International Journal of Computer Science and Electronics Engineering (IJCSEE) Volume 1, Issue 3 (2013) ISSN 2320-401X; EISSN 2320-4028

Evaluating the Performance of Common Background Subtraction Techniques

Altahir Abdalla, Awab Samir, Hamza Abdulbagi, Zain Elabdeen Salah,
Siddig Mohammed and Yasir Massad

Abstract—recognizing moving objects from a video stream considered to be a fundamental and critical task in many computer-vision applications. A common approach is to perform background subtraction, which identifies moving objects from the portion of a video frame that differs significantly from a background model. This paper compares various background subtraction algorithms for detecting a single object. The work considers approaches varying from simple techniques such as static method and frame differencing to more sophisticated probabilistic modeling techniques such as adaptive median filtering and GMM. The evaluating process is based on visual observation of the output of the background subtraction techniques under assessments.

Keywords— Background subtraction, frame differencing, median filtering.

I. Introduction

BACKGROUND subtraction is a class of techniques for segmenting out objects of interest in a scene for applications such as surveillance. There are many studies focused on surveying background subtraction techniques such as [1, 2, 3 and 4]. All these studies agreed that; there are many challenges in developing a good background subtraction algorithm. First, the background subtraction method must be robust against changes in illumination. Second, it should avoid detecting non-stationary background objects such as swinging leaves, rain, snow and shadow cast by moving objects. Finally, its internal background model should react quickly to changes in background.

This work consider approaches varying from simple techniques such as static background subtraction and frame differencing, to more sophisticated probabilistic modeling techniques such as adaptive median filtering and using Gaussian Mixture Models. The evaluating process is based on visual observation of the output of the background subtraction techniques under assessments. The output is presented in an easy manner to simplify the comparison process for the human observer.

Altahir Abdalla Altahir, Lecturer - Department of Computer Engineering – Faculty of Engineering and Techology – University of Gezira – WadMadani – Sudan.

Awab Samir, Hamza Abdulbagi, Zain Elabdeen Salah, Siddig Mohammed, Yasir Massad, Students - Department of Computer Engineering – Faculty of Engineering and Techology – University of Gezira – WadMadani – Sudan.

Introducing the visual observation is due to the lack of evaluating methods, so there is a need for a human intervenes to evaluate the robustness of the methods. Besides the mean objective of the work, the paper illustrates in more details the implemented background subtraction techniques.

The purposed lab work is considered in a full controlled environment, free of camera movement, illumination changes, shadows and other similar scenarios which will lead to errors in foreground extraction. These assumptions are necessary to minimize the errors and make it possible to compare simple approach such as static background subtraction and frame differencing with more advanced approaches such as median filtering and Gaussian Mixture Model for background subtraction.

II. Static Background subtraction

The concept of this method is based on subtracting each pixel value in the newly coming frames from static background [1]. The first process is then followed by comparing the result of the subtraction with a predefined threshold, and then the pixel is labeled as foreground if its value is greater than the step value. Converting the RGB images to grayscale is required in order to minimize the calculations. The formula below describes the method:

$$|\mathrm{It} - \mathrm{Bt}| > Th \qquad (1)$$

Where:

I_t is the newly coming frame, B_t is the background model, and *Th* is the predefined threshold.

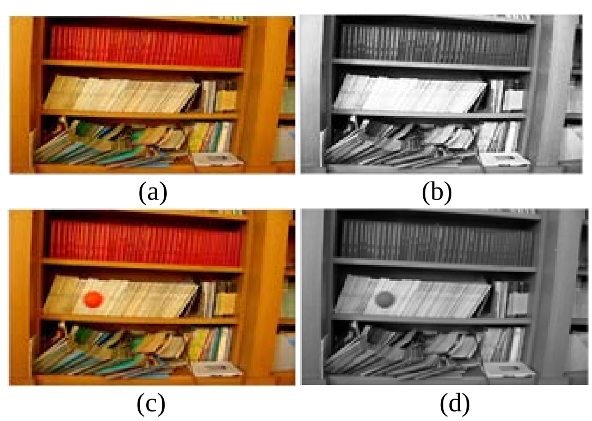


Fig. 1: Static background and its grayscale version with example of a frame with object of interest.

Figure [1] presents the concept of static background subtraction with minimizing the calculation by subtracting grayscale images instead of RGB images. Although such techniques are very fast, the segmentation performance can be quite poor, especially with fluctuating illumination conditions. It is crucial to achieve an accurate component classification in order to secure the quality of further processing steps which is not possible with this technique.

III. Frame Differencing

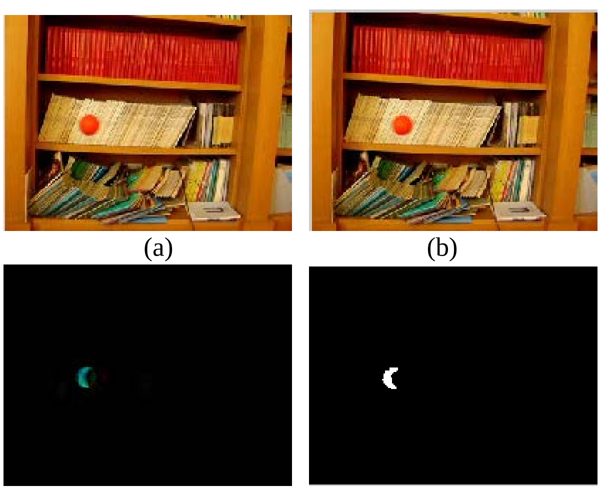
This method is one of the oldest in the field of object detecting from video sequences [1]. Frame differencing is based on subtracting of temporally adjacent frames.

Frame differencing arguably the simplest background modeling technique, it uses the video frame at time t - 1 as the background model for the frame at time t. The formula below describes the frame differencing method:

$$I(x,y)_T - I(x,y)_{T-1} > Th \qquad (2)$$

Where:

$I(x,y)_T$ is the current frame, $I(x,y)_{T-1}$ is the previous frame and *Th* is the predefined threshold.



(a) (b)

(c) (d)

Fig. 2 Frame differencing: (a) Frame number 100. (b) Frame number 101. (c) The subtraction result. (d) The binary image version of the subtraction result.

As mentioned above the Frame differencing uses only a single previous frame, frame differencing may not be able to identify the interior pixels of a large, uniformly-colored moving object. This is commonly known as the aperture problem.

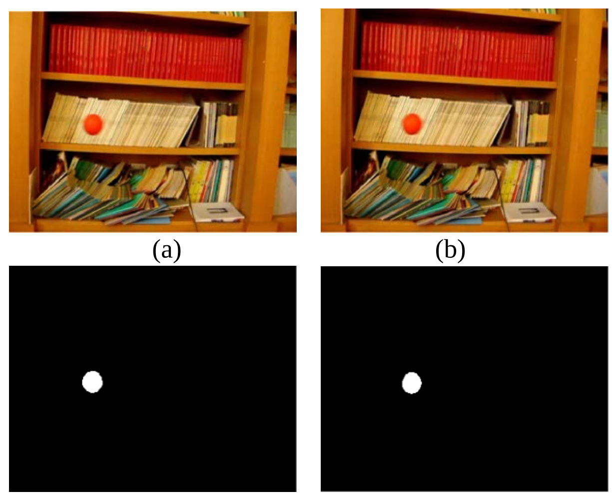
IV. Median Filtering

Median Filtering is one of the most commonly-used background modeling techniques proposed by [5]. The background model is defined to be the median at each pixel location of the frames in the buffer. Assuming that the background is more likely to appear in a scene, the median can be used as the a prefect tool for modeling the background. The background is then represented by the group of the median values in each pixel location. The formula below describes the method:

$$I(x,y)_t - B(x,\ y) > Th \qquad (3)$$

Where

$I(x,y)_t$ is the current frame, $B(x,y)$ is the median value in *(x, y)* location from the first buffered frame to the previous frame . *Th* is the predefined threshold.



(a) (b)

(c) (d)

Fig. 3 Median filtering: (a) Frame number 100. (b) Frame number 101. (c) and (d) The subtraction result.

V. Gaussian Mixture Model

The using of Gaussian for modeling the background is firstly introduces by [6]. Then the method is improved by [7]; through considering a mixture of Gaussian for modeling the background. The improved approach is will explained in [8]. The algorithm by [7] is representative of an adaptive method which uses a mixture of normal distributions to model a multimodal background image sequence. For each pixel, each normal distribution in its background mixture corresponds to the probability of observing a particular intensity or color in the pixel [8]. The pixel value process X is assumed to be modeled by a mixture of K Gaussian densities with parameter sets θ_k, one for each state k:

$$f_{X|k}\,(X|\theta_k) = \frac{1}{(2\pi)^{\frac{\pi}{2}}|\sum k|^{\frac{1}{2}}} e^{-\frac{1(x-\mu_k)^T}{2}\Sigma_k^{-1}(X-\mu_k)} \qquad (4)$$

Where

μ_k is the mean and $\sum_k is$ the covariance matrix of the kth density.

The mixture model models both foreground and background surfaces without distinction. This is why a total of K = 3 Gaussians may be considered a practical minimum to model two background surfaces and one foreground surface in each pixel. (With fewer than two background modes the algorithm is unnecessarily complex and it would be easier to use simple subtraction of an averaged background image. At a minimum the algorithm can work with only one foreground Gaussian because it can be used roughly to model any foreground [6].) Up to K = 7 has been reported in practical applications but it is likely that not much improvement is obtained beyond K = 5 distributions. Once the current state k is estimated, a determination has to be made as to whether it represents a foreground or a background surface [8].

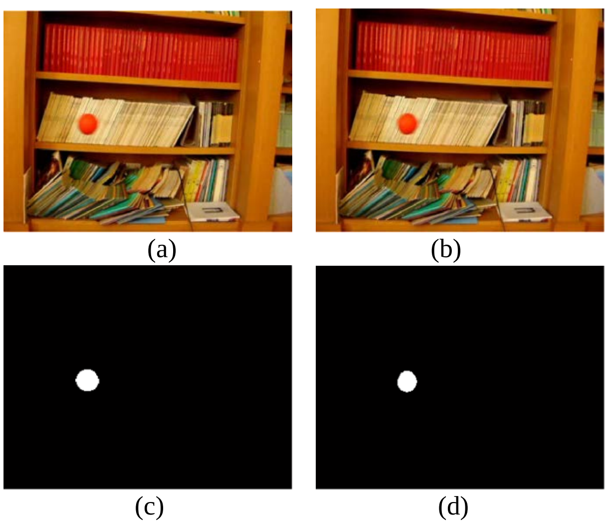


Fig. 4: GMM: (a) Frame number 100. (b) Frame number 101. (c) and (d) The subtraction result.

VI. Discussion

The final result of the implementing of all these techniques is presented in a graphic user interface with a push button to start the video stream.

The user interface is composed of four windows, a window representing the original video and the rest of the show interfaces output algorithms used for comparison. All the methods are compared with the same 500 frame video sequence.

After comparing and applying algorithms we found that the static background is much simple and not suitable for practical applications.

ACKNOWLEDGMENTS

We would like to express our deep appreciation to the family of Department of Computer Engineering-University of Gezira for their unlimited help and support through the all levels of our final year project.

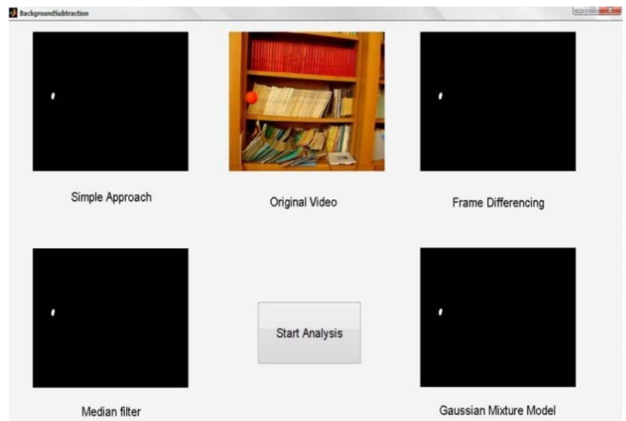


Fig. 5: Graphic user interface for comparing the different background subtraction techniques under the study.

Reference

[1] Massimo Piccardi. Background subtraction techniques: a review. IEEE International Conference on Systems, Man, and Cybernetics - SMC , vol. 4, pp. 3099-3104, 2004.

[2] N. A. OGALE, A survey of techniques for human detection from video, Department of Computer Science, University of Maryland, College Park; http://www.cs.umd.edu /Grad/scholarlypapers/pa pers/ neetiPaper.pdf

[3] Alper Yilmaz, Omer J, Mubarak Shah. Object tracking: A survey. ACM Computing Surveys (CSUR). Volume 38 , Issue 4 (2006). Article No. 13 . ISSN:0360-0300.

[4] Aggarwal, J.K.; Cai, Q. Human motion analysis: a review. Nonrigid and Articulated Motion Workshop, 1997. Proceedings. IEEE. 10.1109/NAMW.1997.609859.

[5] Haritaoglu I., D.Harwood, L.S. Davis "W4: Real-Time Surveillance of People and Their Activities" –IEEE Transaction on Pattern Analysis and Machine Intelligence, Vol. 22, N.8, August 2000.

[6] Wren, C.R.; Azarbayejani, A.; Darrell, T.; Pentland, A.P. Pfinder: real-time tracking of the human body. Transactions on Pattern Analysis and Machine Intelligence. Jul 1997.Page(s):780 - 785 Vol. 19. 10.1109/34.598236.

[7] Stau_er C. and W. E. L. Grimson: Adaptive background mixture models for real-timetracking in: Computer Vision and Pattern Recognition Fort Collins, Colorado June 1999.

[8] P. W. Power & J. A. Schoonees, "Understanding Background Mixture Models for Foreground Segmentation," in Proc. Image and Vision Computing New Zealand, pp. 267-271, Auckland, New Zealand, November 2002.