International Journal of Research in Advent Technology, Special Issue, March 2019
E-ISSN: 2321-9637
Available online at www.ijrat.org

Template Matching Based On Color and Shape for Object Tracking

R.Matheswari[1], Dr.P.Aishwarya[2]
Professor and Head, Department of C.S.E[2],
M.E.S College of Arts[1], Commerce and Science, Atria Institute of Technology[2],
Bangalore.
matheswarir@gmail.com[1], aishwarya.p@atria.edu[2]

Abstract- Template matching plays as essential role in computer vision for pattern matching of objects. The first and foremost step is identification of objects. The proposed research work presents a frame differencing method for foreground detection of objects. Secondly, the method performs template matching of objects based on both color and shape descriptor features of objects to improve the robustness of point matching strategy. A combined frame differencing method with hybrid template matching method has been proposed in the study for performing matching of patterns using both spatial coordinates and the orientation of objects. The research work effectively detects humans and vehicles from sequence of images frames captured over time. From experimental results it is shown that the proposed research work improves template matching of objects and works effectively in case of cluttered background.

Keywords: Background subtraction, pattern matching, object detection, tracking.

1. INTRODUCTION

Object tracking is an active research area in computer vision. The process of detection and tracking objects faces a number of challenging issues like occlusion, dynamic shadows of objects, illumination changes, sudden disappearance of moving objects like passing clouds etc. There are number of object detection methods present and these are classified as either learning module or the template matching module. The complexity of multi-object tracking is caused by many situations: (1) when multiple objects move close to each other or present occlusion, identity management becomes a very complex operation; (2) complex occlusion with static objects in the scene may occur.

Solving the hijacking and data association problems in such situations is challenging especially in complete occlusion situations. The intuitive approach for performing multi-object tracking in videos is the use of multiple independent trackers simultaneously. The research work is implemented on the basis of the below mentioned primary steps.

- Identifying ROI in a frame using background subtraction and frame differencing approach.
- Template matching of objects based on color and shape features
- Such objects are tracked on frame-by-frame basis.

In this paper the objective of the research work is to examine and evaluate a variety of color models used for recognition of multicolored objects according to the following criteria:

1) Ability to adapt dynamically to a changing direction of objects.
2) Variations in the geometry of objects.
3) Variations in lightings and directions.

2. RELATED WORK

In the learning based approach, object signatures (e.g. the features used to describe the objects) are obtained through training using positive/negative samples [1], [2], [3], and object detection is often formulated as a problem of binary classification. In the template based approach, objects are described explicitly by templates and the task of object detection becomes to find the best matching template given an input image. The templates can be represented as intensity/color images [5] when the appearance of the objects has to be considered. Appearance templates are often specific and lack of generalization because the appearance of an object is usually subject to the lighting condition and surface property the object. Therefore, binary templates representing the contours of the objects are often used in object detection since the shape information can be well captured by the templates.

In [6], authors have proposed a mathematical morphological template matching approach for object detection in inertial navigation systems (INS). The major focus of the paper is to detect and track the ground objects. The flying systems equipped with camera were used to capture the photos of ground; to identify the objects. Their method is independent of the altitude and orientation of the object.

In [7], an approach for measuring similarity between visual images based on matching internal self-similarities. A template image is to be compared to another image. Measuring similarity across images can be complex, the similarity within each image can be easily revealed with simple similarity measure, such as SSD (Sum of Square Differences), resulting in local self-similarity descriptors which can be matched across images. Fahad Khan, et.al. [8] proposed the use of color attributes as an explicit color representation for object detection.

In [9], a new shape-based object detection scheme of extraction and clustering of edges in images using Gradient vector Girding (GVG) method is proposed that results a directed graph of detected edges. The algorithm used contains a sequential pixel-level scan, and a much smaller second and third pass on the results to determine the connectiveness.

In [10], Olson et al. considered each image point as a set of 3 elements (*xy*−coordinates and orientation), the Hausdorff distance was employed to the spatial coordinates and orientations separately. To facilitate the matching, the orientation is discretized and DTs are created for each of the discrete orientations, thus the computational complexity is increased.

In [11], the matching score between a template point and an image point is defined as a product of the spatial and orientation distances. The spatial distance is calculated using the conventional DT, e.g. [12], without considering the strength or magnitude of the edges.

3. BACKGROUND MODELLNG AND FRAME DIFFERENCING

Background subtraction is carried out to separate the foreground moving objects from the reference background image. This process is carried on pixel-by-pixel bass. With frame differencing, the method compares two successive frames to find out the difference or changes in pixel values. The frame differencing method is performed for two consecutive frames to identify all the moving objects. For every variations in the pixels, there is a change in the appearance of objects position in the image frames. The algorithm designed using this approach produces the location of moving objects as output. This location is used for obtaining an image template of rectangular size. A series of image templates are produced as foreground changes occurs from its current position.

Algorithm for object detection:

Assumption: The sequence of video frames captured are stored in memory. The current frame taken from the video sequence is assumed as *Fi*

The method receives current frame Fri, called the i^{th} frame as input and the *(i-3)th* frame (*Fri-3*) from the image memory buffer.

The memory buffer here, stores some set of mage frames for further use. The frame differencing function for the *ith* and *(i-3)th* frame is performed. The output image generated is formulated as follows:-,

dfi = Fri-3 – Fri

Here, dfi is the difference between i^{th} and $(i\text{-}3)^{th}$ frame. Hence, the method is said to be more reliable and is able to detect objects irrespective of their speed.

4. COLOR AND SHAPE BASED OBJECT RECOGNITION

A. Color feature descriptor for object recognition

The process of embedding color information with shape feature descriptor provides essential details for recognition of objects. A simple and efficient system generates color histogram for matching objects present in the mage. The other primary factors to be considered are a combined feature vector, illumination variations and compactness of objects. The research work combines color information based on the criteria aforementioned. The benefits of embedding shape and color features with the current framework demonstrate the effectiveness of template matching of objects for an efficient tracking module. A part-based identification with deformable part models and Efficient Subwindow Search (ESS) for identifying the location of objects are used.

The process of comparing image template with another image is essential. Therefore, similarity measure across multiple or sequence of images are too complex. Hence, the method performs matching strategy within each image using simple similarity measure, called as SSD (Sum of Square Differences) model. This further results in local feature descriptor within the image frame for matching across sequence of images. Fig.1 Hybrid algorithm to detect all person called as simpsons using both color and shape in the image frame.

Fig.1. Object identification based on shape, color and parts for classification

B. Shape feature descriptor for object recognition

Identification of objects present in image frame using shape features are widely adopted in surveillance systems. The scale invariant feature transform method is exploited for extracting local image features are effective against messed background. Figure 2 shows the image frame template. Figure 2.b shows the compared image and Figure 2.c shows the identified object that has been superimposed on gray-scale image. Hence, the shape and color feature descriptors can be combined for improved pattern matching of objects. The results explains a combined method of segmentation and multiple detection of objects in an efficient manner.

Fig.2 (a) Original Image pattern (b) Image template compared with another image (c) Identified objects superimposed on gray-scale image.

5. TEMPLATE MATCHING

The template matching method performs matching regions of an image with a template or pattern image. The objects belonging to different categories are stored as a template in the memory. The hybrid method proposed in the study receives input image frames. It then detects the objects present in the successive image frames and compares it with the already stored image template frames. The objects found matching with the image stored in the memory. These type of matching patterns are applied for character recognition, letters, numerals etc. The method is suitable for applying to color and grey-scale image as well. Template matching can be performed on pixel level or it can be applied based on the feature extracted from the image. In feature based approach, the image features are compared with the features of sub-images of the given input image. The spatial variations n mage play a vital role and hence it's obtained using a generalized distance transform (GDT) method. This method performs weights estimation based on the distance transform and relies more on the strong edge pixels. The orientation information is represented as an orientation map (OM) and is calculated using local gradient.

Generalized Distance Transform (GDT).

Let G be a regular grid. The $\Psi : G \rightarrow \mathbb{R}$ represent a function on the grid. The GDT of

Ψ can be formulated as the grid and is expressed as follows:-,

$$D_{\Psi}(p) = \min_{q \in \mathcal{G}} \{ d_s(p, q) + \Psi(q) \}, \qquad (1)$$

Here, for each point p, we find a point q that is close to p, and for which $\Psi(q)$ is small. Also the DT of an edge image using $L2$
$-$ *norm*, $ds(p, q)$ is the Euclidean distance between p and q. *Therefore,* $\Psi(q)$ is defined as follows:-,

$$\Psi(q) = \begin{cases} 0, & \text{if } (q) \in e \\ \infty, & \text{otherwise} \end{cases} \qquad (2)$$

One of the initiative taken to reduce the effect of the weak edge points relies with determining the strong edge points, and is formulated as follows:-,

$$\Psi(q) = \begin{cases} \frac{\eta}{\sqrt{I_x^2 + I_y^2}}, & \text{if } (q) \in e \\ \infty, & \text{otherwise} \end{cases} \qquad (3)$$

Orientation Map (OM). Let $p*$ be the closest edge point to the pixel *p*,

$$p^* = \arg\min_{q\in\mathcal{G}}\{d_*(p,q)+\Psi(q)\} \quad (4)$$

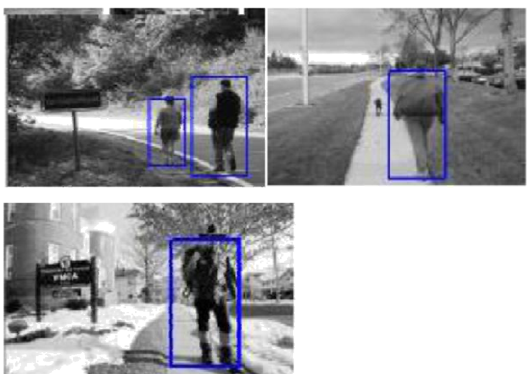
and the orientation value at *p* is defined as,

$$O_\Psi(p) = \arctan(I_{x^*}/I_{y^*}) \quad (5)$$

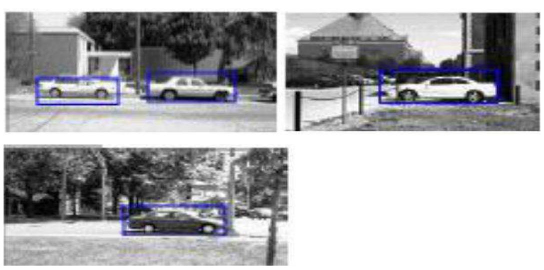
After successful estimation of generalized distance transform and orientation mapping, the overall score of matching template objects are obtained as follows:-,

$$D(T,I)=\frac{1}{|T|}\sum_{t\in T}\sqrt{\alpha D_\Psi^2(t)+(1-\alpha)\sin^2|O_\Psi(t)-o_t|} \quad (6)$$

Initially, the proposed system generated a template for all the input centroid data and is used for matching with the future upcoming image frames. Mathematical correlations are used to match the template.



a)



b)

Fig. 3. A) human detection, B). car detection

Algorithm: Extract the centroid coordinates from object detection module. For each individual pair of centroid, generate a template and estimate the trajectory of moving objects. Template matching method is described below.

For template matched with the new upcoming frame using correlation
if match = found, then
Save the co-ordinate positions
Updation of the template as matched
Else
Discard or ignore the template with assumption that the objects moved out of the scope of the camera. The moving object path s plotted if required.

6. EXPERIMENTAL RESULTS

The following sequence of image frames shows the performance and results obtained using the proposed method. Figure 4 and 5, shown below represents two image frames taken at time t, (i.e ith and (i-3)th sec)

Fig. 4: The frame at time t, marked as the current mage frame obtained from the video sequence.

Fig. 5: Frame obtained at time t-1 or the frame Fri-3 from the video sequence

Fig. 6: The difference image df [i, j] obtained by subtracting the current frame from previous frame

Fig. 7: Multiple moving objects detected with bounding box

Fig. 8: Multiple motion objects identified using edges

Fig. 9: Object to be tracked

Fig. 10: Objects tracked using bounding box approach

The proposed hybrid method is tested using five different sequences taken from 168VJ Clips Dataset. Results are given in Table 1. The estimation of this algorithm is made by the calculation of the rate of good detections moving object (GDR) using the following formula. BDR: Bad Detection Ratio. (7)

GDR = Number of detected moving object / Total moving object (8)

Table 1 Detection accuracy of the proposed system

Video stream	Moving object	Detection	GDR (%)	BDR (%)
V 1	98	95	96.4	3.6
V 2	20	19	95	5
V 3	9	8	89	11
V 4	41	39	95.1	4.9
V 5	4	4	100	0

7. CONCLUSION

The proposed method performs frame differencing, followed by color and shape-based template of objects for obtaining better result. The research work is suitable candidate for detecting and tracking objects of limited number. The proposed hybrid method is capable of recognizing objects with increase in substantial memory consumption. The future work may include embedding the method with advanced optimization methods for minimizing the memory space consumption.

REFERENCES

[1] Viola, P., Jones, M.: Rapid object detection using a boosted cascade of simple features. In: Proc. IEEE Conference on Computer Vision and Pattern Recognition, vol. 1, pp. 511–518 (2001) 202.

[2] Dalal, N., Triggs, B.: Histograms of oriented gradients for human detection. In: Proc. IEEE Conference on Computer Vision and Pattern Recognition, vol. 1, pp. 886–893 (2005).

[3] Zhang, H., Gao, W., Chen, X., Zhao, D.: Object detection using spatial histogram features. Image and Vision Computing 24, 327–341 (2006).

[4[Ferrari, V., Fevrier, L., Jurie, F., Schmid, C.: Groups of adjacent contour segments for object detection. IEEE Trans. Pattern Analysis and Machine Intelligence 30(1), 36–51 (2008).

[5] Agarwal, S., Roth, D.: Learning a sparse representation for object detection. In: Heyden, A., Sparr, G., Nielsen, M., Johansen, P. (eds.) ECCV 2002. LNCS, vol. 2353, pp. 113–127. Springer, Heidelberg (2002).

[6] W. Hu, A.M.Gharuib, A.Hafez, "Template Match Object Detection for Inertial Navigation Systems," Scientific research (SCIRP), pp.78-83, May 2011.

[7] E.Shectman, M.Irani, "Matching Local Self-Similarities across Images and Videos," In IEEE International Conference on Computer Vision and Pattern Recognition, pp. 1-8, 2007.

[8] F. Khan, R. Muhammad , et.al., "Color Attributes for Object Detection," In IEEE International Conference on Computer Vision and Pattern Recognition (CVPR), pp. 3306 – 3313, 2012.[9] H. Moballegh, N. Schmude, and R. Rojas, "Gradient Vector Griding: An Approach to Shape-based Object Detection in RoboCup Scenarios," from:www.ais.uni-bonn.de/**robocup**.de/papers/RS11_Moballegh.pdf

[10] Olson, C.F., Huttenlocher, D.P.: Automatic target recognition by matching oriented edge pixels. IEEE Trans. Image Processing 6(1), 103–113 (1997).

[11] Jain, A.K., Zhong, Y., Lakshmanan, S.: Object matching using deformable templates. IEEE Trans. Pattern Analysis and Machine Intelligence 18(3), 267–278 (1996).

[12] Borgefors, G.: Hierarchical chamfer matching: A parametric edge matching algorithm. IEEE Trans. Pattern Analysis and Machine Intelligence 10(6), 849–865 (1988).