Proceedings of the 5th World Congress on Intelligent Control and Automation, June 15-19, 2004, Hangzhou, P.R. China

第五届全球智能控制与自动化大会
2004 年 6 月 15 － 19 日，中国杭州

Moving Object Tracking Research Based on Active Vision*

Pan Feng, Wang Xuanyin
The State key Lab of Fluid Power Transmission and Control
Zhejiang University
Hangzhou, Zhejiang Province, China
{pf & xywang}@zju.edu.cn

Wang Quanqiang
Zhejiang NanWang Multimedia Technology Co. Ltd.
Hangzhou, Zhejiang Province, China
wangqq@nanwang.com

***Abstract* - In an active vision condition both camera and object may move simultaneously and common motion detection methods could not deal with such situation. To overcome the shortcoming of common motion detection, a background-matching algorithm was proposed to solve such problem. When motion was detected, a Kalman predictor was used to estimate the object's position in the image. Combining increment PID controlling approach with dead zone, the camera can track the moving object steadily and reliably. Experiments proved the effectiveness of the algorithm.**

Index Terms -Active vision, Motion detection, Background matching, Kalman predictor, Increment PID controlling

I. INTRODUCTION

Visual analyses of moving objects have currently become one of the most active research topics in the domain of computer vision. The research mainly focuses on detecting, identifying and tracking objects, and even description or understanding of their behavior from the dynamic images. It has a wide application prospect and potential economic value on smart surveillance, advanced human machine interface, virtual reality, and content-based image storage and retrieval and so on. Especially in the long-distance visual surveillance system, there exists the need for people to sit before the monitor of real- time video. It's a heavy and low-efficiency work and easy to make a mistake. If the system has the ability to detect, identify and track objects smartly, the system will be more reliable and effective.

Various types of methods for object tracking has been proposed in related literature. Lipton [1] proposed a method for extracting moving targets from real-time video stream, classifying them into predefined categories according to image-based properties, and then robustly tracking them. Real-time surveillance system W4 [2] can detect and track multiple people and monitor their activities in an outdoor environment. Sun [3] presented a dynamic fuzzy approach to track an object robustly in a highly noisy background. Yang [4] developed an algorithm combined coarse searching with accurate location to solve object tracking in infrared image system. Wang [5] proposed a difference image based on motion detection algorithm, which can deal with objects overlapped and timely missing problem.

The key point of these technologies is to segment the objects from the background. It is also the premise for succeeding processing. In order to segment the moving objects, it is adopted that the camera stays still while the object is moving. But in practical, the camera's visual field is limited, thus fixing the camera will limit the visual field. If a wide-angle lens camera is used, the images will be usually too distorted to be accepted by users. Camera should be moved to enlarge the visual field. Camera and objects moving synchronously is the most common case in practical application, but it is difficult and a really challenging task to detect the motion of objects in such situation.

An auto object-tracking system based on active vision is developed in this paper. The camera can tilt and pan to scan the scene during detecting moving object. An approach based on background matching is proposed to solve motion detection under camera moving conditions. When background matching is completed, object is extracted through frame difference. The camera's motion parameters are obtained using a Kalman predictor. Then the camera is controlled to tilt and pan, thus the object is adjusted to stay at the center of the visual field. The whole system frame is shown as Fig. 1.

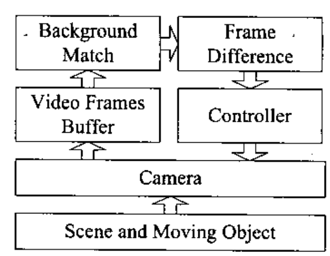


Fig. 1 Moving object detecting and tracking system

II. PRINCIPLE OF BACKGROUND MATCHING

Experiments show that if the camera moves not too fast, the background in the successive images will not vary too much. The background's motion in the images can be regarded as translation. Additionally moving object (the area of the object in the image) is relatively small compared to the whole image. Here a background matching method combined

* This work is supported by Aerospace Technology Support Foundation (No.HT2001-zjdx) and Science and Technology Development Project of Hangzhou City (No.2001121C42).

with frame difference technique is proposed to solve the problem of motion detection under the situation of camera and object moving synchronously. Background matching requires complicated background, which is actually the common situation in reality.

Considering the real-time requirement, background matching based on feature block matching is used, and a simple evaluation function followed by correlation algorithm is used for the feature block selection and matching.

Let the dynamic frames be: I_1, I_2, ..., I_{k-1}, I_k, I_{k+1}, ... In order to make the motion detection more robust, three successive frames are analyzed in a process cycle. Let base frame be I_k, and divide I_k into $n \times n$ blocks W_{ij}^k. It is considered that there may be some plain regions in the image and that block matching in these regions would increase noise to the matching result. So a simple function is used to evaluate the block W_{ij}^k (centered about (x,y)):

$$f = \sum_{(x,y)\in W}(I_k(x,y) - \sum_{(x,y)\in W} I_k(x,y)/n^2)/n^2. \quad (1)$$

If $f > \xi$ (ξ is an appropriate threshold), this block should be marked as a matching block.

For every matching block, it is needed to search frame I_{k-1} for the location of the best-matching block of the same size. The searching window is limited to a $m \times m$ region, which is also centered about (x,y). The displacement vector of the matching block is $r = (\Delta x, \Delta y)$. Matching criteria is the minimum mean square error (MSE):

$$MSE(r) = (\sum_{(x,y)\in W}(I_k(x,y) - I_{k-1}(x+\Delta x, y+\Delta y))^2)/n^2, \quad (2)$$

and $r = (\Delta x, \Delta y) = \arg\min_{(\Delta x,\Delta y)} MSE(r)$.

In order to accelerate the searching process, a three-step search strategy is used.

After the matching process is finished, it is to take statistics of the displacements of all matching blocks. Finding the peak point in the statistics, and the displacement vector, which the peak point corresponds to, can be taken as the background displacement r_{k-1}. And r_k is obtained by using the same method.

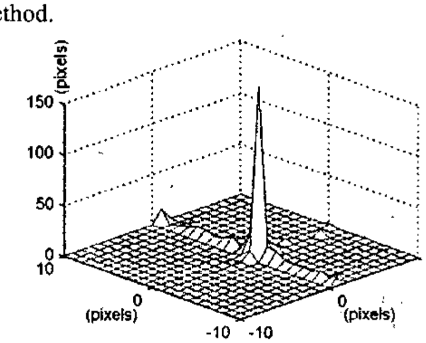


Fig. 2 Background matching

III. Moving Object Detection

The frames difference is defined as:

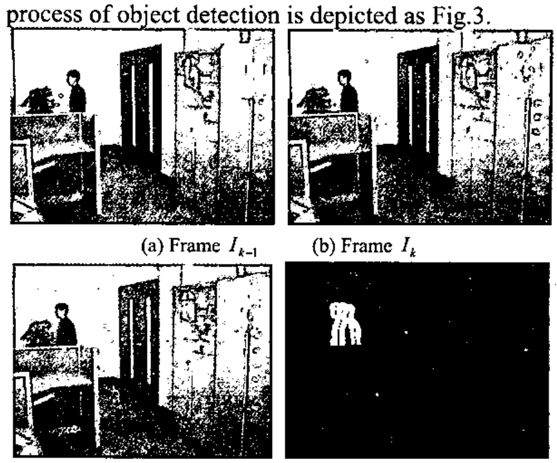
$$\begin{cases} D_{k-1}(x,y) = |I_{k-1}(x+\Delta x_{k-1}, y+\Delta y_{k-1}) - I_k(x,y)| \\ D_k(x,y) = |I_k(x,y) - I_{k+1}(x-\Delta x_k, y-\Delta y_k)| \end{cases}, \quad (3)$$

Object segmentation can be achieved by thresholding the difference image as:

$$B_k(x,y) = \begin{cases} 1 & D_{k-1}(x,y) > \delta \text{ 或 } D_k(x,y) > \delta \\ 0 & others \end{cases}, \quad (4)$$

where δ is an appropriate threshold.

The process of object detection is depicted as Fig.3.



(a) Frame I_{k-1} (b) Frame I_k

(c) Frame I_{k+1} (d) Binary image

Fig. 3 Example of motion detection based on background matching

Usually the binary image B_k contains lots of noisy points, which are caused by image noise and background partly matching. Sub-pixel level background matching could be used to obtain better result, but sub-pixel level background matching would increase the computation burden. In the practical application, morphology filter is implemented to eliminate the influence of noise.

IV. Camera Control

The primary work in this paper is to establish an auto object-tracking system about detecting moving object from the real time dynamic frames, tracking the object and keeping the object at the center of the camera view through servo control.

After motion detection, the object's position is obtained by calculating the barycenter of the area of the object in the binary image.

Let $d(k) = (x_{(k)}, \dot{x}_{(k)}, y_{(k)}, \dot{y}_{(k)})$, where $(x_{(k)}, y_{(k)})$ and $(\dot{x}_{(k)}, \dot{y}_{(k)})$ denote the position and moving speed of the object in the image at time k respectively. Object's moving speed and direction are usually continuous. So a Kalman predictor [6] is used to estimate the position of the object at time $k+1$. The control parameters are calculated according to the estimated result.

The Kalman predictor signal model is defined by the following equation:

$$d(k+1) = Ad(k) + w(k), \quad (5)$$

where A denotes a state transition matrix, w denotes process noise, and it's covariance is Q.

Measurement model is defined as:

$$y(k) = Cd(k) + v(k), \quad (6)$$

where C denotes a measurement matrix, v denotes measurement noise, and it's covariance is R.

Predictor equation is:

$$\hat{d}(k+1|k) = A\hat{d}(k|k-1) + G(k)[y(k) - C\hat{d}(k|k-1)], \qquad (7)$$

where predict gain:

$$G(k) = AP(k|k-1)C^T / [CP(k|k-1)C^T + R(k)], \qquad (8)$$

and predict error covariance:

$$P(k+1|k) = [A - G(k)C]P(k|k-1)A^T + Q(k). \qquad (9)$$

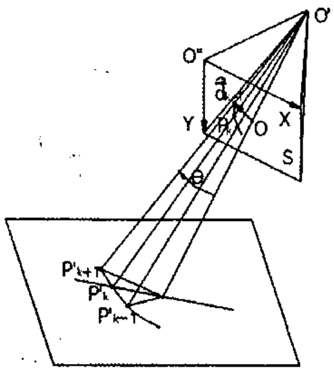


Fig. 4 Camera moving control

In order to adjust the object to stay at the center of the camera visual field, the camera should rotate an angle. The angle ($\hat{\theta}_{(k)}$) is calculated according to the following equation (derived from camera projection model, shown as Fig. 4):

$$\hat{\theta}_{(k)} = (tg^{-1}((\hat{x}_{(k+1|k)} - x_c)/f), tg^{-1}((\hat{y}_{(k+1|k)} - y_c)/f)), \qquad (10)$$

where f is effective focal length, and (x_c, y_c) is coordination of the image center.

In order to control the moving of the camera more robustly, increment digital PID controller with dead zone is applied to the servo control:

$$\Delta p_{(k)} = K_p(\hat{\theta}_{(k)} - \hat{\theta}_{(k-1)}) + K_i\hat{\theta}_{(k)} + K_d(\hat{\theta}_{(k)} - 2\hat{\theta}_{(k-1)} + \hat{\theta}_{(k-2)}), \qquad (11)$$

where K_p is proportional coefficient, K_i is integral coefficient and K_d is differential coefficient.

Let input of the servo controller be $\Delta\hat{p}_{(k)}$, thus:

$$\Delta\hat{p}_{(k)} = \begin{cases} \Delta p_{(k)} & |\hat{\theta}_{(k)}| \geq \phi \\ 0 & |\hat{\theta}_{(k)}| < \phi \end{cases}. \qquad (12)$$

V. SUMMARY

The process of the algorithm is depicted as Fig.5. After the initialization of system, three successive frames from video are captured and stored in the memory buffer. When the camera is scanning (meaning that the camera is moving), background matching is needed. The feature block in the middle frame is selected. Then feature block matching is processed between the frames. Frames difference is carried out according to the feature matching result. If the camera is not in the scanning status, we apply frames difference directly. In order to eliminate the influence of noise, morphology filter is applied to the binary image. Then the position and size of the object is calculated. A Kalman predictor is used to estimate the position of the object, and then tilt and pan the camera through servo control. At the same time, control the camera lens according to the size of the object. If the object is too small, pull the lens nearer. Otherwise pull the lens farther. Thus the object is always at the center of the camera view o an appropriate size. When camera lens' control is finished frames sampling again, and process continues.

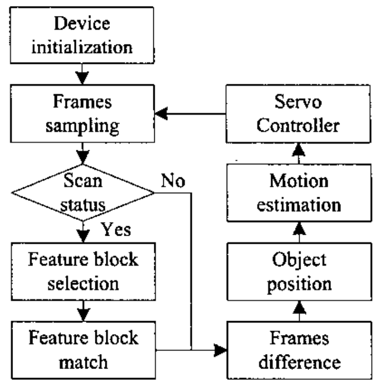


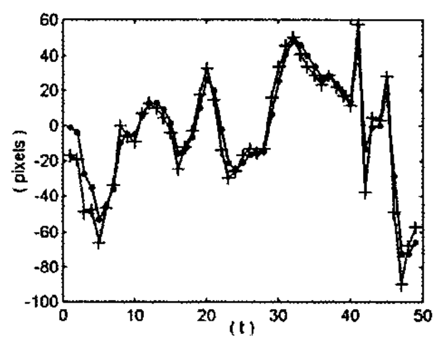
Fig. 5 Algorithm of object tracking

VI. EXPERIMENTS

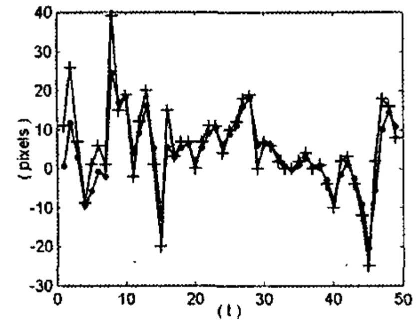
We implemented the algorithm on a PC platform (Pentium 800MHz) with a SDK2000 video capture card. Images were obtained from a dome camera. The camera is mounted on a pan-tilt unit, thus the visual filed of the camera can be changed at a speed of 0~40°/s through servo control. The camera type is VKS914, made by HITACHI Company. The captured images were 320×240 in size with 8 bits per pixel.

Experiments showed that moving object could be detected using background-matching method when the camera was scanning the scene. The camera could track the object through servo control, thus the object was always at the center of the visual field of the camera. We also found that if the background was simple or if the object was too large in the image, the object detection process might fail. It was mainly caused by lack of feature block in simple background and regarding the object as the background. In practical application, these cases can be avoided by placing the dome camera at a reasonable place.

Experiments results are shown in Fig. 6. It can be observed that the position of object obtained using a Kalman predictor is more robust than that obtained using only motion detection.



(a) x dimension of the object



(b) y dimension of the object

('+' denotes position obtained using motion detection only; '.' denotes position obtained using motion detection combined with a Kalman predictor. It is obvious that the '.' curve is more smooth than that of '+' curve, meaning that the control for the moving of the camera using the method proposed in this paper is more steady.)

Fig. 6 Kalman predictor results

VII. CONCLUSION

In this paper an object auto-tracking system based on active vision is proposed. The camera can tilt and pan to scan the scene during moving object detection. An approach based on background matching is proposed to solve motion detection with camera motion. When background matching is finished, object is extracted through frame difference. The camera's motion parameters are obtained through a Kalman predictor. Combining increment PID controlling approach with dead zone, the camera can track the moving object steadily and reliably. Experiments proved the effectiveness and robustness of this algorithm.

REFERENCES

[1] A. J. Lipton, H. Fujiyoshi, and R. S. Patil. "Moving target classification and tracking from real-time video," in Proc. IEEE Image Understanding Workshop, 1998, pp. 129-136.

[2] I. Haritaoglu, D. Harwood and L. Davis. "W4: real-time surveillance of people and their activities. their activities," IEEE Trans Pattern Analysis and Machine Intelligence, 2000, vol. 22, no. 8, pp. 809-830.

[3] Sun Wei, Liu Qizhen and HE Yongbao. "A Dynamic Fuzzy Approach for Tracking Moving Object,". Pattern Recognition and artificial intelligence. 1999, vol. 2, no. 2, pp. 91-198

[4] Yang Haoyun and Zhang Guilin. "Design and Realization of A New Correlation Tracker Algorithm," Journal of Infrared and Millimeter Waves. 2000, vol. 5, no. 9, pp. 77-380.

[5] Wang Gui, Ai Haizhou and He Kezhong. "Difference-Image-Based Multiple Motion Taregts Detection and Tracking," China journal of image and graphics. 1999, vol. 6, no. 4, pp. 470-475.

[6] C. Wang; A. Ohsumi and I. Djurovic. "Model predictive control of noisy plants using Kalman predictor and filter," TENCON '02. Proceedings. 2002 IEEE Region 10 Conference on Computers, Communications, Control and Power Engineering. 2002, vol. 3, pp. 1404-1407.