Foreground Object Motion Detection by Background Subtraction and Signalling Using GSM

S. Priyadharshini
M.E, Final Year,
Arunai Engineering College,
Thiruvannamalai.
Tamilnadu.

S. Dhanalakshmi., M.E
Assistant Professor
Arunai Engineering College
Thiruvannamalai.
Tamilnadu.

Abstract

Intelligent video surveillance systems deals with the monitoring of the real-time environment. It monitors the transient and persistent objects within a specific environment. This is not only designed for security systems and can also be applied for external environmental video surveillance. The basic background subtraction algorithm is used for the detection of moving object. A self-adapting, automatic updating background model is trained to adapt the slow and slight changes of the environment. The foreground object is detected when the subtraction of the current image and the background, which is already trained attains a threshold, the foreground moving object is considered as to be in current view. The mobile phone automatically notifies the control unit through SMS(Short Message Service)or by phone call. Background subtraction technique basically works by feature analysis, pixel differences and so on. Here for feature analysis k-nearest neighbour algorithm (k-nn) is implemented. This proposed system requires little memory and less storage space than the previous. This can be used in implementing mobile based security monitoring system.

Index terms: background subtraction, GSM, support vector machine.

1. Introduction.

1.1 Background modelling and subtraction technique.

Background modeling and subtraction is a technique used for detecting objects in the videos that are captured by fixed camera. It is a major preprocessing step in the field of image processing and computer vision. This uses the pixelwise modeling of the background[1], [2], [3], [4], [9]. It uses the multifeature combination such as color, haar-like features which is integrated and handles the spatio-temporal variations for each pixel. K-nearest neighbour classification algorithm is one of the simplest and rather trivial classifiers is the Route classifier, which memorises the entire training data and performs classification only if the attributes of the test data object match one of the training examples exactly.

The basic background subtraction technique is shown here.

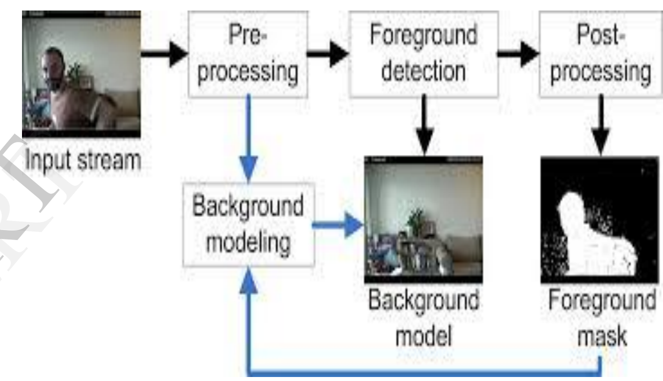

Figure 1. An example for the background subtraction technique.

A more sophisticated approach, *k*-nearest neighbor (*k*NN) classification finds a group of *k* objects in the training set that are closest to the test object, and bases the assignment of a label on the predominance of a particular class in this neighborhood. There are three key elements of this approach: a set of labeled objects, e.g., a set of stored records, a distance or similarity metric to compute distance between objects, and the value of *k*, the number of nearest neighbors. To classify an unlabeled object, the distance of this object to the labeled objects is computed, its *k*-nearest neighbours are identified, and the class labels of these nearest neighbours are then used to determine the class label of the object. The other background subtraction algorithms are used earlier for foreground object detection. The support vector machine (SVM) for a non-linear classification which overcomes the inconsistent and uncorrelated features in [5]. The basic principle of Background subtraction algorithm is that the image's foreground is extracted for futher processing. Image's region of interest are the objects in its foreground. Those objects may be humans, cars or text.

1.2. Related work.

The foremost and the important objective for the background subtraction is to obtain an efficient and effective background model for the foreground moving object detection. The earlier background subtraction algorithm includes frame differences and median filtering based on intensity or colour at each pixel [6], [7]. Then on later years the advanced background modelling used the Density based Background modelling for each pixel defined using PDF(Probability Density Function) based on visual features. The Gaussian mixture models(GMM) developed for the multiple background has a drawback which cannot determine the number of components and it was difficult to add or remove the components into the mixture[12], [13], [17]. None of the GMM's determined the number of Gaussians in a mixture. The YUV colour space explained in [11] using Gaussian distribution for each pixel which cannot handle multimodal density function, it is not robust in changing environments. The real-time applications rely on the fixed number of components to adapt Gaussian. The background model for feature analysis employs the k-nn algorithm. Using the background model which is trained with data obtained from the empty scenes and the foreground regions are identified using the dissimilarity between he trained model and new observation. This procedure is called as Background subtraction. This is to overcome the limitation of combination of the heterogeneous features. By using the texture for the background modelling to handle spatial variations in the scenes that uses filter responses, which is of high computational cost. So, because of this haar-like features are used in order to minimize the errors in background subtraction caused by illumination variations and shadow. Kernal density estimation [8], is a non-parametric density estimation technique which is applied for background subtraction. This method is useful for general density function, it requires many samples for the measurement of accurate estimate of the underneath density function and it is computationally costly. This technique is not appropriate for real-time applications and particularly when multi dimensional features are involved. KDA (Kernel density approximation) in [10] is a multi-modal density representation method. As because each and every parameter of the Gaussian mixture is calculated automatically. This is applied to the real-time computer vision problem for object tracking

2. Moving object detection and alert system.

This section deals with the implementation of background subtraction and signalling the detected foreground object to the corresponding agent using GSM. This exclusively explains the support vector machine for training the background model.

2.1. Support Vector Machine.

There are various kinds of machine learning techniques. They are supervised learning, unsupervised learning and reinforcement. Support vector machines are a kind of supervised learning model associated with the learning algorithm. These algorithms analyse the data by recognising its pattern that are used for regression analysis and classification. The SVM basically takes a input set of data, which predicts the input set for each input given, of two possible output forms. On giving a training example set, each one belonging to any of the two categories, an SVM training algorithm constructs a model that places the new example in any one of the above two categories. [5] An SVM model is the representation of the points in space which is mapped as the examples of different categories are divided separately by a definite gap that they are as far as possible. This enables us to find where or which side the new example mapped into the space falls.

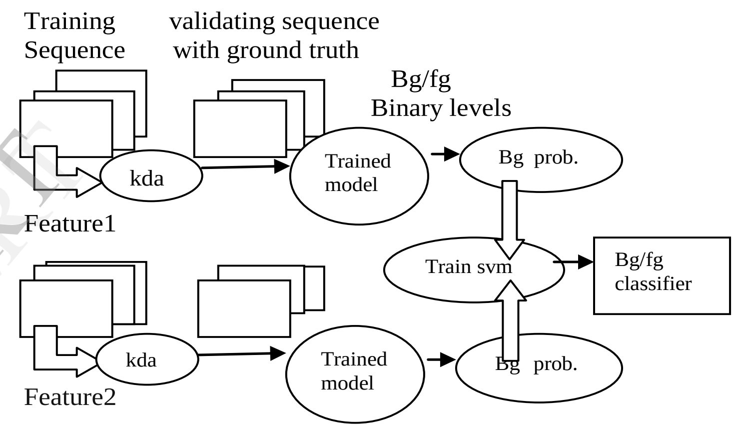

Figure 2. SVM Training model

This SVM in addition to the linear classification is capable of performing knn-linear classification using kernel trick.

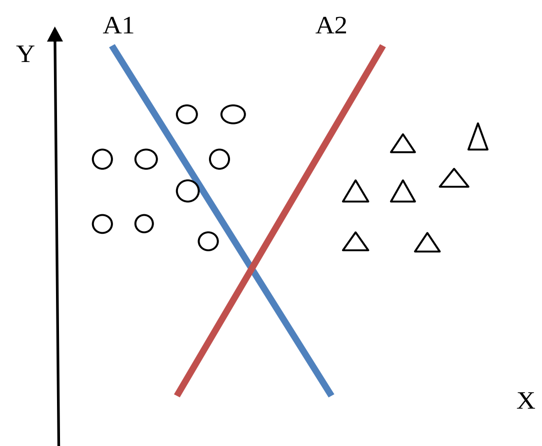

Figure 3. Graph with two different input sets and are separated.

From the above given graph, there are two different input sets shown above and the line is to indicate that they are separated. The blue line(H1) doesn't differentiate it correctly whereas the red line(H2) shows the real division of the two input models.. The new example if anything is given it is analyzed and after pattern recognisation it is grouped under one of the two sets.

A SVM builds a set of hyperplanes in a high or multi-dimensional space can be used for regression and classification. High quality of the separation is attained by generating a functional-margin which provides the greatest distance between the nearest training input data points.

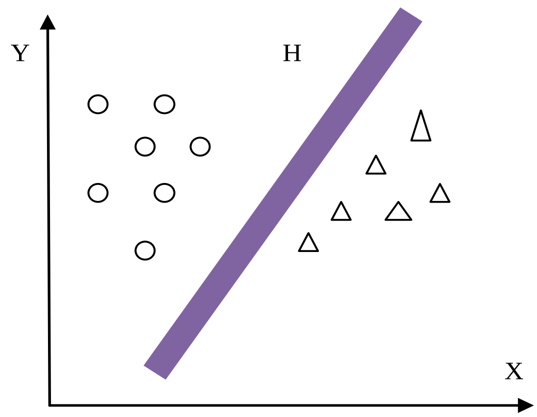

Figure 4. Set of divided input points by functional margin.

On the above diagram H represents the hyperplane which has a wide gap between two different input sets.

3. System architecture.

3.1. Feature analysis .

This feature analysis describes the characteristics and the performance of the integrated features. The haar-like features and the RGB colours are correlated between every pair of the selected features. K-means clustering algorithm is used for background modeling and subtraction technique which has the discriminative and generative models for classification. The performance of the foreground classification is improved by feature analysis. For dimensionality reduction, in image processing and pattern recognition the feature extraction is mainly used. For example consider the input set containing data set with huge amount of data with highly redundant and unwanted information. There the input dataset will be transformed into a reduced feature set called as feature vector. This process of transformation from a dataset into feature set is called as feature extraction. If the extracted features are selected carefully, the feature set is supposed to find only the most relevant information regarding the required input from the whole input set.

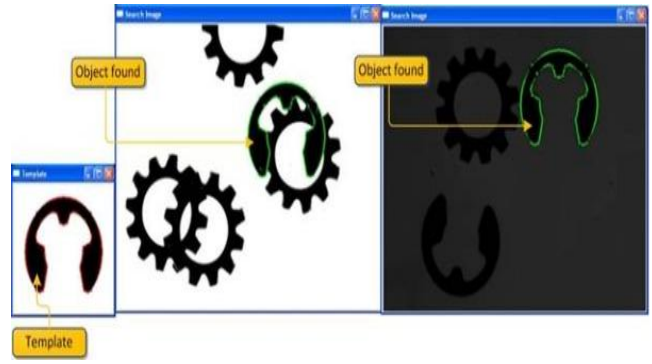

Figure 5. An example for template matching.

The template matching technique can be defined as a technique which finds small regions of an image which is matching the template image. It is widely used in digital image processing. The template image should have a strong feature, only when the feature based approach can be considered. It is computationally more efficient because it doesn't consider the entire image but it works by potential search of huge count of points so that it can identify the location which matches accurately.

3.2. Classification.

After background subtraction, in the classification, every pixel consist of k-Gaussian mixtures, where k can be the total number of features integrated. These distributions constitutes the foreground and background classification for each new frame. The background probability for each feature value is obtained and for each pixel k-probability values are associated. This is represented as k- dimensional vector. These k-dimensional features are gathered from the pixels of foreground and background. The objective of color clustering is to divide a color set into c homogeneous color clusters. Color clustering is used in a variety of applications, such as color image segmentation and recognition. This algorithm classifies a set of data points X into c. Homogeneous groups represented as fuzzy sets $F1, F2, ..., Fc$. The objective is to obtain the fuzzy c-partition $F = \{F1, F2, .., Fc\}$ for both an unlabeled data set $X = \{x1, ..., xn\}$. Fuzzy c-means algorithm for clustering color data is proposed in the present study. The initial cluster centroids are selected based on the notion that dominant colors in a given color set are unlikely to belong to the same cluster.

3.3. Background Detection.

K-means clustering is a method of cluster analysis which aims to partition n observations into k clusters in which each observation belongs to the cluster with the nearest mean. The problem is computationally difficult; however there are efficient heuristic algorithms that are commonly employed that converge fast to a local optimum. These are usually similar to the expectationmaximization algorithm for mixtures of Gaussian distributions via an iterative refinement approach employed by both algorithms. They both cluster centres for data modelling, k-means clustering algorithm finds the clusters of the comparable extend of space, but clusters could be able to take different shapes because of expectation- maximization. Precision-Recall (PR) curves are implemanted, where the precision and recall are defined as

$$\text{Precision}=\frac{n(bt\cap b)}{n(bt)}$$

$$\text{Recall}=\frac{n(ft\cap b)}{n(ft)}$$

where 'ft' and 'bt' are the ground truth sets of foreground and background pixels, and 'f' and 'b' denote the sets of foreground and background pixels obtained by a background subtraction algorithm.

3.4. Global system for mobile communication.

After detecting the difference between the current frame and the template frame, an alert is sent to the central control unit or the coordinated agent using GSM modem. It is a special type of modem which has a SIM card, and subscribes to a mobile operator. This allows the GSM modem connected to a computer to work as a phone widely for internet connection but can also effectively used for sending and receiving SMS (Short Message Service) and MMS (Multimedia Messaging Service).

This supports more protocols of 2.5G and 3G technologies. This modem can be connected to the system using serial, Bluetooth or USB.

This GSM modem supports the protocols such as GPRS and EDGE. This message sending and receiving service plays a major role in it.

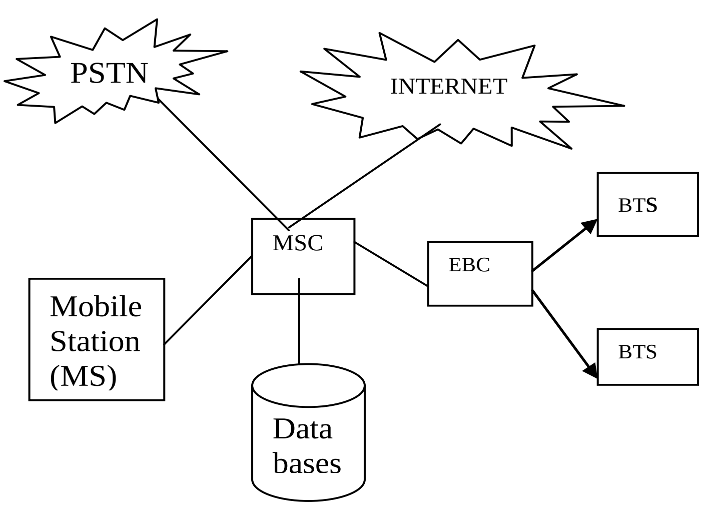

Figure 6. GSM architecture.

GSM uses TDMA (Time Devision Multiplexing) and FDMA (Frequency Devision multiplexing) together. This means that the users share both the time channel and the frequency.The working explanation of this architecture is that the mobile station(MS) communicates with the MSC which is connected with the mobile database. The EBC with the base stations. The base stations may be mobile phone or the computer. This GSM is mainly provided with a port for SIM card. This SIM stands for Subscriber Identity module. This has a unique number and unique identity representing the particular or specific card.

4. System design.

The system architecture is defined as that which explains all the integrated components in a system.This architecture diagram stands as a framework for the description of all the built in components. The main and important components are the object localization and background modelling and subtraction which is detailed below.

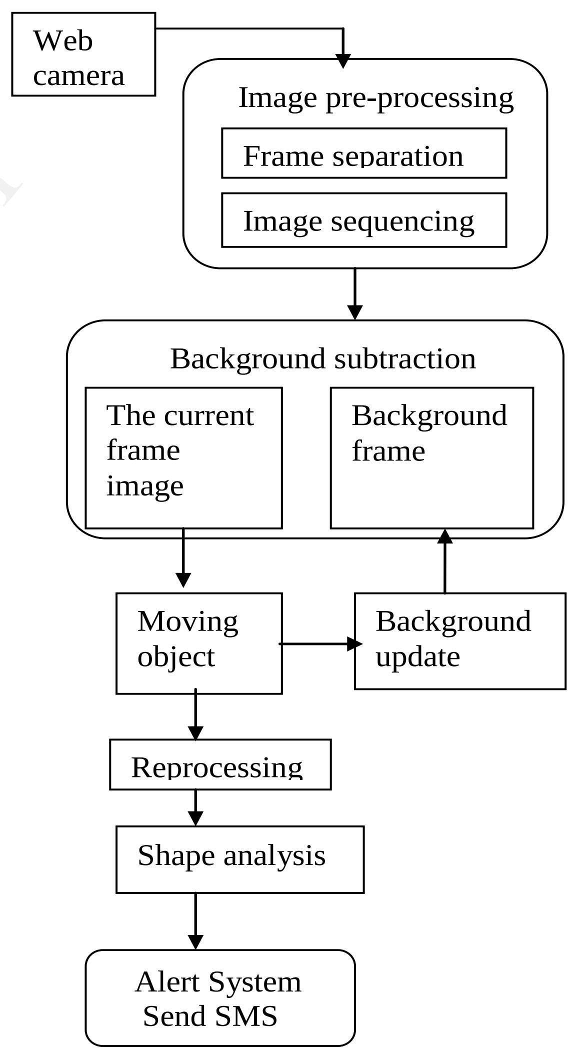

Figure 7. Overview of the system

ISSN: 2278-0181
Vol. 2 Issue 12, December - 2013

The proposed consist of object localization at image pre-processing phase which processes the images captured using web camera.

The video is captured and the frames at which changes occurred are only recorded. The frame separation and image sequencing takes place. The background subtraction is a technique which compares the current frame image and the background frame and captures the difference between the frame. The background is updated in the background model. The moving object is detected and it is reprocessed. It has to go through the shape analysis and the detected image is considered as a moving object and the alert has been sent to the agent or user as SMS.

5. Experiments.

The Experiments has been conducted for the capturing of the moving object. The image is captured only when there occurs changes in the movement of the object. The initial background model is created. The first and the second captured images are compared and the background subtraction algorithm works here and captures the next image only when the changes occurred in the movement of the object of the previous image. When a image is captured a alert is given to the coordinated agent using GSM modem.

The changes in the movement of the object, illumination changes, shadow detection is also analysed. The camera doesn't capture image when there were no motion of the foreground object. The images shown below gives us the details which includes the name of the image and the time at which the image has been captured. It is clear that when the time difference between two images are high it is proved that the image has not been captured when there were no movement or otherwise can be said as no difference between the modelled frame and current frames.

The images are captured when changes there occurs in the first captured image and the following images. These images absorbs the illumination changes. The shadow changes that are captured is shown below. The Figure 8, shows the captured 9 images shows the frames by the time at which they are captured.

The images 1 and 2 are at the time interval of 4 seconds. The interval or time gap between the images are at the range of from 3-5 seconds. So it is very much clear that they have very high frame rates of approximately 17 frames/min.

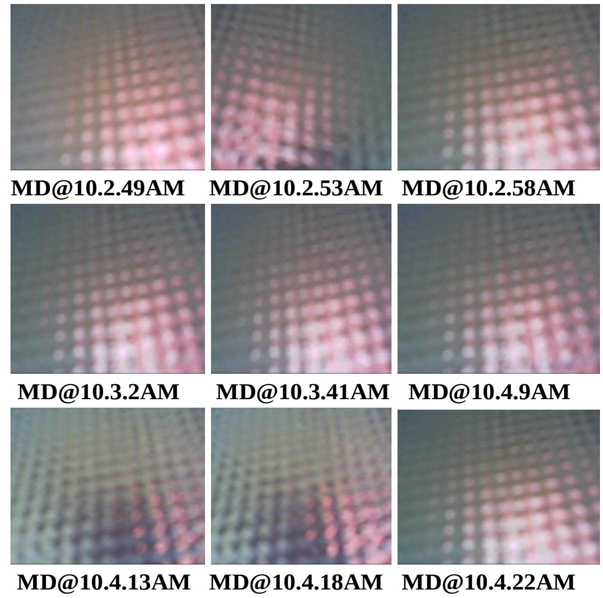

Figure 8. The shadow changes of the images are detected and captured.

6. Future scope.

The alert messages are transmitted for each captured images. These messages are generated for every motion detected and whenever the foreground object's motion detected. This can be further enhanced with the facility of face recognition. By face recognition is it possible to get know about the humans, incase they are detected.

7. Conclusion.

The performance and the rate at which the images are processed and the output are saved is shown. This gives us a clear view about how the images are captured and motion is detected. This can be done with small memory and few hardware requirements. This is of low computational cost.

8. References.

[1]. A. Elgammal, D. Harwood, and L. Davis, "Non-Parametric Model for Background Subtraction," Proc. European Conf. Computer Vision, pp. 751-767,June 2000

[2]. C. Stauffer and W.E.L. Grimson, "Learning Patterns of Activity Using Real-Time Tracking," IEEE Trans. Pattern Analysis and Machine Intelligence,vol. 22, no. 8, pp. 747-757, Aug. 2000.

[3] B. Han, D. Comaniciu, and L. Davis, "Sequential Kernel Density Approximation through Mode Propagation: Applications to Background Modeling," Proc. Asian Conf. Computer Vision, 2004.

[4] D.S. Lee, “Effective Gaussian Mixture Learning for Video Background Subtraction,” IEEE Trans. Pattern Analysis and Machine Intelligence, vol. 27.
.

[5]. I.J. Information Technology and Computer Science, 2012, Published Online May 2012 in MECS.

[6]. Image Classification using Support Vector Machine and Artificial Neural Network by Le Hoang Thai Computer Science Department and Ho Chi Minh City, Vietnam, member of IACSIT

[7] I. Haritaoglu, D. Harwood, and L.S. Davis, “W4: Real-Time Surveillance of People and Their Activities,” IEEE Trans. Pattern Analysis and Machine Intelligence, vol. 22, no. 8, pp. 809-830, Aug. 2000.

[8] 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.

[9]. Motion-Based Background Subtraction using Adaptive Kernel Density Estimation by Anurag Mittal Nikos Paragios.

[10]. Real-Time Vision and Modeling C.E.R.T.I.S. Siemens Corporate Research Ecole Nationale de Ponts et Chaussees Princeton, NJ 08540 Champs sur Marne, France

[11]. Background Modeling and Subtraction of Dynamic Scenes by Antoine Monnet, Anurag Mittal, Nikos Paragios and Visvanathan Ramesh in Real-Time Vision and Modeling Siemens Corporate Research

[12]. Sequential Kernel Density Approximation and Its Application to Real-Time Visual Tracking by Bohyung Han, Member, IEEE Dorin Comaniciu, Senior Member, IEEE Ying Zhu, Member, IEEEand Larry S. Davis, Fellow, IEEE

[13]. Wren, A. Azarbayejani, T. Darrell, and A. Pentland, “Pfinder: Real-Time Tracking of the Human Body,” IEEE Trans. Pattern Analysis and Machine Intelligence, vol. 19, no. 7, pp. 780-785, July 1997

[14]. N. Friedman and S. Russell, “Image Segmenation in Video Sequences: A Probabilistic Approach,” Proc. 13th Conf. Uncertainty in Artificial Intelligence, 1997.

[15]. A. Mittal and D. Huttenlocher, “Scene Modeling for Wide Area Surveillance and Image Synthesis,” Proc. IEEE Conf. Computer Vision and Pattern Recognition, 2000.

[15] R.M. Neal and G.E. Hinton, “A View of the EM Algorithm that Justifies Incremental, Sparse, and Other Variants,” Learning in Graphical Models, M.I. Jordan, ed., pp. 355-368, Kluwer Academic, 1998.

[16] A.D. Jepson, D.J. Fleet, and T.F. El-Maraghi, “Robust Online Appearance Models for Visual Tracking,” Proc. IEEE Conf. Computer Vision and Pattern Recognition, vol. I, pp. 415-422, 2001.

[17]. L. Li, W. Huang, I. Gu, and Q. Tian, “Statistical Modeling of Complex Backgrounds for Foreground Object Detection,” IEEE Trans. Image Processing, vol. 13, no. 11, pp. 1459-1472, Nov. 2004.

[18]. T. Parag, A. Elgammal, and A. Mittal, “A Framework for Feature Selection for Background Subtraction,” Proc. IEEE Conf. Computer Vision and Pattern Recognition, pp. 1916-1923, 2006.

[19]. Real-Time Object Tracking and Classification Using a Static Camera Swantje Johnsen and Ashley Tews.