Human Target Detection and Tracking Based on Machine Vision

Yang Ai, Jing Liu*, Yun Gan, Kun Li

Faculty of Network Science, Haikou University of Economics, Haikou 571127, China

*Corresponding Author

Keywords: Machine Vision, Target Detection, Tracking

Abstract: the Detection and Tracking of Human Motion Target is an Important Content of Human Motion Vision Analysis and One of the Important Fields of Computer Vision Research. It Has a Wide Application Prospect and Great Economic Value in Intelligent Security Monitoring, Advanced Human-Computer Interaction and Detailed Analysis of Human Motion. with the Rapid Development of Artificial Intelligence Technology, It is of Great Significance to Apply Artificial Intelligence Machine Vision Technology to the Construction of Intelligent Hainan. This Paper Focused on the Detection and Tracking of Human Motion Targets under Video Surveillance of Machine Vision. First, the Image is Smoothed and Segmented by Preprocessing. Then, Binarization Background Modeling is Used to Extract the Target, and the Method Combining Histogram and Area Area is Tracked. Finally, by Extracting and Analyzing the Featrues and Movement Trajectories of the Moving Human Body, the High Recognition Rate of the Moving Human Motion Can Be Detected under Machine Vision. 1.Introduction

1. Introduction

Human Body Target Tracking is One of the Core Topics in Computer Vision Research. It Has Very Important Practical Value and Broad Development Prospects in Medical Research, Surveillance, Statistics, Security and Other Fields [1]. in the Colorful Information Society, Video Images Contain a Lot of Information. It Can Be Used as a Carrier of Objective Things, and At the Same Time It Can Vividly Describe and Intuitively Express Objective Things [2]. Therefore, Video Images Are More and More Favored by People Now. the Video Surveillance System Came into Being under This Background. Human Target Detection Refers to Extracting the Region of Interest from the Background Image from the Image Sequence to Make It a Meaningful Entity [3]. the Purpose of Computer Vision is to Realize Perception, Understanding and Interpretation of Scenery Environment, Thus Realizing Computer Simulation of Human Vision [4]. a Good Target Tracking Algorithm Often Needs to Process a Large Amount of Data for Each Frame of Image, Which Requires a Lot of Computing Time. in Real Life, Human Targets Contain Very Important Information, Such as Pedestrians or Cars Running through Red Lights, Speeding, Recognition of People's Actions, and So on. the Intelligent Monitoring System Uses Computer Vision and Digital Image Processing Technology. At the Same Time, Intelligent Video Related Modules Are Added to the Monitoring System Based on Pattern Recognition [5].

The Functions of the Intelligent Video Surveillance System Include Target Detection, Target Feature Extraction, and Target Tracking [6]. a Good Video Target Tracking System Must Be Able to Ensure Real-Time Tracking of Human Targets. the Ultimate Purpose of the Intelligent Video Surveillance System is That the Computer Can Automatically Analyze the Video Image. through Some Computer Vision Technology, the Collected Surveillance Video is Automatically Analyzed and Detected to Obtain the Position, Movement Direction and Behavior of the Human Target. Human Target Tracking Refers to the Estimation of a Target State Based on Spatiotemporal Integration Based on the Initial State of the Target and the Target Model Obtained through Feature Extraction. the Most Critical Technology is Still Human Target Detection and Tracking Technology. Human Target Detection Technology Specifically Refers to the Detection and Extraction of Foreground Objects with Relative Motion from the Background Image of Surveillance Video [8].

  DOI: 10.25236/ISMHI.2019.061

the Moving Foreground Target is Further Divided into a Number of Independent Human Target Regions Based on Characteristics Such as Grayscale, Edges, Texture, and Color. and Use the Computer's Powerful Data Processing Capabilities to Remove the Information That Interferes with the Video Picture [9]. in Video Surveillance, Human Body Target Tracking Can Not Only Provide the Motion Trajectory of the Monitored Human Target, But Also Provide a Reliable Data Source for the Motion Analysis and Scene Analysis of Human Targets in the Scene.

2. Human Target Recognition Scheme

When Acquiring a Video Frame Image, the Image is Inevitably Affected by Noise and Environmental Factors, and It is Likely That the Obtained Image is Not a Qualified Registered Image. Unqualified Original Images May Lead to Excessive Error Matching. to Construct an Image Motion Vector Field, Each Pixel in the Image is Given a Velocity Vector [10]. Motion Analysis of Video Sequence Images Mainly Includes Several Processes of Human Target Detection, Classification, Tracking and Behavior Analysis. the Detection of Human Target is the Foundation of the Whole Motion Detection and Tracking System. Only by Accurately Separating the Target Can the Characteristic Information of the Human Target Be Accurately Calculated. At Any Given Moment, the Points on the Human Target in the Image Correspond to the Points on the Real Three-Dimensional Target Object One by One. during Shooting or Transmission, the Random Interference Signal to Which the Image Signal is Subjected is Image Noise. after the Spatial Domain Image is Converted into the Frequency Domain Image, If the Low Frequency Part is to Be Enhanced, It Can Be Realized by Smooth Filtering [11]. the Reserved Low-Frequency Components Are Restored to the Spatial Domain of the Image for Subsequent Operations. Spatial Domain Enhancement of Video Frames Performs Some Processing on Each Pixel of the Image, Such as Gray Scale Transformation, Sharpening of the Image, Etc. in a Video Sequence, If There is Motion of an Object, the Gray Value in the Motion Region Will Suddenly Change.

In the Video Sequence, the Gray Values of Corresponding Pixel Points in Two Adjacent Frames Are Compared, and Then Differences Are Found. If There is No Human Target in the Image, the Light Flow Vector Changes Continuously in the Whole Image Area. However, If There is a Human Target, That is, There is Relative Movement between the Human Target and the Background Image. Then the Velocity Vector Formed by the Human Target Must Be Different from the Velocity Vector of the Neighboring Background. There is a Certain Gap in Intelligence, Especially Limited by Many Factors in Practical Application, and Further Improvement is Needed. Figure 1 is a Frame of a Motion Detection and Tracking System Using Dense Disparity Variance Technology.

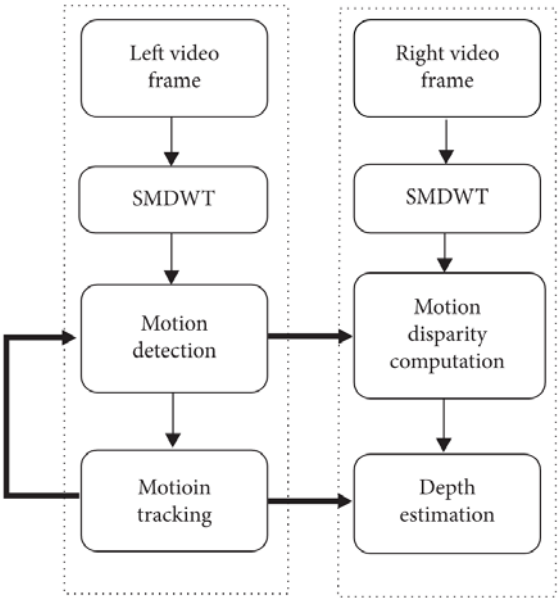

Fig.1 Framework of Motion Detection and Tracking System Using Dense Parallax Variance Technology

In general, the faster the update speed of the background model, the better the effect order of the obtained motion region. The original monitoring technology used a fixed camera and its viewing angle was limited. There are different geometric correction methods in different fields, and usually the geometric deformation of remote sensing images is large. Geometric deformation can be used to correct it. The principle of fidelity of images is not followed, but information of interest is highlighted and useless information is suppressed. If the speed of the target is too slow, the frame difference method will be used to obtain an image in which the target is over-covered and almost completely overlapped, thus making it impossible to detect the human target. We use the context-related features of the human body in the video image sequence and combine the template matching method based on correlation to solve the occlusion problem. Factor analysis is carried out on the indexes of sports events, and the index system is suitable for factor analysis. There is no general theory for image enhancement technology. The final judgment of the observer determines the advantages and disadvantages of a certain enhancement technology. Two or several temporally adjacent images in the image sequence are subjected to differential operation [13]. In the background method, there are many commonly used models, and different models adapt to different environments. The method is simple and easy to implement, has good robustness, and is very reliable in indoor and outdoor background models and widely used.

3. Overall System Analysis

3.1 Human Motion Detection Module

In order to realize all-round monitoring, it is necessary to arrange multiple cameras in the monitoring scene to assist in the work, which not only increases the cost but also makes the system more complicated. When the camera shakes, the original stationary target in the background moves or the foreground target becomes part of the background, the original background model is no longer suitable. Therefore, the background model needs to be continuously updated so as to extract better foreground targets. The primary purpose of image enhancement technology is to make the processed image more suitable for specific applications than the original image. The purpose of tracking is to keep the target within the field of view. Therefore, according to the location of the target centroid, it is necessary to control the pan-tilt head to drive the camera to rotate. In order to change the field of view, the target is always within the field of view. The necessary action points can be quickly located. Table 1 shows the detection rate when the human body is not occluded. Table 2 shows the detection rate with occlusion.

Table 1 Detection Rate Of Unobstructed Human Body

	Detection rate (%)	Discriminant rate (%)
Video 1	94.7	80.5
Video 2	93.12	77.6

Table 2 Detection Rate Of Human Body Being Blocked

	Detection rate (%)	Discriminant rate (%)
Video 1	82.6	75.6
Video 2	78.5	73.2

Because two adjacent frames in a video image have certain continuity, when there is motion of a moving object, the difference of pixel gray values between adjacent frames will increase. The input and output data are randomly selected and submitted to the network. Calculate the output of each neuron in the hidden layer:

$$f(t)=\sum_{j=1}^{N}\sum_{k\in Z}d_k^j\phi_{jk}(t)+\sum_{k\in Z}c_k^N\phi_{Nk}(t) \tag{1}$$

Calculate the response of neurons in the output layer:

$$E_{mi} = \sum_{i=1}^{k} (i\Delta t) \cdot \left| S_{mi} \right|^2 \tag{2}$$

Using the given output data to calculate the error of neurons in the output layer:

$$\vec{E} = \frac{E_{mi}}{\sqrt{\sum_{i=1}^{k} E_{mi}}} \tag{3}$$

3.2 Human Motion Tracking Module

The target to be tracked is detected in the frame to be tracked, and the image contour is binarized. Finally, the two templates obtained from the first two parts are compared. If the matching is successful, the tracking is successful. The algorithm has the probability of misclassifying the background or part of the background into targets, but the misclassification probability is very small. If the pixel features or pixel areas at corresponding positions have certain differences, it is considered that the pixel points or pixel areas at these positions in the video image of the frame have changed. Then these change regions are extracted according to different judgment criteria to form the foreground target motion region. Figure 2 is a human body height estimation model using a reference height.

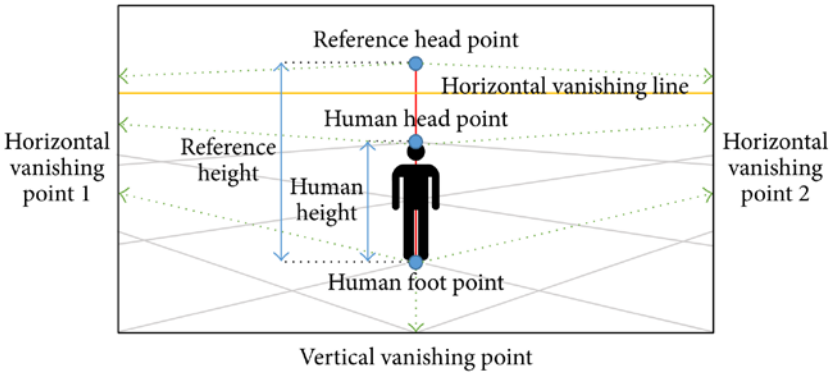

Fig.2 Human Height Estimation Model Using Reference Height

Although there are slight differences between two adjacent frames of images, the color information of the target remains basically unchanged. This is also a major factor for us human beings to identify objects. According to the deviation between the target coordinate position and the center of field of view, the PTZ is controlled to drive the camera to rotate to change the monitoring angle of view. Relevant parameters of collinear equation can be obtained through prediction, and the correction number of pixel points can be obtained by using the least square method to achieve the correction purpose. The collinear equation correction method is an accurate correction method, and has good correction effect on pictures with uneven geographical distribution and static sensor pictures. The target template is not fixed. On the contrary, in order to adapt to the changes of targets in different scenes, the template needs to be updated in time. If the target is not completely occluded and any part of the features can be identified, the target action can be characterized. Then the algorithm can continuously track the human target for a long time.

4. Conclusion

At present, the difficulty of target tracking lies in the contradiction between real-time and accuracy. If we want to obtain high accuracy at the same time, we must sacrifice real-time performance and reduce the accuracy of applications requiring high real-time performance. Human target detection is a key research issue in the field of computer vision, and it is also an active research topic that crosses many disciplines. It plays an extremely important role in video surveillance, military affairs and medical treatment. People have proposed many algorithms to solve

the problem of human target detection and have made great progress. Intelligent video surveillance system has a wide application prospect. It can contain richer and higher-level functions and bring more and more complex problems. In complex background, although the motion detection algorithm based on background difference can effectively filter out noise points of binary segmented images or some isolated points caused by dynamic background interference through mathematical morphology processing and other filtering methods. Constantly break through the difficulties, make the monitoring system more intelligent, and have better real-time and reliability, and finally will certainly be able to develop a more practical intelligent video monitoring system.

Acknowledgement

Funding: This work is supported by the Higher Education Research Project of Hainan Province (Hnky2019-73), the Hainan Provincial Natural Science Foundation of China (No. 619QN246) and the Key Research Project of Haikou University of Economics (HJKZ18-01).

References

[1] Bai Xiaofang, Yang Wei, Chen Peiyu. An Improved Method of Moving Object Detection and Tracking [J]. TV Technology, 2014, 38 (1): 180-182.

[2] Liu Chao, Li Xiuyou, Huang Yong, et al. Optimized multi-model particle filter tracking method for mobile weak target detection [J]. Signal Processing, 2015, 31 (9): 1131-1137.

[3] Liu Yawei, Li Xiaomin. Overview of Target Detection and Tracking Methods in UAV Aerial Video [J]. Aircraft Missiles, 2016 (9): 53-56.

[4] Tingting Wang, State Key Laboratory of Robotics and Systems, Harbin Institute of Technology, Harbin, Heilongjiang, Wang Tingting, et al. Binocular vision-based moving object detection and tracking [J]. Machinery and Electronics, 2015 (6): 73-76.

[5] Ma Chao, Shen Wei, Dong Jingfeng, et al. A specific moving target detection and tracking method in complex backgrounds [J]. Computer Engineering, 2015, 41 (5): 219-223.

[6] Jiang Mingxin, Qiu Tianshuang, Chi Hongbo. A new multi-target tracking algorithm based on the fusion of color and depth information features [J]. Optoelectronics · Laser, 2015 (7): 1342-1348.

[7] Li Miao, Long Yunli, Li Jun, et al. Tracking before detection of oversampling point targets using Dobernuli filter [J]. Optics and Precision Engineering, 2015, 23 (12): 3446-3455.

[8] Tian Helei, Ding Sheng, Yu Changwei, et al. Research on Video Abstraction Technology Based on Object Detection and Tracking [J]. Computer Science, 2015, 43 (11): 297-299.

[9] Fang Guoqiang, Teng Kenan. Method for Mid-range Target Detection and Tracking of Ballistic Missile Based on Sequence Infrared Image [J]. Aerospace Measurement Technology, 2014 (1): 50-56.

[10] Yu Hongbo, Cao Qian, Yang Zhigang. A stealth target tracking method based on extended model before detection [J]. Journal of Naval Aeronautical Engineering Institute, 2014, 29 (6): 569-574.

[11] Qin Yanyuan, Jing Zhongliang, Lei Ming. Multi-sensor target detection tracking and classification algorithm [J]. Computer Simulation, 2014 (9): 364-368.

[12] Jiang Mingxin, Wang Hongyu, Qiu Tianshuang. Multi-target tracking algorithm based on target detection and graph cutting [J]. Journal of Dalian University of Technology, 2014 (6): 632-636.

[13] Liu Weihua, Tan Shuncheng. Reflections on Stealth Target Detection and Tracking [J]. Journal of Naval Aeronautical Engineering Institute, 2014, 29 (5): 410-414.