Vinayak G Ukinkar, Makrand Samvatsar / International Journal of Engineering Research and Applications (IJERA) ISSN: 2248-9622 www.ijera.com
Vol. 2, Issue 3, May-Jun 2012, pp.232-236

Object detection in dynamic background using image segmentation: A review

Vinayak G Ukinkar*, Makrand Samvatsar**
*(Department of Computer Science , Patel College of Science and Technology, Indore)
** (Department of Computer Science , Patel College of Science and Technology, Indore)

ABSTRACT

In computer vision application, object detection is fundamental and most important steps for video analysis. Although several works aimed at detecting objects in video sequences have been reported but due to fast illumination change in a visual surveillance system, many are not tolerant to dynamic background. There are various background subtraction algorithms for detecting moving object but it fails with slow-moving objects or in poor image qualities of videos and does not distinguish shadows from moving objects.

Hence our basic idea is to detect the moving object at foreground and background conditional environment that we can classify each pixel using a model of how that pixel looks when it is part of video frame classes. A mixture of Gaussians classification model for each pixel using an unsupervised technique is an efficient, incremental version of Expectation Maximization (EM) is used for the purpose. Unlike standard image-averaging approach, this method automatically updates the mixture component for each video frame class according to likelihood of membership; hence slow-moving objects and poor image quality of videos are also being handled perfectly. Our approach identifies and eliminates shadows much more effectively than other techniques like thresholding.

Keyword: **Background Subtraction, Image segmentation, Foreground/background extraction, Feature extraction, tracking.**

I. INTRODUCTION

Object detection has become a common task performed in video analysis to deal with surveillance or security. It is commonly used in video surveillances, vehicle auto-navigation, motion capture in sports, child care applications and many more.

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. There are many challenges in developing a good background subtraction algorithm. First, it must be robust against changes in illumination. Second, it should avoid detecting non-stationary background objects such as moving leaves, rain, snow, and shadows cast by moving objects. Finally, its internal background model should react quickly to changes in background such as starting and stopping of vehicle.

Image segmentation is a process used to distinguish objects within images, such as photographs, radar outputs, or x-rays, from their background. For example, given a microscopic image of a blood sample, an image segmentation process could be used to locate and identify all blood cells in the image, recording each cell's position in the sample and even categorizing its blood cell type.

The cells would be stored on a computer as a mathematical model, keeping track of their size and shape. Image segmentation methods can also be used to find and store three-dimensional objects by analyzing multiple images of the same objects from different angles. A challenge, however, arises when the objects being scanned for are not static—in other words, objects that are not always the same shape. For instance, finding and modeling a human liver can be difficult because it has a different shape depending on the specific subject and their age [3].

IMAGE segmentation is the problem of partitioning an image into its constituent components. In wisely choosing a partition that highlights the role and salient properties of each component, we obtain a compact representation of an image in terms of its useful parts. Depending on the end application, the problem of segmentation can be subjective or objective. Digital image processing is being used in many domains today. In image enhancement, for example, a variety of methods now exist for removing image degradations and emphasizing important image information, and in computer graphics, digital images can be generated, modified, and combined for a wide variety of visual effects. In data compression, images may be efficiently stored and transmitted if translated into a compact digital code. In machine vision, automatic inspection systems and robots can make simple decisions based on the digitized input from a television camera. But digital image processing is still in a developing state. In all of the areas just mentioned, many important problems remain to be solved. Perhaps this is most obvious in the case of machine vision: we still do not know how to build machines that can perform most of the routine visual tasks that humans do effortlessly. It is becoming increasingly clear that the format used to represent image data can be as critical in image processing as the algorithms applied to the data. A digital image is initially encoded as an array of pixel intensities, but this raw format is not suited to most tasks. Alternatively, an image may be represented by its Fourier transform, with operations applied to the transform coefficients rather than to the original pixel values. This is appropriate for some data compression and image enhancement tasks, but inappropriate for others. The transform representation is particularly unsuited for machine vision and computer graphics, where the spatial location of pattern elements is critical [2].

Therefore one of the main problems in working with multi resolution representations is to develop fast and efficient techniques. Members of the Advanced Image Processing Research Group have been actively involved in the development of multi resolution techniques for some time. Most of the work revolves around a representation known as a "pyramid," which is versatile, convenient, and efficient to use. We have applied pyramid-based methods to some fundamental problems in image analysis, data compression, and image manipulation.

1.1 Segmentation of objects in image sequences

It is very important in many aspects of multimedia applications. In second-generation image/video coding, images are segmented into objects to achieve efficient compression by coding the contour and texture separately. As the purpose is to achieve high compression performance, the objects segmented may not be semantically meaningful to human observers. The more recent applications, such as content-based image/video retrieval and image/video composition, require that the segmented objects be semantically meaningful. Indeed, the recent multimedia standard MPEG-4 specifies that a video is composed of meaningful video objects. Although many segmentation techniques have been proposed in the literature, fully automatic segmentation tools for general applications are currently not achievable. This is important and challenging area of segmentation of moving objects [1].

Finding moving objects in image sequences is one of the most important tasks in computer vision and image processing. For many years, the "obvious" approach has been first to compute the stationary background image, and then to identify the moving objects as those pixels in the image that differ significantly from the background. Let us call this the background subtraction approach. In earlier work as part of the Road watch project at Berkeley, it was shown that background subtraction can provide an effective means of locating and tracking moving vehicles in freeway traffic.

Moving shadows do, however, cause serious problems, since they differ from the background image and are therefore identified as parts of the moving objects. Moreover, when traffic is slow-moving or stationary, the background image becomes corrupted by the vehicles themselves. These problems arise from an oversimplified view of the task. To classify each pixel of each image as moving object, shadow, or background. The basic idea is that to classify each pixel using a probabilistic model of how that pixel looks when it is part of different classes. For example, a given road pixel in shadow looks different from the same pixel in sunlight or as part of a vehicle. Because the appearance of the pixel in shadow is independent of the object that is casting the shadow, the shadow model for the pixel is relatively constant, like the background model. Furthermore, the probabilistic classification of the current pixel value can be used to update the models appropriately, so that vehicle pixels do not become mixed in with the background model when traffic is moving slowly. Let us show that our approach is successful in coping with slow-moving objects and shadows. There is a large literature on the application of expectation maximization and related techniques to image reconstruction, image segmentation, and motion identification [2].

1.2 Background Subtraction techniques

It is needed to compare various background subtraction algorithms for detecting moving vehicles and pedestrians in urban traffic video sequences. We considered approaches varying from simple techniques such as frame differencing and adaptive median filtering, to more sophisticated probabilistic modeling techniques. While complicated techniques often produce superior performance, our experiments show that simple techniques such as adaptive median filtering can produce good results with much lower computational complexity.

II. LITERATURE REVIEW

Research on video surveillance systems has grown steadily in recent years, but most are not tolerant to the dynamic background. Differencing of adjacent frames in a video sequence has been used for object detection in stationary cameras. However, it was realized that straightforward background subtraction was unsuited to surveillance of real-world situations. Any motion detection system based on background subtraction, needs to handle a number of critical situations such as: gradual variations of the lighting conditions in the scene, small movements of non-static objects such as tree branches and bushes blowing in the wind, noisy image due to a poor quality image source, permanent variations of the objects in the scene, such as cars that park (or depart after a long period), sudden changes in the light conditions, (e.g. sudden clouding), multiple objects moving in the scene both for long and short periods, shadow regions that are projected by foreground objects and are detected as moving objects.

The difference between the current image and the background is used to detect the motion. It does well in motion detection when the background model is good. So for these algorithms, the background model is the key.

Most background modeling approaches tend to fall into the category of pixel-wise models. Early approaches operated on the premise that the color of a pixel over time in a static scene could be modeled by a single Gaussian distribution. Once the pixel wise background model was derived, the likelihood of each incident pixel color could be computed and labeled as belonging to the background or not.

Background modeling approaches consist of two steps: the proper updating of a reference background model, and the suitable subtraction between the current image and the background model. A simple adaptive filter has been used in to update recursively the statistics of the visible pixels. In the Kalman filter is used to model adaptively the background pixel according to known effects of the weather and the time of day on the intensity values. In color and edge information have been used both for background modeling and for subtraction, using confidence maps to fuse intermediate results.

In recent years, a lot of methods are proposed to segment moving objects in dynamic scenes Gaussian mixture model (GMM) [4] is one of the most popular models for moving objects segmentation, which model the color of every pixel in the image with a mixture of Gaussians model. With regard to GMM, some similar methods are proposed. To handle the sharp changes of illumination in scenes, a hierarchical Gaussian mixture model [8] is used in background modeling. In [11], an incremental (on-line) EM algorithm is presented for the learning of Gaussian mixture model.

In optical flow method, the velocity field is used to segment the object. The area with motion field is regarded as object; the other area is regarded as background. This method doesn't need know the scene information, but it has the disadvantages such as high computational cost and vulnerable to noise influence.

III. ANALYSIS OF PROBLEM

In the literature, many approaches for automatically adapting a background model to dynamic scene variations are proposed. Such methods differ mainly in the type of background model and in the procedure used to update the model. Due to the fast illumination changes such as: variations of the light, small movements of non-static objects, noise image in video sequences, and the results can be unreliable. An Accurate detection of moving objects is an important thing in tracking or recognition. A background subtraction method such as Gaussian mixture model (GMM) is one of representative methods used to detect moving objects. Image segmentation identifies the components of the image. It deals with sudden changes in light, changes in background object, presence of noise. Segmentation involves operations such as boundary detection, connected component labeling, thresholding. Our approach identifies and eliminates shadows much more effectively than other techniques like thresholding. Once foreground is extracted a simple subtraction operation can be used to extract the background.

IV. PROPOSED WORK

In proposed work following steps will be carried out.

4.1 Segmentation
4.2 Foreground / background extraction
4.3 Feature extraction and tracking

4.1. Segmentation: It is the process of identifying components of the image. Segmentation involves operations such as boundary detection, connected component labeling, thresholding etc. Boundary detection finds out edges in the image. Any differential operator can be used for boundary detection [4, 5]. Thresholding is the process of reducing the grey levels in the image. Many algorithms exist for thresholding. Segmentation could be used for object recognition, occlusion boundary estimation within motion or stereo systems, image compression, image editing, or image database look-up.

4.2 Foreground extraction: As the name suggests this is the process of separating the foreground and background of the image. Here it is assumed that foreground contains the objects of interest. Some of the methods for foreground extraction are:

Use of difference images -In this method we use subtraction of images in order to find objects that are moving and those that are not. The result of the subtraction is viewed as another grey image called difference image. Three types of difference images are defined [1].

• Absolute accumulative difference image is given by

$f(x,y) = f(x,y) + 1$ ………………if $|g(x,y,t_{i+1}) - g(x,y,t_i)| > T$

• Positive accumulative difference image is given by

$f(x,y) = f(x,y) + 1$ ………………if $g(x,y,t_{i+1}) - g(x,y,t_i) > T$

• Negative accumulative difference image is given by

$f(x,y) = f(x,y) + 1$ ………………if $g(x,y,t_i) - g(x,y,t_{i+1}) > T$

A gap-mountain method: It is applied to identify image blocks that are moving and those that are not moving. The gap-mountain method works as follows- Consider a difference image shown in the adjacent figure. A gap is a sequence of consecutive black pixels and mountain is a sequence of consecutive white pixels. If width of a mountain in a particular row is greater than a preset threshold then we assume that a moving object is present in that row. Similar technique is the algorithm proceeds by dividing the image into smaller sub images (or sub matrices) until each sub matrix contains exactly one object. In the adjacent figure by choosing proper thresholds we can detect the presence of two blocks.

Background extraction: Once foreground is extracted a simple subtraction operation can be used to extract the background [6]. Another method that can be used in object tracking is Background learning. This approach can be used when fixed cameras are used for video capturing. In this method, an initial training step is carried out before deploying the system. In the training step the system constantly records the background in order to ‘learn’ it. Once the training is complete the system has complete (or almost complete) information about the background. Though this step is slightly lengthy, it has a very important advantage. Once we know the background, extracting the foreground is matter of simple image subtraction!

4.3 Feature extraction and object tracking: The next step is to extract useful features from the sequence of frames. Depending on the algorithm, definition of ‘feature’ may vary. The next few sections explore some of the important techniques used for tracking: In the feature based approach discussed by Yiwei Wang, John Doherty and Robet Van Dyck [3] four features namely centroid, dispersion, grayscale distribution, object texture are used for tracking objects. The features are defined as follows:

Centroid = (c_x, c_y) where,

$$c_x = \Sigma(p_{i,j} * i) / \Sigma(p_{i,j}) \quad (5)$$

$$c_y = \Sigma(p_{i,j} * j) / \Sigma(p_{i,j}) \quad (6)$$

$$\text{dispersion} = \left(\Sigma \sqrt{((I - c_x)^2 + (j - c_y)^2)} * p_{i,j} \right) / (\Sigma p_{i,j}) \quad (7)$$

Grey scale distribution of the image is expressed in terms of grey scale range grm, mean of the higher 10% values grh and mean of lower 10% values grl.

Texture of the object tx is defined by the mean of higher 10% values in the wavelet edge image. In current frame under consideration each useful object is assigned the feature vector. Based on the domain knowledge some expected 'tracks' could be generated. Thus the tracking algorithm becomes finding the best track for each object. For this a matrix X of m objects versus n tracks is computed. An element X[i][j] is the number of features that 'match' with the observed features based on some threshold. Further for matrix a threshold for eligible tracks is set. If in a row there is only one track satisfying the eligibility threshold then it becomes the current track of the object. The row and column corresponding to the object and the track is removed from the matrix. For objects with multiple possible tracks, weights are given to the features to evaluate 'cost' for each track. The track that gives the least cost is assigned to the object. Here the weights are given purely on the basis of domain knowledge or based on some heuristic about usefulness of the feature.

General procedure for processing a dynamic background (video) sequence is as follows:

1. Initialize mixture models for each pixel with a weak prior;
2. For each new frame:
   2.1 Update the estimated mixture model for each pixel using incremental EM;
   2.2 Heuristically label the mixture components;
   2.3 Classify each pixel according the current mixture model.

V. CONCLUSION

We proposed new approach for object detection in dynamic background. This approach will suitable for detection of object in indoor and outdoor enviourment and will be robust against fast illumination change in video sequences. It will avoid detecting non-stationary background objects such as moving leaves, rain, snow, and shadows cast by moving objects. And its internal background model will react quickly to changes in background such as starting and stopping of vehicle.

REFERENCE

[1] T. Celik, and H. Kusetogullari, "Solar-Powered automated road surveillance system for speed violation detection," *IEEE Trans. Ind. Electron.*, vol. 57, no. 9, pp 3216-3227, Sep. 2010.

[2] Z. Zhou, Y. Du, and C. Belcher, "Transforming traditional iris recognition systems to work in non-ideal situations," *IEEE Trans. Ind. Electron.*, vol. 56, no. 8,pp 3203-3213, Aug. 2009.

[3] E. Wallace and C. Diffley, " CCTV control room ergonomics," Technical Report 14/98, Police Scientific Development Branch (PSDB), UK Home Office (1988).

[4] H. M. Dee, and S. A. Velastin, "How close are we to solving the problem of automated visual surveillance?," *Machine Vision Applications,* vol. 19, no. 5-6, , pp 329-343, Sep. 2008.

[5] M. Aoki, "Imaging and analysis of traffic scene," in *Proc. IEEE Inter. Conference on Image Processing,* vol. 2, no. 8, 1999, pp 1-5.

[6] M. Brand, V. Kettnaker, "Discovery and segmentation of activities in video," *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol. 22, no. 8, pp 747-757, 2000.

[7] W. Hu, T. Tan, L. Wang, and S. Maybank, "A survey on visual surveillance of object motion and behaviors," *IEEE Trans. On Systems, Man, and Cybernetics, Part C,* vol. 34, no 3, 2004,pp 334-352,2004.

[8] S. Lu, J. Zhang, and D. Feng, "An efficient method for detecting ghost and left objects in surveillance video," in *Proc. IEEE Conference on Advanced Video and Signal Based Surveillance* 2007, pp 540-545.

[9] M. Bhargava, M.S. Chia-Chih Chen, and J.K. Aggarwal, " Detection of abandoned objects in crowded environments," in *Proc. IEEE Conference on Advanced Video and Signal Based Surveillance* 2007, pp 271-276.

[10] S. Calderara, R. Cucchiara, and A. Prati, "Detection of abnormal behaviors using a mixture of Von Mises distributions," in *Proc. IEEE Conference on Advanced Video and Signal Based Surveillance* 2007, pp 141-146.

[11] P. Spagnolo, T.D' Orazio *, M. Leo, A. Distante "Moving object segmentation by background subtraction and temporal analysis" in Image and Vision Computing 24 (2006) 411–423

[12] Yaser Sheikh, Student Member, IEEE, and Mubarak Shah, Fellow, IEEE "Bayesian Modeling of Dynamic Scenes for Object Detection" ieee transactions on pattern analysis and machine intelligence, vol. 27, no. 11, november 2005