A BACKGROUND SUBTRACTION TECHNIQUE FOR OBJECT DETECTION USING SVM

[1]Praveen Kumar B.M, [2]Suma Huddar
[1]Asst. Professor, Dept of Electronics and Communication Engg,
Bapuji Institute of Engineering and Technology, Davangere
[2]Asst. Professor, Dept of Electronics and Communication Engg,
Vishwanathrao Deshpande Rural Institute of Technology, Haliyal
Email: [1]praveen.kumarbm69@gmail.com, [2]sumashuddar@gmail.com

Abstract

The natural technique used for moving object detection in videos captured by a static camera is background modeling and subtraction, and it's a critical preprocessing step in various high-level computer vision applications. However, for this problem there have not been much studies concerning useful multiple features and binary segmentation algorithms. This paper proposed a pixel wise background modeling and subtraction technique using multiple features, where generative (background modeling) and discriminative techniques (background subtraction) were combined for detection. The presented algorithm, multiple features like color, gradient and Haar-like features are integrates for each pixel. The method achieves complete detection of moving objects by involving three significant modules: a multiple feature combination, a background modeling and a foreground and background classification by SVM.

Keywords--Background modeling, Harr-like feature, support vector machine, image processing, support vector machine.

1. INTRODUCTION

Background modeling is often used in different applications to model the background and then detect the moving objects in the scene like in video surveillance, optical motion capture and multimedia. The simplest way to model the background is to acquire a background image which doesn't include any moving object. The real problem arises when static objects start to move. The convention method would produce false alarms in the detection process.A common approach to identifying the moving objects is background subtraction, where each video frame is compared against a reference or background model. Pixels in the current frame that deviate significantly from the background are considered to be moving objects. These “foreground" pixels are further processed for object localization and tracking. Since background subtraction is often the first step in many computer vision applications, it is important that the extracted foreground pixels accurately correspond to the moving objects of interest. Even though many background subtraction algorithms have been proposed, the problem of identifying moving objects in complex environment is still far from being completely solved. Consider the approach, the data is obtained from empty scenes as background model the dissimilarity between the trained model and new observations, the foreground regions are identified. This procedure is called background subtraction. There are many challenges in developing a good background subtraction. First, it must be robust against changes in shadow, illumination changes, and spatial variations of background. A good background model should also react quickly to changes in background and adapt itself to accommodate changes occurring in the background such as moving of a stationary chair from one place to another. It should also have a good foreground detection rate and the processing time for background subtraction .A background can be modeled by considering a color feature only, but it may not give a good result. Therefore in this design a background with multiple components like color, gradient, Harr-like feature. The study of new features or component for background modelling may overcome or reduce the limitations of typically used features, and the combination of several heterogeneous features can improve performance, especially when they are complementary and uncorrelated. According to standard video formats there can be 25 to 30 frames for a second.

In background modelling and subtraction probability density function is quiet common in Model based approaches, where a density function is represented with a compact weighted sum of Gaussians whose number, weights, means, and covariance are determined automatically based on meanshift mode-finding algorithm. In this framework, each visual feature is modelled and every density function is 1D. By utilizing the properties of the 1D mean-shift mode-finding procedure, it can be implemented efficiently and need to compute the convergence locations for only a small subset of data. The three important aspects of our algorithm integration of multiple features, a background modeling, and an foreground/background classification by SVM. The video which is tested is of 320×240 resolution and with 30 frames per second.

The major objective of the paper to introduce a multiple feature integration algorithm for background modeling and subtraction, where the background is modeled with a generative method and background and foreground are classified by a discriminative technique. For classification, an SVM based on the probability vectors for the given feature set is employed. This algorithm demonstrates better performance and the performance is tested quantitatively and qualitatively.

Focus is continuously scoped towards better results.

2. Generic Background Subtraction

Even though there exist many background subtraction algorithms, most of them follow a simple flow diagram shown in Figure 1. The three major steps in a background subtraction algorithm are pre-processing, background modelling and foreground detection. Pre-processing consists of a collection of simple image processing tasks that change the raw input video into a format that can be processed by subsequent steps. Background modelling uses the new video frame to calculate and update a background model. The background model provides a statistical description of the entire background scene. Foreground detection then identifies pixels in the video frame that cannot be adequately explained by the background model, and outputs them as a binary candidate foreground mask. Finally, data validation examines the candidate mask, eliminates those pixels that do not correspond to actual moving objects and outputs the final foreground mask. Domain knowledge and computationally-intensive vision algorithms are often used in data validation. Real-time processing is still feasible as these sophisticated algorithms are applied only on the small number of candidate foreground pixels. Many different approaches have been proposed for each of the processing steps; review some of the representative ones in the following subsections.

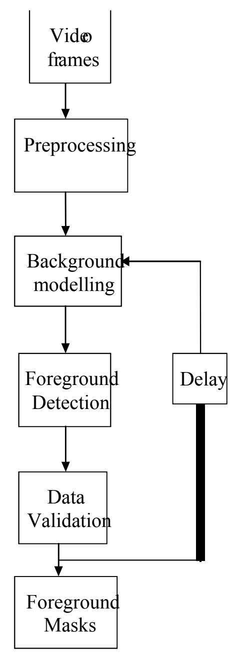

Figure 1. Flow diagram of a generic background subtraction algorithm.

2.1 Preprocessing

In most computer vision systems, simple temporal and/or spatial smoothing are used in the early stage of processing to reduce camera noise. Smoothing can also be used to remove transient environmental noise.

For real-time systems, frame-size and frame-rate reduction are commonly used to reduce the data processing rate. If the camera is moving or multiple cameras are used at different locations, image registration between successive frames or among different cameras is needed before background modelling. Another key issue in pre-processing is the data format used by the particular background subtraction algorithm. Most of the algorithms handle luminance intensity, which is one scalar value per each pixel. However, colour image, in either RGB or colour space, is becoming more popular in the background subtraction. This argue that colour is better than luminance at identifying objects in low-contrast areas and suppressing shadow cast by moving objects. In addition to colour, pixelbased image features such as spatial and temporal derivatives are sometimes used to incorporate edges and motion information. For example, intensity values and spatial derivatives can be combined to form a single state space for background tracking with the Kalman filter. The main drawback of adding colour or derived features in background modelling is the extra complexity for model parameter estimation. The increase in complexity is often significant as most background modelling techniques maintain an independent model for each pixel.

2.2 Background Modeling

Background modelling is at the heart of any background subtraction algorithm. Much research has been devoted to developing a background model that is robust against environmental changes in the background, but sensitive enough to identify all moving objects of interest. We classify background modelling techniques into two broad categories on-recursive and recursive. They are described in the following subsections. The focus only on highly-adaptive techniques, and exclude those that require significant resource for initialization, which store tens of seconds of video to construct initial background models that are characterized by eigen-images or temporal maximum, minimum, and maximum inter-frame differences of all identified background pixels, It(x,y) and Bt(x,y) are used to denote the luminance pixel intensity and its background estimate at spatial location (x,y) and time t.

2.3 Foreground Detection

Foreground detection compares the input video frame with the background model, and identifies candidate foreground pixels from the input frame. Except for the nonparametric model, the techniques use a single image as their background models. The most commonly used approach for foreground detection is to check whether the input pixel is significantly different from the corresponding background estimate:

$$|It(x; y) - Bt(x; y)| > T \tag{1}$$

3. SYSTEM METHODOLOGY

The approach achieves complete detection of moving objects and involves three proposed modules.

1. Multiple Feature Combination.
2. Background Modeling
3. Foreground and Background Classification using SVM.

As shown in Figure 2, construct a statistical representation of the background with multiple features that supports sensitive detection of moving objects in the scene. The background probability of each pixel for each feature is modeled with a Gaussian mixture density function the background modeled for each pixel is associated with k 1D Gaussian mixtures, where k is the number of features integrated. When the background is modeled with probability density functions, the probabilities of foreground and background pixels should be discriminative, but it is not always true. Specifically, the modeled background probabilities between features may be inconsistent due to illumination changes, shadow, and foreground objects similar in features to the background. Also, some features are highly correlated. So, we employ a Support Vector Machine (SVM) for nonlinear classification, which mitigates the inconsistency and the correlation problem among features. The final classification between foreground and background is based on the outputs of the SVM.

Thresholding is the simplest method of image segmentation. From a grayscale image, Thresholding can be used to create binary images. During the Thresholding process, individual pixels in an image are marked as "object" pixels if their value is greater than some threshold value and as "background" pixels otherwise. This convention is known as threshold above. Variants include threshold below, which is opposite of threshold above; threshold inside, where a pixel is labeled "object" if its value is between two thresholds; and threshold outside, which is the opposite of threshold inside. Typically, an object pixel is given a value of “1” while a background pixel is given a value of "0." Finally, a binary image is created by coloring each pixel white or black, depending on a pixel's labels.

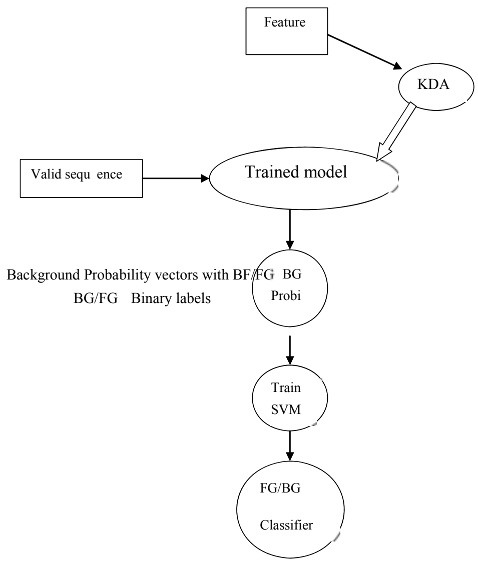

Figure 2. Feature background Subtraction

A. *Multiple feature combination*

3.1 COLOR MODEL One of the fundamental abilities of human vision is color constancy. Humans tend to be able to assign a constant color to an object even under changing of illumination over time or space [5]. The perceived color of a point in a scene depends on many factors including physical properties of the point on the surface of the object. Important physical properties of the surface in color vision are surface spectral reflectance properties, which are invariant to changes of illumination, scene composition or geometry.

3.2 IMAGE GRADIENT An image gradient is a directional change in the intensity or color in an image. Image gradient may be used to extract information from images. In digital image editing, the term gradient or color gradient is used for a gradual blend of color which can be considered as an even gradation from low to high values, as used from white to black in the images to the right. Another name for this is *color progression*.

Mathematically, the gradient of a two-variable function (here the image intensity function) at each image point is a 2D vector with the components given by the derivatives in the horizontal and vertical directions. At each image point, the gradient vector points in the direction of largest possible intensity increase, and the length of the gradient vector corresponds to the rate of change in that direction.

The Sobel operator represents a rather inaccurate approximation of the image gradient, but is still of sufficient quality to be of practical use in many applications. More precisely, it uses intensity values only in a 3×3 region around each image point to approximate the corresponding image gradient, and it uses only integer values for the coefficients which weight the image intensities to produce the gradient approximation.

3.3 HARR-LIKE FEATURE

Figure 3. represents Harr like features which are most commonly used in face recognition, When working with only image intensities (i.e., the RGB pixel values at each and every pixel of image) made the task of feature calculation computationally expensive. A Haar-like feature considers adjacent rectangular regions at a specific location in a detection window, sums up the pixel intensities in each region and calculates the difference between these sums. This difference is then used to categorize subsections of an image. A common haar feature for face detection is a set of two adjacent rectangles that lie above the eye and the cheek region. The position of these rectangles is defined relative to a detection window that acts like a bounding box to the target object. Haar-like features are more robust to illumination changes than color histogram.

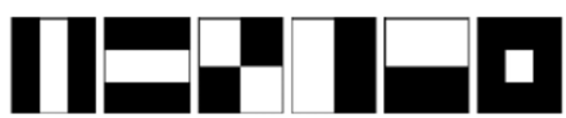

Figure 3. Harr like features

B. *Background Modeling*

Background modelling in surveillance and security applications is very important, because this is the first step in detecting and identifying objects or people in the videos. It is normally used to roughly identify objects and people as the first step in finally detecting and identifying a person. A model of the scene background is built and each pixel in the image is analyzed. A pixel's deviation in colour and/or intensity values is used to determine whether the pixel belongs to the background or the foreground. This information is then grouped together to form regions in the image.

There are various methods to implement this idea, and adopted KDA, where the density functions for each pixel is represented with a compact and flexible mixture of Gaussians.

The sequence training data is composed of n frames and $x_{F,i}(i=1,\ldots,n)$ is the *i*th value for feature. For each feature constructed the density function at each pixel based on Gaussian kernel as follows:

$$f_F(x)=\frac{1}{\sqrt{2\pi}}\sum_{i=1}^{n}(k_{F,i}/\sigma_{F,i})\ \exp\{[(x-x_{f,i})^2)/(2\sigma^2_{F,i})]\}\ ..\ (2)$$

where $\sigma_{F,i}$ and $k_{F,i}$ are the bandwidth and weight of the i*th* kernel, the main issue is the bandwidth selection, even though there is no optimal solution for bandwidth selection for nonparametric density estimation, the initial kernel bandwidth of the i*th* sample feature F is given by: $\sigma F,i=\max(\sigma min, med_{t=i-tw,\ldots,i+tw}|xF,t - xF,t-1|)$…………..(3) where t_w is the temporal window size and σ_{min} is the predefined minimum kernel bandwidth to avoid a too narrow kernel

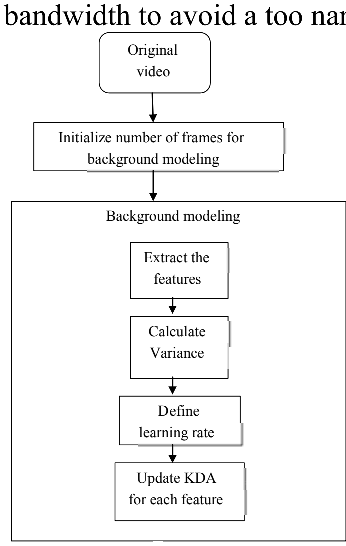

Figure 4. Flowchart for Background modeling

C. Foreground and background Classification by SVM

Once the background is modeled is modeled, each pixel is composed of *k* Gaussian mixtures, where *k* is the number of features integrated. the background probability for a each feature is computed by (2), and *k* probability values, which are represented as a vector and denote foreground and background pixels, by $y_j(j=1,\ldots,N)$,where N is the number of data points. An SVM is employed, where foreground and background data points are trained in order to classify the background and foreground pixels. Note that a universal SVM is used for all sequence not a separate SVM for each pixel.

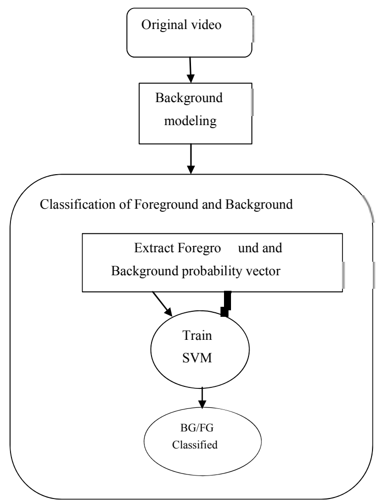

Figure 5. Flowchart for BG/FG Classification.

4. EXPERIMENTAL RESULTS

The experiments are conducted for video having only one and two objects in motion and results are compared with the other background subtraction algorithm. If the background is not properly subtracted the foreground will result in segmentation and gives the less efficiency in detection of object in motion. In Figure 6,shows the original video in which two objects are in motion and the results of different background subtraction algorithm and the Figure 7, gives the result of object identification and the accuracy of the background subtraction.

Sampled Frame		
SVM Method		
Number of objects detected	2	2

Figure 6. Result of Background Subtraction using SVM

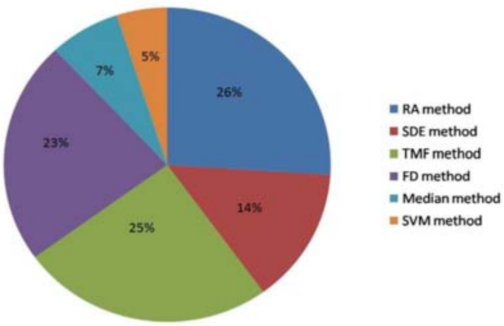

Figure 7. Percentage of variation=changes in the values of detected objects/frame (frame rate=30frames per second)

CONCLUSION

Multiple feature integration algorithms for background modeling and subtraction, where the background is modeled with a generative method and background and foreground are classified by a discriminative technique. KDA is used to represent a probability density function of the background for RGB, gradient, and Haar-like features in each pixel, where 1D independent density functions are used for simplicity. For classification, an SVM based on the probability vectors for the given feature set is employed.

REFERENCES

[1] *Ahmed Elgammal, Ramani Duraiswami, David Harwood,* and *Larry S. davis*, "Background and Foreground Modeling Using Nonparametric Kernel Density Estimation for Visual Surveillance", Proceedings of the IEEE, VOL 90, NO. 7, JULY 2002.

[2] *D.Neelima* and *D.Santhi Priya*, "An Extended and Multifeatures Background Subtraction with SVM", IJESAT, Volume-2, Issue-5.

[3] *Z. Zivkovic* and *F. van der Heijden*, "Efficient Adaptive Density Estimation Per Image Pixel for Task of Background Subtraction," Pattern Recognition Letters, vol. 27, no. 7, pp. 773-780, 2006.

[4] *P. Viola* and *M. Jones,* "Rapid Object Detection Using a Boosted Cascade of Simple Features," Proc. IEEE Conf. Computer Vision and Pattern Recognition, pp. 511-518, 2001.

[5] *I. Haritaoglu,* D. Harwood, and *L.S. Davis*, "W4: RealTime Surveillance of People and Their Activities," IEEE Trans. Pattern Analysis and Machine Intelligence, vol. 22, no. 8, pp. 809-830, Aug. 2000.

[6] *K. Kim, T.H. Chalidabhongse, D. Harwood* and *L. Davis*, "Real-Time Foreground-Background Segmentation Using Codebook Model," Real-Time Imaging, vol. 11, no. 3, pp. 172-185, 2005.

[7] *Anurag Mittal* and *Nikos Paragios* ,"Motion-Based Background Subtraction using Adaptive Kernel Density Estimation".

[8] *Kin-Yi Yam, Wan-Chi Siu, Ngai-Fong Law* and *Chok-Ki Chan*, "Fast video object detection via multiple background modeling", Proceedings of 2010 IEEE 17th International Conference on Image Processing.

[9] *Ahmed Elgammal*, *David Harwood* and *Larry Davis*, "Non-parametric Model for Background Subtraction" .

[10] *Thanarat Horprasert, David Harwood* and *Larry S. Davis*, "A Robust Background Subtraction and Shadow Detection".

[11] *Manjunath Narayana*, Allen Hanson, and Erik Learned-Miller, "Background Modeling Using Adaptive Pixelwise Kernel Variances in a Hybrid Feature Space".

[12] Niel Joubert"Background modelling and subtraction for object detection in video".

[13] Praveen M Ullagaddi and Dr. A.P.Jagadeesh Chandra, "A surveillance system for secured area using Motion Detection and Tracking".

[14] Marko Heikkila and Matti Pietikainen "A Texture-Based Method for Modeling the Background and Detecting Moving Objects".

[15] S.V.N. Vishwanathan, M. Narasimha Murty "A Simple SVM Algorithm".

[16] Seunghoon Park "Adaptive Feature Extraction with Haarlike Features for Visual Tracking".