International Academic Journal of Science and Engineering

International Academic Journal of Science and Engineering
Vol. 4, No. 1, 2017, pp.153-162.
ISSN 2279-0039

International Academic Journal of Science and Engineering

www.iaiest.com

Tracking Objects in Video Frames with Moving Camera

M. Javaherian[a], E. Yakhtifard[b], B. Abedi Ravan[c*]

a. Department of Physics, University of Zanjan, Zanjan, Iran.
b. Faculty of engineering, Department of Computer, Islamic Azad University Zanjan Branch, Zanjan, Iran.
c. Faculty of Basic Sciences, Shahid Sattari Aeronautical University of Science and Technology, Tehran, Iran.

Abstract

Image segmentation and tracking objects are essential in image processing used in segregating an image into regions of interest. Different kinds of methods have been proposed for image segmentation such as region growing approaches, watershed, and mean shift procedure. One of the challenges in image analysis, especially in tracking, is working on images taken by moving camera. The background in a moving camera constantly changes, and detecting moving object makes tracking very difficult. In this paper, the concept of edge detection, the histogram of simple and complex images, and similarity between sequences of frames are exploited to facilitate the process of segmentation. Comparing our hybrid algorithm with other methods reveals that our method shows less detection errors than the others with lower temporal expenses in performance.

Keywords: Image segmentation, Hybrid algorithm, Edge detection, Histograms, Boundary displacement, Separation of background, Simple and complex background.

1. Introduction

Capability of recognizing moving objects from a sequence of video frames is one of the crucial challenges in the field of many vision systems which is applicable in video surveillance, target detection and tracking objects *in the military sphere*, human-machine interface, and traffic monitoring [1]. In the field of segmentation approaches, three kinds of algorithms were proposed which are as follows: region-based methods, edge-based methods, and hybrid approaches based on the two mentioned algorithms. To detect moving objects real-time nonparametric modeling approach is introduced which works based on a General Purpose Graphics Processing Unit (GPGPU). The quality of obtained results highlights that this method is designed to work in a variety of conditions [2]. However, these strategies increase the computational cost and lower the convergence rate of real-time processing. On the other hand, these kinds of algorithms make some approximations that influence the efficiency of detection. The accuracy of recognition can be improved by computing bandwidth of the kernels in each step and updating the procedure [3]. In other methods, in order to detect multiple moving objects, a model has been proposed to work based on background subtraction and then, tracking targets in consecutive video frames according to feature based approach [1]. In the field of studying tracking objects from surveillance videos, background modeling is a usual strategy that increases the proficiency rate [4]. To carry out the model background, three parameters were introduced: minimum intensity, maximum intensity, and maximum intensity difference in a sequence of frames [5]. Pixels with intensities greater than a criterion, which is determined by maximum intensity difference, is grouped as foreground pixel. In other scientific articles, to analyze and model background, a single Gaussian model (SGM) [6] and/or a mixture of Gaussian distributions are employed [7]. To model the background and foreground of each frame and track moving objects, Elgammal et al. (2002) proposed a method to estimate nonparametric kernel density [8]. In other applications, Hidden Markov Models (HMMs) [9], or combining Markov–Gauss model (MGM) with non-local means [10] are used to model the pixel intensity. Sometimes, working on the background-likelihood model using background subtraction method is very powerful in image segmentation which is rooted in co-occurrence of variations [11]. In recent studies for background modeling, a three-component system is proposed including pixel level, region level, and frame level [12]. To find changes between two frames, a method is developed to detect objects by integrating intensity and texture difference [4, 13]. Choosing the type of segmentation algorithms for video frames depends on several parameters such as fixed or moving camera, and features of target. When the camera is fixed, there are no significant changes in background. However, the background undergoes changes in a fast manner when we deal

with moving the camera. Similarity between consecutive video frames can be employed to simplify the segmentation process. We follow an approach that says the slight differences occurred at the successive frames can be related to boundary displacement components. So, on the basis of edge detection algorithm, the boundary of objects is determined, and image is segregated into regions.
Our hybrid segmentation algorithm is designed to detect moving target in the case of moving camera. In the first step, the edge detection method using Sobel's algorithm [14] is employed. Subtracting dilated image from eroded one in consecutive frames is processing steps to delineate the boundaries of regions. The histograms of simple and complex images are employed to threshold and then extract the regions of interest from the background. When the camera is not fixed and the image consists of the complex background, the registration method is applied on image to align the image with the next one. The Flowchart of our hybrid algorithm is presented in Figure 4.

This article is organized as follows: Sections 2-6 describe our hybrid algorithm in different cases of encountering with simple and complex backgrounds. The results are presented in Section 7 and they are compared with the results of the other methods. Finally, the conclusions are given in Section 8.

2. Edge detection

In our hybrid algorithm, to work with each kind of image, the image is converted to gray-level image. To enhance the quality of boundaries and adjust the image contrast, the sharpening filter is applied on the image. Using Sobel edge detection procedure, boundaries of objects are specified [15]. The result of applying Sobel edge detection on gray-level image is represented in Figure 1.



Figure1. The result of Sobel edge detection.

The output image should include the continuous boundaries surrounding the regions of interest. Edges or lines may be broken at some points. To adjoin boundary pixels in the blank space and construct connectivity, the bridge morphological function is employed to enclose regions. Now, pixels within the enclosed regions (objects) determined by edges are segregated from the background.



Figure2. The eroded image (a), and the dilated image (b).

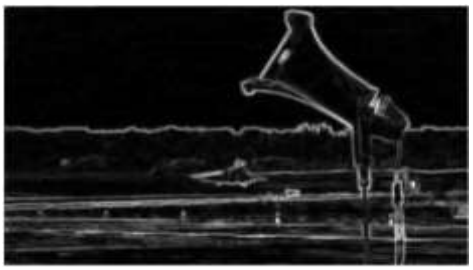


Figure 3. The result of subtracting the eroded image from the previous dilated one in the gray scale.

The dilation (expanding the features appeared in binary image) and erosion (sharpening lines and boundaries) are procedures (see Figure 2) which are combined to apply in a sequence of frames including edges.

Subtracting eroded gray level image from previous dilated gray level image helps to determine the differences between two successive frames [1] as shown in Figure 3. The subtracted image plus the result of the edge image gives the resultant image with determined boundaries.

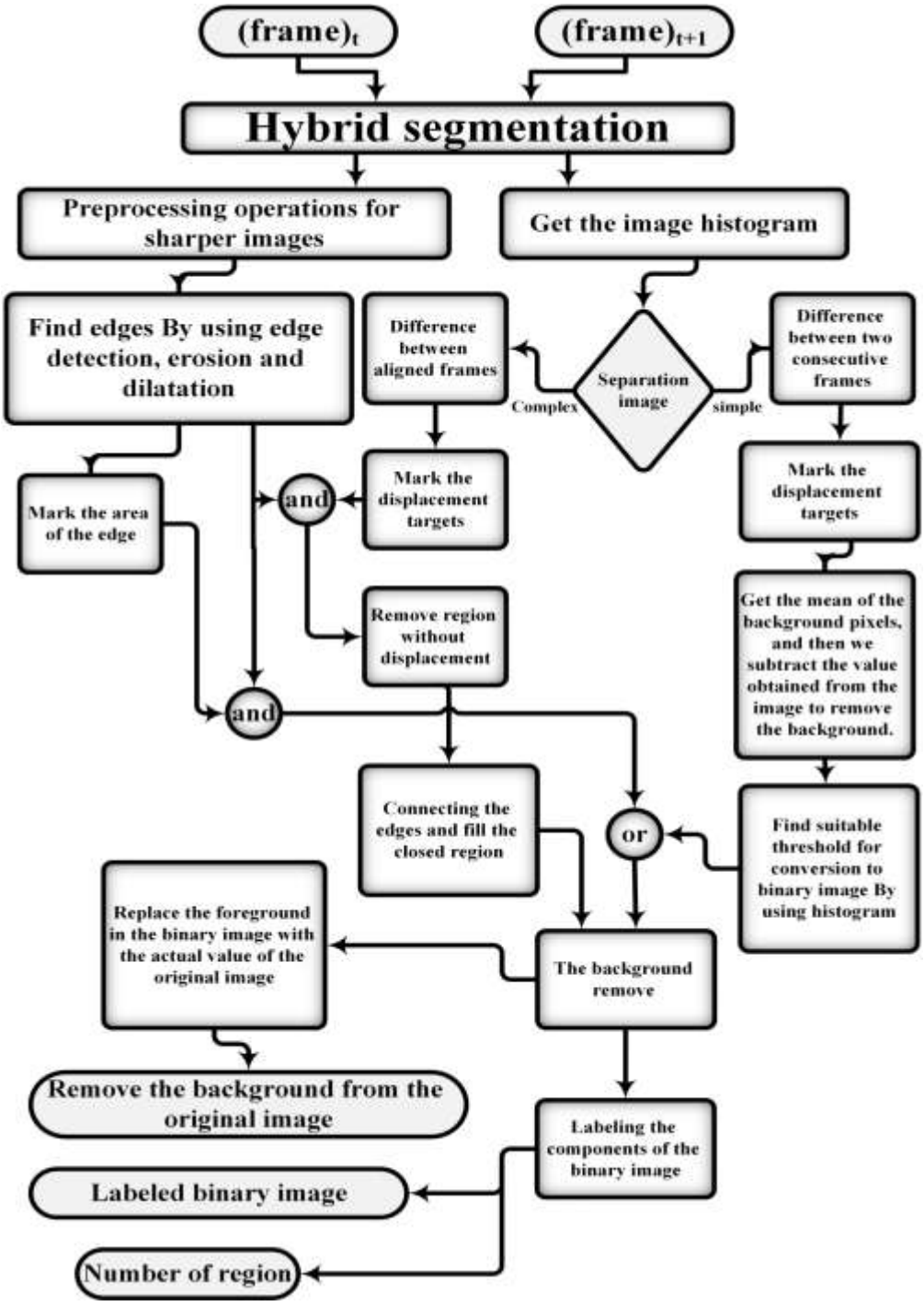


Figure 4. Flowchart of our hybrid segmentation algorithm

3. Classifying simple and complex images

Subtraction process applied on successive frames shows the displacement components of the image. If the camera is fixed and the background is unchanged, the first frame is used as the reference frame for subtracting eroded image from dilated one. When we deal with moving camera, the subtraction between two consecutive frames is computed. The subtraction procedure depends on the type of background. Backgrounds in images are divided into two categories: backgrounds of simple images, and backgrounds of complex images. Images with lower details are considered as simple images (e.g., an airplane flying in the blue sky). On the other side, if image of interest involves more details, the image is categorized as a complex one (e.g., an airplane moving on the runway). To obtain the characteristics of the images and classify them into simple or complex group, the image data (intensity of pixels) is extracted using histogram. Figure 5 (left panel) displays an example of an image with simple background. It seems that one or two Gaussian functions can be fitted in the histogram of simple images. In other words, the number of some kind of a few gray-level intensities takes the greater values, while the majority of pixels take values in lower levels with slight variations in the histogram. The peak values of these kinds of histograms are greater than the maximum points revealed in the histograms of complex images. As we see in Figure 5 (right panel), the peak value belongs to intensity of background pixels.

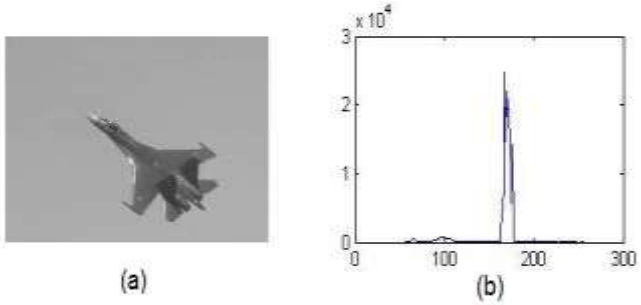


Figure 5. Image with simple background (a), and related histogram (b).

On the contrary, in the histogram of complex image, the range of variations and fluctuated trends are more than observed for simple images. In fact, the abundance of gray-scale colors, which is abnormally distributed in the lower part of the histogram and the range of values between the peaks and valleys, does not vary as large as the range of values defined for simple images. The background of a complex image and its histogram are shown in Figure 6. In fact, the histogram of unknown image gives a criterion to choose way of segmentation. Figure 7 shows the comparison between histograms of simple and complex backgrounds in one selected image.

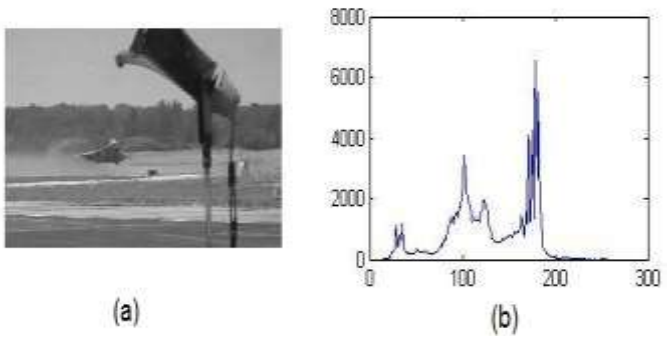


Figure 6. Image with the complex background (a), and related histogram (b).

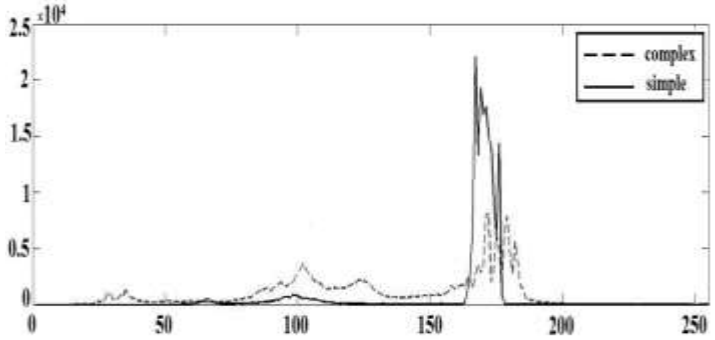


Figure 7. The comparison between simple and complex background in one image.

4. Detecting boundaries displacements in frames with a simple background

Using histograms, the background is considered to be uniform in simple images. So, the displacement components can be obtained by subtracting two successive frames. Similar regions in two successive frames are removed and the remaining regions are representative of differences between consecutive frames.

On the other hand, to remove the background from a simple image, the mean of pixels' intensity of the background is computed using both the histogram of image and the edge pixels of rectangular regions including objects. Then, the intensity of all pixels within the image is subtracted from the obtained mean value. By converting the output to a binary image, the background pixels are separated from the foreground components (Figure 8). This robust approach is applicable for all kind of images to achieve the binary images without any backgrounds.

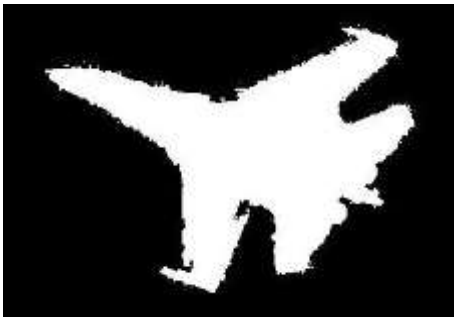


Figure 8. A binary format of segmented object obtained from consecutive frames using histogram.

In the next step, the image of boundaries obtained from the previous step and the current image are added on a pixel-by-pixel basis. In this case, information about the displacement and the edges are integrated into one image. Finally, the background is eliminated from the segmented image with the binary format and only the foreground components with their original intensities are kept.

5. Displacement boundary detection in frames with a complex background

There are many components in the complex background, so that differences between two consecutive frames show lots of details. To solve this problem, the registration process is applied on all of the images to align frames with a previous one. Image registration is the name of approach for overlaying images (two or more) recorded from the same scene with slight displacements occurred with passage of time or incurred by moving camera.

Image registration method usually involves procedures of detecting features, matching corresponding features in both consecutive frames, designing mapping function, and extracting geometric transformations to align the frames with each other [16]. In Figure 9, an example of aligning a frame using the previous one, and the subtraction results are shown.



Figure 9. The first frame (a), the aligned second frame (b), and the image obtained from subtracting these two consecutive frames (c).

When the aligned frame is subtracted from the previous one, the differences in the output image show less detail. In other words, by using alignment of the second image to the first one, some unnecessary movements are removed. Now, the obtained image from the previous steps including boundaries is added to the result of this level.

In the same way described in Section 4, logical operators (i.e., or/and) are employed to integrate the results of displacements and the image output including edges. Finally, the image is converted to the binary format.

6. Separating segmented region of interest

In the binary format of the resultant image, the background is removed, and the remaining regions in the foreground appear as pixels with high-value brightness. For separating the segmented

regions, unique labels must be assigned to each region. To complete the segmentation process, all regions take the value of pixels' brightness in the original image. Figure 10 displays the result of image segmentation process.

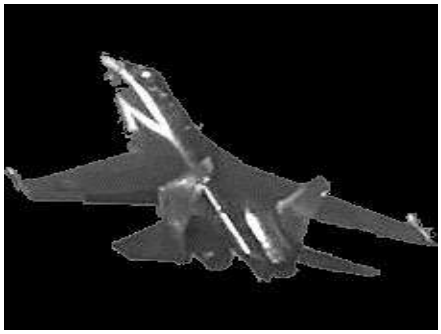


Figure 10. The segmented image.

7. Results

The segmentation results of our hybrid method are promising and can be compared with other known presented algorithms such as K-means, Region Growing procedure, Fast C-means, and Fuzzy C-means (FCM) (for example, see [17]).

The segmentation errors of applying our method and above-mentioned algorithms on the images with both simple and complex backgrounds are compared in Figures 11, and 12. After applying various algorithms on images, the number of images wherein the codes were not able to find (segment and track) the object of interest in frames is counted. Then, the probability of detection error is computed.

Figure 11 shows the comparison of detection errors between different algorithms. It can be seen that our hybrid segmentation algorithm has less error than other algorithms.

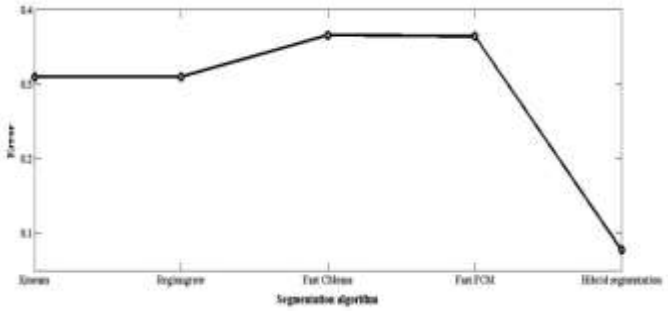


Figure 11. Comparing the probability of detection error of our algorithm and other methods applied on images with simple backgrounds.

According to Figure 12, the comparison of detection errors probability between our hybrid method and other algorithms represents that our algorithm has been the better results than the others when applied on image sequences with complex background.

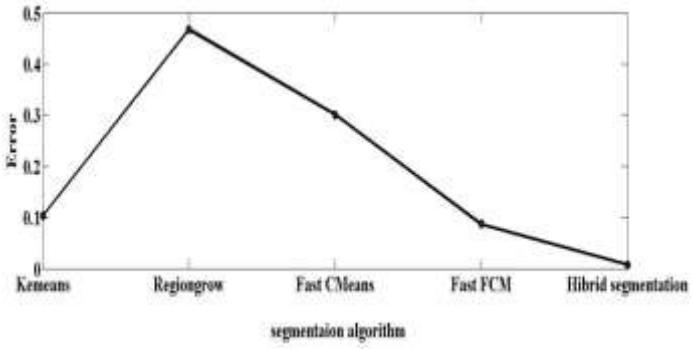


Figure 12. Comparing the probability of detection error of our algorithm and other methods applied on images with complex backgrounds.

Moreover, the execution time of the proposed method and the other algorithms are compared in Figures 13, and 14.

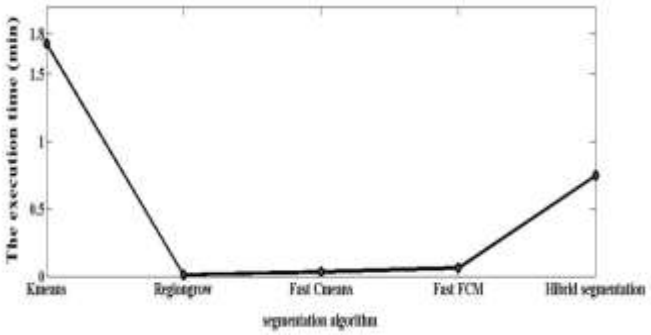


Figure13. The time of running different programs applied on images with simple backgrounds.

We applied on our hybrid segmentation and tracking algorithm on various kinds of video frames. The objects of interest in consecutive frames were one airplane and car in images with simple and complex backgrounds. Our code was not able to track more than one object in frame sequences.

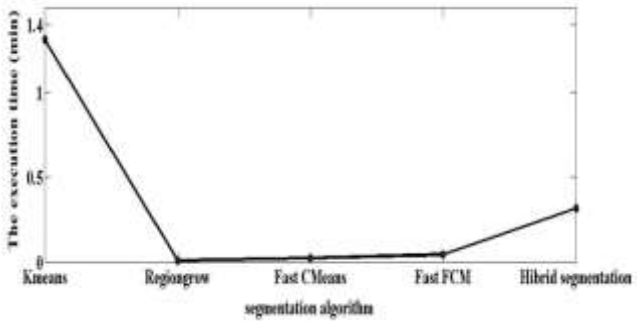


Figure14. The time of running different programs applied on images with complex backgrounds.

8. Conclusions

The hybrid segmentation algorithm works based on the edge detection, intensity-based histograms, morphological operations and image registration. Some of the video surveillance algorithms detect and track moving targets based on criteria that can be determined by user (e.g., the threshold of the background, the size of the target). In this case, they are not automatic and only depend on the values of the input parameters that can be changed during frames. The proposed hybrid algorithm automatically finds the parameters such as background thresholds to segment and track the target. Another advantage of this algorithm is its ability to adjust/synchronize itself with moving cameras recording images with the simple or/and complex backgrounds. Furthermore, the execution time or convergence rate of our algorithm is comparable with the other fast methods such as K-means, and FCM.

Now, we work on our code to improve that for tracking object of interest in frames with more than moving targets. One solution is extraction of invariant features from the targets using image moments.

References

[1] Gopala Krishna, M. T., Ravishankar, M., Ramesh Babu, D. R. (2011). Automatic Detection and Tracking of Moving Objects in Complex Environments for Video Surveillance Applications. IEEE International Conference on Electronics Computer Technology (ICECT), vol. 1, 8-10 April, pp. 234-238.

[2] Cuevas, C., Berjon, D., Moran, F., and Garcia, N. (2012). Moving Object Detection for Real-time Augmented Reality Applications in a GPGPU. IEEE Transactions on Consumer Electronics, vol. 58, no. 1, pp. 117-125.

[3] Cuevas, C., Berjon, D., Moran, F., and Garcia, N. (2013). GPGPU Implementation of an Improved Nonparametric Background Modeling for Moving Object Detection Strategies. IEEE International Conference on Consumer Electronics (ICCE), 11-14 January, pp. 23-24.

[4] Zhang, J., Chen, C. H. (2007). Moving Objects Detection and Segmentation in Dynamic Video Backgrounds. IEEE Conference on Technologies for Homeland Security, 16-17 May, pp. 64-69.

[5] Haritaoglu, I., Harwood, D., Davis, L. S. (1998). W4: A Real Time System for Detecting and Tracking People. IEEE Computer Society Conference on Processing Computer Vision and Pattern Recognition, pp. 962-967.

[6] Wren, C. R., Azarbayejani, A., Darrell, T., Pentland, A.P. (1997). Pfinder: Real-time Tracking of Human Body. IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 19, pp. 780-785.

[7] Stauffer C., Grimson, W. E. L. (2000). Learning Patterns of Activity Using Real-Time Tracking. IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 22, no. 8, pp. 747-757.

[8] Elgammal, A., Duraiswami, R., Harwood, D., and Davis, L. S., (2002). Background and Foreground Modeling Using Nonparametric Kernel Density Estimation for Visual Surveillance. Proceedings of the IEEE, vol. 90, no. 7, pp.1151-1163.

[9] Kato, J., Watanabe, T., Joga, S., Rittscher, J., Blake, A. (2002). An HMM-Based Segmentation Method for Traffic Monitoring Movies. IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 24, no. 9, pp. 1291-1296.

[10] Ayasso, H., Djafari, A. (2010). Joint NDT Image Restoration and Segmentation Using Gauss-Markov-Potts Prior Models and Variational Bayesian Computation. IEEE Transactions on Image Processing, vol. 19, no. 9, pp. 2265-2277.

[11] Seki, M., Wada, T., Fujiwara, H., Sumi, K. (2003). Background Subtraction based on Co-occurrence of Image Variations. IEEE Computer Society Conference on Computer Vision and Pattern Recognition, vol. 2, pp. 65-72.

[12] Toyama, K., Krumm, J., Brumitt, B., Meyers, B. (1999). Wallflower: Principles and Practice of Background Maintenance. International Conference on Computer Vision, Corfu, Greece.
[13] Li, L., Leung, M. K. H. (2002). Integrating Intensity and Texture Differences for Robust Change Detection. IEEE Transactions on Image Processing, vol. 11, no. 2, pp. 105-112.
[14] Gonzalez, R. C. and Woods, R. E. (2007). Digital Image Processing, 3rd Edition, Pearson Prentice Hall.
[15] Shih, F. Y. (2010). Image Processing and Pattern Recognition: Fundamentals and Techniques, John Wiley & Sons, Inc., Hoboken, New Jersey.
[16] Zitova, B., Flusser, J. (2003). Image registration methods: a survey. Image and Vision Computing, vol. 21, pp. 977-1000.
[17] Yousefzadeh, M., Javaherian, M., Safari, H. (2015) Segmentation of Photospheric Solar Images by Using *c*-Means, *k*-Means, and FCM Algorithms, Iranian Journal of Astrophysics and Astronomy, vol. 2, pp. 69-76.