Proceedings of the IEEE
International Conference on Automation and Logistics
August 18 - 21, 2007, Jinan, China

Object Tracking Algorithm Based on Camshift Algorithm Combinating with Difference in Frame

Hongxia Chu[1], Shujiang Ye[1], Qingchang Guo[2], Xia Liu[2]

Department of Electronic Engineering
Heilongjiang Institute of Technology[1]
Harbin 150001, Heilongjiang Province, China
chx0420@163.com

College of automation
Harbin Engineering University[2]
Harbin 150001, Heilongjiang Province, China
chx0420@163.com

***Abstract* - Object tracking algorithm based on Camshift algorithm which is combinating with difference in frame is put forward to quickly and exactly track movement object. Firstly, we confirm the movement region of object by difference in frame, and we again confirm centroid of movement region which is centered to initialize tracking curve. Secondly, the feature is abstracted in the region and the target is tracked by Camshift algorithm. The method overcomes shortcoming of artificial orientation and divergence problem for traditional Camshift during tracking. We prove the validity of the algorithm by experiment at last.**

***Index Terms - Difference. Frame. Object tracking.* Camshift**

I. INTRODUCTION

The moving object detection is an important subject in the field of application vision research. In real life, much useful vision information is included in movement. Human vision can see not only movement object but also static one, but people are sometimes much more interested in the movement object in many situations such as the detection of traffic flow and the inspection of important site's safety and the control and guide of aviation and warplane and assistant driving and so on. The research target of the moving object is sequential images which contain much information than single frame image. So it is significative for the research of the moving object detection and tracking system.

The research on the moving object detection and tracking is separated into two kinds:

1) Cinema head shifts by following movement object and always keeps target in the center of image [1][2][3];

2) Cinema head is fixed and movement object only is tracked in vision field [4].

Existing algorithms consist of the algorithm based on template match and the algorithm based on beset picture and optical flow and so on[5][6]. The object is all matched directly in these algorithms and only the static single picture in sequence is processed, but these algorithms don't take into account relativity among sequential images.

The algorithm studies on moving object detection and tacking in static scene. Namely, the method of difference in frame is used in moving object detecting and Camshift algorithm is applied in the moving object tracking. Different image between two frames is still calculated during the tacking by using Camshift algorithm in order that movement scope of moving object is ascertained. So the divergence of tracking window is also prevented.

II. OBJECT POSITIONING

A. Theory of Difference in Frame

Difference method is the most direct method to examine movement information from serial image. The method uses grey degree and grads of space-time in picking up movement information. The method calculates difference between former frame and back frame by comparing point-by-point grey value.

The formula of difference in frame can be written in (1) and (2):

$$D_k(x, y)=|f_k(x, y)\text{-}\mathrm{f}_{k\text{-}1}(x, y)| \tag{1}$$

$$R_k(x,y)=\begin{cases}0 & \textit{background} \quad \text{if } \mathrm{D_k}(x,y)>T \\ 1 & \textit{object} \quad \text{if } \mathrm{D_k}(x,y)\le T\end{cases} \tag{2}$$

Where D_k is image after difference, R_k is image after binarization, f_k is current frame image, $f_{k\text{-}1}$ is previous frame image, T is threshold.

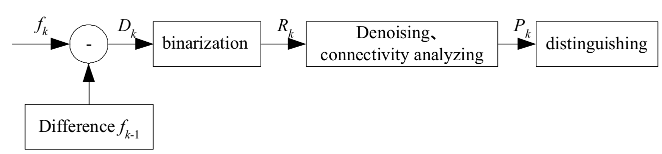


Fig. 1 Illustrating difference in frame

Difference method in frame is used in object detecting and abstracting by applying diversity between continuous two frames or several frames in video sequence. The process of a most essential difference method in frame is shown in Fig.1. First, Equation (1) calculates difference between No.k frame image and No.k-1 one and gains image D_k after difference. Then the image is binarizated in equation (2). When a certain pixel value is bigger than a given threshold, we think this pixel as foreground's pixel which is perhaps a point on the object. Otherwise, we suggest it background's pixel. We obtain image R_k after image D_k is binarizated. Finally, we analyze connectivity about image R_k and get image P_k. When the area of some connective region is bigger than a given threshold, we detect object and think this region is possessed by the object.

The merit of video detection and abstracting based on difference method in frame is shown as follows:

1) Algorithm can be easily realized.

2) Complexity degree of programmer is lower.

3) It easily implements real-time watching.

4) The method is not too sensitive to the change of beam

1-4244-1531-4/07/$25.00 © 2007 IEEE.

in scene because interval between frames is commonly shorter.

Shortcoming of difference method in frame is shown in the following:

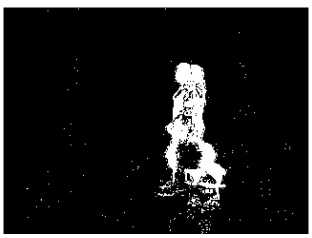
1) The positioning of moving object is not too exact. Area of detection is sportive variation district in the forward and behind frame.

2) Movement velocity of target is a bit slow because interval between frames is commonly shorter.

Experiment on difference in frame is shown in Fig. 2. Where Fig. (2.a) and Fig. (2.b) is respectively continuous two frame images. Fig. (2.c) is image after difference.

(2.a) Previous frame image

(2.b) Current fame image



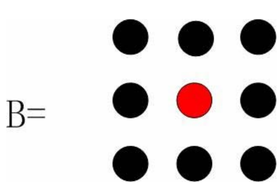
(2.c) Image after difference

Fig. 2 Difference experiment

Noise is existing in the binarization image because of the interference of noise. Morphologic method is introduced to wipe off noise. Center of mass of movement object is confirmed at last.

B. *Morphologic*

A little noise appears in target because of the interference of noise. Noise need to be removed in order that object positioning can be well done. The method of eroding and expanding on binary image in morphologic is used to accomplish the task.



Expanding: $\{z \in E \mid B_z \cap X \neq \Phi\}$

Eroding: $\{z \in E \mid B_z \subset X\}$

Where X is binary image, B is grid of view picture, Bz is a image which is obtained by moving structure element B from center(which is red) to position z.

Experiment on morphologic is presented in Fig.3. Fig.3 is the denoised image.

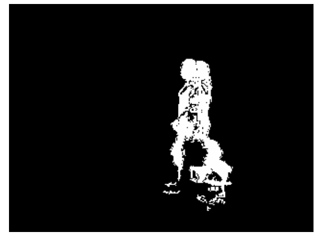


Fig. 3 Image after denosing

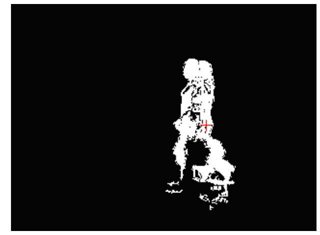


Fig. 4 Centorid calculating image

C. *Confirming Centroid*

Centroid of movement region is confirmed by moment method in the paper. Formula is shown as follows:

The zeroth moment is:

$$M_{00} = \sum_x \sum_y I(x, y) \tag{3}$$

The first moment for x and y:

$$M_{10} = \sum_x \sum_y xI(x, y) \tag{4}$$

$$M_{01} = \sum_x \sum_y yI(x, y) \tag{5}$$

The mean search window location is in (6)

$$x_c = \frac{M_{10}}{M_{00}}, y_c = \frac{M_{01}}{M_{00}} \tag{6}$$

Experiment about centroid computing can be seen in Fig. 4. The position of centroid is marked by the red cross point. Coordinate of centroid is (130, 198) in the experiment. Long shaft and short shaft of ellipse is ascertained according to twoth moment of binarization image. Long shaft and short shaft is reduced by percent 20 because variational region of binarization image is relatively larger.

III. CAMSHIFT ALGORITHM

Recently, tracking algorithm based on Camshift algorithm is more and more noticed by means of its favorable performance in reality and robust. Camshift algorithm is widely used in face tracking about apperceiving user interface now. Camshift algorithm makes use of color information in region to track object and adopts a non-parameter technique and searches movement target by clustering method[7].

To be brief, Camshift algorithm applies target's hue character to find the size and location of movement object in video image. In the next frame video image, search window is initialized by current position and size of movement object. The process is repeated to realize continuous object tracking. Initial value of search window is set as current position and size before search.

Search window searches near the area that movement object perhaps appears, so a mass of search time may be saved and Camshift algorithm has favourable reality. At the same time, Camshift algorithm finds movement object by colour matching. Color information has no much change during movement object moving, so it also has favorable robustness.

A. Probability Distributing Chart

Transformation formula between RGB space and HSV space is shown in (7a)-(7c):

$$H=\begin{cases} \frac{\pi}{3}\left(\frac{b-r}{\max(r,g,b)-\min(r,g,b)}\right) & \text{if } g=\max(r,g,b) \\ \frac{\pi}{3}\left(2+\frac{b-r}{\max(r,g,b)-\min(r,g,b)}\right) & \text{if } b=\max(r,g,b) \\ \frac{\pi}{3}\left(4+\frac{b-r}{\max(r,g,b)-\min(r,g,b)}\right) & \text{if } r=\max(r,g,b) \end{cases} \tag{7a}$$

$$S=\frac{\max(r,g,b)-\min(r,g,b)}{\max(r,g,b)} \tag{7b}$$

$$V=\max(r,g,b) \tag{7c}$$

Where, H represents hue heft, S represents hue saturation heft, V represents brightness heft, r, g, b respectively represents red and green and blue heft.

For each image, the method of calculating hue histogram is[8] : the total of image's pixel is given as n. Histogram's hue level is m. We also define $c(x_i^*)$ is histogram bin index associated to the pixel $No.x_i$, where i is from 1 to n. The value q_u of unweighted histogram is computed in (8), where u is from 1 to m:

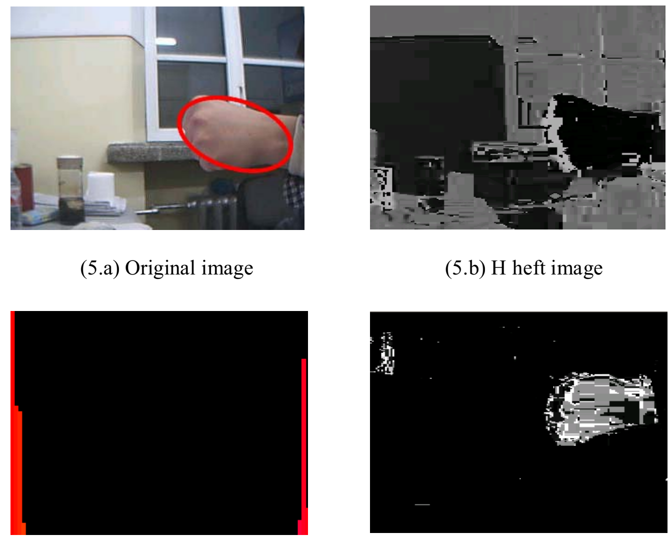
$$q_u=\sum_{i=1}^{n}\delta\left[c(x_i)-u\right] \tag{8}$$

Where, δ is unit impulse transfer function.

The process of histogram normalization is shown in (9):

$$\left\{p_u=\min\left(\frac{255}{\max(q)}\right)q_u,255\right\}_{u=1,\ldots,m} \tag{9}$$

According to (9), value of the histogram and each chroma level is normalized from [0, max(q)] to the new range [0, 255]. Then back-projection is used to associate the pixel values in the image with the value of the corresponding histogram bin. Namely, pixels with the highest probability of being in the histogram will be mapped as visible intensities in the image. So color probability distribution image can be seen in Fig. 5. Where Fig. (5.a) is Original image, Fig. (5.b) is H heft image, Fig. (5.c) is color histogram of refered area, Fig. (5.d) is probability distribution image.



(5.a) Original image (5.b) H heft image

(5.c) Hue histogram image referred (5.d) probability distribution image

Fig. 5 Color probability distribution image

B. Camshift Tracking Algorithm

Camshift tracking process is shown in the following:

1) Regions that target is tracked are confirmed by difference method in frame.

2) Calculating the zeroth moment and first moment

The zeroth moment is calculated in (10)

$$Z_{00}=\sum_{x}\sum_{y}I(x,y) \tag{10}$$

The first moment for x and y is computed in (11a) and (11b)

$$Z_{10}=\sum_{x}\sum_{y}xI(x,y) \tag{11a}$$

$$Z_{01}=\sum_{x}\sum_{y}yI(x,y) \tag{11b}$$

Where, $I(x, y)$ is pixel value at point(x, y). Change scope for x and y is the range of search window.

3) Compute the centroid (x_c, y_c) of search window location in (12).

$$x_c=\frac{Z_{10}}{Z_{00}},\ y_c=\frac{Z_{01}}{Z_{00}} \tag{12}$$

4) Reinstall the size of search window S as a color probability distribution function the region of search window.

5) Repeat step *2)*, *3)*, *4)* to converge(the change of centroid is smaller than given threshold)

Zeroth moment describes distribution area of target in image. Color probability distribution image is discrete grey image that maximum value is 255. Therefore, relation in size S and 255 for search window is written in (13):

$$S = 2\sqrt{Z_{00}/256} \tag{13}$$

Considering symmetry, S is selected as odd number which approaches result. Long axis and short axis and orientation angle of tracking object is calculated by computing second moments. Second moments for x and y is defined in (14a) and (14b) and (14c).

$$Z_{20} = \sum_x \sum_y x^2 I(x,y) \tag{14a}$$

$$Z_{02} = \sum_x \sum_y y^2 I(x,y) \tag{14b}$$

$$Z_{11} = \sum_x \sum_y xyI(x,y) \tag{14c}$$

The orientation angle θ of the object Long axis is confirmed in (15):

$$\theta = \frac{1}{2}\arctan\left(\frac{2\left(Z_{11}/Z_{00} - x_c y_c \right)}{\left(z_{20}/Z_{00} - x_c^2 \right) - \left(z_{02}/Z_{00} - y_c^2 \right)} \right) \tag{15}$$

Supposed parameter a and b and c in (16):

$$a = \frac{z_{20}}{Z_{00}} - x_c^2$$
$$b = \frac{Z_{11}}{Z_{00}} - x_c y_c \tag{16}$$
$$c = \frac{z_{02}}{Z_{00}} - y_c^2$$

The distance of Long axis and short axis from the image is given in (17a) and (17b):

$$l = \sqrt{\frac{(a+c) + \sqrt{b^2 + (a-c)^2}}{2}} \tag{17a}$$

$$w = \sqrt{\frac{(a+c) - \sqrt{b^2 + (a-c)^2}}{2}} \tag{17b}$$

IV. Experiment

We validate the validity of the algorithm by tracking movement hand. Result is shown in Fig. 6 and Fig. 7. Fig. 6 is the result of tracking algorithm based on Camshift algorithm combinating with the difference in frame. Fig. 7 is the result of conventional tracking algorithm.

Fig. 6.1

Fig. 6.2

Fig. 6.3

Fig. 6.4

Fig. 6 Tracking algorithm based on Camshift algorithm combinating with the difference in frame

Fig. 7.1

Fig. 7.2

Fig. 7.3

Fig. 7.4

Fig. 7 *Conventional tracking algorithm*

We prove the validity of the algorithm by upper experiment. The algorithm can quickly and exactly track movement object. It doesn't lose target information and tracking region is also convergence.

V. Conclusion

We proposal object tracking algorithm based on Camshift algorithm combinating with the difference in frame. Movement object is accurately positioned by difference in frame, which overcomes shortcoming of artificial orientation for traditional Camshift. Approximate sport scope of object is confirmed by frame difference during tracking. The algorithm also overcomes radiation's shortcoming during tracking. But there are some aspects to need to be improved in the future, such as target moving too quick and tracking information losing, etc.

REFERENCES

[1] Pan feng, Wang xuanyin, Xiang Guishian, Liang dongtai. A kind of new tracking algorithm of movement object [J]. optics engineering, 2005,32(1):43-46,70
[2] P.Nordlund, T.Uhlin, Closing the loop: detection and pursuit of a moving object by a moving observer [J]. Image and Vision Computing. 1996, 14: 265~275.
[3] Clarke J C, Zisserman A. Detection and tracking of independent motion [J]. Image and Vision Computing. 1996, 14: 565~572.
[4] St.Edmunds College. Active contour models for object tracking [J]. Computer Science Tripos, Part II, 2003,(5)
[5] Hogg D, Model-based vision: A program to see a walking person [J]. Image and Vision Computing. 1995, 1: 5~20.
[6] Rowe S, Blake A. Statistical mosaics for tracking [J]. Image and Vision Computing. 1995, 14: 549~564.
[7] Harrington S. Computer graphics, a programming approach [M]. New York, USA: McGraw Hill Company, 1987: 376-388P.
[8] Allen John G, Xu Richard Y D, J in Jesse S. Object tracking using Camshift algorithm and multiple quantized feature spaces [A]. In: Pan-Sydney Area Workshop on Visual Information Processing VIP2003[C], Sydney, Australia, 2003: 1-5P.