

## Semi-Automatic Moving Objects Segmentation and Tracking Base on Background Subtraction Using Fuzzy C-Means

Moch Arief Soeleman<sup>1,2,a</sup>, Mauridhy Hery Purnomo<sup>1,b</sup>, Mochamad Hariadi<sup>1,c</sup>, Kondo Kunio<sup>3,d</sup>, Masanori Kakimoto<sup>3,e</sup> and Mikami Koji<sup>3,f</sup>

<sup>1</sup>Department of Electrical Engineering Institut Sepuluh Nopember Surabaya, Indonesia

<sup>2</sup>Department of Computer Science Dian Nuswantoro University Semarang, Indonesia

<sup>3</sup>School of Media Science Tokyo University of Technology Tokyo, Japan

<sup>a</sup>arief22208@gmail.com, <sup>b</sup>hery@ee.its.ac.id, <sup>c</sup>mochar@ee.its.ac.id,

<sup>d</sup>kondo@stf.teu.ac.jp, <sup>e</sup>kakimotoms@stf.teu.ac.jp, <sup>f</sup>mikami@stf.teu.ac.jp

**Keywords:** semi-automatic, fuzzy c-means, moving object segmentation.

**Abstract.** In this paper we present a semi-automatic technique for moving objects segmentation and tracking. Clustering as a segmentation method has been used for many applications. The problem of clustering in many segmentation methods is a need of high performance and low computational cost. We proposed the Fuzzy C-Means method for clustering moving objects. To evaluate the performance, we compare FCM against K-Means and SOM algorithm. Semi-automatic used by human for create image ground truth for measure performance by MSE and PSNR. Based on experiment the MSE of Fuzzy C-Means is lower than K-Means and SOM. Also PSNR of FCM is higher than K-Means and SOM. The result proved that Fuzzy C-Means is promising to cluster pixels in moving objects segmentation.

### Introduction

In the recent years, moving objects segmentation has become increasingly important and challenging problem in some video application areas such as automated visual surveillance, content-based video analysis, video annotation, and human-computer interaction. The results of video segmentation process are important for further analysis such as feature extraction, recognition, and classification.

The basic goal of moving objects segmentation process in video is to divide an image sequence into mutually distinctive regions to which can subsequently enclose meaningful labels, where a set of decompose regions with identical and homogeneous attributes such as pixel intensity, color, motion. Several problems of moving objects segmentation have been discussed in many literatures, it can be classified into three categories according to their primary approaches: temporal differencing [1]; motion optical flow [2] [3]; and background subtraction.

The segmentation of semantic video object is an ill-imposed problem [4]. The human only can understand the meaning of semantic video segmentation. The common approach for the first process for analysing the image sequence in video is through background subtraction in [5], that compose in maintaining an update model of moving objects from the background components. The useful of background subtraction is being able to segment video streams from the background. Compared to other approaches, such as optical flow in [6], this approach is computationally affordable for real-time applications.

Based on the advantages of background subtraction in [7], this study decides to use temporal frame difference in moving object segmentation. In this paper, we evaluate the performance of clustering algorithm for segmentation in moving objects.

### Related Work

A robust foreground segmentation algorithm have proposed in [8], this approach is used to categorize between background and foreground with employed a multiple threshold and refine color information then using morphological process.

Meanwhile, the reliable foreground segmentation approach have proposed in [9], the authors combines temporal image analysis with the reference background image to overcome the problem of outdoor daylight scenes that cause change intensities on the background reference image for moving object segmentation.

### Review Related Theory

In this section, we explain how segmentation moving objects using clustering technique between K-Means, Fuzzy C-Means, and Self-Organization Map algorithm are presented.

### K-Means Algorithm

The k-means algorithm is a hard clustering technique that partitions the objects into  $k$  clusters, until each objects are being clustered to one and only one membership. For minimize the sum of squared distances between each data point and the empirical mean of the corresponding cluster.

#### Algorithm 1: K-Means Algorithm

1. Select objects to be  $k$  initial centroid randomly
2. Calculate the distance of each centroid to the each object using distance or similarity metric; assign each object to the cluster with the nearest centroid point.
3. Compute the new centroid point.
4. Return to the step 2 if the current centroid are different to the previous.

### Fuzzy C-Means Algorithm

Fuzzy C-Means is a popular fuzzy unsupervised clustering. FCM produces a membership matrix, which contains the degree of membership of a pixel to all the clusters in [10]. FCM attempts to minimize the sum of square error (SSE).

#### Algorithm 2 : Fuzzy C-Means Clustering

1. Get the input data from an image
2. Choose the number of cluster and the value of  $\varepsilon$  ( $\varepsilon > 0$ )

3. Compute partition matrix  $u_{ik} = \frac{1}{\sum_{j=1}^c \left( \frac{d_{ik}}{d_{jk}} \right)^{\frac{2}{m-1}}}$

4. Update the cluster center  $c_j = \frac{\sum_{k=1}^n u_{ik}^m x_k}{\sum_{k=1}^n u_{ik}^m}$

Repeat step 3 to 4 where  $\|c^k - c^{k+1}\| < \varepsilon$ .

### Self-Organization Map

Self-Organization Map (SOM) is an iterative algorithm in [11] [12] and one of the widely used algorithm for clustering. SOM comprises of two stages: the competitive stage and the cooperative stage.

### Algorithm 3: Self-Organizing Map algorithm

1. Initialize the learning rate  $\alpha$ , radius of the neighbor function  $r$  and random values for the initial weight  $w_{ij} \in W$
2. Repeat until  $\alpha$  reaches 0
  - a. For  $k=1$  to  $n$
  - b. The competitive stage: for all  $w_{ij} \in W$  find the winning neuron that minimize  $d_{ij} = \|x_k - w_{ij}\|$
  - c. The cooperative stage: update each unit  $w_{ij} \in W$ :  $w_{ij} = w_{ij} + \alpha h(w_{winner}, w_{ij}, r) \|x_k - w_{ij}\|$
  - d. Decrease the value of  $\alpha$  and  $r$

### Model of Moving Object Segmentation and Tracking

In this section we explain how to segmentation and tracking moving object using clustering techniques modelling. Moving object segmentation and tracking in each frame is performed by several steps which is described in the following subsections.

#### Background Subtraction

For the background subtraction approach, we convert the frame from RGB into HSV colour space and this allows change detection invariant to modifications of illumination strength. It is intended to recognize the intensity different of current image and the background image [13].

#### Clustering using K-Means, Fuzzy C-Means and SOM Algorithm

In this stage the three methods for moving object segmentation and tracking which in paralel process among K-Means, FCM and SOM have been applied. Three different dataset have been used in each experiment for moving object segmentation.

#### Morphology

Morphology is used to manipulate the features in images based on shape. Dilatation, erosion, opening, and closing are the basic operation of morphology. It is stated in [14] that the purpose of opening to filter the detail and simplifies images by rounding corner from inside the objects.

#### Region of Interest Cropping for measurement

Semi-automatic in this experiment is processed to create image ground truth, where as human operate to cropping region of interest image reference in moving object clustering for comparing performance with moving object segmented to calculate MSE and PSNR.

### Experimental Results

Experimental results of the algorithms implemented for clustering moving objects have been produced for several image sequences. These experiments are conducted on a series of real video. We describe three different sequences that represent typical situations critical for video surveillance systems. All video processing is used for moving objects where the objective of the system is clustering and segmented the objects in the building. We used MATLAB 2010b and ram on PC M 370 with processor i3, 2.40GHz, RAM 4.00 GB

In this experimental results of the algorithms implemented for clustering moving objects have been produced for several image sequences.

- 1) Sequence SC : Sequence Student Campus is a home-made indoor sequence manually labelled, consisting of 996 frames of 352 x 240 spatial resolution.

- 2) Sequence Walk1 : Sequence Walk1 of the database CAVIAR Project<sup>1</sup> is labeled and comprise 611 frames of 484 x 288 in spatial resolution.
- 3) Sequence Walk2: Sequence Walk2 of the 2<sup>nd</sup> database CAVIAR project is labelled and comprises 700 frame of 388 x 288 in spatial resolution.

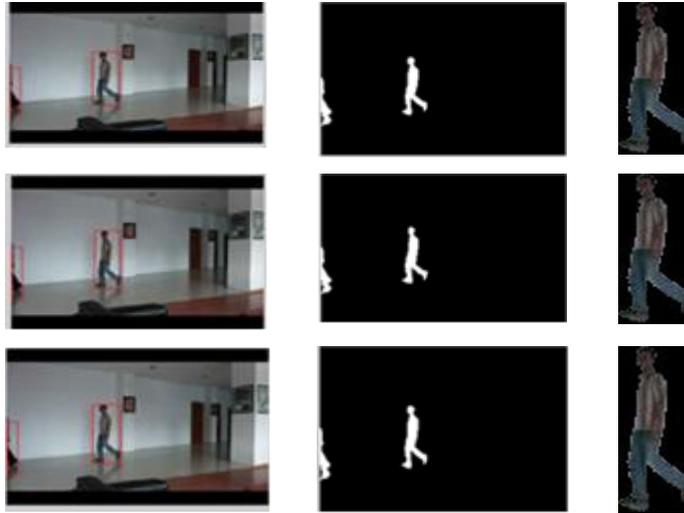


Fig 1., Result of segmentation moving object from top to down using K-means, FCM and SOM (Sequence SC)

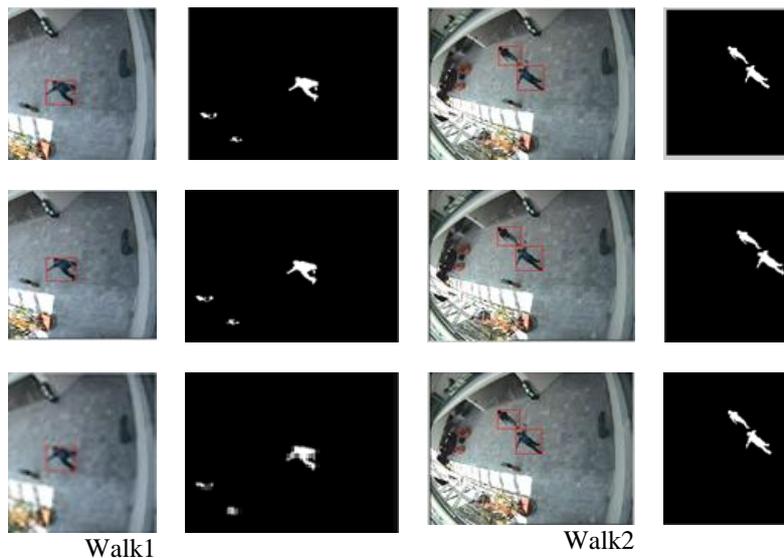


Fig. 2, Result of segmentation moving object from top to down using K-Means, FCM, and SOM (Sequence Walk1 and Walk2)

### Performance Evaluation.

Mean Square Error (MSE) and Peak Signal to Noise Ratio (PSNR) are used to measure the performance of our moving object segmentation. MSE and PSNR were used to compute the different value of segmented image and ground truth image. The lower value of MSE the better is image segmentation. Otherwise, the higher value of PSNR the better is image segmentation [15].

$$MSE(X,Y) = \frac{1}{MN} \sum_{i=1}^M \sum_{j=1}^N [X(i,j) - Y(i,j)] \quad \text{Eq.1}$$

<sup>1</sup> <http://homepages.inf.ed.ac.uk/rbf/CAVIAR/>

$$PSNR(X,Y)=10.\log_{10}\left(\frac{\max^2}{MSE(X,Y)}\right) \quad \text{Eq.2}$$

where  $X$  is the ground truth image,  $Y$  is the segmentation image of size  $M \times N$  and  $\max$  is the maximum possible pixel value of the image.

Table 1 shows the MSE and PSNR of three video. FCM produces the best segmented result Fig. 2 and Fig. 3, show the MSE and PSNR of dataset walk1 respectively.

Table 1. Average MSE and PSNR for K-means, FCM and SOM

Dataset	Average MSE			Average PSNR		
	K-Means	FCM	SOM	K-Means	FCM	SOM
Student	9.913	9.886	9.899	8.68	8.69	6,68
Walk1	13.827	13.170	13.973	6.93	7.14	6.87
Walk2	9.799	9.976	9.809	8.66	8.67	8.66

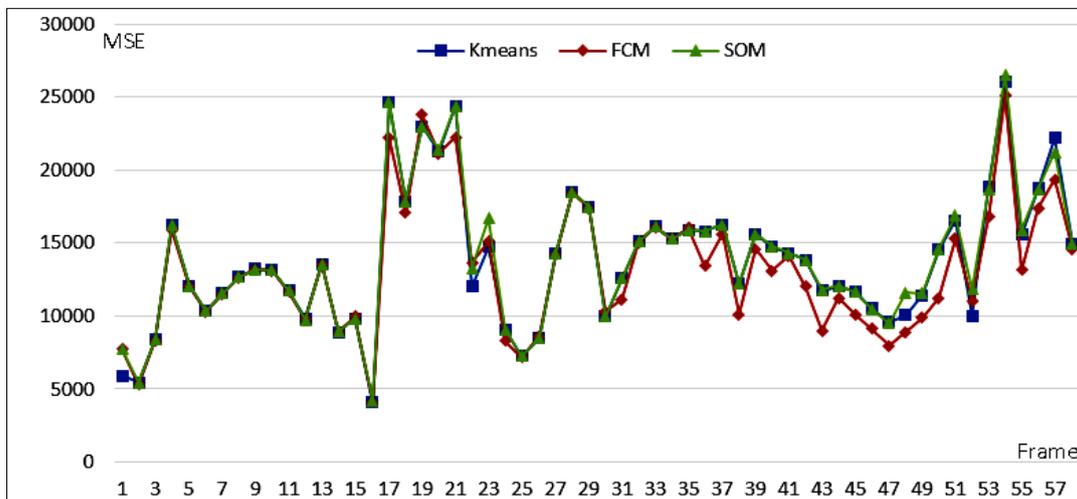


Fig. 3, MSE of Walk1 Dataset using K-Means, FCM, and SOM

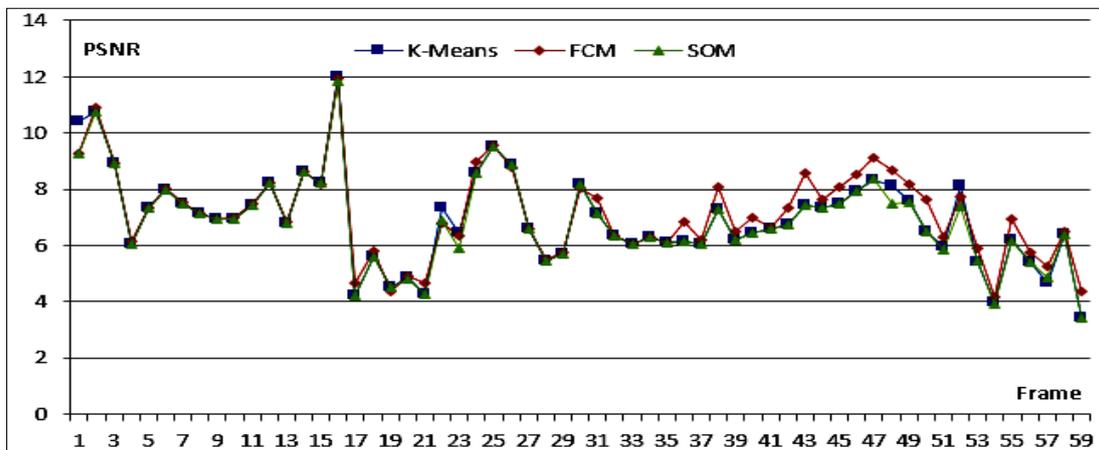


Fig. 4, PSNR of Walk1 Dataset using K-Means, FCM, SOM

Table 2 show the processing time of FCM, K-means and SOM on three video dataset. The computational time of FCM is lower than K-means and SOM.

Table 2. Time Process (second) Average for K-means, FCM and SOM

Dataset	K-means	FCM	SOM
Student	4.09	1.19	8.22
Walk1	6.70	2.39	9.93
Walk2	4.65	1.46	11.8

## Summary

We have presented semi-automatic for clustering using FCM for background subtraction in moving object segmentation. The result shows that the performance of object moving segmentation using FCM is better than with K-Means and SOM algorithm. FCM produced lower MSE and higher PSNR compared to K-Means and SOM. The processing time of FCM faster than K-means and SOM.

## Acknowledgements

This work have supported by - Indonesian Higher Education Directorate General, Ministry of Education and Culture, Republic of Indonesia for this financial support at Tokyo University of Technology, Japan.

## Bibliography

- [1] K. Ng and E. Delp, "Object tracking initialization using automatic moving object detection," in *Proc of SPIE/IS&T Conference on Visual Information Processing and Communication*, January 2010.
- [2] Hui.Y, Yilin.C, Yanzhuo.M, Donglin.B and Zhaoyang.L, "Motion feature descriptor based moving objects segmentation," *High Technology letters*, vol. 18, no. 4, pp. 84-89, 2012.
- [3] L. Wixson and M. Hansen, "Detecting salient motion by accumulating directional-consistent flow," in *ICCV Conference*, Corfu, Greece, 1999.
- [4] Bovic, *The hand book of image and video processing*, Academic Press, 1998.
- [5] K. Srinivasan, K. Porkumaran and G. Sainarayanan, "Improved background subtraction techniques for security in video applications," in *3rd International Conference on Anti-counterfeiting Security, and Identification in Communication*, 2009.
- [6] M. M. A. e. Azeem, "Modified background subtraction algorithm for motion detection in surveillance systems," *Journal of American Arabic Academy for Science and Technology*, vol. 1, no. 2, pp. 112 - 123, 2010.
- [7] L. Maddalena and A. Petrosino, "A self-organizing approach to background subtraction for visual surveillance applications," *IEEE Transactions on Image Processing*, vol. 17, no. 7, pp. 1168 - 1177, July; 2008.
- [8] H. Kim, R. S. I. Kitabara, T. Toriyama and K. Kogure, "Robust foreground segmentation from color video sequences using background subtraction with multiple threshold," in *Proc. KPJR*, 2006.
- [9] P. Spagnolo, T. D. Orazio, M. Leo and A. Distanto, "Moving object segmentation by background subtraction and temporal analysis," *Journal Image and Vision Computing*, vol. 24, pp. 411 - 423, 2006.
- [10] S. Shambharkar and S. Tirpude, "Fuzzy C-Means clustering for content based image retrieval system," in *Proc of International Conference on Advancements in Information Technology with workshop of ICBMG*, 2011.
- [11] T. Kohonen, "Self-organized formation of topologically correct feature map," *Biological Cybernetics*, vol. 43, no. 1, pp. 59-69, 1982.
- [12] T. Kohonen, "Self-organizing map," *Springer Verlag*, 1997.

- [13] B. Sugandi, H. S. Kim, J. K. Tan and S. Ishikawa, "Tracking of moving object by using low resolution image," in *International Conference on Innovative Computive, Information and Control*, 2007.
- [14] A. Amer, "New binary morphological operations for efective low-cost boundary detection," *International Journal of Pattern Recognition and Artificial Intelligence*, vol. 17, no. 2, 2002.
- [15] Z. Qu, "Two allgorithms of image segmentation and measurement method of particle's parameters," *An International Journal Applied Mathematic and Information Science*, no. 1, pp. 105 - 109, 2012.