2012 Fifth International Symposium on Computational Intelligence and Design

Object Tracking Based on Local Feature Matching

Qinkun Xiao Xiangjun Liu and Mina Liu
Department of Electronics Information Engineering
Xi'an Technological University
Xi'an, China, 710032
xiaoqinkun@tsinghua.edu.cn

***Abstract*—The feature extraction was analyzed and matching algorithms of object tracking, then we develop an object tracking method based on local feature matching in this paper. First we introduce about the existing feature extraction algorithms and the basic idea of Shape Context. Then we propose the tracking algorithm based on local characteristics. In this algorithm, we evaluate the similarity between the two feature points by histogram distance statistics based on χ^2-test and we determine the tracking trajectory by votes and prediction to the matching points. It is better matched in target deformation, rotation and occlusion by using this method. The experiment results show that our object matching method is effective.**

***Keywords-* shape context; feature matching; votes and prediction; object tracking**

I. INTRODUCTION

With the development of Computer vision technology, Signal process and Multimedia, Video Surveillance System is applied widely. So it is necessary to research the effective approaches about object tracking to reduce the human participation and develop the intelligent video surveillance [1]. Object tracking use the characteristics of dot, line, area block and so on to match the feature, thus we can track the target by feature points matching. The typical feature matching algorithms as: Mean-shift, Particle filter and Template matching. Mean-shift algorithm [2] use the color features, and it makes small amount of calculation and good real-time. But the weakness is: when the image is gray or less texture information, the result of object tracking is not well, and if the object is sheltered or obvious scale variation, it is too easy to lost the target. As a kind of nonlinear filtering algorithm based on Bayesian estimation, Particle filter [3]has the unique advantages in processing the parameter estimation and filter for the Gaussian nonlinear time-varying system. As a new kind of algorithm, particle filter is hard to be used in universal questions, and this algorithm has problem in convergence, besides it is bad real-time. For Template matching [4], in order to judge the matching degrees, it needs to calculate the similarity of adjacent pixels. So, for this kind of algorithm, it is easy in calculation, but we need to calculate the similarity for the whole image and when the target is deformation, rotation or occlusion, it is hard to track the target. For all the problems we discuss above, we need to find a new method to improve the object tracking. the research about object tracking algorithm based on local feature matching which we proposed has good practical significance.

This paper puts forward to a new kind of object tracking algorithms based on local feature matching. We use Shape Context [5] to extract the mutual position of target edge, and match the feature points by votes and prediction. Using this algorithm we find that we can match the feature points better when the target is deformation, rotation or occlusion. The basic idea of Local feature matching is: we use Shape Context to extract the contour information and features of the target, then we compare about the similarity of feature points between two goals, at last we use prediction and votes to conclude the center of the target.

II. OBJECT TRACKING ALGORITHM

For the common feature extraction algorithm about like SUSAN algorithm, Harris algorithm and so on, when target is occlusion or deformation，it is hard to get the exact feature information about the targets and because of this, we prone to false targets match or we even cannot find a match. But feature extraction based on local feature matching can resolve this problem well.

A. Feature Extraction Based on Shape Context

Outline description can be the most direct feature for the object, thus a certain number of discrete points on outline can represent shape information of the object. Shape Context, through extracting the boundary information and calculating its similarity to describe the characteristics.

Based on edge detection for one image, we can get a mount number of discrete boundary points as $P_i=\{p_1,p_2,\ldots p_n\}$, the value of n has a relative large impact for the final result. A shape is represented by these points sampled from the internal or external contours of the object. For any point p_i, the vector group which is comprised by the remaining $n-1$ points and the point p_i can well reflect in the shape information of the point.

For the convenience of description and calculation, we introduce a Log-polar coordinate system similar to the radar scan polar coordinates. Shown in figure 1:

978-0-7695-4811-1/12 $26.00 © 2012 IEEE
DOI 10.1109/ISCID.2012.106

The entire space is divided into 12×5 bins. So the feature information of the pixel p_i is a log-polar histogram of the coordinates of the rest points, setting measured by the reference point as the origin. The outline of any object can be indicated by a matrix size of n×60.

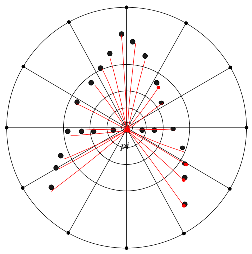


Figure 1. Schematic of the Feature Extraction Based on Shape Context

B. Feature Matching and Object Tracking

This part based on Shape Context feature information, We establish feature database as an intermediary to match with feature information of the target. Then we can find the center and boundary information of the tracked target, thus we can determine the target.

1) Establish the Feature Database

For example, we select a few images, masks (only edge map inside the mask) information like people, then we train to get the SC feature and build a codebook of local feature as D ($D = \{ce_k\}$, $k=1,...,n$). In the codebook, we record the SC feature and the relative position of edge points with center, and other information.

2) Feature Matching

After edge detection and training, we get the SC feature about the image frame, and then we evaluate the similarity between the target feature points and the codebook by COST [6]. The COST for Shape Context is defined as C_s. C_s is the value of histogram distance statistics based on χ^2-test. We define $g(k)$ and $h(k)$ as histogram value for the two points between two images. The C_s calculated as follow:

$$C_s = \frac{1}{2}\sum_{k=1}^{K} \frac{[g(k)-h(k)]^2}{g(k)+h(k)} \tag{1}$$

The C_s value smaller, the similarity is greater.

3) Vote to predict the target center

The SC information about detected point A is defined as $\{f_i\}$, and $P(ce_k|A)$ is defined as the matching probability between ce_k from the codebook and f_i, and the probability of center c for target o predicted by A is defined as $P(o, c|ce_k, A)$. So we can get the probability of the center of object as:

$$P(o,c) = \sum_{i,k} P(o,c \mid ce_k, A)P(ce_k \mid A)P(A) \tag{2}$$

$P(o, c)$ gives a vote map V for different locations c for the object class o, the maximum scores of vote map are the possible object center expressed as *hypo_center*, and meantime we can get the object boundary *bbox_list* form the point to point voting map.

It is different from the traditional tracking algorithm that we evaluate the similarity with histogram value of Shape Context. It is better to settle about object deformation, rotation, parallels move and so on. Furthermore, vote and prediction method used to determine the estimated target position can achieve good matching results.

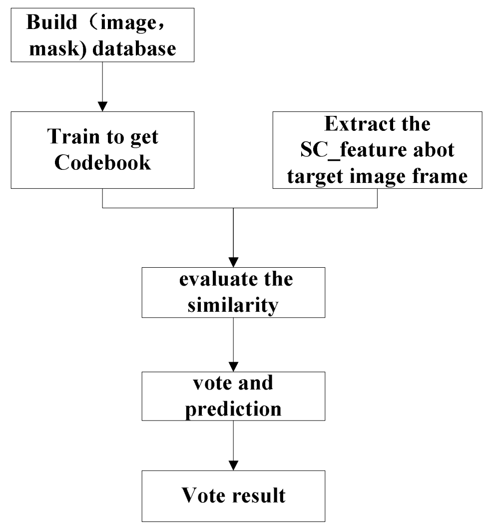


Figure 2. Tracking Flowchart

III. EXPERIMENTAL RESULTS AND ANALYSIS

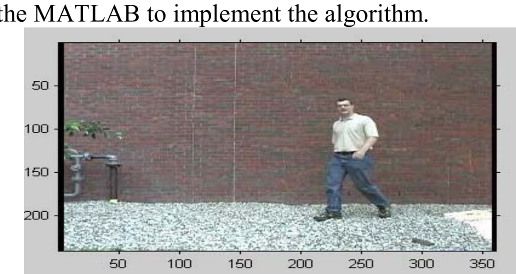
For the similarity, we choose the first 7 images and masks in the database to show the experiment. We evaluate the similarity between the detected image and the 7 images from database. The result is showing as follow:

	1	2	3	4	5	6	7
Cost	0.25	0.35	0.32	0.27	0.23	0.29	0.33
T=0.2	95%	15%	23%	75%	99%	83%	24%
T=0.3	100%	84%	87%	100%	100%	100%	85%

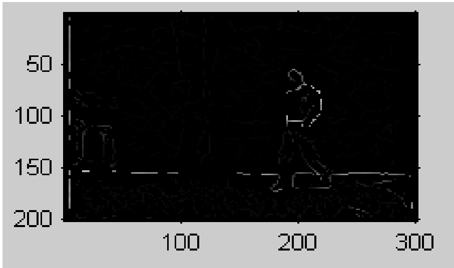
Figure 3. Computing the similarity

Showing in the fig.3, *Cost* means the average value of the histogram distance statistics value based on χ^2-test, and T means the threshold, when T=0.3, we can see the similarity is higher.

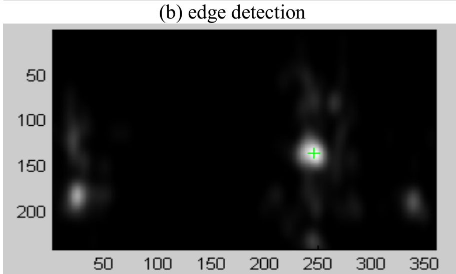
For feature extraction method proposed in this paper, we use the MATLAB to implement the algorithm.



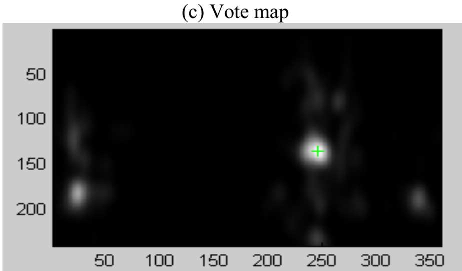
(a) Image



(b) edge detection



(c) Vote map

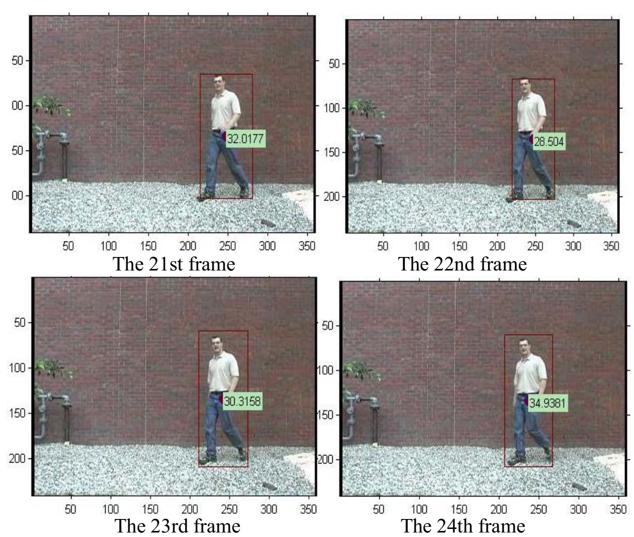


(d) Vote result

Figure 4. Object tracking experiment

Showing in fig.4, first, we do edge detection for the image, then train to get SC_feature and evaluate the similarity between the detected point and codebook. (c) Map is vote map, "+" means the maximum scores of vote map. (d) map shows the vote result.

Showing in fig.5, we can see the object tracking method is effective.



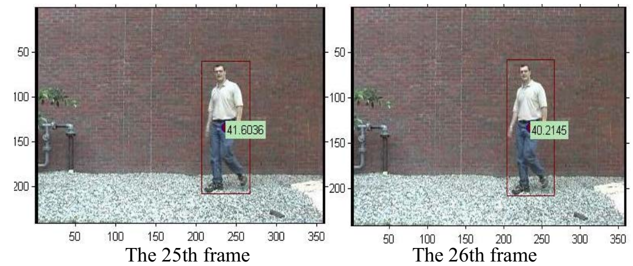


Figure 5. Tracking Result

The fig.6 gives us the trajectory of object center, "*"means the detected object center which we use the proposed method in the paper and the "+" shows the true trajectory of object center. Showing in the picture, it can well reflect the moving trajectory when we use this method to track the target.

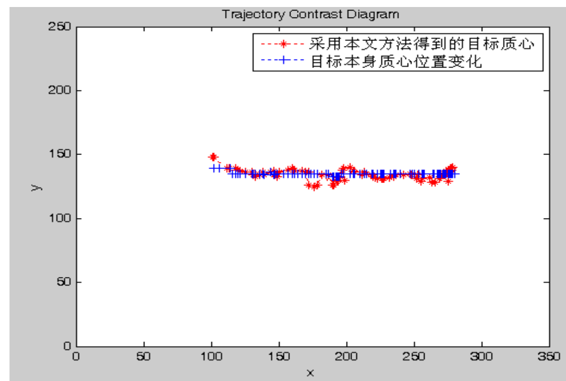


Figure 6. Trajectory of object center

IV. CONCLUSION

This paper studies about the local feature tracking algorithm based on Shape Context feature extraction, this algorithm is better matched in target deformation, rotation and kept out. The novelty of this paper is that:

(1) It is different from the traditional target tracking algorithms that this paper introduces the histogram value similarity measurement mechanism based on Shape Context; it is better matched in object deformation, rotation.

(2) Based on similarity measurement, we use the voting method to predict and estimate the location of the object. Experiments show that this method can be a good match results.

ACKNOWLEDGMENT

For the financial supports, the authors are sincerely indebted to the National Natural Science Foundation of China under the grant numbers60972095, and also to the Program for based on the dynamic three-dimensional scene graph model re trieval researc and Research Team of Xi'an Technological University, respectively.

REFERENCES

[1] Segen J, pingali S. A camera-based system for tracking people in real time. *the 13th International Conference on Pattern Recognition. IEEE Computer Press*, 1996,pp.63-67.

[2] Comaniciu D, Ramesh V, Meer P. Kernel-based object tracking. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2003, 25(5),pp. 564-575.

[3] N. Johnson, D. C. Hogg etc. Learning the distribution of object trajectories for event recognition, *Image and Vision Computing,*1996,pp.14:609-615.

[4] Feng Z R, Lu N, Jiang P. Posterior probability measure for image matching. *Pattern Recognition*, 2008,pp. 41(7):2422-2433.

[5] S. Belongie, J. Malik, and J. Puzicha, Shape Context: A New Descriptor for Shape Matching and Object Recognition, *Neural Information Processing Systems* 2001, pp.13:831-837.

[6] S. Belongie and J.malik edc. Matching with shape contexts, *IEEE Workshop on Content-based access f Image and Video-Libraries*,2000.