Density Based Multi feature Background Subtraction Using Support Vector Machine

[1]G. Saivamsidhar, [2]Smt S.J.R.K PadminiValli

[1]M.Tech Student, [2]Assistant Professor

[1,2]Department Of Computer Science and Engineering

[1,2]RVR & JC College of Engineering and Technology

Guntur, Andhra Pradesh, India- 522 019

Abstract: - Video surveillance takes place normally by using CCTV cameras (Closed Circuit Television) for monitoring or surveillance for intruder detection in case of emergencies in hospitals, shopping malls, banking sectors, personal purpose automation and so on. Later Video fusion approach also used for monitoring such systems. Due to time delay, we cannot get the update information for every minute or second and so it is not possible to detect the intruder in an appropriate time. These systems use the moving average algorithm to store the monitored images. Firstly, the basic principle of moving object detecting is given by the Background Subtraction algorithm. Then, a self-adaptive background model that can update automatically and timely to adapt to the slow and slight changes of natural environment is detailed. When the subtraction of the current captured image and the background reaches a certain threshold, a moving object is considered to be in the current view, and the mobile phone will automatically notify the central control unit or the user through phone call, SMS (Short Message System) or other means.

Keyword: *Video surveillance, SMS, image, Background Subtraction algorithm.*

I. INTRODUCTION

Closed circuit television (CCTV) has been increasingly deployed across Queensland public authorities for a variety of visual surveillance purposes. These purposes span incident monitoring, detection and deterrence, contributing to the safety of the public, personnel and property. The use of CCTV extends to operational areas such as traffic monitoring on roads, rail and at sea, and the provision of evidence in Queensland courts. Effectively managing CCTV records as public records may present significant challenges. Issues which may need to be addressed by public authorities include the proliferation of proprietary visual surveillance systems and encodings, overcoming poor picture quality owing to the lack of operational standards, the divergence in business processes to manage the records, and a lack of recognition of the total cost of ownership in managing records throughout their lifecycle.

Background modeling is an important component of many vision systems. Existing work in the area has mostly addressed scenes that consist of static or quasi-static structures. When the scene exhibits a persistent dynamic behavior in time, such an assumption is violated and detection performance deteriorates. In this paper, we propose a new method for the modeling and subtraction of such scenes. Towards the modeling of the dynamic characteristics, optical flow is computed and utilized as a feature in a higher dimensional space. Inherent ambiguities in the computation of features are addressed by using a data dependent bandwidth for density estimation using kernels. Extensive experiments demonstrate the utility and performance of the proposed approach.

Increased computational speed of processors has enabled application of vision technology in several fields such as: Industrial automation, Video security, transportation and automotive. Background subtraction forms an important component in many of these applications. The central idea behind this module is to utilize the visual properties of the scene for building an appropriate representation that can then be utilized for the classification of a new observation as foreground or background. The information provided by such a module can then be considered as a valuable low level visual cue to perform high level object analysis tasks such as object detection, tracking, classification and event analysis.

The magnitude of the deviation between the predicted and actual observation can then be used as a measure of change. Predictive mechanisms of varying complexity have been considered in the literature. Several authors [16, 17] have used a Kalman-filter based approach for modeling the dynamics of the state at a particular pixel.

II. RELATED WORK

Many methods have been used for background modeling. Although most of these methods deal only with axed camera, they provide a good starting point for a moving camera scene. Simple methods include averaging the pixels at a particular location, taking the median of all the values at a location, and calculating spatially weighted values in order to reduce the effect of outliers.

Managing Closed Circuit Television:

In recent years a combination of perceptions and fears of increased street crime and advances in technology has seen an upsurge in the use of closed circuit television (CCTV) as a tool in tackling crime in public places. Many private companies and a number of local government authorities have initiated trials in the use of CCTV, and the technology is also being used in a number of ways in the public transport system. Because CCTV is relatively new, it is still not clear how effective it is in deterring or reducing crime. Research evidence so far suggests that it can be an effective strategy in situational crime prevention at a local level, but only as one of a range of crime prevention strategies. It appears from the research that CCTV may be effective in addressing property crime and some types of assault and robbery. Evidence also suggests that the benefits of CCTV surveillance fade after a period of time, and that displacement may occur, that is, the crime may simply move to other areas away from the CCTV surveillance, or there may be a shift to different sorts of crime which are less susceptible to CCTV surveillance.

For these reasons, CCTV on its own can do little to address long term crime prevention. CCTV should only be considered as one part of an integrated crime prevention strategy and should be installed on a trial basis subject to rigorous evaluation as to its effectiveness. These Guidelines have been developed by the NSW Government to provide a policy framework and a set of underlying principles to assist agencies considering CCTV as a possible response to local community safety concerns.

Efficient hierarchical:

Detecting moving objects by using an adaptive background model is a critical component for many vision-based applications. Most background models were maintained in pixel-based forms, while some approaches began to study block-based representations which are more robust to non-stationary backgrounds. In this paper, we propose a method that combines pixel-based and block-based approaches into a single framework. We show that efficient hierarchical backgrounds can be built by considering that these two approaches are complementary to each other. In addition, a novel descriptor is proposed for block-based background modeling in the coarse level of the hierarchy.

A Texture-Based Method:

Various visual features may be used to model backgrounds, including intensity, color, gradient; motion, texture, and other general filter responses. Color and intensity are probably the most popular features for background modeling, but several attempts have been made to integrate other features to overcome their limitations. There are a few algorithms based on motion cue [18], [21]. In [20], spectral, spatial, and temporal features are combined, and background/foreground (Fig.1) classification is performed based on the statistics of the most significant and frequent features. Recently, a feature selection technique was proposed for background subtraction.

ISSN-2320-7884 (ONLINE)
ISSN-2321-0257 (PRINT)

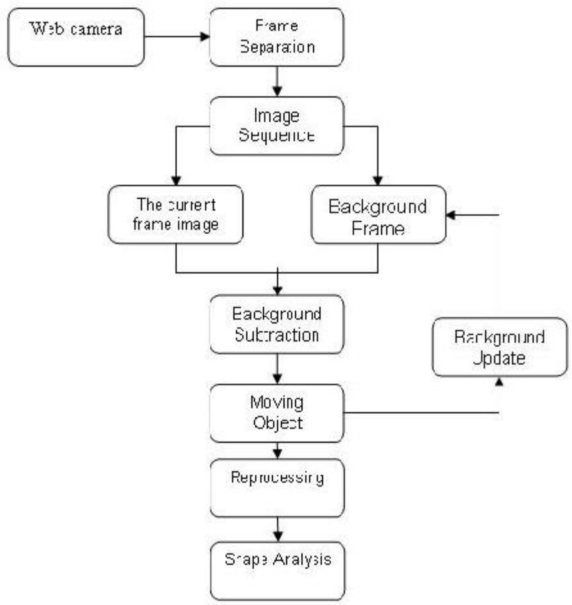

Figure 1. System Architecture

III. KERNEL DENSITY APPROXIMATION

Model-based approaches involving probability density function are common in background modeling and subtraction, and we employ Kernel Density Approximation (KDA) [3], [7], where a density function is represented with a compact weighted sum of Gaussians whose number, weights, means, and covariances are determined automatically based on mean shift mode-finding algorithm. In our framework, each visual feature is modeled by KDA independently and every density function is 1D. (Fig.2) By utilizing the properties of the 1D mean-shift mode-finding procedure, the KDA can be implemented efficiently because we need to compute the convergence locations for only a small subset of data.

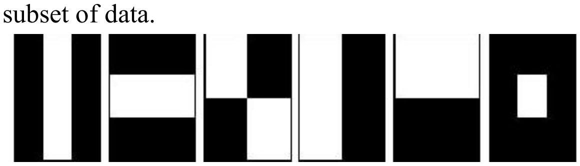

Figure 2. Background Modeling.

The background probability of each pixel for each feature is modeled with a Gaussian mixture density function.1 There are various methods to implement this idea, and we adopt KDA, where the density function for each pixel is represented with a compact and flexible mixture of Gaussians. The KDA is a density approximation technique based on mixture models, where mode locations (local maxima) are detected automatically by the meanshift algorithm and a single Gaussian component is assigned to each detected mode.

$$P(X1) = \frac{1}{n}\sum_{i=1}^{m} \eta(\mathbf{x}|txi, \textstyle\sum[t])$$

Model update is obtained by simply updating the buffer of the background values in fifo order by selective update in this way, "pollution" of the model (7) by foreground values is prevented. However, complete model estimation of $\sum t$ also requires the estimation of (which is assumed diagonal for simplicity). This is a key problem in KDE. In [7], the variance is estimated in the time domain by analyzing the set of differences between two consecutive values.

Clustering:

Cluster analysis is an unsupervised learning method that constitutes a cornerstone of an intelligent data analysis process. It is used for the exploration of inter-relationships among a collection of patterns, by organizing them into homogeneous clusters. It is called unsupervised learning because unlike classification (known as supervised learning), no a priori labeling of some patterns is available to use in categorizing others and inferring the cluster structure of the whole data. Intra-connectivity is a measure of the density of connections between the instances of a single cluster.

Hierarchical Clustering:

The hierarchical methods group data instances into a tree of clusters. There are two major methods under this category. One is the agglomerative method, which forms the clusters in a bottom-up fashion until all data instances belong to the same cluster. The other is the divisive method, which splits up the data set into smaller cluster in a top-down fashion until each cluster contains only one instance. Both divisive algorithms and agglomerative algorithms can be represented by dendrograms. Both agglomerative and divisive methods are known for their quick termination. However, both methods suffer from their inability to perform adjustments once the splitting or merging decision is made.

Density-based Clustering:

Density-based clustering algorithms try to find clusters based on density of data points in a region. The key idea of density-based clustering is that for each instance of a cluster the neighborhood of

ISSN-2320-7884 (ONLINE)
ISSN-2321-0257 (PRINT)

a given radius (*Eps*) has to contain at least a minimum number of instances (*Min Pts)*. One of the most well known density-based clustering algorithms is the DBSCAN [9]. DBSCAN separate data points into three classes

- Core points. These are points that are at the interior of a cluster. A point is an interior point if there are enough points in its neighborhood.
- Border points. A border point is a point that is not a core point, i.e., there are not enough points in its neighborhood, but it falls within the neighborhood of a core point.
- Noise points: A noise point is any point that is not a core point or a border point.

Grid-based Clustering:

Grid-based clustering algorithms first quantize the clustering space into a finite number of cells (hyper-rectangles) and then perform the required operations on the quantized space. Cells that contain more than certain number of points are treated as dense and the dense cells are connected to form the clusters. Some of the grid-based clustering algorithms are: STatistical INformation Grid-based method STING Wave Cluster, and Clustering in Quest CLIQUE.

Classification:

Once an appropriate mechanism for density approximation is built, the next step is to determine a classification mechanism for the observed data (Fig.3). Classification may be performed by thresholding on the probability of a new observation to belong to the background. However, two observations need to be taken into account:

• The threshold should be adaptive and determined based on the uncertainty or spread of the background distribution at a particular pixel (called *entropy* in information theory).

• Any available prior information about the foreground distribution should be utilized.

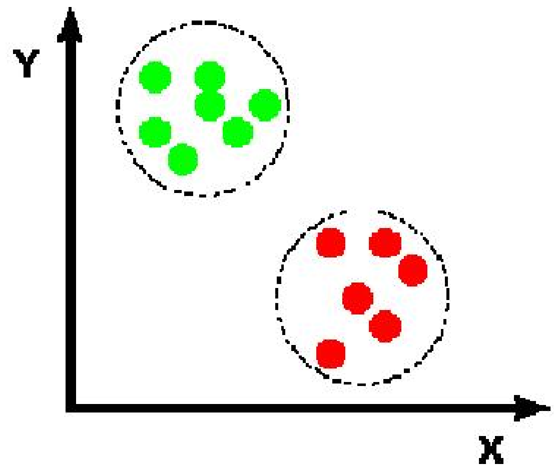

Figure 3. Classification

IV. EXPERIMENTAL EVALUATION

In video surveillance systems, stationary cameras are typically used to monitor activities at outdoor or indoor sites. Since the cameras are stationary, the detection of moving objects can be achieved by comparing each new frame with a representation of the scene background. This process is called background subtraction and the

Figure.4 Background Subtraction Modeling

Scene representation is called the background model. Typically, background subtraction forms the first stage in an automated visual surveillance system (Fig.4). Results from background subtraction are used for further processing, such as tracking targets and understanding events.

Illumination changes:

Gradual change in illumination, as might occur in outdoor scenes due to the change in the location of the sun;

• Sudden change in illumination as might occur in an indoor environment by switching the lights on or off, or in an outdoor environment by a change between cloudy conditions.

• Shadows cast on the background by objects in the background itself (e.g., buildings and trees) or by moving foreground objects.

Motion changes:

• Image changes due to small camera displacements (these are common in outdoor situations due to wind load or other sources of motion which causes global motion in the images);

• Motion in parts of the background, for example, tree branches moving with the wind or rippling water.

K-Means Algorithm

The continuous *k*-means algorithm is faster than the standard version and thus extends the size of the datasets that can be clustered. It differs from the standard version in how the initial reference points are chosen and how data points are selected for the updating process. In the standard algorithm the initial reference points are chosen more or less arbitrarily. In the continuous algorithm reference points are chosen as a random sample from the whole population of data points. If the sample is sufficiently large, the distribution of these initial reference points should reflect the distribution of points in the entire set. Another difference between the standard and continuous *k*-means algorithms is the way the data points are treated. During each complete iteration, the standard algorithm examines all the data points in sequence. In contrast, the continuous algorithm examines only a random sample of data points. If the dataset is very large and the sample is representative of the dataset, the algorithm should converge much more quickly than an algorithm that examines every point in sequence.

- For each center we identify the subset of training points (its cluster) that is closer to it than any other center;
- The means of each feature for the data points in each cluster are computed, and this mean vector becomes the new center for that cluster.

These two steps are iterated until the centers no longer move or the assignments no longer change. Then, a new point x can be assigned to the cluster of the closest prototype.

Experiment results:

Two types of experiments have been performed on the modified K-means. The first type focuses on filtering out spike noise, while the second tests the algorithm's ability to remove Gaussian noise with different blur radii.

a.Removal of Spike Noise:

A flattened image with multicolor Illuminated Line Segment based Markers and noise in the form of spurious black pixels can be observed(Fig 5). The noise data points are circled for the purpose of clarity. The same image is shown after the modified Kmeans algorithm has cleaned the data. Further experiments on synthetic data from a standard ping-pong ball style marker show that the modified K-means algorithm also is capable of cleaning this data successfully.

b.Removal of Gaussian Noise:

The modified K-means algorithm is then used to remove this noise and recapture the data. The results of the experiments show that the Gaussian noise is completely removed regardless of the radius. It shows that the number of data points recaptured naturally decreases as the radius of the Gaussian noise

increases. However, it is also shown that the degradation of performance occurs gradually, as oppose to abruptly, when the radius is increased up to 2.5 pixel.

Median Filter:

Various authors have argued that other forms of temporal average perform better than that shown in Lo and Velastin in proposed to use the median value of the last *n* frames as the background model. Cucchiara *et al.* in [4] argued that such a median value provides an adequate background model even if the *n* frames are sub sampled with respect to the original frame rate by a factor of 10. In addition, [4] proposed to compute the median on a special set of values containing the last *n,* sub-sampled frames and *w* times the last computed median value. This combination increases the stability of the background model.

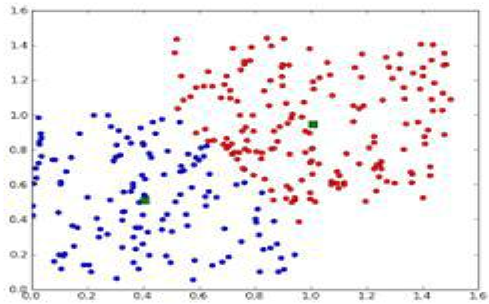

Figure 5. K-means Clustering

The main disadvantage of a median-based approach is that its computation requires a buffer with the recent pixel values. Moreover, the median filter does not accommodate for a rigorous statistical description and does not provide a deviation measure for adapting the subtraction threshold.

Image Variations

Their main statement is that neighboring blocks of pixels belonging to the background should experience similar variations over time. Although this assumption proves hue for blocks

belonging to a same background object(Fig 6) (such as an area with tree leaves), it will evidently not hold for blocks at the border of distinct background objects (this is likely the cause of several false detections shown in [12], appearing at the borders of different background objects). Instead of working at pixel resolution, it works on blocks of N x N pixels treated as an N2 component vector. This trades off resolution with better speed and stability.

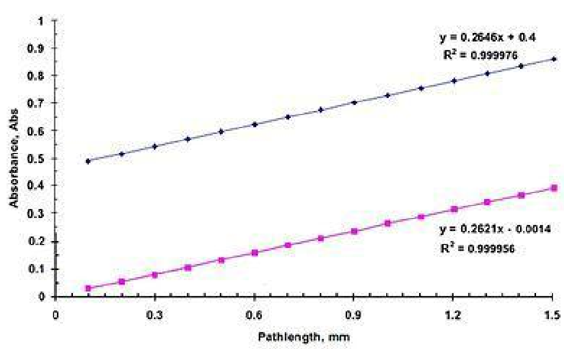

Figure.6 Background Subtraction Verification

- For each block, a certain number of time samples is acquired; the temporal average is fust computed and the differences between the samples and the average are called the image variations;
- The N2 x N2 covariance matrix is computed with respect to the average and an eigenvector transformation is applied reducing the dimensions of the image variations from N2 to K. A neighboring block, ***U,*** is considered, with its current input value; the corresponding current Eigen image variation is computed, called *z (u,i)*
- The L-nearest neighbors to *tu* in the eigenspace, zed, are found and zu expressed as their linear interpolation

Alerting System:

After detecting the changes in video frames, we are alerting the central control unit or the user through SMS using the GSM Modem. A GSM modem is a wireless modem that works with a GSM wireless network. A wireless modem behaves like a dial-up modem. The main difference between them is that a dial-up modem sends and receives data through a fixed telephone line while a wireless modem sends and receives data through radio waves. A GSM modem is a specialized type of modem which accepts a SIM card, and operates over a subscription to a mobile operator, just like a mobile phone. From the mobile operator perspective, a GSM modem looks just like a mobile phone. When a GSM modem is connected to a computer, this allows the computer to use the GSM modem to communicate over the mobile network. While these GSM modems are most frequently used to provide mobile internet connectivity, many of

them can also be used for sending and receiving SMS messages.

V. CONCLUSION

We have introduced 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. KDA is used to represent a probability density function of the background for RGB, gradient, and Haar-like features in each pixel, where 1Dindependent density functions are used for simplicity. A feature selection algorithm can be seen as the combination of a search technique for proposing new feature subsets, along with an evaluation measure which scores the different feature subsets. The server will pass the small message like"Intruder Found". After receiving the text message the owner can view the detected image

ISSN-2320-7884 (ONLINE)
ISSN-2321-0257 (PRINT)

by using GPRS supported mobile using. This entire application was deployed in web logic server so it will give response to client requests.

References

1. Abramson. On bandwidth variation in kernel estimates- a square root law. *The Annals of Statistics*, 10:1217– 1223, 1982.

2. P. Anandan. A computational framework and an algorithm for the measurement of visual motion. *IJCV*, 2(3):283–310, January 1989.

3. J.L. Barron, D.J. Fleet, and S.S. Beauchemin. Performance of optical flow techniques. *IJCV*, 12(1):43–77, February 1994.

4. L. Breiman, W. Meisel, and E. Purcell. Variable kernel estimates of multivariate densities. *Technometrics*, 19:135–144, 1977.

5. H. Chen and P. Meer. Robust computer vision through kernel density estimation. In *ECCV*, pages I: 236–250, Copenhagen, Denmark, May 2002.

6. D. Comaniciu. Nonparametric information fusion for motion estimation. In *CVPR*, Madison, Wisconsin, June 2003.

7. G. Doretto, A. Chiuso, Y.N. Wu, and S. Soatto. Dynamic textures. *IJCV*, 51(2):91–109, February 2003.

8. B. Han and L.S. Davis, “Adaptive Background Modeling and Subtraction: A Density-Based Approach with Multiple Features” Intelligent Video

9. Surveillance Systems and Technology, Y. Ma and G. Qian eds., ch. 4, pp. 79- 103, CRC Press, 2010.

10. 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.

11. 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.

12. C. 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.

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

14. 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. An investigation of five procedures for detecting nested cluster structure. In, Data Science, Classification, and Related Methods, edited by C. Hayashi, N. Ohsumi, K. Yajima, Y. Tanaka, H. Bock, and Y. Baba. Tokyo: Springer-Verlag.

17. Jerzy W. Grzymala-Busse, Ming Hu, A Comparison of Several Approaches to Missing Attribute Values in Data Mining, Rough Sets and Current Trends in Computing : Second International Conference, RSCTC 2000 Banff, Canada.

18. Guha, S, Rastogi, R., Shim K. (1999), "ROCK: A Robust Clustering Algorithm for Categorical Attributes", In the Proceedings of the IEEE Conference on Data Engineering.

19. Guha, S., Rastogi, R., Shim K. (1998), "CURE: An Efficient Clustering Algorithm for Large Data sets", Published in the Proceedings of the ACM SIGMOD Conference.

20. Hinneburg A. and Keim D. (1998), Anefficient approach to clustering in large multimedia data sets with noise, In Proceedings of the 4th International Conference on Knowledge Discovery and Data Mining, pages 58-65 (1998).

21. Huang Z. (1998), Extensions to the k-Means Algorithm for Clustering Large Data Sets with Categorical Values, Data Mining and Knowledge Discovery 2, 283–304 (1998).