

# Object Matching using Skeletonization based on Hamming Distance

M. Sankari

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

C. Meena

Avinashilingam University for Women,  
Coimbatore, INDIA.

## ABSTRACT

Identification of an object from a dynamic background is a challenging process in computer vision and pattern matching research. The proposed algorithm identifies moving objects from the sequence of video frames which contains dynamically changing backgrounds in the noisy environment. In connection with our previous work, here we have proposed a methodology to perform skeletonization of an object and identifies it. The recent vehicle recognition methods could fail to recognize an object and produce more false acceptance rate(FAR) or false rejection rate (FRR). This paper recommends a method for object identification using weighted distance to extract features. Experimental results and comparisons using real data demonstrate the pre-eminence of the proposed approach.

## Keywords

Distance transformations, false acceptance rate, false rejection rate, Skeletonization, Traffic video sequences, weighted distance.

## 1. INTRODUCTION

In visual surveillance model, estimating the dynamic background and detecting the object from the noisy environment is a computationally challenging problem. Our main target is to identify the object from the multi model background using background subtraction, shadow removal, noise removal techniques and skeletonization of an object. Skeleton is an important shape descriptor for object representation and recognition since it can preserve both topological and geometrical information. It has important applications in object representation and recognition, such as image retrieval and processing. Our main target is to identify the object from the multi model background. For that we need to detect and extract the foreground object from the background image. After detecting the foreground object there may a large number of possible degradations that an image can suffer. Common degradations are blurring, motion and noise. After detecting the object image, converting it into its spatial frequencies, developing a point spread function (PSF) to filter the image with, and then converting the filtered result back into the spatial domain to see if blur was removed. After removing the blur skeletonization technique is used to identify an object. This can be done in several steps. The entire process is automatic and uses computation time that scales according to the size of the input video sequence.

## 2. OVERVIEW OF THE RELATED WORK

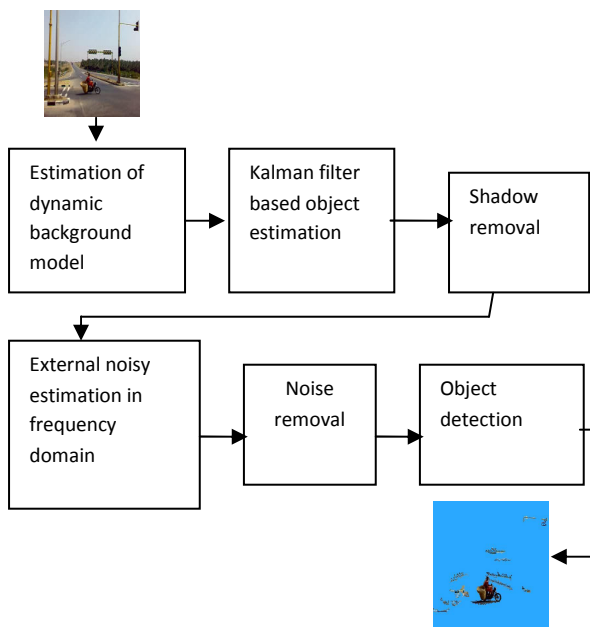
Many works have been proposed in the literature in order to attain a solution to an efficient and reliable background subtraction. To detect moving objects in a dynamic scene, adaptive background subtraction techniques have been developed [1] [2] [3]. In [4], a work has presented on a novel background subtraction algorithm that is capable of detecting objects of interest while all pixels are in motion[5]. Shapes can be segmented by considering their boundary or interior. Approaches that are based on the boundary usually define segment borders at object-surface concavities as a solution to an efficient and reliable skeletonization. Applications which utilize curve skeletons for shape analysis and processing include segmentation [6] [7], object matching and retrieval [8], skeletonization algorithms based on the distance map are available in the literature [9]. They used to produce skeletons whose pixels are correctly labeled according to their distance from the input image. Approaches based on Voronoi techniques preserve topology, but heuristic pruning measures are introduced to remove unwanted edges. Methods based on Euclidean distance functions can localize skeletal points accurately, but often at the cost of altering the object's topology [10]. The responsibility of object recognition is to identify whether an object of a specific class is present in an image or not. Mathematical morphology can be used for shape description, typically in connection with region skeleton construction and this approach will unify many other approaches to skeletonization [11]. Thinning is a widely used transformation in the fields of image analysis and shape recognition. The outcome of the thinning process is a skeleton. The purpose of reducing images to their skeletons is to reduced amount of data as well as the shape analysis can be done more easily on skeletons. Different skeletonization techniques like morphological transform are used to get a skeleton [12][13][14][15]. The responsibility of object recognition is to identify whether an object of a specific class is present in an image or not. All thinning algorithms can be classified as either iterative or non-iterative. In iterative methods, thinning algorithms produce a skeleton by examining and deleting contour pixels through an iterative process in either sequential or parallel way. Parallel algorithms may also be further classified according to their performance, i.e., in 4-, 2-, or 1-subcycle manners.

In our paper, we propose a novel simple method that exploits all these features, combining them so as to efficiently identify an object using skeletonization. The resulting method proves to be accurate and reactive and, at the same time, fast and flexible in

the applications. The paper is organized as follows. The proposed morphological skeleton transform is described in section III. Section IV contains experimental results and discussions and conclusions are given in section V.

### 3. 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 [5]. 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 detect the objects in the noisy environment.**

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

#### 3.1 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].

#### 3.2 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

#### 3.3 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.

#### 3.4 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.

#### 3.5 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.

#### 3.6 Skeletonization

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.

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

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,

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

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 phases 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.

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

**3.7 Object matching**

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 probability 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.

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

Where  $d'$  is discriminability factors,  $\mu_1$  is a mean value when training set present,  $\mu_2$  is a mean value when test set present and  $\sigma$  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 [18].

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

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  $0 \leq WD \leq 0.20$  and  $WD > 0.20$ , respectively .

**4. IMPLEMENTATION AND PERFORMANCE ANALYSIS**

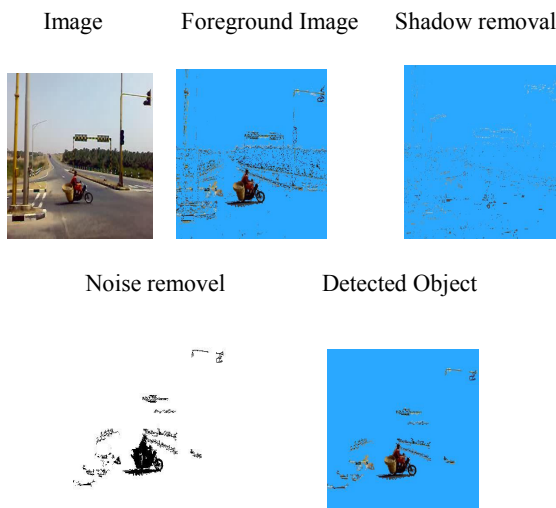
Real life traffic video sequence is used to demonstrate the vehicle tracking from traffic video sequences using the proposed framework. All the videos chosen for vehicle tracking have same light intensity and have been taken during day time. We convert the color video frames to gray scale images. Automatic monitoring visual surveillance system implementation needs to

detect and identify vehicles using automatic background extraction and skeletonization. Digital camera used to take shots. The camera placed over the highway directly.

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.

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, 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. We have illustrated the approach on a variety of 2D shapes. This is shown in Fig. 3.



**Figure 3. Representation of before and after Skeletanization process of Vehicle images.**

## 5. CONCLUSION

The experimental results of using this approach lead to detect moving vehicles efficiently. This algorithm has been implemented and evaluated experimentally using natural traffic images. It gives promising and effective results. In this approach the background subtraction and edge detection are used. It is mainly focused on the autocorrelation method. And then an adaptive algorithm is applied to autocorrelation. It could be improved and used as a basis for automatic traffic monitoring. Failure detection resulted from occluding large vehicles with small ones and the far moving vehicles that appear as a point in the image. The skeletonization and Object matching processes are effectively carried using the proposed. In near future, rotation invariant feature extraction and vehicle object matching methods will be suggested for classifying two-dimensional invariant transformation vehicle objects.

## 6. REFERENCES

- [1] M. Cristani, M. Bicegi, and V. Murino, "Integrated Region- and Pixel-based Approach to Background Modeling", Proceedings of the MOTION, 2002.
- [2] Eng, J. Wang, A. Kam, and W. Yau, "Novel Regionbased Modeling for Human Detection within High Dynamic Aquatic Environment," Proceedings on CVPR, 2004.
- [3] P. KaewTraKulPong and R. Bowden, "An Improved Adaptive Background Mixture Model for Real-time Tracking with Shadow Detection," In Proc. 2nd European

- Workshop on Advanced Video Based Surveillance Systems, 2001.
- [4] Dongxiang Zhou, Hong Zhang and Nilanjan Ray, "Texture Based Background Subtraction," Proc. IEEE Int. Conf. on Information and Automation, Zhangjiajie, China, pp. 601-605, 2008.
- [5] 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.
- [6] T. K. Dey and J. Sun, "Defining and computing curve-skeletons with medial geodesic function," in Symp. on Geom. Proc., pp. 143–152, 2006.
- [7] O. K.-C. Au, C.-L. Tai, H.-K. Chu, D. Cohen-Or, and T.-Y. Lee, "Skeleton extraction by mesh contraction," ACM Trans. On Graph, vol. 27, no. 3, pp. 44:1–44:10, 2008.
- [8] H. Sundar, D. Silver, N. Gagvani, and S. Dickinson, "Skeleton based shape matching and retrieval," in Proc. IEEE Int. Conf. on Shape Modeling and Applications, p. 130, 2003.
- [9] C. Arcelli and G. Sanniti di Baja, "A width-independent fast thinning algorithm, IEEE Trans PAMI, 7(1985) 463-474.
- [10] L. Dorst, Pseudo-Euclidean skeletons, Proc. 8th Int. Conf. Pattern Recognition, Paris, France(1986) 286-288
- [11] R. L. Ogniewicz. Discrete Voronoi Skeletons. Hartung-Gorre, 1993.
- [12] J.M.Reinhardt and W. E. Higgins, "Efficient morphological shape representation," IEEE Trans. Image Process., vol. 5, no. 6, pp. 89–101, Jun. 1996.
- [13] P. Dimitrov, C. Phillips and K. Siddiqi , Robust and Efficient Skeletal Graphs, Proc.Computer Vision and Pattern Recognition Conf., Vol. 1, 2000, 417-423.
- [14] P.E. Trahanias, "Binary shape recognition using themorphological skeleton transform", Pattern Recognition 25(1992) 1277–1288.
- [15] V. Vijaya Kumar, et. al., "A new skeletonization method based on connected component approach," International Journal of Computer Science and Network Security, vol.8 no.2, pp.133-137, 2008.
- [16] M. Sankari, C. Meena," (IJACSA) International Journal of Advanced Computer Science and Applications, Vol. 2, No. 6, pp.77-83, June 2011.
- [17] Richard O Duda, Peter E.Hart and DavidG.Stork, "Patter Classification", John wiley Sons, 2001.
- [18] Bremananth R, A. Chitra, "Rotation Invariant Recognition of Iris", Journal of Systems Science & Engineering, Vol.17, No.1, pp.69-78, June 2008