OBJECT TRACKING METHOD USING BACK-PROJECTION OF MULTIPLE COLOR HISTOGRAM MODELS

Jung-ho Lee, Woong-hee Lee, Dong-seok Jeong

Multimedia Lab., Department of Electronic Engineering
Inha University, Incheon, South Korea

ABSTRACT

Automated object tracking system is needed for unmanned observing and proper recording of important places. In this paper, we propose an object tracking method which uses back-projection of color histogram with multiple models. We can make some representative models of an object from its color histogram distribution. 3D Labeling is introduced to eliminate unsuitable histogram blobs and color histogram models are composed from survived blobs. The position of interested object could be estimated with the back-projection image of each model. The proposed method can reduce the miss tracking even if object enters similar colored region in complex video scenes.

1. INTRODUCTION

Various approaches for tracking the object in the moving picture have been proposed in the literature [1][2][3][4][5][7]. The tracking methods which use the background information are good for the fixed background [3][4][7]. But it is not suitable for the swaying scene. Color histogram back-projection is a low complexity, active vision algorithm for finding objects in complex scenes. This algorithm is little affected by the movement of camera [1].

In this paper, we propose an object tracking method using back-projection with multiple color histogram models. Multiple models have many advantages for object tracking in real video. Existing methods using back-projection need human's intuition when initially selecting the color region of interest [1]. But proposed method can extract this region automatically if object's area was known. And if there exists only single color histogram model, the object could be lost during the tracking process due to the similar colored background area. Multiple models can prevent these kinds of drawbacks.

We used 3D labeling to separate the color histogram of an object into several clusters. Each cluster is transformed into back-projection model. Figure 1 shows the block diagram of proposed tracking method.

We apply this method to the actual videos of soccer game and crowded street of several levels. Experimental results on actual scene show that the performance of our method is robust enough.

The paper is organized as follows. We first describe in section 2 the color histogram back-projection. 3D labeling method, which is used to separate color histogram into several parts are described in section 3. Section 4 presents the method used for finding blobs from the back-projection image and the criteria for finding object's position. Finally, we discuss the experimental results and the future directions in section 5 and 6.

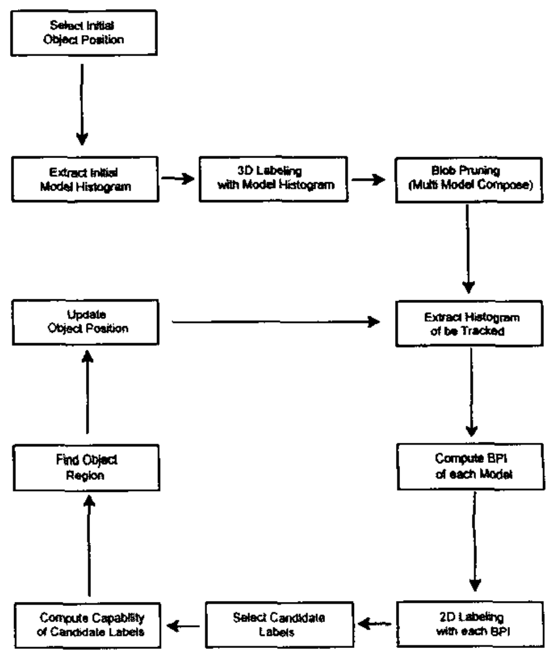


Figure 1. The block diagram of proposed method.

2. BACK-PROJECTION of COLOR HISTOGRAM

2.1 Color Histogram Back-projection image

The color histogram back-projection is used in many fields associated with color feature extraction. A grayscale image is

0-7803-7761-3/03/$17.00 ©2003 IEEE

obtained by color histogram back-projection. This image shows the distribution of interested color in the scene. We present below the formula of back-projection. This concept is built on the assumption that object's histogram has narrow range and frame's histogram is widely distributed.

$$\xi(c) = \frac{M(c)}{I(c)} \quad (1)$$

$$BPI(c) = 255 \times \xi(c) \quad (2)$$

***M(c)*: object's histogram value of color c**

***I(c)*: frame's histogram value of color c**

***BPI(c)*: Back-Projection Image of color c**

To get the back-projection image, we should know histograms of the object and of the whole frame. Let ***M*** be the histogram of the object of interest, and ***I*** be the histogram of the whole frame. Compute the ratio as shown in equation (1). This is the confidence measure of color *c*'s contribution to the characteristics of the model. Then a gray-scale back-projection image is constructed from the frame by replacing each pixel with color c by its scaled confidence measure, as in equation (2). In the back-projection image, the model appears as a bright "blob" [8].

Figure 2 shows an example where the object is the person who wears the pink shirt from the frame of actual outdoor scene.

Figure 2. An example of back-projection image
(a) Original scene (b) Back-projection image

2.2 Single and Multiple model based approach

The existing method is based on the single model framework, but has some problems [1]. First of all, the area of an object which represents its color has to be selected. This means human interaction should be involved in tracking process. We can compose the whole area of an object into the model by existing method. But it could not represent the color characteristics of an object successfully. Second, some areas in background could have object's primary color. Although it doesn't matter when very small area has that color, it could induce tracking errors or loss of tracking when the primary color area is dominant in background area, i.e. larger than the object's size.

In general, an object could have several representative colors. So, it is reasonable to make multiple back-projection models for an object. Also it could be done without human interaction. If the tracking system detects the appearance of an object from the scene, its color histogram can be obtained. And several color histogram back-projection models can be composed from the histogram distribution.

Multiple back-projection images can be generated from each model. Although one or some back-projection images could have problems, we can find correct object's position from the other back-projection images.

3. OBJECT TRACKING

3.1 Extracting Models with 3D Labeling

The object in the scene is a kind of color-distributed plate. Any color can be distributed in that region. But in general, it turns out that only limited number of color categories are distributed [6].

Usually color is expressed with 3D component, and color histogram is expressed in four-dimensional space like Figure 3. Three axes are Red, Green, and Blue, respectively. The histogram value is represented by the density of a bin.

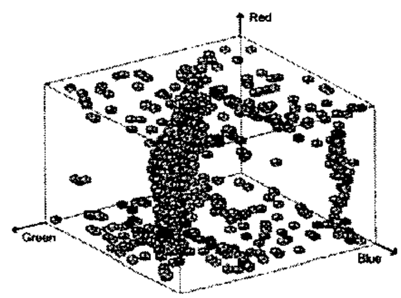


Figure 3. Color histogram in three-dimensional space. The density of each bin represents the histogram value.

We grouped the color histogram of an interested object into several categories to build multiple models. Each blob contains the representative characteristics of the object. And the back-projection model is constructed from these blobs

The linear quantization of color histogram is very simple, but it can not categorize the representative colors of the object. Also, it can't properly represent the characteristics of the object, because some similar colors could belong to neighboring lumps. We have to efficiently divide the color histogram into several lumps to represent the characteristics of the object.

Therefore, we propose an efficient method which classifies the color of an object into some groups. This method employs the labeling concept into the color histogram. The color histogram is represented in 3-D form as shown in Figure 3. We extend the union-find labeling algorithm into 3-D space to use it on the color histogram [9][10]. It is similar to the ordinary labeling method except the propagating direction. 3-D labeling propagates toward six directions, which is different from the 2-D labeling that propagates toward two directions.

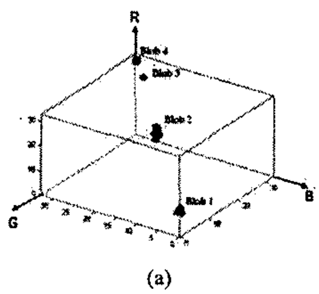
After the labeling of color histogram in three-dimensional sense, we get several blobs of histogram bin. Each histogram blob has different value. And to process efficiently, we eliminate insignificant models which have small histogram value.

$$Sum_i = \sum_c H_i(c) \qquad (3)$$

In equation (3), $\boldsymbol{Sum_i}$ means summation of all histogram values of ith histogram blob. If $\boldsymbol{Sum_i}$ is smaller than predefined threshold value, ith blob will not be considered as a model. Generally, the two or three largest blobs are adopted as models in the actual videos. Models are constructed from the survived blobs using equation (4). $\boldsymbol{M_i}$ represents ith model.

$$M_i(c) = \frac{H_i(c)}{\sum_j H_i(j)} \times 100 \qquad (4)$$

Figure 4 shows the process of categorization by eliminating blobs that has small value fewer than 10% for entire histogram.



(b)

Figure 4. Small blobs of color histogram are eliminated during model construction.
(a) Labeled color blobs. (b) Selected color models

3.2 Determining the Position of Object

In general, we have one to three histogram models. We can generate the back-projection image of each model and find the candidates of object's position from the each back-projection image. We use labeling algorithm to find the candidate positions from the back-projection image. The labeling procedure needs extensive computations but its performance is improved [10].

We determine object's position from the candidate positions with the following criteria.

- Size of object
- Position of the object in the previous frame
- Speed and Direction of movement

Firstly, we choose candidate regions from back-projection images. And we give score to each candidate according to the position of object in the previous frame. The speed and direction of movement limits the range of candidate regions in the frame. Then we can compose several combinations from candidates of each back-projection image according to its score. If combined area exceeds 150% of interested object size, this region will not be considered. And if some regions are occluded because of wearing similar color (Soccer game), the direction of movement will be principal criteria to determine the right position.

4. EXPERIMENTAL RESULTS

We apply the proposed method to the actual video scenes which include soccer game video and three different levels of crowded street scenes. In our experiments, the tracking failure occurs if the area of an actual object region is smaller than the found region or is more than twice larger than it.

Table 1 shows the tracking results of our method compared with the existing (Johnson's) method [1]. Although Johnson's method requires human interaction for the representative color selection of interested object, our proposed method shows the better success rate. Street1 is the scene of very crowded street, Street2 is of normal crowded street, and Street3 is less crowded street, respectively.

Table 1. Experimental results

	Street 1	Street 2	Street 3	Soccer
Number of tested objects	6	8	5	5
Number of tested frames	200	400	300	200
Success rate of Johnson's method (%)	63.6	73.3	86.9	92.0
Success rate of proposed method (%)	92.2	88.0	91.3	94.7

As can be seen in Table 1, the experimental results on actual scene show that the performance of our method is more steady and robust than that of the existing method in all cases. Figure 5 and Figure 6 shows examples of object tracking. The rectangular indicates the estimated object region and the white wave curve indicates the route of the object. The Figure 5 illustrates that the proposed method works well even for the very complex scene.

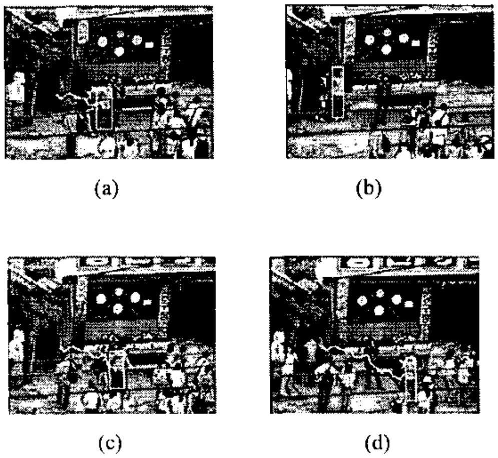


Figure 5. Tracking results of crowded street scene.
(a) Frame 1 (b) Frame 40 (c) Frame 80 (d) Frame 120

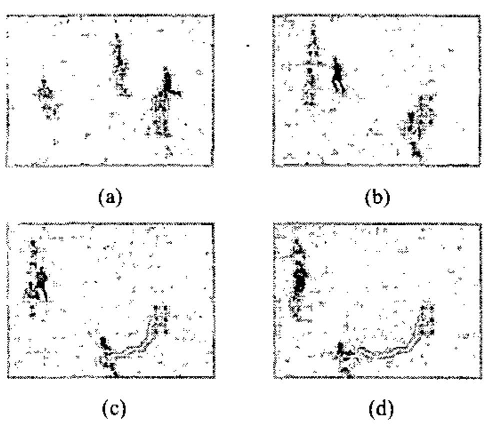


Figure 6. Tracking results of soccer game.
(a) Frame 1 (b) Frame 30 (c) Frame 75 (d) Frame 80

5. CONCLUSION

In this paper, we propose a new tracking algorithm using back-projection of color histogram. It involves 3-D labeling of object's color histogram to build models. Each derived model is used for constructing the back-projection images. We can find the blob candidates of the object from the back-projection images using the labeling algorithm. Some criteria like object's size, position, and speed are employed to find the correct position of a target.

The proposed method shows satisfactory performance of more than 90% accuracies even foe the complex scenes. The method can track people at the outdoor environment or in soccer game. All these situations involved unfixed cameras, different lighting and different objects being tracked. The proposed system achieves better performance than that of the existing methods even without human intervention.

We will consolidate the color histogram with more features such as the texture in future work.

6. REFERENCES

[1] Johnson I Agbinya and David Rees, "Multi-Object Tracking in Video", *Real-Time Imaging*, Vol. 5, pp. 295-304, 1999.

[2] J.K.Aggarwal and Q.Cai, "Human Motion Analysis: A Review", *Computer Vision and Image Understanding*, Vol.73 No.3, pp. 428-440, Mar. 1999.

[3] Yoshinori Ohno, Jun Miura and Yoshiaki Shirai, "Tracking Players and a Ball in Soccer Games", *Proc. of the 1999 IEEE Int. Conf. on Multisensor Fusion and Integration for Intelligent Systems*, Taipei, Taiwan, R.O.C., pp. 147-152. Aug. 1999.

[4] Chris Stauffer and W.E.L Grimson, "Adaptive background mixture models for real-time tracking", *Computer Vision and Pattern Recognition 1999(CVPR99)*, Vol. 2, Fort Collins, Colorado, June 1999.

[5] François Bremond and Monique Phonnat, "Tracking Multiple Nonrigid Objects in Video Sequences", *IEEE Trans. on Circuits and Systems for Video Technology*, Vol. 8, pp. 585-591, Sept. 1998.

[6] Henry R. Kang, *Color Technology for Electronic Imaging Device*, pp. 3-11, SPIE Press, 1997.

[7] Isaac Cohen and Gérard Medioni, "Detecting and Tracking Moving Objects for Video Surveillance," *Proc. of the 1999 IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, Vol.2, pp. 319-325, June 1999.

[8] Robert M. Haralick and Linda G. Shapiro, *Computer and Robot Vision*, Vol. 1, pp. 14-16, Addison-Wesley, 1992.

[9] Ramesh Jain, Rangachar Kasturi and Brian G. Schunck, *Machine Vision*, pp. 44-47, McGraw Hill, 1995.

[10] Linda Shapiro, *Computer Vision*, pp. 56-63, Prentice Hall, 2001.