A Study on the Motion Vector Prediction Schemes for AVS

Honggang Qi, Wen Gao, Siwei Ma, and Debin Zhao*

Institute of Computing Technology
Chinese Academy of Sciences

ABSTRACT

The motion vector prediction (MVP) is an important part of video coding. There have been numerous workings on the topic done by researchers before. In this paper, a continue study on MVP of video coding based on the workings of predecessors is made. The video sequences with various motion characteristics are further investigated. The characteristics of motion vectors of objects in video scenes are discussed briefly. Then, summarizing these characteristics, two MVP schemes for a new coding standard, Audio and Video Standard (AVS), are proposed. In these schemes, current block's MV can be predicted based on statistical correlation of MVs of spatial contiguous neighbor blocks. A correlation criterion is employed to measure how correlated between two MVs. With the correlation criterion, the correlated MVs of neighbor blocks are determined. Then, the predicted MV of current block can be obtained with some simple algebraic operations on determined MVs. The two proposed schemes, as the alternative ones of median predictor, are suitable for different video sequences with different motion characteristics, respectively. The experimental results show that the bit rate savings are achieved with these schemes in most of typical video sequences, compared with the median predictor implemented in AVS.

Keywords: Audio and Video Standard, motion vector prediction, video coding

1. INTRADUCTION

The motion vector prediction (MVP) is an important issue in a video standard. MVP acts on three basic parts of an encoder: 1) predicting the starting point of fast motion estimation (FME) searching; 2) predicting the currently encoded block's MV; 3) deriving MVs for special modes without MVs information in bitstream. First, the goal of fast motion estimation is finding the approximate optimal MV by searching fewer points than full search. Usually, an MV is predicted for the initial search center of FME, and then the possible points around the initial search center are searched according to FME algorithm. The majority of FME algorithms often fall into local optimal point which may be not real optimal MV for current block. If the MVP predicts a point closer to the global optimal point, the probability of encountering local optimal points is reduced. Thus, the failure may seldom occur. The accurate MVP is also helpful for reducing the time complexity of motion search, and thus reduces the complexity of encoder. In contributions[12, 13, 18], MVP is adopted for FME. Second, in inter-picture coding, the MV information must be coded into bitstream for the right reconstruction of the inter-pictures in decoder. The motion vector coding has a significant influence to the video sequence coding efficiency. Especially, at low bitrate, the proportion of bits for MV is large in whole bitrates. The MV coding influences the coding efficiency more significantly than the texture coding. In order to compress the bits for coding MV, the MV difference (MVD) is coded into bitstream instead of the MV by oneself. MVD is the difference of an actual MV and a predicted MV. In the decoder end, the MV is retrieved through MVD adding the predicted MV. Because the predicted MVs got by encoder and decoder are identical, the decoded MV in the decoder is also identical to that estimated by motion estimation (ME) in the encoder. In the way of coding MVD, if the MVP is accurate enough, the actual and predicted MVs are close. The value of MVD approaches to zero. And thus, the fewer bits are consumed at MVD. Third, MVP can derive MVs for the special inter-prediction modes without MV information in bitstream. These special modes use the MVs derived by MVP for motion compensation instead of coding MVs into bitstream[7, 8, 16]. An accurately derived MV makes smaller residuals between current block and prediction block. Since the deriving MVs are not coded into bitstream, it contributes to the bit saving. At low bitrate, these modes are often pitched on for the total bit-consumings for them is less than those for other modes in large quantization step size. The above three important roles of MVP in video coding motivates us to improve it. Many works in MVP or relative MVP have been done by pioneer researchers for last several decades. Presently, the median predictor is widely applied in many important coding

*{hgqi, wgao, swma, dbzhao}@jdl.ac.cn

Visual Communications and Image Processing 2005, edited by Shipeng Li,
Fernando Pereira, Heung-Yeung Shum, Andrew G. Tescher, Proc. of SPIE Vol. 5960
(SPIE, Bellingham, WA, 2005) · 0277-786X/05/$15 · doi: 10.1117/12.633434

standards, such as H.263[6]/H.264[1]. It obtains the median values of corresponding components of MVs from the neighbor blocks as the components of predicted MV. The scheme is simple but is not suited for the ununiform motion scene. As an alternative solution for median predictor, the contribution[5] proposes another one which selects the MV of the one closer to a reference block from two neighbor blocks as the predicted MV. In the contribution[10], besides the median prediction, the linear prediction, the mean prediction and the weighted mean prediction, are also employed to locate the initial searching point of ME. The mean prediction is only the average of the MVs from the neighbor blocks. The weighted mean prediction is an improvement of the mean prediction. The scheme assigns a weighted factor for each neighbor block according to the relation between MVs of current block and its neighbor blocks before coding. And then, calculates the weighted mean of MVs from the neighbor blocks with these weighted factors while coding. The contribution[11] uses conditional probabilities of finite state machine uniquely determined by a dynamic state vector quantization code book to estimate statistical mean for MVP. Considering the inaccurate prediction with uncorrelated MVs of neighbor blocks, the contirbution[12] joins a predefined threshold to guarantee the validate MVs to be used for median predictor. Considering the condition of multiple reference frames in H.264, the contirbution[8] scales each neighbor MV, which refers different frame from that of current block, to the same one according to the temporal distances of the reference frame. The method improves the reliability of neighbor MVs. Ismaeil, etc. The contirbution[14] adds the temporal prediction in MVP to compensate the insufficiency of spatial prediction. The idea of spatio-temporal prediction is not initiated by them, but they bring a new temporal prediction method based on a projection technique. The spatio-temporal predictor can achieve small gains over single spatial prediction at the cost of higher complexity in both computation and memory. It is may be unpractical to apply in video coding standard. However, the temporal prediction finds good application in motion compensation temporal filtering (MCTF). The performance of the temporal prediction outperforming that of the spatial prediction is showed in the MCTF with multiple temporal hierarchies[17]. In the contirbutions[2, 3, 4], the S. D. Kim and J. B. Ra introduce a flag indicating which neighbor MVs are used to predict the current MV in bitstram. The neighbor MV obtaining MVD with minimum bit amount is selected as the predicted MV of current block in encoder. Then, two-bit flag is transmitted to signal the decoder that the indicated MV is the predicted one. Considering the median prediction with the better performance in uniform motions scenes, the two schemes are mixed into one. A threshold is employed for determining proper scheme adaptively. The predictor is good at coding efficiency but the complexity is high, too. Except for reducing the MVD, in the contirbution[7, 16], a predictor is employed in the direct mode for B picture and the skip mode for P picture. In the paper, the predictor is employed for not only predicting the current MV but also deriving the MVs for blocks with special modes. The improved results show the MVP applying in the special modes saves the bitrate for coding MVs efficiently.

The AVS is an up to date standard in China just like MPEG-2[9], MPEG-4[15], and H.264[1] in the world. The basic framework of AVS is similar to that of these popular video standards. However, the details in each technique of AVS are different. Studying on some methods used in MPEG-2/4 and H.26L and referencing from some others in relative topic papers, we propose two simple MVP schemes for AVS. The two proposed schemes predict MV based on statistic analysis to spatial characteristics of video pictures. Although only one of them is used in AVS reference software now, they present similar performances in coded bit saving. As the characteristics of different video sequences are different, each scheme is more suitable for some kinds of typical video sequences. These predictors can all be used in all existing video standard.

The paper is organized as follow: In the next section, the theory of MVP is discussed. The third section describes the three proposed predictors based on the theory of MVP. And then, the experimental results are presented in the fourth section. Finally, conclusion is followed.

2. THE THEORY OF MOTION VECTOR PREDICTION

In present video coding standard, a video picture is coded in multiple macroblocks (MB). The size of macroblock is 16x16 pixels. A MB may be divided into blocks with smaller sizes for hierarchical motion compensation (MC). The way of video coding based on block motion estimation and compensation is called block matching algorithm (BMA). In video sequences with motion scenes, MVP of BMA is to predict a MV for current block using the known information. The information can be retrieved from the previously encoded blocks which have finished motion estimation and can provide its MVs. A motion object in a picture usually covers more than one block. The object is divided into several parts by the neighbor blocks. Some neighbor blocks may contain the parts of a motion object locating in current block. In this case, the motion direction of the current block and its neighbor blocks should tend to be identical to that of the motion object in ideal condition. Thus, the MVs of current block and its neighbor blocks should be identical. Based on the assumption, the MVs of blocks around current block is useful for predicting the MV of current block. However, in practice, the situation is not simple like that. Since the neighbor blocks around current block may contain other objects not located in current block, with the ME, the MVs of the neighbor blocks acted on by multiply objects may be different from the current block. In this situation, the MVP is invalid. This is a disadvantage of video coding based on BMA. However, at least, they are still highly correlated with MV of current block. In the coding way of MVD, anyway, a predicted MV for current block must be obtained. The MVs of spatial neighbor blocks are the preference for current MV prediction. The Fig.1, taking some examples, shows the MVs of blocks and the motion objects in a picture of video sequence. In the figure, the MV of current block 5 is identical to the motion object located in it as there are no other objects disturbing the MV of current block tending to the object motion. The MVs of the neighbor blocks 2, 3 and 6 are also identical to the MVs of the parts of the object located in them. Whereas, to the neighbor blocks 1 and 4, the other motion objects may result in that the MVs of blocks are different from that of the motion object.

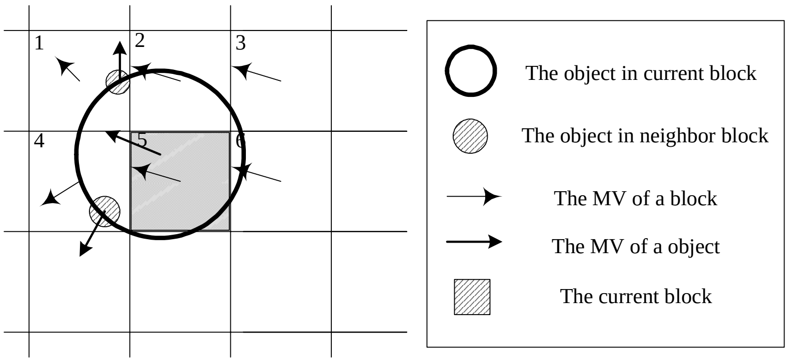

Fig.1. The relation of MVs of blocks and motion objects

Object motion can be classified into rigid motion and non-rigid motion. In the case of rigid motion, the motion is simple translation. Its inner motion velocity of object is identical everywhere. The characteristic of rigid motion help us to predict the MV of a part of a motion object by that of the other part of the motion object accurately. However, in the case of non-rigid motion, the motion includes rotation and distortion more than translation. The inner motion velocity of a non-rigid motion is not uniform. Unlike the translation, the rotation and deformation are not easy to be simply predicted by the correlation of MVs of parts of the object. The complex motion situation makes the correlation of MVs of parts of the object weak. This is the reason why the compression ratio of sequences with non-rigid motion is low. This is the other disadvantage of video coding based on BMA. The situation of MVs of the inner parts of the rigid motion object and the non-rigid motion object is demonstrated in Fig.2. Although the correlation of inner parts of non-rigid motion object is weak, the MVs of the parts around the current parts are still the best references for predicting a close MV for current part. The median predictor is good at rigid motion for its uniform motion of rigid objects. However, as for the non-rigid objects, since the irregular motion velocities of inner points of the objects, the median predictor is invalid. Thus, we are aiming to find other predictors to avoid the shortage of median predictor.

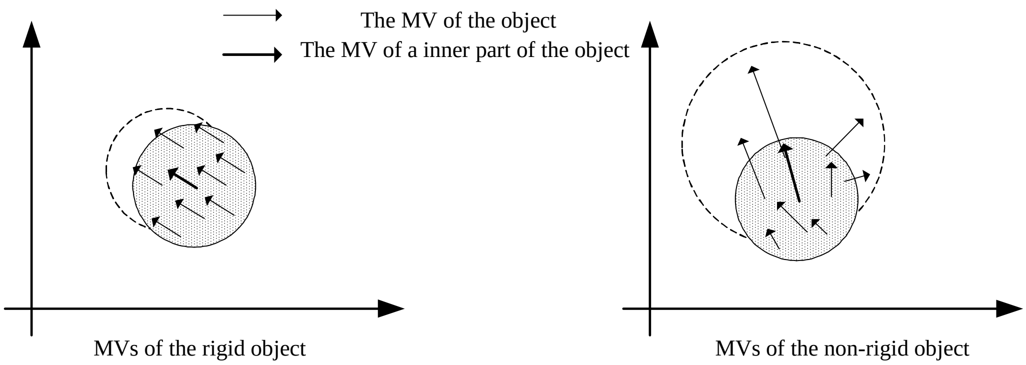

Fig.2. The MVs of every part in the rigid object and the non-rigid object

Whether a motion object is rigid or non-rigid, it should have a MV which can represent the motion direction of the object. The MV is called main MV. To rigid objects, the main MV is identical to the MVs of every part of the object. To the non-rigid objects, although the MVs of every part are different, there is a main MV for the non-rigid objects which reflects the motion direction of non-rigid objects roughly. The main MV of a motion object is preponderantly determined by the majority of MVs of parts in the object, thus the majority of MVs should approach to the main MV. In BMA, a motion object may cover several blocks including current block. Whatever are rigid objects or non-rigid objects, the MV of current block, i.e. current MV, should be approximate to the main MV in probability without regard to some unexpected conditions. In view of it, we think that the main MV of an object is a good prediction for current MV. The main MV can be expressed as the vector mean or median of the majority of similar MVs. Meanwhile, the minority of MVs far apart from the majority of MVs are discaded. In order to obtain the similar MVs, a criterion is employed to decide how close two MVs in distance. The distance between two vectors is usually calculated by Euclidean distance. The Euclidean distance is expressed as:

$$\sqrt{((x_1 - x_2)^2 + (y_1 - y_2)^2)} \tag{1}$$

where x_1, x_2, y_1, y_2 are components of the vector (x_1, y_1) and (x_2, y_2). The Euclidean distance formula is often used to calculate the distance between two points in the two dimension space. The Euclidean distance expression is a square root form. In order to reduce computing complexity, a simplified form of Euclidean distance is used as the correlation criterion of MVs in MVP

$$|x_1 - x_2| + |y_1 - y_2| \tag{2}$$

where the sum of the absolute items substitutes the square root of the square sum items. The result of formula (2) can also indicate the approximate Euclidean distance between two vectors like the formula (1). Larger is the result, farther the distance is. The distance between two MVs can reflect their correlation, so the criterion is called correlation criterion. With the correlation criterion, the candidate MVs are dealt with further. Then, the main MV for prediction of current MV can be obtained with some simple algebraic operations on the dealt MVs.

In coding standards, the codec complexity is very important for practice applications. Thus, for simplifying complexity, the fewest but representative blocks are used for MVP. Also, it is not the MVs of all neighbor blocks can be obtained before predicting the current MV. Only those, having been encoded previously, can be obtained. The blocks selected for prediction are called candidate blocks and their MVs are called candidate MVs. Commonly, three representative blocks are employed by predictor. The spatial positions of current block and three candidate blocks may be clearly learned from Fig.3. In the figure, the block E denotes the current block to be predicted. The candidate block A contains the pixels of the left of block E; the candidate block B contains the pixels of just above block E; the candidate block C contains the pixels of upper right of the block E. The candidate block D contains the pixels of upper left block E. the block D is used only if the block C is not available. Motion situations of the four candidate blocks around the current block can contain that of current block basically. If a motion object locating in the current block is larger than the current block, the other parts are likely to fall into some of the candidate blocks. Thus, the MVs of neighbor blocks may estimate the current MV well. However, the parts of the motion object sometimes fall into the current block and the unprocessing blocks. The MVs of unprocessing blocks cannot be obtained while coding current block. In this situation, the spatial predicted scheme is failing, and the current MV is effectively predicted by nothing but the temporal prediction. However, the temporal prediction increases both the memory and time complexity. It is a high cost trading off the coding gain and complexity. Thus, the temporal prediction is not usually employed in some video coding standards. Some schemes about temporal prediction are merely limited in research.

In multiple reference frame prediction, the candidate MVs from different reference frames is also a reason of inaccurate prediction. The MVs from different reference frames contain the temporal motion information which disturbs the prediction based on spatial domain. The unprocessed MVs are valueless for the predictors merely based on spatial prediction. The temporal disturbing information can be approximately removed by scaling the MVs from different reference frames in the same frame referenced by current MV according to the temporal distance between reference frames. The temporal distance between two consecutive frames is fixed. It can be calculated by the inverse of frame rate. The accurate prediction can be obtained with scaled MVs.

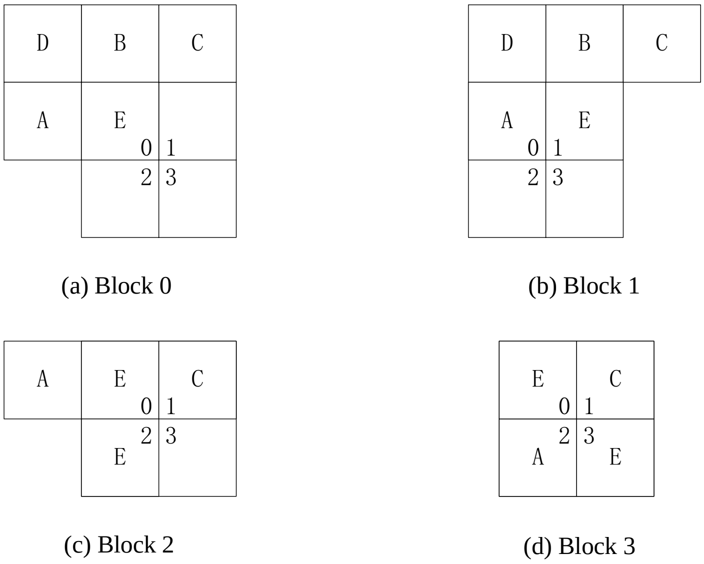

Fig.3. the location of current block and its neighbor blocks

3. THE METHODS OF MOTION VECTER PREDICTION

In terms of the above analysis, we propose two MVP schemes for AVS. They are Average of Closest MVs Scheme and Median MV Scheme. Their prediction process is very similar. They all employ the correlation criterion for calculating the correlations of candidate MVs. However, the minor difference is how to deal with the MVs obtained by the correlation criterion.

3.1 Average of Closest MVs Scheme (AOC)

After the selection of candidate blocks, three candidate blocks around the current block are obtained. Their MVs for MVP along with them are also obtained. Assumed an area is composed of current block and its candidate blocks. Although MVs of the blocks in the area may be different, the area still has a main motion direction. The main MV of the area should approach MVs of the majority of blocks in the area. The current MV in the area is very likely to approach the MVs of the majority of blocks in the area. Thus, the correlation between main MV and the current MV is strong. In MVP, The current MV should approach no fewer than two of three candidate MVs. Moreover, considering the exception of a block with weak motion correlation to current MV, the two candidate blocks with closest MVs are employed as the prediction for current MV. Thus, only the two candidate blocks with shortest distance are selected for prediction and the left one is discarded. According to above description, the scheme is expressed with the following formula:

$$P,Q = \underset{\substack{i,j=1,2,3,i\neq j \\ M_i\in\{A,B,C(D)\}}}{\arg\min} \left(\left|M_i(x)-M_j(x)\right| + \left|M_i(y)-M_j(y)\right|\right) \tag{3}$$

where M denotes the candidate MV; the M_i and M_j denote MVs two different blocks; and x and y are the horizontal and vertical components of vector M respectively. The pair of MVs with smallest result from three candidate MVs is obtained with formula (3), which is just the pair of correlated MVs. After selecting two closest candidate MVs, a decision which one is closer to the current MV between the two MVs can not be made yet. The problem can be solved by an eclectic decision, which uses the mean value of the two cloest MVs as the main MV of the area. The decided MV represents the motion direction of the majority of MVs in the local area composed of the current block and its neighbor blocks. The mean value of the two MVs is the mean values of their corresponding components. The main MV is just the prediction of current MV.

3.2 Median MV Scheme

The method is inspired by the median predictor. Median predictor uses the median value of the corresponding components of candidate MVs as the components of predicted MV, whereas, in this method, the median of candidate vectors instead of the median of corresponding components of candidate vectors becomes the predicted MV. The median is the best choice in probability without any prior knowledge. The median of three MVs should be the MV in the middle of other two MVs. Thus, the median MV is not farthest from any one of other two MVs. In other words, the other two MVs are the pair of three MVs with farthest distance. The median MV can be obtained excluding the pair MVs with farthest distance. After the selection of candidate MV, the correlations of each pair of three candidate MVs are calculated with the correlation criterion. Then, the pair of MVs with maximum result is obtained through the following formula:

$$P, Q = \underset{\substack{i,j=1,2,3, i \neq j \\ M_i \in \{A,B,C(D)\}}}{\arg\max} \left(\left| M_i(x) - M_j(x) \right| + \left| M_i(y) - M_j(y) \right| \right) \tag{4}$$

Excluding the pair of MV with maximum result, the left one just becomes the predicted MV. The method considers that the MV employed as the elements of MVP is better than its components as there exists the correlation between the two components. The object motions in scenes are measured by the whole MV instead of its separated components and the components just are the analysis of a MV. Thus, a median MV is used as the predicted MV. The method, like the median predictor, is suitable for video sequences with characteristic in relatively uniform motion. It can achieve as well predicted results as the median predictor.

4. EXPERIMENTAL RESULTS

The two proposed MVP schemes and median predictor are implemented into the AVS-M reference software WM2.0. Several typical sequences with different quantization parameters (QP) and sizes are test. The test conditions are listed in Table 1. The comparisons are performed among these schemes with fixed quantization parameters (QP). The test results of three predictors in different QPs are compared in Table 2 and Table 3. PSNR and bitrate of sequences in CIF or QCIF are showed in the two tables. A comparison can be observed that the proposed predictors achieve better performances than median predictor for provided test sequences in most of QPs. The improvement is the integral behaviors of three aspects, ME, MVD and special modes, which MVP acts on. These sequences feature in the fast and complexity local motions. As for these sequences, the median predictor is not as good as the proposed predictors.

Table 1. The condition for MVP test

Frame to Be Encoded	200
Sequence type	IPPP…
Frame rate (fps)	30
Number of Reference Frames	2
RDO	On
QP	28，32，34，36，40，44
Search rang	CIF:32;　QCIF:16

5. CONCLUSION

In this paper, two simple MVP schemes are proposed for usage in AVS. The schemes are designed based on the analysis of spatial motion characteristics of objects. The correlation of current block and its neighbor blocks is fully explored to improve the coding efficiency for some typical video sequences. The process of MVP is the simple algebraic operations on the few candidate MVs. Meanwhile, the proposed schemes predict the current MV spatially. Thus, they are as simple in codec complexity as the median predictor. As a general scheme, the proposed MVP schemes are not limited to usage in AVS. It is capable of substituting for any existing predictor in video coding standards.

Table 2. The test results of three predictors for sequences in QCIF

Sequence	QP	Median predictor		AOC predictor		Median MV predictor	
		PSNR (db)	Bitrate (bps)	PSNR (db)	Bitrate (bps)	PSNR (db)	Bitrate (bps)
Foreman	28	37.1419	195.902	37.1069	195.997	37.1135	195.572
	32	35.0712	122.722	35.0628	123.168	35.0789	122.996
	36	33.1961	80.3940	33.1830	80.9628	33.2121	80.4276
	40	31.2094	54.3384	31.2749	54.3636	31.2307	53.9556
Container	28	37.3015	59.8056	37.2920	59.6952	37.3057	59.4912
	32	35.2851	34.5744	35.2947	34.5108	35.2808	34.4964
	36	33.1924	21.8940	33.1981	21.8292	33.1901	21.9912
	40	31.2435	14.3676	31.2262	14.2164	31.2242	14.3424
Tempete	32	32.7497	337.235	32.7467	337.303	32.7556	337.134
	34	31.7153	271.390	31.7269	271.658	31.7308	270.748
	40	28.2109	120.928	28.2244	120.926	28.2302	120.779
	44	26.1315	72.8160	26.1337	72.0872	26.1267	72.8056
Paris	32	33.69329	167.881	33.7308	169.003	33.7242	168.361
	34	32.62646	141.082	32.6394	141.596	32.6370	141.034
	40	29.01461	73.1484	29.0377	73.3368	29.0108	73.2372
	44	26.84480	46.1244	26.8718	46.2012	26.8572	46.0272

Table 3. The test results of three predictors for sequences in CIF

equence	QP	Median predictor		AOC predictor		Median MV predictor	
		PSNR (db)	Bitrate (bps)	PSNR (db)	Bitrate (bps)	PSNR (db)	Bitrate (bps)
Foreman	28	38.3570	630.912	38.3473	631.688	38.3422	629.284
	32	36.4820	411.888	36.4915	413.566	36.4936	412.037
	36	34.7761	280.529	34.7837	281.712	34.7789	280.990
	40	32.9589	193.400	32.9818	193.937	32.9810	192.682
Stefen	28	36.9023	1677.06	36.9096	1680.25	36.9049	1676.86
	32	34.5340	1025.15	34.5341	1027.60	34.5343	1025.14
	36	32.3953	639.421	32.4041	640.702	32.4027	639.131
	40	30.1501	394.798	30.1531	395.860	30.1565	394.490
Tempete	28	36.2535	1741.40	36.2502	1739.61	36.2554	1740.80
	32	33.8552	1044.67	33.8534	1044.89	33.8540	1044.52
	36	31.7585	620.844	31.7570	619.199	31.7487	620.100
	40	29.6272	352.477	29.6273	352.057	29.6292	353.297

ACKNOLOGMENTS

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

REFERENCES

1. "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
2. S. D. Kim and J. B. Ra, " An efficient motion vector coding scheme based on minimum bitrate prediction", IEEE Trans. Image Processing, vol. 8, no. 8, pp. 1117-1120, Aug. 1999
3. Sung Deuk Kim, Jong Beom Ra,"Improved Motion Vector Coding Scheme", m1748, 38th MPEG meeting, Sevilla, Spain, Feb. 1997
4. Sung Deuk Kim, Jong Beom Ra, "A New Motion Vector Coding Technique", m2100, 39th MPEG meeting, Bristol, UK, Apr. 1997
5. Gisle Bjøntegaard, Tom-Ivar Johansen, Michael Horowitz, "Technical solutions relating to RFB achievement," Document JVT-F058, 6th Meeting: Awaji, Island, JP, 5-13 December, 2002
6. ITU-T Recommendation H.263, "Video Coding for Low Bitrate Communication", June 1996
7. Alexis M. Tourapis, Feng Wu, Shipeng Li, "Direct macroblock coding for predictive (P) pictures in the H.264 standard", SPIE Visual Communications and Image Processing (VCIP), vol. 5308, pp 364-371, January 2004
8. Alexis M. Tourapis, Feng Wu, Shipeng Li, "Motion vector prediction with reference frame consideration", SPIE proceeding on Applications of Digital Image Processing, vol. 5203, August 2003
9. ITU-T Recommendation H.262 . ISO/IEC 13818-2 (MPEG-2): Generic coding of moving pictures and associated audio information. Part 2: Video ITU-T and ISO/IEC JTC1, November 1994
10. 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", Acoustics, Speech, and Signal Processing, 1997. ICASSP-97, vol.4, pp2861-2864: 21-24 Apr 1997
11. F. Kossentini, Y. Lee, M. Smith, and R. Ward, "Predictive RD-Constrained Motion Estimation for Very Low Bit Rate Video Coding", Special Issue of IEEE Journal on Selected Areas in Communication, 15 (9), December 1997
12. 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, pp. 1025-1029, Oct. 1999
13. C. Hsieh, P. C. Lu, J. S. Shyn, and E. H. Lu, "Motion estimation algorithm using interblock correlation", Electron. Lett., vol. 26, pp. 276–277, Mar. 1990
14. 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
15. Coding of Moving Pictures and Associated Audio ISO/IEC 14 496-2, MPEG-4 Committee Draft ISO/IEC TC1/SC29/WG11, Oct, 1998
16. Tourapis, A.M., Feng Wu, Shipeng Li, "Direct mode coding for bipredictive slices in the H.264 standard", IEEE Trans. Circuits Syst. Video Technol., vol. 15, pp. 119- 126, Jan. 2005
17. D. S. Turaga and M. van der Schaar, "Temporal prediction and differential coding of motion vectors in the MCTF framework", ICIP, September 2003
18. H.Y.Chung, N.H.C.Yung and P.Y.S.Cheung, "Fast motion estimation with search center prediction", Optical Engineering - The Journal of SPIE, Vol. 40, No. 6, June 2001, pp.952-963