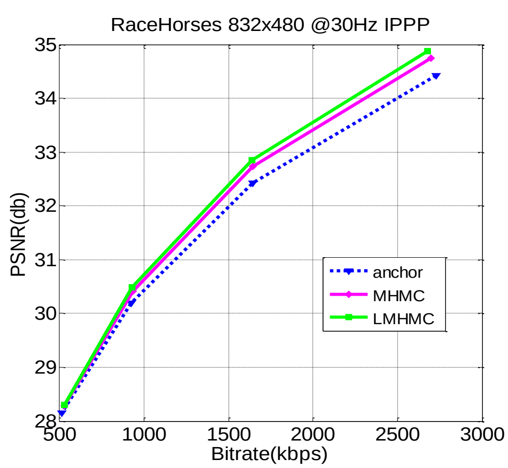

Figure 5. Rate distortion curves with different motion compensation schemes for the WVGA sequence RaceHorses

5. CONCLUSION

In this paper, we propose a new low-cost multi-hypothesis motion compensation (LMHMC) scheme. It's different from conventional motion compensation as prediction block is not only related with motion vector searched through motion estimation process but also related with motion vectors of neighboring blocks to provide more accurate prediction block. In LMHMC, two motion vectors are derived, the first one is derived through motion vector prediction process and the second is searched through joint motion estimation process. Each motion vector points to a prediction block and the final prediction block of the current block is the averaging of these two prediction blocks. For entropy coding process, only the difference between the second motion vector and its predictor is coded. By adding LMHMC as an additional mode in MPEG internet video coding (IVC) platform, the B-D rate saving is up to 10% and the average saving is about 5%.

6. ACKNOWLEDGE

THIS WORK WAS PARTLY SUPPORTED BY THE GRANT OF NATIONAL SCIENCE FOUNDATION OF CHINA 61370115, AND SHENZHEN BASIC RESEARCH PROGRAM OF JC201104210117A, JC201105170732A, AND JCYJ2012061450301623.REFERENCES

REFERENCES

[1] Flierl, M. and Girod, B., "Generalized B pictures and the draft JVT/H.264 video compression standard," IEEE Trans. Circuits Syst. Video Technol., vol. 13, pp.587 -597 (2003).

[2] Sullivan, G. J., "Multi-hypothesis motion compensation for low bit-rate video coding," Proc. IEEE Int. Conf. Acoustics, Speech, and Signal Processing (ICASSP), pp.437 -440 (1993).

[3] Flierl, M., Wiegand, T. and Girod, B., "Rate-constrained multi-hypothesis motion-compensated prediction for video coding," Proc. IEEE Int. Conf. Image Processing, vol. 3, pp.150 -153 (2000).

[4] Wang, R. G., Zhang, X. G., Ma, S. W., Chen, J. W., "Internet Video Coding Test Model (ITM) Version 4.0," ISO/IEC JTC1/SC29/WG11 N13353, Geneva, Switzerland, January (2013).

[5] Wiegand, T., Sullivan, G. J., Bjntegaard, G. and Luthra, A., "Overview of the H.264/AVC video coding standard," IEEE Trans. Circuits Syst. Video Technol., vol. 13, no. 7, pp.560 -576 (2003).

[6] Chen, Z., Xu, J., He, Y. and Zheng, J., "Fast integer-pel and fractional-pel motion estimation for H.264/AVC," Journal of Visual Communication and Image Representation, vol. 17, pp.264 -290 (2006).

[7] Sullivan, G. J., Ohm, J., Han, W. –J. and Wiegand, T., "Overview of the High Efficiency Video Coding (HEVC) standard," IEEE Trans. Circuits Syst. Video Technol., vol. 22, pp.1648 -1667 (2012).

[8] Ismaeil, I., Docef, A., Kossentini, F. and Ward, R., "Efficient motion estimation using spatial and temporal motion vector prediction," Proc. Int. Conf on Image Processing(ICIP), vol. 1, pp. 70-74 (1999).

[9] Lv, H., Wang, R. G., Chen, L., Ma, S. W., Chen, J. W. and Gao, W., "A direction and distance based motion vector prediction method for internet video coding," ISO/IEC JTC1/SC29/WG11 M27965, Geneva, Switzerland, January (2013)