

# Small Target Detection and Recognition based on Template Matching

Tiansheng Zhao

School of Faculty of Information Technology, Beijing University of Technology, Beijing 100124, China;

tiansheng05@126.com

**Keywords:** unit entropy, small target detection, template matching, background suppression

**Abstract:** A small target detection and recognition algorithm based on template matching is proposed to quickly detect and recognize small targets in a video whose size are from  $10 \times 10$  to  $40 \times 40$  pixels in complex background. Existing approaches cannot detect and recognize small target whose size is less than  $40 \times 40$  pixels in images taken by common cameras, because there is no obvious contour and texture in such small target images. The proposed algorithm combines image preprocessing, background subtraction, motion prediction, unit entropy and traditional template matching techniques to form a complete algorithm for tracking small targets. The proposed algorithm reduces the running time by narrowing the search space and reducing the number of matching points. Experimental results showed that the algorithm can accurately and fast recognize the small target with any angle rotation, while SIFT(scale-invariant feature transform) algorithm cannot calculate sufficient feature points to recognize such small targets.

## 1. Introduction

In recent years, small target detection and recognition is becoming more and more significant in the computer visual field. There is a growing demand for surveillance system to monitor small moving objects, for example, monitoring the existence of pedestrians on the railway and other animal invasion. If an exception occurs, surveillance system will recognize the moving objects and give an alarm immediately. However, automatic detection and recognition gets more and more difficult when the targets are getting smaller, due to the small targets have very little feature information.

For pictures taken by common cameras, existing object detection and recognition algorithms are based on shapes, edges, textures[1, 2, 3] or other features [4, 5]. Among these approaches, methods based on SIFT [4,5,6] are most accurate in 2D models. However, they need sufficient pixels to build feature space, and the computing cost is huge. They are not suitable for detecting small target in real time application. Some statistical features can be used to characterize small images, e.g. histogram and entropy of the images, because they do not rely on the size of the image. But statistical features are too general and without position information, so they are hardly applied to object recognition.

To solve the above mentioned problems, we propose a simple and efficient method in this paper, which combines image preprocessing, background subtraction, motion prediction, unit entropy and traditional template matching techniques to track small targets. The main contributions of this paper are summarized as follows: (1) Discrete wavelet transform is applied to remove the noise in the background to reduce the computational cost. (2)Background subtraction is used to extract moving objects. (3)The motion prediction method is used to analyze the moving target position of the next frame, in order to reduce the search area during the tracking process. (4)A new algorithm of template matching based on unit entropy is proposed for recognizing small targets whose size are from  $10 \times 10$  pixels to  $40 \times 40$  pixels.

The rest of this paper is organized as follows: In section 2, the main idea of the proposed method is introduced to detect small target. The experimental results and discussion are shown in section 3. The conclusion is given in Section 4.

## 2. The Proposed Method

The proposed method detects and recognizes small targets based on template matching. Firstly, the video image is grayscale processed and the noise is removed from the background. Secondly, moving object was extracted based on background subtraction. Thirdly, according to the motion prediction technique, the moving target position of the next frame is predicted and the target search area is determined. Finally, the template matching algorithm based on unit entropy was applied to realize the tracking of moving small targets.

### 2.1 Image Preprocessing

Color image in digital processing has great flaws. In order to reduce unnecessary information in the original image, the color image is converted into a gray image. This process is called grayscale processing of the image. Image gray-scale processing is to remove color images of color information. For an image, there are three color components R, G, and B for each pixel. Gradation is the R, G, B, respectively, given a specific value, this value is called the gray value. According to the color space of YUV, the physical meaning of the Y component is the brightness of the point, and the luminance level is reflected by the value.

The gradation Y is for the luminance information of a certain pixel in the image data. For the color image data, the brightness information can be obtained by the following formula:

$$Y = 0.3R + 0.59G + 0.11B \quad (1)$$

Y is the brightness value of the pixel, R is the value of the red element of the pixel, G is the value of the green element, B is the blue element of the pixel value. Through the above formula, we will image the color information into a gray value.

Wavelet threshold denoising is widely used because of its simplicity, little computation and obvious denoising effect. Wavelet soft threshold denoising[7, 8] is applied in this paper.

### 2.2 Moving Object Extracting

This paper will be based on the background model of the background subtraction, used to detect static scene in the surveillance camera moving small targets. In order to improve the background subtraction accuracy of moving small targets, background modeling has become a key factor in target recognition.

Background modeling is a process of creating a background image without moving objects, separating the moving object and the background from successive multi-frame images to create a background image. Background modeling and updating is the basis for extracting foreground moving objects. At present, there are many algorithms for background modeling, such as mean method, inter-frame difference method, mixed Gaussian method, Bayesian method and codebook method. The application environment in this paper is that the background change is not violent and the moving target is simple. Therefore, the paper chooses the mean filter method and improves the algorithm.

The basic idea of the mean method is to store continuously acquired L-frame sequence images into the buffer, and average the L-frame image corresponding pixels, and use the mean as the background model of the next frame. When the value of N is bigger, the reliability and stability of the background are better, but it needs to store the data of N frames into memory and consume a lot of space. In this paper, we do not need to save each frame image, just save the background B of the previous frame and the weighted average of the current frame. After acquiring a frame, we update the background. We suppose the function expression  $B_N$  is the background model established by the system when the Nth frame image is collected, N is the average frame number, and  $f(i)$  is the i-th frame image.

The background model is

$$B_N = \frac{1}{N} \sum_{i=1}^N f(i) \quad (2)$$

Then the updated background model after acquiring the next frame is

$$B_{N+1} = \frac{1}{N+1} \sum_{i=1}^{N+1} f(i) = \frac{1}{N+1} \left( \sum_{i=1}^N f(i) + f(N+1) \right) = \frac{1}{N+1} (B_N \times N + f(N+1)) \quad (3)$$

The above describes the background image acquisition process. The background subtraction is the

subtraction of the current frame image and the background image, and the region whose difference is greater than a certain threshold is the moving target region. In this paper, the gray value of the target region which is saved as the original target template, and then the tracking of the moving target is completed based on the template.

### 2.3 Moving Object Predicting

The basic idea of motion prediction is based on the coordinate position of two moving objects, find the coordinate difference and get the direction of the target movement, and because the time difference between the two frames is known, we can calculate the moving speed of the object. Based on the velocity and direction of motion and the time difference between frames, we can compute the approximate position of the moving object in the next frame. In this way, the search strategy can be reduced in the matching process to reduce the matching point and improve the computing efficiency.

Assume the object is moving at a uniform speed and the center point of the moving object in the previous frame is  $F(x_{n-1}, y_{n-1})$ . The center point of the moving object in the current frame is  $C(x_n, y_n)$ . Time interval between frames is  $t$ , The predicted approximate point of the moving object in the next frame is  $P(x_{n+1}, y_{n+1})$ .

$$X_{n+1} = x_n + \frac{(x_n - x_{n-1})}{t} \times t = 2x_n - x_{n-1}, \quad y_{n+1} = y_n + \frac{(y_n - y_{n-1})}{t} \times t = 2y_n - y_{n-1} \quad (4)$$

When we predict the position of moving object in the next frame, we assume that the object is moving at a uniform velocity, but there are situations in which the object moves too fast or accelerates. In order to improve the adaptability of the algorithm, it is necessary to enlarge the search area. So, if the predicted point is  $P(x_{n+1}, y_{n+1})$ , the search area will be explored from a rectangle area whose side length is  $2M \times 2N$  (a small target image with  $M \times N$  pixels) and centered at  $P(x_{n+1}, y_{n+1})$ . Here, this rectangle area is called as region of interest center (ROIC). Then we explore all candidate target areas placing their top-left corners at each point in the ROIC as shown in Fig1.

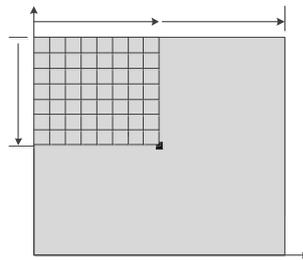


Fig1. Exploring candidate target areas placing their top-left corners in ROIC

### 2.4 Small Target Matching

With local entropy[9] can be used to characterize the small targets. Assume a small target image with  $M \times N$  pixels is partitioned into  $P \times Q$  sub-images, as shown in Figure 2. The gray value of point  $(i, j)$  is  $f(i, j)$ . Each sub-image is regarded as a unit whose size is  $\lfloor M/P \rfloor \times \lfloor N/Q \rfloor$  pixels.

The starting point of the unit  $(i, j)$  is  $(x_i, y_j)$ ,  $i=0, 1, 2, \dots, \lfloor M/P \rfloor$ ;  $j=0, 1, 2, \dots, \lfloor N/Q \rfloor$ .

The gray value of the unit  $(i, j)$  is the sum of all pixels' gray values:

$$F_{ij} = \sum_{x=x_i}^{\lfloor M/P \rfloor} \sum_{y=y_j}^{\lfloor N/Q \rfloor} f(x, y) \quad (5)$$

Then the unit entropy  $H_{ij}$  for the unit  $(i, j)$  is

$$H_{ij} = - \sum_{i=1}^P \sum_{j=1}^Q P_{ij} \log_2 P_{ij} \quad (6)$$

And  $P_{ij}$  is the probability distribution of the gray in the unit  $(i, j)$ .  $P_{ij}$  is

$$P_{ij} = \frac{F_{ij}}{\sum_{i=1}^P \sum_{j=1}^Q F_{ij}}, \quad 0 \leq P_{ij} \leq 1, \quad \sum_{i=1}^P \sum_{j=1}^Q P_{ij} = 1 \quad (7)$$

The gray relativity of the information entropy of all the units in the two images is compared, and the correlation coefficient can represent the similarity between the candidate target and the target template image. We set a threshold to judge the end of the search. We think it is probably the true target when the calculated correlation coefficient exceeds this threshold.

The formula for calculating the correlation coefficient is described as follows:

$$R = \frac{\sum_{i=1}^P \sum_{j=1}^Q [H_c(i, j) - E_c][H_t(i, j) - E_t]}{\sqrt{\sum_{i=1}^P \sum_{j=1}^Q [H_c(i, j) - E_c]^2 \cdot \sum_{i=1}^P \sum_{j=1}^Q [H_t(i, j) - E_t]^2}} \quad (8)$$

We suppose the formula expression  $H_c(i, j)$  is the unit entropy of the candidate target,  $H_t(i, j)$  is the unit entropy of the target template.  $E_c$  is the average value of all the unit entropy of the candidate target.  $E_t$  is the average value of all the unit entropy of the target template.

### 3. Evaluations

The proposed algorithm is evaluated in a PC. The PC is equipped with Dual-core CPU and 8G memory, the CPU is Intel(R) Core(TM) i5, 3.2 GHz. The algorithm is implemented in C++ code. We design a set of experiments to evaluate our proposed algorithm for the target images with scale, rotation, comparing computing time cost and recognition rate to those of the algorithm SIFT.

In the PC, we implemented the proposed algorithms and SIFT algorithms, and used an aircraft video taken in uncontrolled environment to test the algorithms. There are 1019 frames in this video and the image sizes in the video were 600×448 pixels. The aircraft is changing from big to small as it flies from near to far sky area.

Fig2 showed the recognition results of the proposed algorithm and SIFT. The target in the left image is the target template extracted from the image based on background subtraction. The middle image rotated with a great angle. The right image showed the feature points extracted by SIFT. The proposed method can still recognize the target while SIFT cannot detect the target. Because there is only a few feature points, it is insufficient for SIFT to detect the target.



Fig. 2 Comparison of recognition results by the proposed algorithm and SIFT

We applied the proposed algorithm to recognize the aircraft target in the 157 frames in which the target size was from 10×10 pixels to 40×40 pixels, respectively. The recognition rate of the proposed algorithm was 61.23%, while the recognition rate of SIFT is 0. Fig3 showed the time consumption for detecting the target by the two algorithms, respectively.

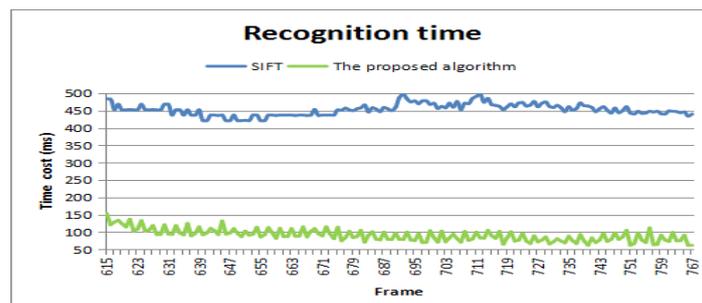


Fig. 3 Computation consumption

#### 4. Conclusions

A small target detection and recognition algorithm based on template matching is proposed to quickly detect and recognize small targets in a video. Wavelet transform and soft-threshold de-noising are applied to improve the computing efficiency, which can also avoid influence of noise and illumination changing. Moving object was extracted as target template based on background subtraction and updated the target template during the recognition process, which can adapt to the changing moving objects and scene changes in video. According to the motion prediction technique, the moving target position of the next frame is predicted to narrow the search space and reduce the number of matching points. Then the target is matched based on unit entropy of the image. Comparing with SIFT, the proposed algorithm is more suitable for detecting and recognizing small targets whose size are from  $10 \times 10$  pixels to  $40 \times 40$  pixels. The experimental results show our algorithm is adaptable to scale changing, small-angle rotation and small illumination changes.

#### 5. References

- [1] Kim S, Lee J. Scale invariant small target detection by optimizing signal-to-clutter ratio in heterogeneous background for infrared search and track[J]. *Pattern Recognition*, 2012, 45(1):393-406.
- [2] Bai X, Zhou F. Analysis of new top-hat transformation and the application for infrared dim small target detection[J]. *Pattern Recognition*, 2010, 43(6):2145-2156.
- [3] Leung H, Young A. Small target detection in clutter using recursive nonlinear prediction[J]. *IEEE Transactions on Aerospace & Electronic Systems*, 2000, 36(2):713-718.
- [4] Cao Y, Liu R M, Yang J. Small Target Detection Using Two-Dimensional Least Mean Square (TDLMS) Filter Based on Neighborhood Analysis[J]. *Journal of Infrared, Millimeter, and Terahertz Waves*, 2008, 29(2):188-200.
- [5] Xun Y, Dehua L I, Fei H, et al. Detection of small target in infrared image sequences based on attention mechanism[J]. *Engineering Journal of Wuhan University*, 2006, 39(6):108-112.
- [6] Chen Weishi, NingHuansheng, and Li Jing. Detecting and Tracking of Small Moving Target in Avian Radar Images. In *IASP*: 473-476, 2010.
- [7] Boccignone G, Chianese A, Picariello A. Small target detection using wavelets[J]. 1998, 2:1776.
- [8] Yuan Huijing. A New Denoising Method for Small Target Detection[J]. *Acta Photonica Sinica*, 2004, 33(05):609-612.
- [9] Wu, J. and Qiu, Z.D. "A hierarchical palmprint identification method using hand geometry and grayscale distribution features." *18th International Conference on Pattern Recognition(ICPR)*. vol. 4, pp. 409-412, 2006.