Motion analysis using the normalization of motion vectors on MPEG compressed domain

N. W. KIM, T. Y. KIM, and J. S. CHOI

Department of Image Engineering, Graduate School of
Advanced Imaging Science, Multimedia, and Film, Chung-Ang University
221 Huksuk-dong, Dongjak-ku, Seoul, 156-756, Korea
TEL : +82-2-820-5295, FAX : +82-2-825-1584
E-Mail: mysope@imagelab.cau.ac.kr

Abstract : In this paper, we propose a method that converts motion vectors on MPEG coded domain as a uniform set, independent of the frame type and the direction of prediction, and directly utilizes these normalized motion vectors for understanding video contents. This frame-type-independent motion vectors are utilized as feature information for image retrieval or moving object tracking on compressed domain. By simulation, we evaluate the effectiveness of the proposed method and compare its performance to the conventional method.

1. Introduction

According to the explosive increase of video data, active research is underway in video paring algorithm for effective video analysis and video indexing algorithm for rapid access and retrieval to parsed video data [1][2].

The content of shot segmented by video parsing is essentially characterized by two factors: underlying scene, and the camera work. Especially, since motion is an important component of the underlying scene as well as the camera work present in a shot, the motion-based video analysis has received large attention in video databases research [3-7].

In this paper, we propose a method for effective motion analysis using motion vector information in MPEG compressed domain. Previous methods have a problem that the number of normalized motion vectors is rarely large enough to represent each frame, because they have only considered motion vectors single-directionally predicted from one reference frame[8][9]. In this paper, we have presented a new technique for motion analysis with a lot of normalized motion vectors obtained by bi-directional prediction structure from two reference frames on MPEG compressed domain.

Experimental results show that the performance of our algorithm is superior to that of conventional method.

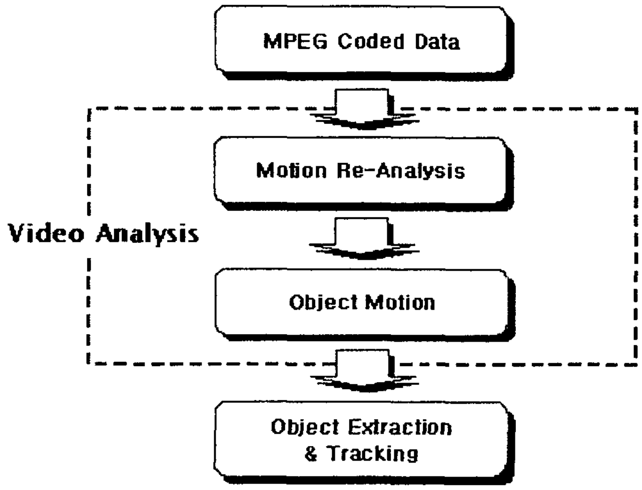


Fig. 1. Block diagram of the proposed algorithm.

2. Motion analysis using the normalization of motion vectors

A macroblock(MB) can have zero, one, or two motion vectors depending on its frame type. Moreover, these motion vectors can be forward-predicted or backward-predicted with respect to reference frames. For effective motion analysis, we therefore require a more uniform set of motion vectors, independent of the frame type and the direction of prediction.

2.1 Conventional method

Let us consider two consecutive reference frames, R_i and R_j. Let the B frames between the them be denoted by $B_1,\ldots,B_n$, where n is the number of B frames between these two reference frames (typically, n=2). From the mutual relation among these frames, we can represent each motion vector as a backward-predicted vector with respect to the next frame, independently of frame type [8].

The first step is to derive the flow between the first reference frame R_i and its next frame B_l using the forward-predicted motion vectors of B_l. Intuitively, if an MB in the B_l frame, $(B_l)_{u,v}$, is displaced by a motion vector (x, y) with respect to an MB in the R_i frame, $(R_i)_{u,v}$, then it is logical to conjecture that the latter MB is displaced by a motion vector (-x,-y) with respect to $(B_l)_{u,v}$. If $(B_l)_{u,v}$ does not have a forward-predicted motion vector with respect to R_i, we look at successive B frames until we find a frame, say B_k, in which the corresponding MB has a valid forward-predicted motion vector, and we use the inverse of that vector. Since this vector is predicted from k frames earlier, we scale it down by a factor of k. Therefore, the flow for an R_i frame is obtained by a motion vector with respect to an MB in the B_l frame.

Similarly, using the backward motion vectors of frame B_n with respect to R_j, the flow is derived for frame B_n. There is no need to invert the motion vectors here, since the flow vectors essentially are backward-predicted vectors. The flow for the conventional algorithm is shown in Fig. 2.

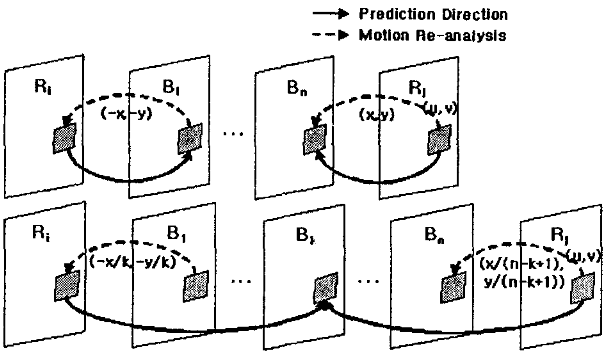


Fig. 2. Re-construction of motion vectors by the conventional method in R_i, B_n frames.

But, this algorithm is not suitable for bi-directional prediction framework like MPEG video sequence, because it considers only single-directional prediction as the flow estimation for R_i, B_n frame.

2.2 Proposed method

In this paper, our approach is to implement a new algorithm for bi-directional motion analysis using forward-predicted motion vector from R_i to R_j, free from the limitation of single-directional motion analysis.

The numerical formula is expressed by the equation (1) and (2). Let the forward-predicted motion vector of the current MB in B_n be denoted by $\overrightarrow{B_nR_i}$, and let the forward-predicted motion vector of the corresponding MB in R_j be denoted by $\overrightarrow{R_jR_i}$. If, we denote the flow of the MBs between B_n and R_j by $\overrightarrow{B_nR_j}$, then we have the relationship

$$\overrightarrow{B_nR_j} = \overrightarrow{B_nR_i} - \overrightarrow{R_jR_i} \qquad (1)$$

from which $\overrightarrow{B_nR_j}$ can be obtained easily. Contrary to the conventional method which only uses backward-prediction motion vectors to obtain the flow $\overrightarrow{B_nR_j}$, the proposed method estimates more accurate motion flow $\overrightarrow{B_nR_j}$ using forward-predicted motion vector $\overrightarrow{R_jR_i}$ as shown in Fig. 3.

And, let the backward-predicted motion vector of the current MB in B_1 be denoted by $\overrightarrow{B_1R_j}$. Then, we calculate $\overrightarrow{R_iB_1}$ as

$$\overrightarrow{R_iB_1} = -\overrightarrow{R_jR_i} - \overrightarrow{B_1R_j}\,. \qquad (2)$$

We derive more elaborate motion flow in B_n, R_i frame by using the numerical formula (1), (2) above. As a result, adding this algorithm to the conventional method, we can obtain much more normalized motion vectors than ones extracted by the conventional method in R_i and B_n frames. The uniformed motion vector set estimated by such method is large enough to be used as feature factor to represent each frame. This result is showed by Fig. 4.

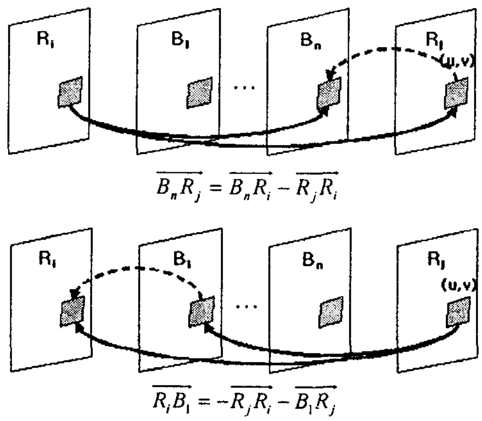


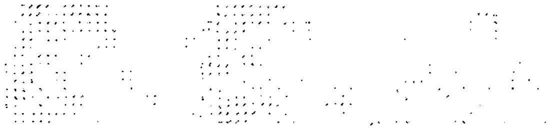
Fig.3. Re-construction of motion vectors by the proposed method in R_i, B_n frames.

3. Simulation Results

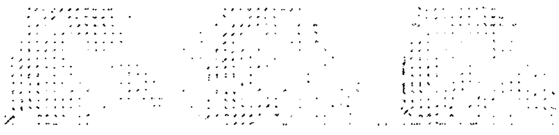
In this section, we have investigated the validity of proposed method and evaluated its performance.

Experimental database consists of MPEG-1 coded video sequences, which have a various camera work and object motion.

(a) Samples of original sequence



(b) Motion flow representation by the conventional method



(c) Motion flow representation by the proposed method

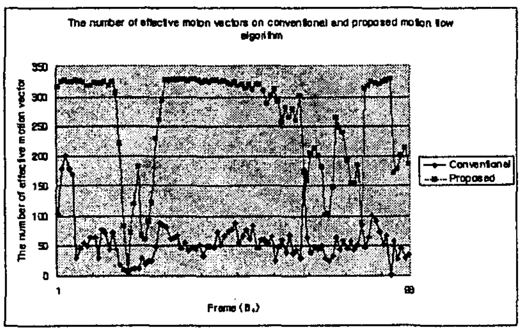
Fig. 4. Motion flow analysis (R_i, B_1 and B_2 frames from left-side to right-side).

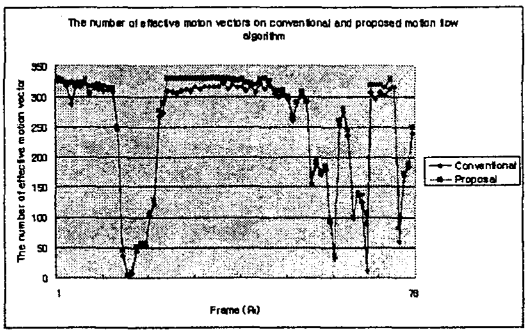
Fig. 5 depicts the comparison of the numbers of normalized motion vectors between the conventional and the proposed motion analysis method. The number of normalized motion vectors in B_n, R_i obtained by the proposed method has increased over 3 times, 10%, respectively. This result shows that most of the MBs in B_n frames have forward-predicted motion vector from R_i rather than backward-predicted motion vector from R_j. B_n Also, we can lead to effective extraction of moving object by using the vector space in the proposed method. It is because we are able to analyze each motion flow in all frames (including R_i, B_n) of MPEG sequence with the proposed algorithm, opposite to the conventional method Using our frame-type-independent framework, the number of object-detected frame from sequences with moving objects has increased over 50% as seen in Table I. This result is caused by explosive increase of the number of normalized motion vectors in B_n, R_i.

Fig. 6 represents the result after applying the proposed motion flow to object tracking algorithm. Let the center position of objects extracted from motion frame be denoted by P. Then, let overlapped times for the same P point extracted from moving sequence be denoted by M. Fig. 6-(a) shows when overlapped object center position, M is equal to 0, that is, moving object is tracked from all frames on MPEG domain, without respect to overlapped times of object center position. Also, Fig. 6-(b) and Fig. 6-(c) show the object trajectory in M=3 and M=5, respectively.

TABLE I. Comparison of the number of object detected frames.

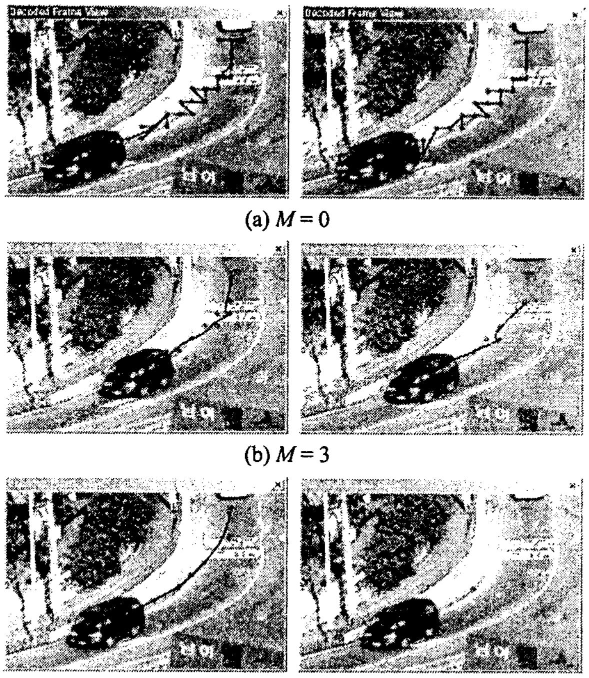
Sequence	Spatial Resolution	The number of total frame	Detected frame (conventional)	Detected frame (proposed)	Improved ratio (%)
Boat	336 x 208	83	18	28	35
Truck	352 x 240	298	54	120	55
Van	352 x 240	130	44	116	62





(a) The number of normalized motion vectors in B_n (b) The number of normalized motion vectors in R_i

Fig. 5. Comparison of the number of normalized motion vectors.



(a) $M = 0$ (b) $M = 3$ (c) $M = 5$

Fig. 6. Object tracking with overlapped object center position, M.
(Left-in the vector space by the proposed method, Right-in the vector space by the conventional method)

Fig. 6 illustrate that the more M increases to a definite level, the more moving objects are exactly tracked in the proposed vector space, but it is not true in the conventional vector space any more. Such performance by the proposed method shows the fact that the number of increased object-detected frames affects accurate tracking algorithm of moving objects, as seen in Table. I.

4. Conclusions

In this paper, we have proposed a new motion analysis algorithm using the normalization of motion vectors on MPEG compressed domain. Using the increase of the number of normalized vectors obtained from the proposed method, we can achieve a good result for extraction or tracking of moving objects by improving the ratio of object-detected frames in sequences. Our method will be well used not only in tracking moving objects, but in extracting global camera motion and feature information for video indexing or retrieval.

Acknowledgement

This research was supported in part by the Ministry of Education, Seoul, Korea, under the BK21 Project, and by the Ministry of Science and Technology, Seoul, Korea, under the NRL Project (2000-N-NL-01-C-285) and by the Institute of Information Technology Assessment (IITA), Taejon, Korea.

Reference

[1] Y. Nakajima, K. Ujihara and A. Yoneyama, "Universal scene change detection on MPEG-coded data domain," *in Proc. SPIE Visual Comm. and Image Proc.*, vol. 3024, pp. 992-1003, 1997.

[2] B.L. Yeo and B. Liu, "Rapid scene analysis on compressed video," *IEEE Trans. on Circuits and Systems for Video Technology*, vol. 5, no. 6, pp. 533-544, 1995.

[3] N.V. Patel and I.K. Sethi, "Video shot detection and characterization for video databases," *Pattern Recognition*, vol. 30, no. 4, pp. 583-592, 1997.

[4] H.L. Eng and K.K. Ma, "Bidirectional motion tracking for video indexing," *Multimedia Signal Processing*, pp. 153-158, 1999.

[5] J.G. Kim, H.S. Chang, J.W. Kim and H.M. Kim, "Efficient camera motion characterization for MPEG video indexing," *IEEE International Conference on Multimedia and Expo*, vol. 2, pp. 1171-1174, 2000.

[6] O. Sukmarg and K.R. Rao," Fast object detection and segmentation in MPEG compressed domain," *Proceedings of TENCON 2000*, vol. 3, pp. 364-368, 2000.

[7] A. Yoneyama, Y. Nakajima, H. Yanagihara and M. Sugano, "Moving object detection and identification from MPEG coded data," *Proc. of International Conference on Image Processing*, vol. 2, pp. 934-938, 1999.

[8] R. Milanese, F. Deguillaume and A. Jacot-Descombes, "Video segmentation and camera motion characterization using compressed data," *SPIE-Multimedia Storage and Archiving Systems II*, vol. 3229, pp. 23-31, 1997.

[9] V. Kobla and D. Doermann, "Compressed domain video indexing techniques using DCT and motion vector information in MPEG video," *in Proc. SPIE Conference on Storage and Retrieval for Image and Video Database*, vol. 3022, pages 200-210, 1997.