Research on Human Behavior Recognition

Yongliang Xiao[a], Xiangbao Li[b], Wenbin Liu[c] and Canwei He[d]

School of Information Technology and Management, Hunan College of Finance and Economics, Changsha 410205, China

[a]xylroc@163.com, [b]251306550@qq.com, [c]2913510@qq.com, [d]24253006@qq.com

Keywords: Human behavior recognition, Human tracking, Artificial intelligence, Computer vision

Abstract: In recent years, human behavior recognition has become a research hotspot in the field of computer vision and pattern recognition, which have some important application prospects in intelligent monitoring and smart home. The process of human behavior recognition is quite complex, especially in complex environments, which involves many aspects such as feature extraction, human tracking and behavior recognition. This paper summarizes some main aspects of human behavior recognition and describes the future research directions.

1. Introduction

Various emergencies have emerged in the world such as the "911" terrorist attacks in the United States, the Madrid bombings, and the Turkish bombings. People all over the world are paying more and more attention to the safety of their residence. The demand for intelligent monitoring in key areas is also becoming more and more significant. Many important locations are equipped with surveillance cameras, which can provide critical investigative information in criminal offenses. People want computers to have certain thinking functions of human beings. It is even expected to take the analysis of the outside things in place of the human eye and brain. With the rapid development of the global economy, various computer-related technologies are constantly improving. People use computers to solve certain problems in their work, especially in terms of vision. Video based human behavior recognition is an important branch of the application field of public safety intelligent systems.

At present, these monitoring systems only retain the current scene through video recording, and cannot automatically monitor the behavior in real time. This is mainly because the current video surveillance system requires manual participation to complete. Due to the large number of surveillance areas, it is impossible to realize all manual monitoring, and it is impossible to make timely judgments on various emergencies. The security is relatively poor. At the same time, the changes of noise, brightness and illumination and the change of the viewing angle further increase the difficulty of human behavior recognition.

2. Related Research

In recent years, human behavior recognition technology has become a research hotspot in the field of computer vision and pattern recognition. It has broad application prospects and potential economic value in intelligent video surveillance, content-based video retrieval, sports event analysis, human-machine interface, and virtual reality. However, due to the complexity of the real environment, the variability of human posture in sequences, and the inaccurate definition of human behavior patterns, human behavior recognition is becoming a very challenging research topic.

Many countries have begined to deploy video surveillance programs. The program mainly uses video analysis, computer communication, information fusion and other technologies to automatically monitor and deal with battlefields and other sensitive scenes. Then monitor the sensitive area through, and automatically detects the alarm when the sensitive area appears significant moving object. The video surveillance program includes a number of advanced real-time monitoring algorithms, such as

using mobile cameras to track moving targets; further dividing moving targets into common targets and key targets to set different warning levels; analyzing and understanding human gait behavior and simple behavior of multiple moving targets, etc. Human behavior recognition technology emerges as the times require, through the camera to obtain external motion information, to achieve human target detection, human target tracking, human behavior feature extraction and human behavior understanding and analysis.

Human behavior recognition is a key that must be solved in social public places with achieve intelligent monitoring. Although many scholars have done in-depth research in these areas, human behavior recognition research in complex environments is a complex problem across multiple disciplines. This is mainly because the human behavior patterns have different meanings in different environments, and it is impossible to accurately define abnormal behaviors. At the same time, due to the diversity and variability of human behavior, the human behavior recognition in complex environments needs to be further improved.

3. Human Tracking

In the video surveillance system, stable tracking of moving human targets is a basis for subsequent human behavior analysis and understanding. The moving human tracking generally considers the various environments in which the human targets are located, and determines certain features that can fully represent the moving human targets. These commonly used features mainly include texture, shape, and color. Generally, the initial position of the human target is first determined, and then the subsequent video frames are analyzed according to some significant features, and then matched with the features of the previous video frame, and the dynamic information of the moving human target is obtained. The characteristics of the motion trajectory complete the requirements for the positioning of the moving human, and continuously track the target of the moving human. Therefore, the main task of moving human target tracking is to select the appropriate tracking and matches strategy to achieve the tracking of the moving human, especially to achieve stable tracking of the moving human target in complex scenes.

3.1 The region-based model

The basic idea of the region-based motion human tracking method is that we first use the motion estimation algorithm to obtain a template containing the moving human. The template can be more than a human target, and the shape of the template can be regular or irregular. Then we use the various tracking techniques to track the moving human target in the moving human video sequence.

Paper [1] proposed a method for human tracking through multi-region joint particle filter to solve the problem of occlusion of human motion in video sequences. Paper [2] considered that the color image of the human motion target detection and segmentation algorithm is not easy to obtain the complete frame of the human, and only the two-dimensional coordinates of the human joint can be obtained. They proposed to use the sensor to capture the depth of the human image, and this information can be used to reconstruct the human skeleton, which greatly improves the three-dimensional image of the joint point in the human skeleton model. Specifically, this paper uses a human skeleton model to get a three-dimensional information description based on coordinate transformation strategy.

3.2 The active contour models

The active contour model is to use the boundary contour of the moving human to describe the human target. Firstly, the boundary of the moving human target is obtained by the image segmentation method, and the boundary is used as the initial template of the human target, and then in the subsequent video sequence, the boundary analysis of the human target is continuously updated, and finally the contour of the human is gradually consistent with the true contour of the target. In fact, the process is a continuously optimized process.

Paper [3] analyzed the advantages and disadvantages of the Snake active contour model. For the target tracking problem of the traditional Snake in the design and automatic initialization algorithm, the four-neighbor search strategy is used instead of the eight-neighbor search method, the calculation time of tracking is greatly reduced. Paper [4] considered that the human tracking is easy to be affected by factors such as illumination, complex environment and occlusion. To solve this problem, this paper proposes a method of obtaining depth images for human target tracking using Kinect technology. Paper [5] considered that the standard particle filter algorithm for human target tracking has the characteristics of nonlinear and non-Gaussian models, and is a typical system-based tracking problem. However, the particle update process is strictly dependent on the choice of parameters and cannot handle changes in the curve topology.

3.3 The classifier models

The use of classification methods to track the human is a new idea that has emerged in recent years, in which the human motion tracking problem is equivalent to a classification problem. The classifier can clearly separate the human from the background, and constantly update the parameters of the classifier to achieve tracking of the target. These methods select appropriate human motion information and background information, which can greatly improve the adaptability of the algorithm. At the same time, for human motion tracking in complex environments, the stability and tracking effect are significantly improved. However, this method has some difficult problems, such as how to select the feature and update the parameters of the classifier.

Paper [6] proposed a novel moving target tracking algorithm to improve the robustness of target tracking. The main contribution of the algorithm is to convert the tracking problem into two categories of target and background classification problems, and then use the positive and negative data samples in each frame image to obtain the support vector machine parameters, and finally use the appropriate parameter with adjustment optimization strategy to make a classifier. Paper [7] viewed the target tracking problem in video sequences from the perspective of two types of pattern classification techniques, and proposed a target tracking algorithm based on Adaboost learning technology. Paper [8] proposed to use the fusion texture contour and color feature strategy in feature selection step, then incorporate the Camshift algorithm into the online Adaboost algorithm to obtain a more accurate position of the moving target. Considering a large change in the posture, shape and clothing of the human, paper [9] proposed a method based on statistical strategy for detecting and tracking moving human targets. To improve pedestrian detection and tracking accuracy, a pedestrian tracking algorithm based on Kalman filter is proposed [10].

The actual monitoring environment is quite complex, such as dynamic changes of weather and illumination, multi-moving targets with interactive motion and occlusion, so the accurate tracking of human motion has not been well solved, which makes human behavior recognition very difficult.

4. Human behavior recognition

Human behavior recognition has become a challenging research topic in the field of computer vision due to the combination of human behavior analysis and understanding of advanced technologies in image processing, pattern recognition, artificial intelligence and computers.

Paper [11] proposed the use of behavioral motion energy images and motion history images to represent the behavior of each human. Both are calculated from the sequence image according to the change of each pixel before and after. The video motion information is projected onto an image for dimension reduction processing, while retaining the motion information, and then template matching is used to determine the degree of matching between the test sequence and the reference sequence to obtain the category attribute of the behavior. Sheikh et al. used 13 joint point trajectories to represent the behavior of the human. The advantage of this method is that it can analyze various movements of different human bodies in detail, but the disadvantage is that it needs to accurately detect and track the various components of the human. Paper [12] proposed a human behavior recognition model based on improved canny operator and neural network to improve the correct rate of human behavior

recognition. Paper [13] extracted a pose feature with translation, rotation and zoom invariance from human motion video, which can retain the local pose of the human. Paper [14] proposed a human behavior recognition method based on deep learning. First, all images in the training sample set and the test sample set are preprocessed, and the target motion foreground is extracted by the Gaussian mixture model. Secondly, a sample set is built for various target behaviors in the training sample set. Finally, combining with the network model obtained by deep learning, the classification identifies various behaviors in the test sample set.

To solve the problem that the energy-like graph is susceptible to the movement time and position of the human, paper [15] proposed a human behavior recognition method based on the energy-like graph pyramid gradient histogram fusion feature and multi-class Adaboost classifier. Paper [16] proposed a human behavior recognition method based on Fourier-Hidden Markov Model. Paper [17] used the convolution neural network algorithm in deep learning to automatically extract the acceleration characteristics, and then combines the decision tree algorithm to realize the recognition of human behavior. Paper [18] used the classification of motion pose descriptors and the word bag model to classify human behavior.

At present, most of the behavior recognition methods only identify simple behaviors, and mainly judge the similarity of low-level visual features of behaviors, ignoring the duration of motions, and failing to consider some occlusion situations. The behavior of the representative is different. In fact, complex behavior is composed of a set of coherent actions, not only related to the content of the action, but also related to the order of each action and the duration of each action. Therefore, it is necessary to recognize the behavior according to the visual similarity, chronological order and time span of the behavior.

In addition, the meaning of human behavior often depends on the objects involved in the behavior, which in turn makes it necessary to have a deep understanding of the actions. Most of the current behavior recognition methods only identify simple behaviors, and there are few studies on long-term motion sequences containing different actions. At the same time, it is rare to study the loss of the motion behavior and the inconsistent order of the complex long sequence behavior. In addition, existing methods can only identify simple, pre-defined behaviors, without the ability to automatically learn new behaviors.

5. Conclusion

Video is generally present at the same time as other related texts and pictures, with a deep semantic background, and most videos cannot be directly downloaded. How to search and identify human behavior directly in mobile intelligent terminals is a subject to be further studied. There is still much work to be done in human behavior recognition in complex environments. We believe that with the continuous development of related disciplines, the research and application of human behavior recognition will certainly achieve greater results.

Acknowledgments

This work was supported by Hunan Provincial Natural Science Foundation of China No. 2017JJ2015.

References

[1] Y. R. Wang, J. F. Liu, G. J. Lu, et al, People tracking based on multi-regions joint particle filters [J]. Acta Automatica Sinica, 2009, 35(11):1387-1393.

[2] X. Chen, Q. H. Meng, Human action r ecognition based on skeleton joints tracking [J]. Journal of Henan University of Science and Technology: Natural Science, 2015(2):43-48.

[3] T. Fang, Z. Yand, C. L. Shen, A moving object tracking method based on active contour models [J]. Information and Control, 2008, 37(1):99-102.

[4] H. F. Yang, G. S. Zhang, X. B. Wang, et al, Human object tracking in depth image based on Snake algorithm [J]. Journal of Tianjin University of Technology, 2014, 30(5):41-45.

[5] J. Cao, Q. H. Zeng, J. H. Wang, Particle filter tracking algorithm based on geometric active contours[J]. Journal of Computer Applications, 2011, 31(5):1205-1208.

[6] Z. Y. Wang, K. D. Zhang, Y. Wu, et al, Target tracking in infrared image sequences by combining SVM and AdaBoost [J]. Journal of Image and Graphics, 2007, 12(11):2052-2057.

[7] J. P. Jia, F. Z. Zhang, Y. M. Chai. Adaboost object tracking algorithm [J]. PR&AI, 2009, 22(3):475-480.

[8] Z. W. Lv, K. J. Wang, H. Y. Li, et al, Online Adaboost target tracking algorithm combined fused with Camshift [J]. Journal of Central South University, 2013, 44(S2):232-238.

[9] K. Hashimoto, H. Kataoka, Y. Sato, et al, Human tracking with statistical shapes model and rough pose estimation by random forest [J]. Journal of the Japan Society for Precision Engineering, 2015, 81(12):1162-1167.

[10] Z. P. Wan, Pedestrian tracking algorithm based on human body characteristics identification and Kalman filter [J]. Electronics Optics & Control, 2016, 23(11):97-102.

[11] A. Bobick, J. Davis, the recognition of human movement using temporal templates [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2001, 23(3): 257-267.

[12] Y. G. Hong, Human action recognition based on improved Canny operator and neural network [J]. Computer Engineering and Applications, 2013, 49(8):156-159.

[13] J. X. Cai, G. C. Feng, X. Tang, et al, Human action recognition by leaning pose dictionary [J]. Acta Optica Sinica, 2014(12):173-184.

[14] H. Fan, J. Xu, Y. Deng, et al, Behavior recognition of human based on deep learning[J]. Geometrics and Information Science of Wuhan University, 2016, 41(3):10-15.

[15] S. W. Ma, L. N. Liu, Q. Fu, et al, Using PHOG fusion features and multi-class Adaboost classifier for human behavior recognition [J]. Optics and Precision Engineering, 2018, 26(11):2827-2837.

[16] W. Hu, Human behavior recognition based on Fourier-hidden Markov model [J]. Electronic Design Engineering, 2018, 26(7): 185-188.

[17] Z. M. Wang, C. Zhang, X. Heng, Research on new human behavior recognition method based on CNN and decision tree [J]. Application Research of Computers, 2017, 34(12):3569-3572.

[18] X. Yao, Behavior recognition based on motion posture descriptor and bag-of-words model [J]. Computer Era, 2018(05):18-20.