

Motion Detection and Clustering Using PCA and NN in Color Image Sequence

Mourad Moussa*¹, Nesrine Bdioui², Ali Douik³

¹Department of informatics Science Faculty of Gafsa, University of Gafsa, Tunisia

^{2,3}Department of applied informatics National engineering school of Sousse, University of Sousse, Tunisia

*Corresponding author, e-mail: Mourad.enim@yahoo.fr

Abstract

This paper presents a motion detection method with the use of the Principal Component Analysis. This method is able to detect and track moving objects in a sequence of images. The tested sequence is segmented within the meaning of movement. In this paper, the concept of extracting significant information from a large number of data is adopted to provide an effective method for tracking moving objects on the video image. The principal components are different in term of getting significant information, the nature of motion (the nature of information) is responsible of this difference, the algorithm in this paper distinguish the motion nature and choose the appropriate components to give a best segmentation.

Keywords: neural networks, PCA, motion analysis, motion detection, clustering

Copyright © 2018 Universitas Ahmad Dahlan. All rights reserved.

1. Introduction

The detection and tracking of objects in an images sequence is one of the issues that arises in many image processing applications such as traffic monitoring systems [1] or surveillance systems [2]. The difficulty of such application is accentuated in environments without constraints where the monitoring system should be adapted to the high variability of the objects as well as motion detection problems. Motion detection plays a fundamental role in the tracking algorithms, in so far as almost all begin with motion detection. Since PCA (Principal Component Analysis) is generally used for extracting significant information from a large number of dataset, this study adopts the concept of PCA to provide an effective method for tracking the moving objects on the video image. The proposed algorithm allows the motion detection; areas and moving objects, with the use of principal component analysis. It is a region segmentation within the meaning of movement of an image sequence. Sometimes, the movement is so slow that the algorithm of detection cannot detect this motion. The proposed motion detection algorithm is a recursive algorithm that detects each time for ten frames, the area where the pixels are in motion, whatever the nature of motion, fast or slow.

The problem is to segment the motion from sequences of images in the case of a stationary camera. This can be solved by threshold difference images [3]. However, the choice of this threshold results from a post hoc analysis of the results by an outside observer. In the case of a very noisy sequence, the same threshold is no longer appropriate. Indeed, several methods have been proposed for segmenting a moving object in a sequence of images including segmentation by subtracting the background. As a set of static pixels, the background can be removed from a given image in order to extract the pixels in movement. The variation of the different techniques depends mainly on the type of used background model [4-10]. The background image without moving objects [11-12] is needed to be used as reference information. But most of the time, it is not possible to obtain a sequence of images without moving objects, for example in traffic monitoring. In [13-16], methods for generating a background model from a sequence of images with moving objects are presented. Another approach is to make the motion detection problem, in a Bayesian framework, in terms of energy minimization. This energy will then be minimized by the algorithm of minimum cut calculation "graph cuts" [17].

In this paper, first, a brief overview of previous related approaches is presented. Then, we propose a motion segmentation algorithm that successfully detect fast and slow motions. Finally, performance evaluation, analysis, and discussion are carried out.

2. Our Approach

Because the PCA is generally used for extracting significant information from a large number of dataset, therefore, this approach uses the concept of PCA to detect the trace of the moving object by converting the video image domain to the PC domain. Our algorithm is a recursive algorithm which is designed to define each time the moving area using the principal component analysis on ten images. For a sequence of several frames, the recursive algorithm takes each time subsequences of 10 frames and makes a call to the motion detection algorithm using the PCA to segment the movement of the scene, in order to be a real time algorithm.

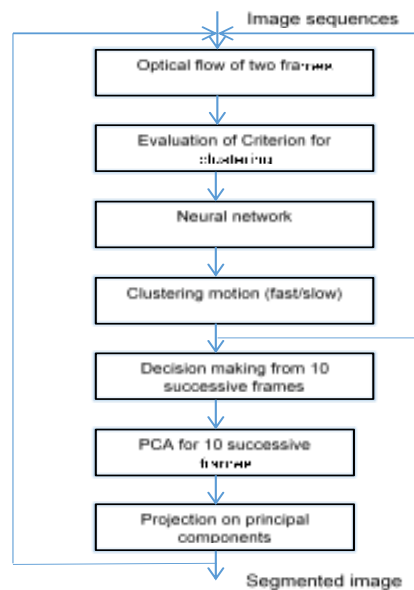


Figure 1. Block diagram of the proposed algorithm

The segmentation of the movement depends on its nature, fast moving or slow. The proposed algorithm should then do this distinction. For this, we use criteria for distinguishing extracted from optical flow of movement. The fast/slow classification is done by using a network of neurons apprentices. The block diagram of the algorithm is presented by Figure 1.

2.1. Optical Flow

The optical flow is a visual displacement field that helps to explain variations in a moving image in terms of moving image points. Calculating an apparent movement between two images corresponding to the estimation of parameters of a transformation affecting the image points where we associate for each image pixel (x, y, t) , a vector (v_x^t, v_y^t) representing the apparent speed of the pixel (x, y) at time t . the vector (v_x^t, v_y^t) represents the projection on the image plane of the velocity vector (v_x^t, v_y^t, v_z^t) of objects in the scene with respect to image reference (o, x, y, z) . It is calculated from the temporal variation of the image function $I(x, y, t)$.

We use the optical flow between two images representing motion in our algorithm and determine the velocity vector (v_x^t, v_y^t) for each pixel of the image we then define two matrices u

for speeds v'_x and v for speeds v'_y for all pixels (x, y) belonging to image. From the speeds matrices u and v , we measure criteria that we use it soon to distinguish the nature of motion slow or fast. Figure 2 shows the optical flow between two images with movement. The field of vectors corresponds to the displacement of the pixels of the image. An overview of different methods to calculate the optical flow is presented in [18].

2.2. Classification Criterion

Using the matrices velocities u and v , we calculate the following measures: The sum of absolute differences, the Euclidean distance, the Hausdorff distance and difference of energies. These measures are used as criteria for the classification of slow or fast motion. They are brought in a network of neurons that will be our tool for classifying movements of the scene.



Figure 2. Vector field of optical flow of variation of image pixels; (a) and (b) Two images presenting motion, (c) First image presenting the vector field of optical flow of variation.

Indeed, the amounts thus calculated measure the movement of image pixels. The movement is considered fast for a certain value for each of the calculated measures. The sum of absolute differences is determined by calculating the absolute difference of speeds for each pixel, these differences are summed to create a measure of movement in an image.

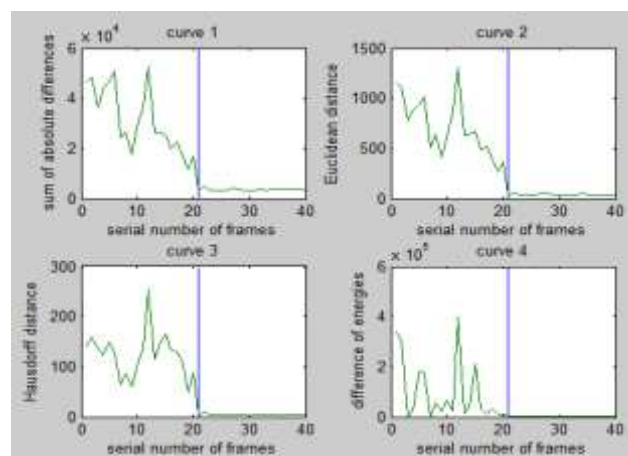


Figure 3. Classification criteria measured for a series of quick movements followed by a series of slow movements

Euclidean distance is the known distance that defines a Cartesian space. The Hausdorff distance is also a topological tool that measures the distance of two subsets of a metric space underneath. As to the difference of energy is the difference of the sums of the squares of each matrix velocity elements. We choose twenty frames contain fast motions, and the same for the slow motions. Figure 3 shows the measurements calculated for the series of quick movements followed by the series of slow movements. We note that rapid movements take the great values of the measurements compared to the slow movements. So we can get the critical values that distinguish the nature of movement. This task is performed by a neural network apprentice introduced into our algorithm, which are able thanks to this criterion to classify a motion as fast or slow.

3.3. Neural Network

The network has nine neurons and one hidden layer. The network has as inputs the classification criteria measured previously it classifies according to these inputs a slow or fast motion. Figure 4 shows our neural network in which the output is a two dimensional variable that we convert it to a Boolean variable; 1 indicate that the motion is fast and 0 indicate that it is slow. We train the network with a training set of forty pairs of slow and fast images. For twenty pairs of images taken for validation, classification gives nineteen-ranked movements, that to say an error rate of 0.05. As it is shown in Figure 5 the best validation performance of the neural network, when it is training, is obtained at epoch 10 where the mean squared error $8.8191e-08$.

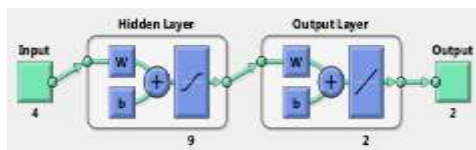


Figure 4. The neural network

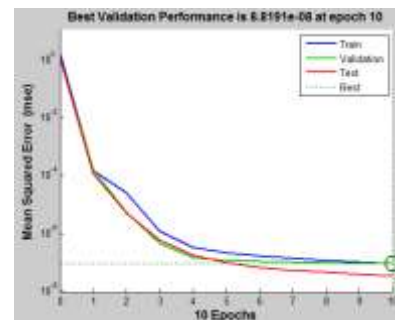


Figure 5. Performance plot of the neural network

3.4. Principal Component Analysis

The video data are initially represented by a function defined in a three dimensions space: two spatial dimensions (x, y) and a time dimension (t). At each point in this space is assigned a gray level (or a color component vector) at a point (x, y) at time t . The various semantic entities (background, moving objects) are then a subset of points in this space. To identify them, they should be aggregated into classes of points with common characteristics. It goes that the number of points to consider is very important, especially if we want to consider more than two frames to detect moving objects. This is why, the approach of building a model of background is so common: the only points to be considered are those of the current frame, while the model of the background is supposed to summarize all past observations. We believe it is best to keep a less synthetic knowledge of the past because the relevant information to extract is not always the same. We therefore intend to select a representation space more suited to the sequence itself rather than each of the frames and which allows to take into account the movement without changing the initial information. Also in view of the use of data analysis techniques, the sequence is no longer considered as a function but as a set of individuals: the pixels that we see when we look at the sequence. To avoid having to make a detailed analysis, it is not the objects that we follow but it is a fixed position which is considered on the surface of the image. From each pixel we will hold multiple values of the gray levels over time. Each pixel is an individual characterized by a set of parameters. Individuals are marked in a p -dimensional space.

Obviously to stay within reasonable processing times, almost real-time, we must make a reduction in the mass of information. To preserve the information that will allow the best to discriminate the points and build classes, we chose to use a principal components analysis. We consider that we want to treat a set of 10 observations, which are realizations of a random p -dimensional variable, which p is the size of an image of the scene, it is the number of pixels in the image. These observations are arranged in an array of data X , which contain p rows and 10 columns, each column is formed by an observation. The PCA is a statistical technique which aims to simplify a dataset by expressing it in a new coordinate system such that the greatest variances are observed on the first coordinates. This reduces the dimensionality of the search space, keeping only the first dimensions of the projection space obtained [19]. We consider the data matrix X . If we call C the variance-covariance matrix associated with X , then the main axis directions are given by the eigenvectors of C as in Equation (1).

$$C = \frac{1}{n} X^T X \quad (1)$$

The axis on which it is observed the larger variance is defined by all eigenvector associated with the largest eigenvalue in absolute value. Figure 6 represents the steps of the principal component analysis. The proposed motion detection algorithm is to create a data matrix from a set of images and apply the PCA [20] to release a basis of eigenvectors which we can keep only those who explain the better the variance of the image database.

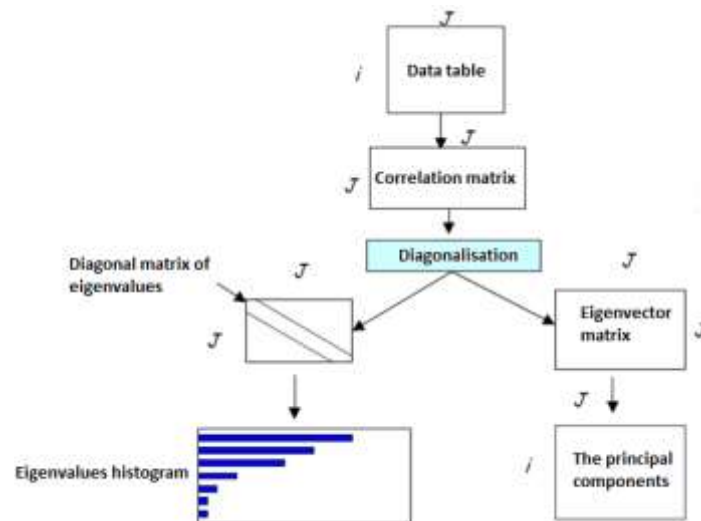


Figure 6. Steps of the principal component analysis

Each image is then projected into the space of reduced dimension. We consider a sub-sequence of ten frames, where each frame is associated with a size of the data representation space, wherein we apply PCA in order to represent the data in a low dimensional space wherein the points representing a coherent movement are close. The method must be as insensitive as possible to the various conditions in which acquisitions are made: we focus more on the gray level variations rather than the grayscale of the pixel itself. We can, from the start, delete a dimension of the data representation space by choosing to fill the data matrix with time derivatives at any point, then we are in a space of nine dimensions [21]. (Y is data matrix represented in this space). Moving areas appear clearly when we project the sequence of the 10 frames on the first three principal axes. The difference between a static area and a moving area is accentuated on these axes. According to the eigenvalues histogram and also to the principal components, we find that for a slow motion, due to the small quantities of information

comparing to a fast motion, we can consider only the two first principal axes, these axes can show the moving areas. Figure 7 compares two eigenvalues histograms for a slow and fast motion. The x axis represents the nine axes and the y axis represents the eigenvalues. As we can see, the information is concentrated almost in the first component for a slow motion, so the variance is greater for this component comparing to the others. For a fast motion, it is true the greatest variance is taken by the first component but the first others have a near values of variance. So we superpose the three components to build the mask of segmentation, in which we project our images, for a fast motion and just two for a slow motion.

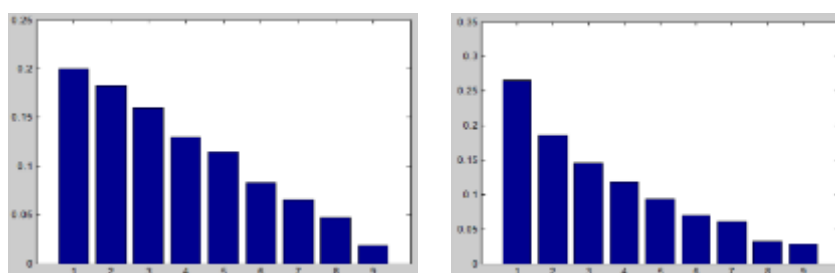


Figure 7. Eigenvalues histograms of a slow motion frames (on the right) and a fast motion frames (on the left)

4. Experimental results and discussion

The experiment is designed to test the proposed method for detecting and tracking the objects moving (the man, the chair, the phone) in conference room in the image sequence, whatever the motion is fast or slow. The tested sequence, contain 131 frames in which a man is entering to the scene, sitting on the chair, calling in the phone and then leaving the scene. The sequence is segmented and the movements are detected. The result of the motion extraction is shown in the Figure 8. The x axis represents the serial number of the frames and the y axis indicates the number of motion pixels.

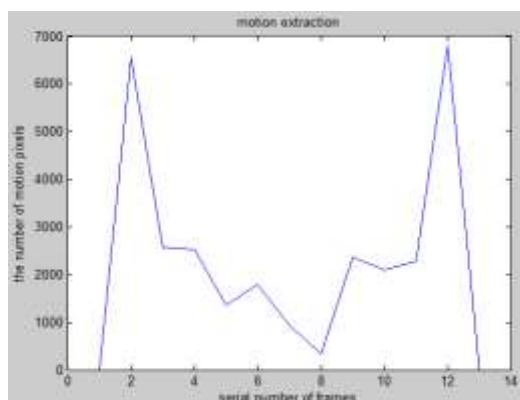


Figure 8. Motion pixels extraction

Two significant peaks can be evidently observed in the graphic, this peaks correspond to the entering and the leaving of the man in which we have a large area of motion pixels. When the man is talking on the phone the motion is very slow and we have a small number of motion pixels. The results of experiment indicate that the proposed method can track the moving object successfully whatever the motion is slow or fast. Although the noise will cause some pseudo moving events in the video images, the proposed method still can track the interested target and

filter out the noise by checking the size of moving regions. We represent in Figure 9 (a) and (c) respectively fast and slow subsequences motion.



(a)



(b)



(c)



(d)

Figure 9. (a) Fast motion subsequence (b) Principal components of fast motion (c) Slow motion subsequence (d) Principal components of slow motion

5. Conclusion

This paper presents a new method for motion detection and tracking using the concept of extraction significant information, the Principal Component Analysis. The information provided by principal components depends on the type of motion. This method distinguishes the nature of

motion and uses Neural Network to classify it, and according to this it uses the appropriate principal components to segment motion. The proposed method is tested on images sequences and gives a satisfying result.

References

- [1] Zhang R, Ge P, Zhou X, Jiang T, Wang R. An Method for Vehicle-Flow Detection and Tracking in Real-Time Based on Gaussian Mixture Distribution. *Advances in Mechanical Engineering*; 2013: ID 861321.
- [2] Haibin W, Jianfeng H, Xiaoning Y, Jinhua Y, Sumei H. A Robot Collision Avoidance Method Using Kinect and Global Vision. *TELKOMNIKA (Telecommunication, Computing, Electronics and Control)*. 2017; 15(1): 4-17.
- [3] Jain R, Nagel H-H. On the analysis of accumulative difference pictures from image sequence of real word scenes. *IEEE Transaction. Pattern Anal. Machine Intellelligent*. 1979; 1(2): 206-214.
- [4] Brutzer S, Höferlin B, Heidemann G. *Evaluation of background subtraction techniques for video surveillance*. IEEE conf. comput. Vision pattern recognit. 2011; 1937-1944.
- [5] Haritaoglu I, Harwood D, Davis L.S. Real-time surveillance of people and their activities. *IEEE Trans. Pattern Anal. Machine Intell*. 2000; 22(8): 809–830.
- [6] Piccardi M, Jan T. *Mean-shift background image modeling*. Int. Conf. on Image Processing (ICIP'04).
- [7] Djelouah A, Franco J-S, Boyer E, Le Clerc F, Pérez P. *Modélisation probabiliste pour la segmentation multi-vue*. France. 2013; Congrès des jeunes chercheurs en vision par ordinateur (ORASIS).
- [8] Zhong J, Sclaroff S. *Segmenting Foreground Objects from A Dynamic Textured Background Via A Robust Kalman Filter*. Int. Conf. on Computer Vision. Beijing. 2003: 44–50.
- [9] Ming Y, Jiang J, Ming J. *Background Modeling and Subtraction Using a Local-Linear-Dependence-Based Cauchy Statistical Model*. Proc. VIIth Digital Image Computing: Techniques and Applications. Sydney. 2003; 469-478.
- [10] Elgammal A, Harwood D, Davis L. *Non-parametric Model for Background Subtraction*. 6ème European conference on Computer Vision (ECCV 2000), Dublin Ireland, 2000.
- [11] Kamal S, Cherrad B, Kobzili E, Chouireb F. A Real-Time Implementation of Moving Object Action Recognition System Based on Motion Analysis. *Indonesian Journal of Electrical Engineering and Informatics (IJEEI)*. 2017; 5(1): 44-58.
- [12] Rima T, Indah AS, Yonathan FH, Arik K, Ari K. Double Difference Motion Detection and Its Application for Madura Batik Virtual Fitting Room. *TELKOMNIKA (Telecommunication, Computing, Electronics and Control)*. 2015; 13(4): 1446-1455.
- [13] Sidibé D, Strauss O, Puech W. *Automatic Background Generation from a sequence of Images Based on Robust Mode Estimation*. Proc. SPIE 7250 SPIE-IS&T Electronic Imaging, Digital Photography. V, San Jose, California, US. 2009.
- [14] Cho SH, Kang HB. *Panoramic Background Generation using Mean-shift in Moving Camera Environment*. Int. Conf. on Image Processing, Computer Vision, and Pattern Recognition (IPCV). 2011: 829-835.
- [15] Zhang R, Gong W, Yaworski A, Greenspan M. *Nonparametric on-line background generation for surveillance video*. Pattern Recognition (ICPR). Tsukuba. 2012.
- [16] Abu PA, Chu VS, Fernandez P. A Monte-Carlo-based Algorithm for Background Generation. *Philippine Information Technology Journal*. 2013; 6(1): 4-10.
- [17] Zhang Guang-hua, Xiong Zhong-yang, Li Kuan, Xing Chang-yuan, Xia Shu-yin. A Novel Image Segmentation Algorithm Based on Graph Cut Optimization Problem. *TELKOMNIKA (Telecommunication, Computing, Electronics and Control)*. 2015; 13(4): 1337-1342.
- [18] Jahangir Alam SM, Hu G. Mosquito Tracking by Image Segmentation of Optical Flow Field. *TELKOMNIKA Indonesian Journal of Electrical Engineering*. 2014; 12(11): 7798-7807.
- [19] Barron JL, Fleet DJ, Beauchemin SS. Performance of Optical Flow Techniques. *Int. J. Computer Vision*. 1994; 12(1): 43-77.
- [20] Rima Tri W, Fitri D. Efficient Kernel-based Two-Dimensional Principal Component Analysis for Smile Stages Recognition. *TELKOMNIKA (Telecommunication, Computing, Electronics and Control)*. 2012; 10(1): 113-118.
- [21] Li Q, Yuan Y, Chen J. Segmentation for Fabric Weave Pattern using Empirical Mode Decomposition based Histogram. *Indonesian Journal of Electrical Engineering and Computer Science (IJEECS)*. 2013; 11(4): 1996-2001.