

Movement Detection and Tracking Using Video Frames

Josue Hernandez¹, Hiroshi Morita², Mariko Nakano-Miytake¹,
and Hector Perez-Meana¹

¹National Polytechnic Institute, Av. Santa 1000, 04430 Mexico D. F. Mexico

²The University of Electro-Communications, 182-8585, Tokyo, Japan
mariko@infinitum.com.mx, hmperezm@ipn.mx

Abstract. The use of image processing schemes as part of the security systems have been increasing, to detect, classify as well as to track object and human motion with a high precision. To this end several approaches have been proposed during the last decades using image processing techniques, because computer vision let us to manipulated digital image sequences to extract useful information contained in a video stream. In this paper we present a motion detection algorithm using the movement vectors estimation which are subsequently filtered to obtain better information about real motion into a given scenes. Experimental results show that the accuracy of proposed system.

Keywords: Motion Vectors, Movement Detection, Surveillance System, Surveillance system development.

1 Introduction

The advance of electronic and computer technologies increase the ability to perform video analysis to extract useful information from a video frames to carry out [1] motion detection and characterization [2], remote sensing, and pattern recognition, among others [3]-[5]. Pattern recognition is a research area that has been amply studied during the last years, it use information contained in video sequences; In many applications of pattern recognition, approaches with recognition capability are usually based on a corpus of data which is treated either in a holistic manner or which is partitioned by application of prior knowledge [5] like shape, velocity, direction, texture, magnitude, behavior and so on from different kind of objects around the area of interest.

Several approaches have been proposed to solve the problem of detecting and tracking motion during the last several years [6]-[16]. Some of them are describe in the following paragraphs [6]. Reference [7] proposes a motion detection approach based on the MPEG image compression algorithm in which the estimation of detection motion and the moving object direction only use the information contained in the MPEG motion vectors and the DC coefficients of the DCT directly extracted from the MPEG bit stream of the processed video. Evaluation results of this method shows that it can handle several situations where moving objects are present in the scene observed with a mobile camera. However, the

efficiency of the moving object detection depends on the quality of MPEG motion vectors [8-12]. This method also has a constraint in the context of content-based video browsing and indexing that should be taken into account when the MPEG encoder is selected. Reference [9] proposed an algorithm that intends to extend the capabilities to MPEG2 streams, allowing tracking objects selected by a user throughout the sequence. The way in which the tracking has been realized, is through exploitation of the motion information already present in the bit stream and generated by the decoder, another important feature of this scheme is that it is not necessary to decode the full sequence, if the user wants to access only part of the video stream [10]. This proposed scheme performs well in assisting the information insertion/retrieval process. However, no segmentation or filtering techniques are used for the extraction and tracking of the objects, since it relies exclusively on the motion information already provided by the MPEG encoder. Evaluation results show that the algorithm performs well though it is slightly dependent on the object shape. Hariharakrishnan et al [8] proposes an algorithm in which the tracking is achieved by predicting the object boundary using motion vectors, followed by contour update, using occlusion/disocclusion detection. An adaptive block-based approach has been used for estimating motion between consecutive frames. Here an efficient modulation scheme is used to control the gap between frames used for object tracking. The algorithm for detecting occlusion proceeds in two steps. First, covered regions are estimated from the displaced frame difference. Next these covered regions are classified into actual occlusions and false alarms using the motion characteristics. Disocclusion detection is also performed in a similar manner [8].

This paper proposes an algorithm for detection and tracking of movement of objects and persons based on a video sequence processing. Evaluation results show that proposed scheme provides a fairly good performance when required to detect relevant movements and tracking the motion of objects under analysis.

2 Proposed Movement and Tracking Detection System

The proposed system, firstly estimates the motion vectors, using an input video frames, which are then filtered to reduce distortion due to noise and deficient illumination. Next using the estimated motion vectors the movement trace is estimated to determine if it is a relevant or irrelevant motion. Finally if the movement is relevant, its tracking is carried out until the object left the restricted zone.

2.1 Motion Vector Estimation

The motion estimation is carried out dividing, firstly, the actual image at time, $t-1$, into a non-overlapping macro-blocks of 16×16 pixels. Next in the image frame at time t , the algorithm looks for the region that closely matches the macro-block under analysis [2]. Taking in account that the time difference in time between two consecutive images in a given frame is relatively small, only is necessary to carry out the

analysis in a region slightly larger than the given macro-block. Here the distance and direction that minimizes the criterion given by [5]

$$MAE(i, j) = \frac{1}{MN} \sum_{k=0}^{M-1} \sum_{l=0}^{N-1} |C(x+k, y+l) - R(x+i+k, y+j+l)| \quad (1)$$

Where $C(x+k,y+l)$ is the $(k,l)th$ píxel of macro-block (x,y) of the actual frame and $R(x+i+k,y+j+l)$ is the macro-block in the position $(x+i,y+j)$ in the reference frame, where $-p \leq i \leq p$ $y - p \leq j \leq p$. Then the vector motion magnitude and direction is obtained from to actual position (x,y) in the actual frame, to the position $(x+i,y+j)$ in which the MAE (i,j) is minimum. This process is illustrated in Fig. 1.

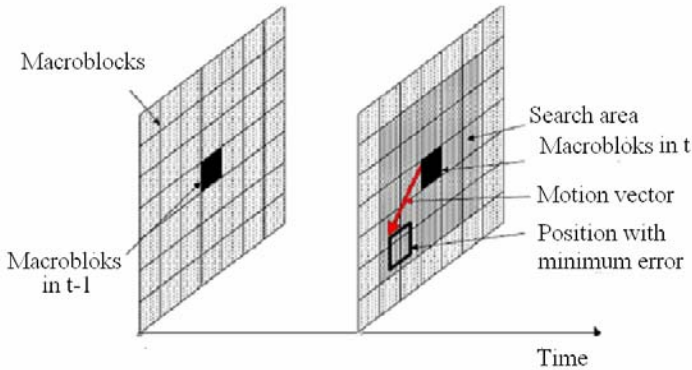


Fig. 1. Motion vector estimation process

Several methods have been proposed for motion vector estimation, which are based on the MAE (Mean Absolute error) minimization (1), all of them providing similar results although their computational complexity presents important differences. Among the Hierarchical estimation method presents the less computational complexity and then is one of the most widely used motion vector estimation method. Here to reduce the computational complexity, the image is decimated by 2 using low pass filters, reducing in such way the image size. Thus the position (x,y) in the original image becomes the position $(x/2,y/2)$ in the image corresponding to the first decomposition level. Subsequently a second decomposition level is applied using a low pass filter, such that the original (x,y) point becomes the point $(x/4,y/4)$ in the second decomposition level image [5].

After the second decomposition level is performed, the motion vector estimation starts in the second decomposition level, in which the image size is 1/16 of the original image size, with a macro-block of size 4x4. Assuming that the point (u_2,v_2) corresponds to the minimum MAE of the actual macro-block, the search of the motion vector in the decomposition level 1 is carried out in the macro-blocks of 8x8 whose search start in $(x/2+2u_2,y/2+2v_2)$, with a search range of [-1,1] around the origin píxel. Finally the motion vector estimation in the level decomposition 0 is carried out with macro-blocks of size 16x16, with search starting at point $(x+2u_1,y+2v_1)$ with a search range equal to [-1,1], around the origin píxel. Here

(u_1, v_1) corresponds to the minimum MAE at the level decomposition 1. These motion vectors consist of horizontal and vertical components which are used to estimate the motion vector magnitude and angle. Finally using these data it is possible to determine if the movement is relevant or not. Figure 2 illustrates this process.

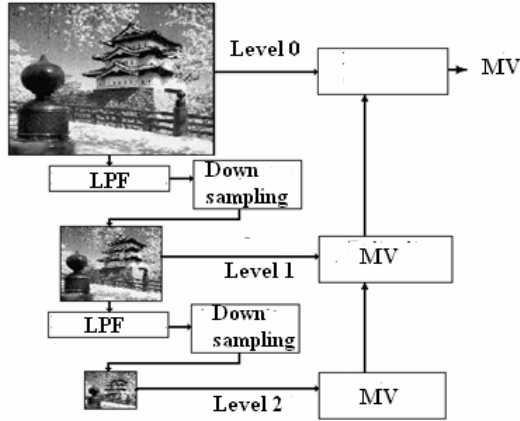


Fig. 2. Hierarchical motion vector estimation algorithm

Several times the estimated motion vectors only represent abrupt illumination changes in the image or small vibrations of the camera used to capture the image sequence, resulting in wrong movement estimation if these distortions are not cancelled or at least reduced. To this end, the estimated motion vectors are filtered using a one-dimensional median filter as shown in Fig. 3, before the overall movement estimation. This allows to cancel most motion vectors that do not provide useful information, such as those due to illumination changes, background movement, etc. This process is shown in Fig. 3.

Because the amount of noisy motion vectors is relatively small in comparison with the correctly estimated motion vectors, as well as due to the fact that the noisy vectors are continuous or similar among them, as happens with the actual motion vectors, the noisy motion vectors can be easily distinguished from the remaining ones. On the other hand, because the motion vectors are estimated one by one it is not necessary to create a temporal register, to be used during the filtering process [5]. Thus after the noisy vectors have been eliminated, the information generated by objects moving in non-relevant directions is cancelled, based on the fact that the movement direction is different to that of the security zone. This fact allows the proposed algorithm to measure the relevance level, according to their position in the scene. Thus a motion vector is eliminated if its position does not change after several video sequences or if the motion direction is opposite the restricted zone, as shown in Fig. 4.

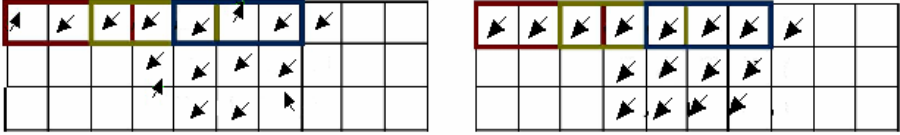


Fig. 3. Motion vectors filtering process

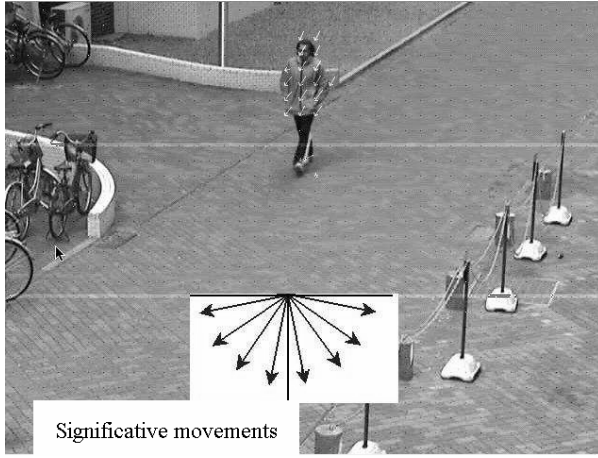


Fig. 4. Relevant movement detection in the direction of the restricted zone

2.2 Trajectory Estimation

Once the motion vectors have been estimated, it is necessary to estimate other important parameters such as: the speed, direction and movement tracking of the moving object present in the video sequence. To this end firstly it can be used the information provided by the motion vector to estimate the movement angle to classify the movement in, either, relevant or not relevant. To this end, we can take in account that a relevant movement is that in which the people enter or intend to enter into the restricted zone. This fact implies that, according with the video camera position, the motion vectors angles must be between 200 and 340 degrees. This fact takes in account that a movement is not relevant if it takes place inside the restricted zone, or if its direction is from inside to outside the restricted zone.

Other important factor that must be considered to obtain an accurate estimation of the motion vectors is the camera position, because it allows to obtain constant motion vectors, with similar magnitude during all trajectory, avoiding detection errors. This fact also allows that a person magnitude may be represented using among 9 to 13 motion vectors. A correct distance also allows adding some divisions to measure the movement importance and be able to track the position of a given object during a determined time; as well as to add several counters to determine the number of



Fig. 5. Tracking ability of proposed algorithm

persons that are generating relevant movements, as well as the direction of these movements.

2.3 Movement Tracking

Other important issue is the tracking of the generated relevant movement, which means that not only is important to determine if the movement a relevant movement has been produced, but also to have the ability to track the movement of the person generating it until the go out of the restricted zone. The min difficulty to track the object movement using a sequence of successive images is the object localization in the image sequences, especially when the movement of them is relatively faster than the images rate. For these reason, the tracking systems usually uses moving models that intend to predict the variations of the image due to possible movements present in the object trajectory. Although these methods provide fairly good results, they may be computationally complex. Thus to reduce the computational complexity, the object is divided in $N \times M$ non overlap blocks. Next using the motion vector estimation described above, the movement of each corner is estimated. Once the motion vectors are obtained, the tracking of object motion in each frame is obtained as the resultant vector of all individual motion vectors in such frame. Finally to obtain a smoother trajectory the resulting vector in the previous step is filtered using a low pass filter.

3 Evaluation Results

The proposed system was evaluated using computer simulation using a Power Mac G4, with a data bus speed of 167MHz, a CPU of 1.25GHz ad 1.25GB of RAM memory. The video sequences used to evaluate the proposed algorithm were previously recorded with a resolution of 640x480 pixels per image. The obtained results show that the proposed algorithm is able to accurately detect the relevant movements and correctly tract the person u object motion.



Fig. 6. Detection performance of proposed algorithm, here in the fourth and fifth image there are a relevant movement

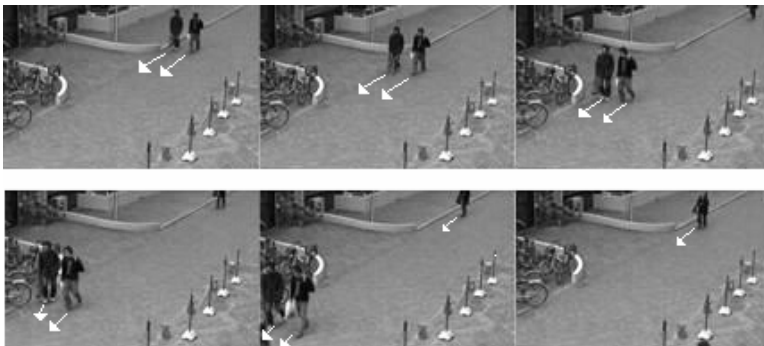


Fig. 7. Two persons moving to the restricted zone

4 Conclusions

This paper proposed an automatic movement detection system, using a video sequence, based on the motion vector estimation, which are filtered to eliminate the noisy and distorted vectors, due to illumination variations and background movement. Proposed algorithm is also to discriminate between relevant and non relevant movements which allows only take in account the movements whose direction is from outside to inside the restricted zone. Using also a motion vector estimation the algorithm is also able to tract the trajectory of a given person whose movement is from inside to outside of the restricted zone.

Acknowledgements

The authors thanks the National Science and Technology Council, to The Science and Technology Institute of the D. F. Mexico, and to the National Polytechnic Institute of Mexico for the financial support provide during the realization of this research.

References

1. Jones, B.T.: Low-Cost Outdoor Video Motion and Non-Motion Detection. *Processing of Security Technology*, 376–380 (1995)
2. Paladino, V.: Introduction of Video Compression Under the Standard MPEG-2. *The Institute of Electronic Engineer, Spain*, 3–24 (2005)
3. Richardson, I.E.: H.264 and MPEG-4 Video Compression, Video Coding for the Next-Generation Multimedia, pp. 27–41. Wiley, UK (2004)
4. Watkinson, J.: *The MPEG Handbook: MPEG-1, MPEG-2, MPEG-4*. Focal Press (2001)
5. Tudor, P.N.: MPEG-2 Video Compression: Tutorial. *Journal of Electronics and Communication Engineering*, 1-5 (December 1995)
6. Zhang, Z.: Mining Surveillance Video for Independent Motion Detection. In: *IEEE International Conference on Data Mining*, pp. 741–744 (2005)
7. Favalli, L., Mecocci, A., Moschetti, F.: Object Tracking For Retrieval in MPEG2. *IEEE, Trans. on Circuit and Syst. for Video Technology*, 427–432 (2000)
8. Hariharakrishnan, K., Schonfeld, D., Raffy, P., Yassa, F.: Video Tracking Using Block Matching. In: *IEEE, International Conference on Image Processing*, pp. 945–948 (2003)
9. Yoneyama, A., Nakajima, Y., Yanagihara, H., Sugano, M.: Moving Object Detection from MPEG Video Stream. *Systems and Computers in Japan* 30(13), 1–11 (1999)
10. Avidan, S.: Support Vector Tracking. *IEEE* 26(8), 1064–1071 (2004)
11. Nguyen, H., Smeulders, A.: Fast Occluded Object Tracking by a Robust Appearance Filter. *IEEE Trans. on Image Processing* 26(8), 1099–1103 (2004)
12. Lin, M., Tomasi, C.: Surfaces with Occlusions from Layered Stereo. *IEEE* 26(8), 1073–1098 (2004)
13. Sebastian, T., Klein, P., Kimia, B.: Recognition of Shapes by Editing Their Sock Graphs. *IEEE Trans. on Image Processing* 26(5), 550–554 (2004)
14. Coding Audiovisual Objects Part 2, International Standard Organization / Int. Electronics Communications (ISO/IEC), 14496
15. Hernández García, J., Pérez-Meana, H., Nakano Miyatake, M.: Video Motion Detection using the Algorithm of Discrimination and Hamming Distance. In: Ramos, F.F., Larios Rosillo, V., Unger, H. (eds.) *ISSADS 2005. LNCS*, vol. 3563, pp. 321–330. Springer, Heidelberg (2005)
16. Gryn, J.M., Wlides, R., Tsotsos, J.K.: Detecting Motion Patterns via Directional Maps with Applications to Surveillance. *Computer Vision and Image Understanding* 113, 291–307 (2009)