# An Improved Motion Vector Prediction Scheme for Video Coding

Dliu[1], Dbzhao[1], Qwang[1], Wgao[2]

[1]Department of Computer Science, Harbin Institute of Technology, Harbin 150001, China
[2]Institute of Computer Technology, Chinese Academy of Science, Beijing 100080, China

{dliu, dbzhao, qwang, wgao}@jdl.ac.cn

**Abstract.** The motion vector prediction (MVP) is an important part of video coding. In the original median predictor, if the neighbor blocks of current block are intra-mode coded, their motion vectors (MVs) will be set to zeros for MVP of current block. This is not very precise for sequences with strong motion. This paper propose an improved motion vector prediction (MVP) scheme for H.264. In the proposed scheme, when there are intra-mode macroblocks beside current block, more MV of the neighbor inter-mode block is utilized instead of zero MVs of intra-mode macroblocks for MVP of current block. The experimental results show that the improved scheme achieves better coding efficiency than the original median predictor. Meanwhile the point obtained by the proposed MVP scheme is closer to the global minimum point, the following fast motion estimation (FME) computation complexity is reduced.

**Keywords:** motion vector prediction, video coding

## 1 Introduction

The MVP plays an important role in the video coding standard. A good MVP can give a good starting point for following fast motion estimation (FME) in inter-prediction. FME would like to obtain the best matching block for current block from the reference frames by searching a few points. The majority of FME algorithms often fall into local optimal point which may not be really optimal MV for current block. If the MVP predicts a point closer to the global optimal point, a high quality reference block for current block is obtained, the coding performance of the picture is improved, and the computation complexity of FME is reduced. The MVPs of neighbor blocks can influence current block's MVP. MVP of current block use MVs of neighbor blocks as reference MV, the MVP accuracy of current block is influenced by the accuracy of neighbor blocks' MV. A good MVP of neighbor block can obtain a good estimated MV for itself and can give a good reference MV for current block’s MVP. Then, the coding performance of whole picture is improved if a few influencing blocks obtain more accurate MVs.

At present, the median predictor is widely used in many important coding standards, such as H.263/H.264 [1][2]. It obtains the median values from the MVs of neighbor blocks. As can be seen form Fig. 1, MV of block E is predicted using MVs of block A B C when all the blocks are available. The scheme is simple, but when meeting inhomogeneous motion scene, its efficiency is not very good. Besides the median prediction, the weighted mean prediction is employed in [3]. It assigns every neighbor block a weighted factor. [4] adds the temporal prediction in MVP to compensate the insufficiency of spatial prediction. [5] joins a predefined threshold to guarantee the validate MVs to be used for median predictor. In [6], [7] and [8], S. D. Kim and J. B. Ra introduce a flag indicating which neighbor MVs are used to predict the current MV in bitstream. All the proposed MVP methods following the median predictor want to improve MVP accuracy in inhomogeneous motion scene and can achieve a good MVP value. But when there are intra-mode macroblocks beside current block, all the proposed MVP methods look the reference MVs of the intra-mode macroblocks as zero when predicting MV of current block. However some other useful reference MVs of neighbor inter-mode blocks have not been utilized, this is a waste of surrounding inter MV information, and the MVP accuracy of current block will be influenced.

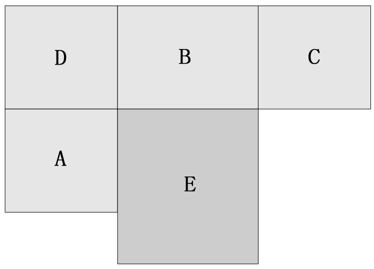


Fig.1. Block E and the neighbor blocks

The influence of intra-mode macroblocks to the MVPs of other macroblocks in the picture is large. Four surrounding macroblocks will be influenced by one intra-mode macroblock. Fig. 2 give an illustration, block I are separately at up-right, up, up-left, left of macroblocks F, G, H, J. When predicting MVs of macroblocks F, G, H, J, MV of block I will be used as a reference MV. If block I is intra-mode macroblock, MV of macroblock I will be looked as zero, MVPs of macroblocks F, G, H, J are all influenced. The more the intra-mode macroblocks in a picture, the more the macroblock’s MVPs will be influenced. Table 1 shows in the sequences with strong motion, the percentage of intra-mode macroblocks is not small. In sequence stefan and foreman, it is separately up to 12% and 9% in all kinds of macroblocks. The percentage of macroblocks, whose MVPs are directly influenced by the intra-mode macroblocks, is separately about 48% and 36%. In this condition, lots of macroblocks' MVPs are not accurate enough, its following FMEs can not fall into global optimal point then can not bring the optimal MV and the best matching block for current macroblock. The MVPs accuracy of some other macroblocks in the picture will be influenced too because of poor MVs accuracy of the macroblocks beside intra-mode macroblocks. The coding performance of whole picture will be greatly influenced by the intra-mode macroblocks.

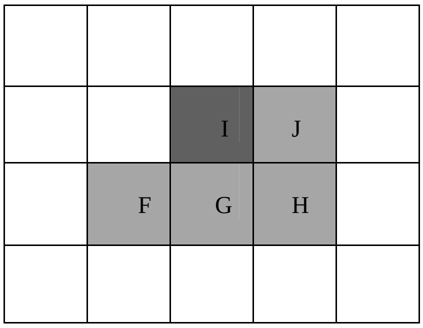


Fig.2. The location of macroblock I (intra-mode macroblock) and the macroblocks influenced by macroblock I

Table 1. Percentage of intra-mode macroblocks and influenced macroblocks(QP=32)

| | Percentage of intra-mode macroblocks | Percentage of influenced macroblocks |
|---|---|---|
| Stefan | 12% | About 48% |
| Foreman | 9% | About 36% |

As mentioned in [9], if one neighbor macroblock of the current block is intra coded, the motion vector (MV) of current block is not easy to predict and greater computation cost is needed. In this case, [9] will select one between two FME methods according to whether there are intra-mode macroblocks beside current block. But [9] has not considered how to improve the MVP accuracy to reduce the computation complexity of FME for current block when there are intra-mode macroblocks in surrounding area.

In this paper, we propose an improved MVP scheme. When there are intra-mode macroblocks beside current block, some abundant MV of neighbor inter-mode block is utilized instead of the zero MVs of intra-mode macroblocks for MVP of current block. More inter-mode MV information is utilized to achieve a better prediction performance and bring less FME computation complexity. The coding efficiency is enhanced, the computation complexity is reduced.

The remainder of the paper is organized as follows. In Section 2, the detailed description of the method is given, when there are intra-mode macroblocks beside current block, the original zero MVs of intra-mode macroblocks can be optionally replaced by MV of block D (see Fig. 1) in different cases for MVP of current block. Section 3 provides the experimental results and performance comparison. Finally, section 4 concludes this paper.

## 2 Improved MVP scheme

To give clearer illustration, some average performance comparisons are showed in two typical sequences “Stefan” and “foreman”, which have strong motion scene. They are both 300 frames CIF sequences with the frame rate of 30f/s. The experiment conditions are showed in Table 4, depicted in section 3. QP is set to 32.

There are four macroblocks beside the intra-mode macroblock as shown in Fig.2, its MVPs are influenced by the MV of intra-mode macroblock. We name this kind of areas as intra-mode macroblock areas. The average performance of the intra-mode macroblock areas and average performance of all macroblocks in the sequence are compared. Table 2 shows that, the average PSNR of intra-mode macroblock areas is smaller than the average PSNR of all macroblocks in sequence.

Table 2. PSNR comparison (QP=32)

| | Average PSNR of intra-mode macroblock areas(dB) | Average PSNR of all macroblocks in sequence(dB) |
|---|---|---|
| Stefan | 32.20 | 32.44 |
| Foreman | 34.61 | 34.82 |

The computation complexity can not be simply evaluated by the count of search points, it is because that the block matching size in H.264 is variable and the partial SAD (sum of absolute difference) computation is used. The computation of a SAD is early terminated when the partially accumulated SAD has exceeded the minimum known SAD. So we use the count of AD (absolute difference) operations as the evaluation of a ME algorithm’s complexity. An AD operation includes the subtracting between one pixel in the current block and corresponding pixel in the candidate block, getting the absolute value of difference, adding the absolute value to the SAD.

Table 3 shows the mean AD counts per macroblock in intra-mode macroblock areas is greater than the mean value of all the macroblocks in the sequence. This means larger computation is needed for macroblocks in the neighbor of intra-mode macroblocks.

Table 3. AD counts comparison (QP=32)

| | Mean AD counts of intra-mode macroblock areas($*10^3$) | Mean AD counts of all macroblocks in sequence($*10^3$) |
|---|---|---|
| Stefan | 124.82 | 94.43 |
| Foreman | 33.54 | 25.47 |

All the blocks have the same coding method, the reason of small PSNR and high computation complexity in macroblocks beside intra-mode macroblocks is that its MVPs are influenced by the MVs of intra-mode macroblocks. Based on the above analysis, the details of the proposed algorithm can be depicted as follows:

In the first case, there is only one intra-mode macroblock beside the block E. When predicting the MV of block E, the MV of the intra-mode macroblock will be looked as zero in the original median predictor. In our proposed scheme, if block D is the intra-mode macroblock, the MV of block D is looked as zero. MVs of blocks A, B, C are used for MVP of block E, the proposed method is the same as the original median predictor. If block A is the intra-mode macroblock (see Fig. 3), the MV of macroblock A will be set to the MV of block D. If macroblock B or macroblock C is the intra-mode block, the proposed scheme is the same as the condition when block A is the intra-mode macroblock (see Fig. 3).

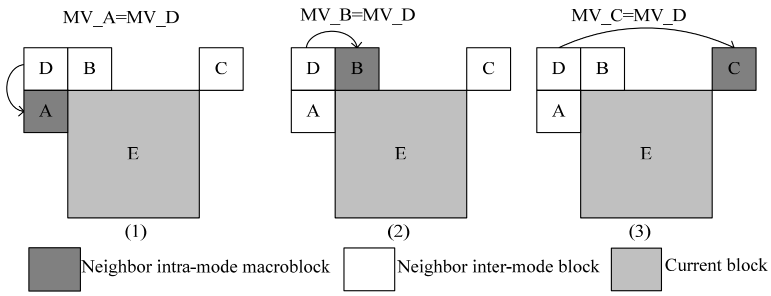


Fig.3. Only one intra-mode macroblock beside block E

In the second case, there are two intra-mode macroblocks beside the block E. When predicting the MV of block E, the MVs of the two intra-mode macroblocks will be looked as zero in the original median predictor. In our proposed scheme, if block A and block D (or block B and block D, or block C and block D) are the intra-mode macroblocks, MVs of block A and block D (or block B and block D, or block C and block D) are still zero, the same as original median predictor. Because there are always one block D in the two intra-mode blocks. MV of block D is looked as zero. No more redundant neighbor MV can be used. When block A and block B are the intra-mode macroblocks(see Fig. 4), block D and block C are inter-mode blocks. In predicting MV of block E, MV of block D can be utilized by setting its MV to the MV of block A or MV of block B. In the two intra-mode macroblocks, only one macroblock’s MV can be set to MV of block D, the other macroblock’s MV is still zero. Which intra-mode macroblock’s MV remained zero have no influence on the MVP value in the median predictor. When block A and block C (or block B and block C) are intra-mode macroblocks, the proposed scheme is the same as the condition when block A and block B are intra-mode macroblocks (see Fig. 4).

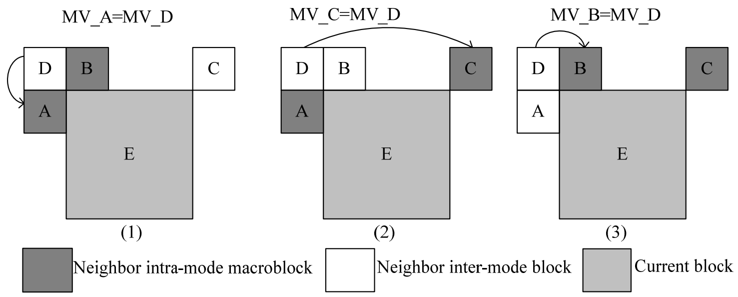


Fig.4. Two intra-mode macroblocks beside block E

In the third case, there are three intra mode macroblocks beside the block E. In the original median predictor, there are at least two zero reference MVs in the surrounding blocks of block E, MV of block E is zero. In the proposed scheme, when block D is located in the three intra-mode macroblocks, the proposed method is the same as the original median predictor. When block D is not located in the three intra-mode macroblocks (see Fig. 5), blocks A, B, C are intra-mode macroblocks. The MVP value of block E is set to the MV of the block D, which is not intra-mode macroblock and have an inter MV.

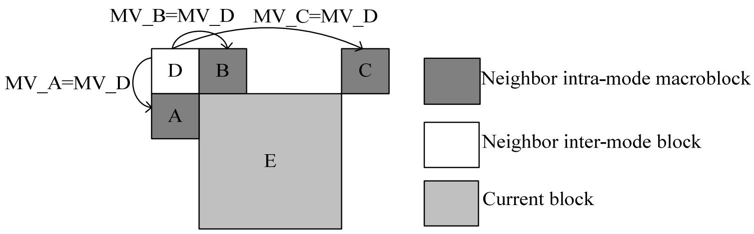


Fig.5. Three intra-mode macroblocks beside block E

## 3 Experimental results

To evaluate the coding efficiency, our experiments compare the performance of the proposed MVP scheme with that of median predictor, which is widely used in current video coding standards. The proposed method is implemented based on the H.264/AVC reference software JM76. The two test sequences, stefan and foreman, are both 300 frames CIF sequences with the frame rate of 30f/s. The test conditions are showed in Table 4.

Table 4. Test conditions

| | |
|---|---|
| MV resolution | 1/4 pel |
| Hadamard | ON |
| RD optimization | ON |
| Search Range | ±16 |
| Reference Frames | 2 |
| GOP structure | IBPBP |

Table5 shows that the number of intra-mode macroblocks is reduced by about 28% using proposed MVP scheme than before in the sequences. It is thought that after implementation of the proposed MVP scheme, less macroblocks fall into intra-mode than before.

As shown in table 6, for macroblocks in the neighbor of intra-mode macroblocks, the average coding efficiency of our scheme is better than median predictor, the maximum PSNR gain is up to 0.19dB (QP=32).

The test results of two predictors in different QPs are compared in Table 7. PSNR and bitrate of sequences are showed in the two tables. A comparison can be observed that the proposed predictor achieves better performances than median predictor in most of QPs. Especially at low bit rate, the luma PSNR gain is up to 0.18dB.

Table8 shows that the computation complexity of the proposed algorithm is reduced. In blocks beside intra-mode macroblocks, the speeding up ratio is about 24%. In the whole sequence, the speeding up ratio is about 12%.

Table 5. Percentage of intra mode macroblocks (QP=32)

| | Original median predictor | Proposed MVP scheme |
|---|---|---|
| Stefan | 12% | 9.5% |
| Foreman | 9% | 7.1% |

Table 6. Average PSNR comparison in intra-mode macroblock areas(QP=32)

| | Original median predictor | Proposed MVP scheme |
|---|---|---|
| Stefan | 32.20 | 32.39 |
| Foreman | 34.61 | 34.79 |

Table 7. The test results of two predictors for sequences in CIF (300 frames)

| Sequence | QP | Original median predictor | | Proposed MVP scheme | |
|---|---|---|---|---|---|
| | | PSNR(dB) | Bitrate(bps) | PSNR(dB) | Bitrate(bps) |
| Stefan | 24 | 38.71 | 1861.80 | 38.77 | 1859.65 |
| | 28 | 35.63 | 1002.00 | 35.69 | 1000.20 |
| | 34 | 31.05 | 358.10 | 31.13 | 354.05 |
| | 40 | 26.72 | 143.38 | 26.83 | 138.43 |
| | 44 | 23.99 | 81.47 | 24.17 | 76.74 |
| Foreman | 28 | 37.22 | 376.34 | 37.27 | 374.32 |
| | 34 | 33.69 | 165.23 | 33.75 | 163.34 |
| | 40 | 30.33 | 76.47 | 30.42 | 73.53 |
| | 44 | 28.11 | 48.67 | 28.21 | 46.35 |
| | 48 | 25.59 | 30.73 | 25.76 | 28.43 |

Table 8. AD counts comparison(QP=32)

| | | Mean AD counts of intra-mode areas(*$10^3$) | Mean AD counts of all blocks(*$10^3$) |
|---|---|---|---|
| Stefan | Median predictor | 124.82 | 94.43 |
| | Proposed predictor | 98.23 | 83.57 |
| Foreman | Median predictor | 33.54 | 25.47 |
| | Proposed predictor | 26.32 | 22.42 |

## 4 Conclusion

In this paper, we firstly review the median predictor and other MVP schemes proposed by researchers before. Then we propose an improved MVP scheme for H.264. For the block whose MV is to be predicted, the scheme exploits its more neighbor MV information and uses these information adaptively. The proposed MVP method is simple, but it is effective. It gives a better coding efficiency and brings less FME computation complexity. It can be applied in current wireless video communication. In the future, the proposed method can further exploit the spatial-temporal MV information, in that it gets the more accurate derivation of MVP than original median predictor.

**Acknowledgements**

The study work has been supported in part by the National Natural Science Foundation of China under grant No.60333020.

## References

[1] ITU-T Recommendation H.263: Video Coding for Low Bitrate Communication (June 1996)

[2] Draft ITU-T recommendation and final draft international standard of joint video specification (ITU-T Rec.H.264/ISO/IEC 14 496-10 AVC). In Joint Video Team (JVT) of ISO/IEC MPEG & ITU-T VCEG, JVT-G050 (2003)

[3] Y. Lee, F. Kossentini, R. Ward, M. Smith: Prediction and Search Techniques for RD-Optimized Motion Estimation in a Very Low Bit Rate Video Coding Framework. ICASSP-97, vol.4 (Apr 1997) 2861-2864

[4] Ismaeil. I., Docef, A., Kossentini, F., Ward, R.K.: Efficient Motion Estimation Using Spatial and Temporal Motion Vector Prediction. Proc. of the IEEE Inter. Conf. on Image Proc.. Koke, Japan (October 1999)

[5] J. B. Xu, L. M. Po and C. K. Cheung: Adaptive motion tracking block matching algorithms for video coding. IEEE Trans. Circuits and Syst. Video Technol., vol. 97 (Oct. 1999) 1025-1029

[6] Sung Deuk Kim, Jong Beom Ra: A New Motion Vector Coding Technique. 39th MPEG meeting, m2100. Bristol, UK (Apr. 1997)

[7] S. D. Kim and J. B. Ra: An efficient motion vector coding scheme based on minimum bitrate prediction. IEEE Trans. Image Processing, vol. 8 (Aug. 1999) 1117-1120

[8] Sung Deuk Kim, Jong Beom Ra: Improved Motion Vector Coding Scheme. 38th MPEG meeting, m1748. Sevilla, Spain (Feb. 1997)

[9] Li Zhang, Wen Gao: Hybrid Algorithm with Adaptive Complexity for Interger Pel Motion Estimation of H.264. Conference of ICME (Mar 2005)