Distance Metrics based Vehicle Object Identification in Dynamic Vision

M.Sankari
Department of Computer Applications,
Nehru Institute of Engineering and Technology,
Coimbatore, INDIA.

C. Meena
Avinashilingam University for Women,
Coimbatore, INDIA.

Abstract— **Vehicle object identification is a challenging issue in the visual surveillance. In recent years, video monitoring and surveillance system have been widely utilized for traffic monitoring and management. In this paper, we propose an algorithm to identify the moving objects from the sequence of video frames which contains dynamically changing backgrounds in the noisy environment. Reference to our previous surveillance skeletanization works, here we have proposed a methodology to identify an object using distance metrics such as Hamming and Euclidian distance metrics. The recent vehicle recognition methods could fail to identify the objects and produce more false acceptance rate (FAR) or false rejection rate (FRR) however our research recommends a method for object identification using weighted distance to extract features in order to obtain robust identification in the noisy environment.**

Keywords- Distance transformations, Euclidian Distance, false acceptance rate, false rejection rate, Hamming Distance, Skeletanization, Traffic video sequences, weighted distance.

I. INTRODUCTION

In visual surveillance model, estimating the dynamic background and detecting the object from the noisy environment is a computationally challenging problem. Contribution of this paper is to identify the object from the multi model background using distance metrics. For that we need to detect and extract the foreground of the object from the background image. After detecting the foreground object there may a large number of possible degradations that an image can suffer. After removing the blur, skeletanization technique is used to identify an object. The key issue of the pattern identification problem is the relation between inter-class and intraclass variability. That is, classes can be efficiently differentiated only if the variability between features of a given class is less than the variability between other classes. In the vehicle identification process, variability of vehicle intra-class features is less than the inter-class variability. However, inter-class variability is limited because different classes hold the same basic set of features. The proposed vehicle identification system initially acquires different possibility of patterns of the vehicle images in the dissimilarity distance of capturing. Subsequently, background subtraction, vehicle segmentation, and skeletanization are performed to make the vehicle identification suitable for feature extraction process. Finally, the vehicle discriminator design phase classifies the vehicles. The entire process is automatic and uses computation time that scales according to the size of the input video sequence. The overall organization of this paper is as follows: Section 2 summarizes the literature on the previous research. The algorithm of proposed vehicle identification system is illustrated in section 3. Experimental details and results are given in 4. Section 5 consists of our concluding remarks

II. OVERVIEW OF THE RELATED WORK

Detection of moving objects in video can be achieved in three main approaches: Temporal difference, optical flow, and background subtraction. In temporal difference, the image difference of two consecutive image frames are obtained [6][7]. However, this approach has some limitations such as visual homogeneity requirement and its effectiveness depends on the speeds of moving objects [3]. Optical flow method was developed to obtain effective background modification, which bases on the detection of intensity changes [3]. However, illumination change due to weather or sun-light reflections decreases its effectiveness. It is also computationally inefficient [3]. The third method, background subtraction, is the mostly seen method in the literature for effective motion tracking and moving object identification [4, 5]. In background subtraction, background can be static, in which a fixed background is obtained beforehand and used in the entire process or dynamic, in[1], background is dynamically updated with changing external effects like weather. Static background may not be effective in most applications, many methods include dynamic background subtraction. In [8], the background is detected dynamically by using dynamic threshold selection method.

III. PROPOSED WORK

The proposed system extracts foreground objects such as people, objects, or events of interest in variety of noisy environment. This is an extension work of our previous method [1,2]. The schematic flow of the proposed algorithm is shown in Fig.1.

Typically, these systems consist of stationary cameras placed in the highways. These cameras are integrated with, intelligent computer systems that perform preprocessing operation from the captured video images and notify human operators or trigger control process. The objective of this real-time motion detection and tracking algorithm is to provide low-level functionality for building higher-level recognition capabilities.

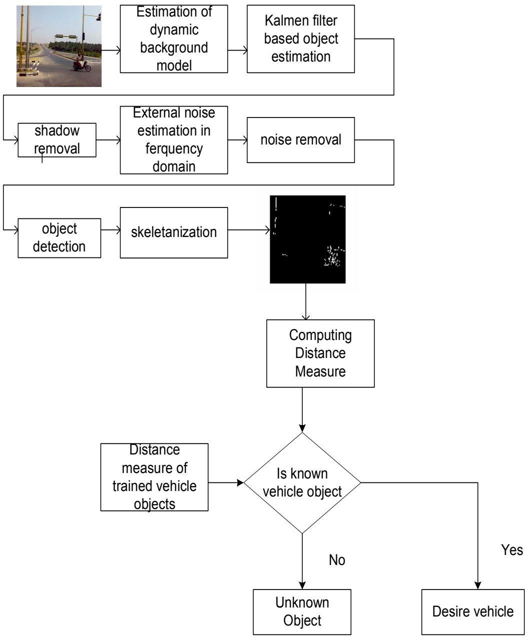

Figure 1. Schematic flow of the proposed algorithm to identify an objects in the noisy environment.

The objective of this work is to identify an object from a dynamic background using skeletanization which provides low-level functionality for building higher-level recognition capabilities.

A. Preprocessing

Preprocessing is the key step and the starting point for image analysis, due to the wide diversity of resolution, image format, sampling models, and illumination techniques that are used during acquisition. In our method, preprocessing step was done by statistical method using adaptive median filter. The resultant frames are then utilized as an input for the background subtraction module. Image *I(x,y)* at time *t.* Basically, impulse noise is a major artifact that affects the sequence of frame in the surveillance system. For this reason to estimate the background in the noisy environment we have proposed an adaptive median filter (AMF). The AMF can be used to enhance the quality of noisy signals, in order to achieve better forcefulness in pattern recognition and adaptive control systems. These noise pixels are then substituted by the median pixel value of the pixels in the neighborhood that have passed the noise labeling test [16].

B. Foreground Detection

In this module estimated background and foreground mask images are used as an input for further processing. Thus, we use grayscale image sequences as input. Elements of the scene and the sizes of the traffic objects (vehicles and pedestrians) are unknown. The Foreground detection is done by using accumulative difference method. The basic idea in background adaptation is to integrate the new incoming information into the current background image using a Kalman filter.

C. Shadow Removal

Shadows appear as surface features, when they are caused by the interaction between light and objects. This may lead to problems in scene understanding, object segmentation, tracking, recognition, etc. Because of the undesirable effects of shadows on image analysis, much attention was paid to the area of shadow detection and removal over the past decades and covered many specific applications such as traffic surveillance. In this paper, 8-neighborhood gray clustering method is used to define the precise shadow and remove it.

D. Noise removal

In regular practice due to the camera noise and irregular object motion, there are some noise regions existed in both the object and background regions. In our method we have incorporated Gaussian noise with the acquired image and propose a solution to see how the background subtraction module would behave while the traditional background algorithms are not providing the significant results. The focus is on the background subtraction module because image noise mostly impacts the foreground extraction process. If the foreground objects are not detected well, the rest of the modules will possibly fail at their tasks. The proposed method is compared with Autocorrelation, Wiener filter, Lucy-Richardson filtering methods.

E. Detection

After post-processing, the image is compared with the one of the original frames (usually, the first frame). If the pixels are less than certain threshold, then they are ignored. Otherwise, they are replaced by the pixels of original image. This resulting image will be consisting of the moving object ignoring the background and hence satisfying our requirement.

F. Skeletanization

The word morphology refers to the scientific branch that deals the forms and structures of animals/plants. Morphology in image processing is a tool for extracting image components that are useful in the representation and description of region shape, such as boundaries and skeletons.

In our method, sequential thinning algorithms, contour points are examined for deletion in a predetermined order Thinning is an iterative shrinking process: each pixel is analyzed. Morphological operators employ a method that deletes simple points. Deleting a simple point does not change the connectivity properties of the set. If certain criteria are satisfied, that pixel is deleted. Pixels removal continues until no changes. Skeletonization is performed using the following steps:

Step 1: Compute the digital Euclidean distance transform of the object and form a distance map. Let the pixels within a two-dimensional digital image $I(x, y)$ be divided into two classes – object pixels and background pixels.

$$I(x, y) \in \{O, B\} \qquad (1)$$

The distance transform of this image, $I_{dm}(x, y)$ then labels each object pixel of this binary image with the distance between that pixel and the nearest background pixel. Mathematically,

$$I_{dm}(x, y) = \begin{cases} 0 & I(x,y) \in \{B\} \\ \min(\|x - x_0, y - y_0\|), \forall I(x_0, y_0) \in B & I(x,y) \in \{O\} \end{cases} \qquad (2)$$

where $\|x, y\|$ is a two-dimensional distance metric.

Step 2: The distance transform finds the distance between all non-boundary foreground points in an image and their nearest boundary points. The distance between points may be defined using a 4-connected neighborhood or an 8-connected neighborhood.

S = the set of object pixels

where S is a Connected Component if for each pixel pair $(x_1, y_1) \in S$ and $(x_2, y_2) \in S$ there is a path passing through X-neighbors in S (X = 4, 8). S may contain several connected components. The result of applying a distance transformation is a distance map.

Connected Component Algorithm: Two passes over the image.

Phase 1:

Scan the image pixels from left to right and from top to bottom.

For every pixel P of value 1 (an object pixel), test top and left neighbors (4-neighbor metric).

- If 2 of the neighbors equals 0: assign a new mark to P.

- If 1 of the neighbors equals 1: assign the neighbor's mark to P.

- If 2 of the neighbors equals 1: assign the left neighbor's mark to P and note equivalence between 2 neighbor's marks.

Phase 2:

Divide all marks in to equivalence classes (marks of neighboring pixels are considered equivalent).

Replace each mark with the number of its equivalence class.

Step 3: A point within the object which does not have at least one background point as an immediate neighbor cannot be removed, since this would create a hole. Therefore, the only potentially removable points are on the border of the object. Once a border point is removed, only its neighbors may become removable All the boarder points are arranged according to the ascending ordering of distance. This set is denoted as S.

```
if(P is removable)
  if(P is not an end point)
    Remove P
   For all neighbors S of P
    If S is removable
       Insert S into an array
       Else mark P as a skeletal point
    End
  End
  End
```

G. Object matching

1. Hamming distance

In the object matching process, to authenticate a genuine vehicle, object feature sets are treated as trained sets and stored in the encrypted file. Verification objects are represented as test sets. At any instance when no test set is present, the propability density for a trained set is normal, i.e., $p(x/\omega_1) \sim K(\mu_1, \sigma^2)$, when the test set is present, the density is $p(x/\omega_2) \sim K(\mu_2, \sigma^2)$ [17]. Decision threshold X will determine the probability of a hit and of a false alarm. The detector employs a threshold value for determining whether the test set is present or not. The discriminability is defined as in Eq.3.

$$d' = \frac{|\mu_2 - \mu_1|}{\sigma}, \qquad (3)$$

where d' is discriminability factors, μ_1 is a mean value when training set present, μ_2 is a mean value when test set present and σ is a standard deviation of feature set. Object trained and test set are compared to determine if the test set belongs to the intra class or inter class of object set. Object matching is computed based on minimum distance between two object feature sets that is defined as, $\min\{WD(VFC(x_{trained}), VFC(x_{test}))\}$, where represents weighted distance in between two object feature sets x_1 and x_2. The weighted distance is calculated using Hamming distance operation that provides faster object matching process as defined in Eq. 4 [9].

$$WD(VFC(x_{trained}), VFC(x_{test})) = \frac{(VFC(x_{trained}) \oplus VFC(x_{test}))}{N}, \qquad (4)$$

where N is the number of bits in the object feature set $\oplus$ denotes Hamming distance operation. The Weighted Distance (WD) represents the number of error bits between two object classes. The distance between intra (same class) and inter(different class) object features are discriminated by the constrains $0 \leq \text{WD} \geq 0.20$ and WD $>$ 0.20, respectively .

2. Euclidean Distance

In order to distinguish a candidate's vehicle class, vehicle feature is represented as a 1D feature vector of size 64 per candidate. Feature set $FS_i = \{f_1, f_2, ...f_{64}\}$ is used for the recognition process. These sensitive features are the mean value of resultant Gabor kernel convolution operation. During the testing, standard deviation values of known vehicle patterns are calculated from the set of known mean features. In this proposed approach, the Euclidean norm distance measurement (ENDM) discriminator is employed to separate intra- and interclass variations of iris patterns. To authenticate a known vehicle , vehicle feature sets are treated as known sets that are already stored in the encrypted file and verification candidates' vehicle features are represented as unknown sets. The same candidate's vehicle feature codes may vary due to external noises, lighting etc. During any instant when no unknown set is present, the probability density for a known set is normal, that is, $p(x \mid \omega_1) \sim K(\mu_1, \sigma^2)$ when the unknown set is present, the density is $p(x \mid \omega_2) \sim K(\mu_2, \sigma^2)$. Any decision threshold X^* will determine the probability of a known and unknown vehicle object. As per signal detection theory, ENDM is described as,

$$ENDM(FS_i) = \sum_{i=1}^{n} (FS_i \mu_2 - FS_i \mu_1)^2 / FS\mu_i^2 \qquad (5)$$

$$FS\sigma_i = \left[\left[NM \sum_{i}^{N} \sum_{j}^{M} I(x_i, y_i)^2 - \left(\sum_{i}^{N} \sum_{j}^{M} I(x_i, y_j) \right)^2 \right] \div \left[N(N-1)M(M-1) \right] \right]^{1/2} \qquad (6)$$

where $FS_i\mu_2$ is an unknown vehicle feature code, $FS_i\mu_1$ is a known vehicle feature code and $FS\sigma_i$ is the standard deviation of the known feature code. In the proposed approach, the threshold value of the decision boundary of ENDM ranges from 0·0 to 0·6. If the boundary is raised up to 0·6, the probability of a hit is present, i.e., a genuine candidate has been recognized, otherwise it is an unknown vehicle.

IV. Implementation and Performance Analysis

Implementation has been done in Java. This system was implemented on an Intel Pentium IV 280 GHz PC. We have tested the system on image sequences on different scenarios like traffic junction intersection, highways and other real-time situations.

Real life traffic video sequence is used to demonstrate the vehicle tracking from traffic video sequences using the proposed framework. All the videos chosen for vehicle tracking have same light intensity and have been taken during day time. We convert the colour video frames to gray scale images. Automatic monitoring visual surveillance system implementation needs to detect and identify vehicles using automatic background extraction and skeletanization. Digital camera used to take shots. The camera placed over the highway directly.

Automatic monitoring visual surveillance system implementation needs to detect vehicles using automatic background extraction. Background subtraction is the main step for vehicle detection. Digital camera used to take shots. The camera placed over the highway directly. It shots eight frames per second. After applying the threshold, φ to the absolute difference we got the binary moving objects hypothesis mask. There is no noise information then the Wiener and other filters do a poor job at realizing the original, non- degraded, image. However, the Lucy-Richardson filter works really well, despite having no information about the noise in the image.

In the real-life noise information, the Wiener filter gives better result than other filters. Fig.2 shows the foreground detected objects obtained after background subtraction, shadow and noise. The automatic background extraction results are very good and promising. The most effective parameters that are playing a main role for automatic background extraction are the threshold level. This threshold is used to extract the moving vehicles from the background. Matlab built-in function has been employed for the evaluation of the threshold.

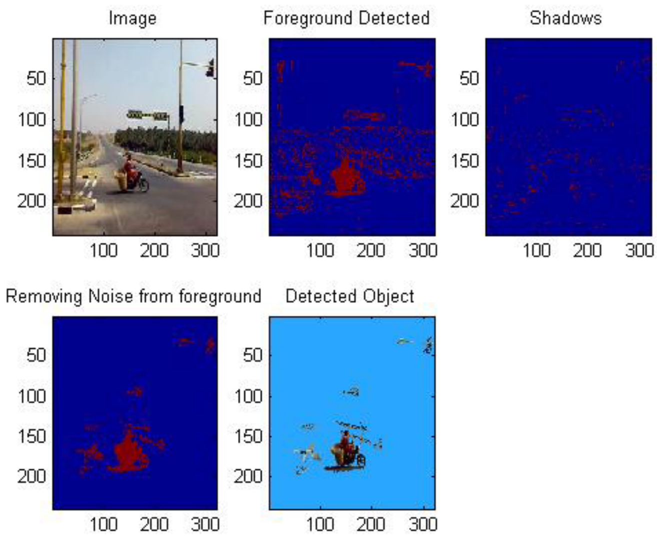

Figure 2. Sequence of steps in the foreground object detection process.

We have introduced a novel algorithm for computing pixels of skeletal image graphs which is robust, accurate, computationally efficient, and invariant to Euclidean transformations. The essential idea is to combine a divergence computation on the gradient vector field of the Euclidean distance function, with a thinning process that preserves topology.

V. CONCLUSION

The experimental results of this approach lead to detect moving vehicles efficiently in dynamic background. This paper mainly focuses on identification of an object using distance metrics in dynamic background traffic images such as buses and two wheelers. Results show that this approach can be adopted robustly in the automatic traffic monitoring system. Failure detection resulted from occluding large vehicles with small ones and the far moving vehicles that appear as a point in the image. The present approach clearly specifies that at iterative step three, in all cases, a good shape is obtained automatically.

REFERENCES

[1] M. Sankari, C. Meena "Adaptive Background Estimation and object detection applying in Automated Visual surveillance" International Journal of Computer science and Information Security, July 2010, Vol. 8, No. 4. pp.275-279.

2] M. Sankari, C. Meena, “Estimation of Dynamic Background and Object Detection in Noisy Visual Surveillance” Int. Jour. of Adv. Comp. Sci. and Appl., Vol.2, No. 6, pp.77-83, June 2011.

[3] X. Ji, Z. Wei, Y. Feng, Effective vehicle detection techniques for traffic surveillence systems, J. Vis. Commun. Image R. 17, 647-658, 2006

[4] J. Zhou, D. Gao, D. Zhang, Moving vehicle detection for automatic traffic monitoring, IEEE Trans. Vehic. Tech., 56(1), 51-59, 2007

[5] J. K. Cai Q.i Aggarwal, Tracking human motion in structured environments using a distributed camera system, IEEE Trans. Pattern Anal. Mach. Intell. 21 (12), 1241-1247, 1999

[6] M. T. Lo´pez, Ferna´ndez-Caballero, A., Mira, J., Delgado, A. E., & Ferna´ndez, M. A. Algorithmic lateral inhibition method in dynamic and selective visual attention task: Application to moving objects detection and labelling. Expert Systems with Applications, 31(3), 570–594, 2006

[7] M. Dubuisson, & A. Jain, Contour extraction of moving objects in complex outdoor scenes. International Journal of Computer Vision, 14, 83–105, 1995

[8] M. Fathy and M. Y. Siyal, A window-based image processing technique for quantitative and qualitative analysis of road traffic parameters, IEEE Trans. Veh. Technol., 47, (4),1342–1349, 1998.

[9]. Bremananth R, Chitra A, “New methodology for a person identification system”, S¯adhan¯a Vol. 31, Part 3, June 2006, pp. 259–276.