Research on Responsive Information Interaction Design Based on Facial Recognition Technology

Chen Ma

Research Center for Smart Policing and Big Data Technology, China People's Police University, Langfang, Hebei, 065000, China
mulanchenchen@sohu.com

Abstract. The recognition and tracking down moving figures have always been a hot research topic in the field of computer vision. With the development of artificial intelligence in recent years, the application of deep neural networks involves various fields. Compared with the previous solutions, some problems even have better solutions based on deep neural networks. This article mainly discusses targeting people based on face recognition and realizing movement tracking down the target people. The research content of this article can be divided into two aspects. The first aspect is the identification of the target person. Since the overall characteristics of the target individual are not recognizable, and the overall recognition easily leads to errors, this article uses the method of identifying the target person based on the face. This article introduces the application of deep neural networks of facial recognition, and finally selects the SeetaFace faces recognition system that does not depend on any third-party library functions and is open source. The second aspect are the tracking of sports figures. When the algorithm recognizes that the similarity between the image and the target face reaches a certain level, it will automatically switch to tracking the overall torso of the moving person. In the tracking part, a variety of algorithms are compared and compared with experimental data. Experiments have verified that the convergence of face recognition and tracking algorithms is feasible, and this solution not only prevents the interference and occlusion of similar colored objects of the tracking process, but also ensures the stability of the tracking process and is robust., Accuracy and real-time performance have also achieved good results.

Keywords: Face recognition, target tracking, Kalman filter.

1. Background and significance
With the rapid development of the global economy, the gradual improvement on the level of science and technology, and the continuous enhancement of my country's national strength, the fields of transportation, security, banking, construction, power, storage and even military construction are increasing the demand for security and video surveillance, and the requirements are also increasing [1]. Higher and higher. At the same time, due to the development of computer hardware technology, the performance of computers in terms of processing speed, storage capacity, and parallel computing has
improved, which provides the possibility for the rapid processing of intelligent real-time image sequences. That is, with the rapid development of information technology brought about by the information explosion, the processing and analysis of surveillance video has a wide range of market demands in the fields of civil, commercial, and even national defense security and military applications. As a result, the field of computer vision has developed rapidly in recent years, and a more researched application direction has gradually emerged, intelligent video surveillance (IVS) [2]. IVS realize the control of the video surveillance system based on the analysis, processing and understanding of video information, and uses machine vision technology to realize the collection of video information, so that the video surveillance system has better intelligence and robustness.

The use of intelligent video surveillance system is very wide, mainly related to the subject knowledge of image processing, computer vision, pattern recognition, artificial intelligence, etc. It has a wide range of application prospects of civil, commercial and military fields. For example, in terms of home security, surveillance cameras are installed to record people who have appeared in residential areas, and an alarm system is combined to prevent theft during unmanned hours.

2. Research status of face recognition

2.1. Historical development of face recognition problems

Face detection and recognition is a classic problem in the field of pattern recognition, involving technologies and knowledge in multiple fields. The research on face detection and recognition technology can be traced back to the work of Frenchman Galton in the nineteenth century. As early as 1888 and 1910, he published two articles on how humans use people in Nature. Facial features are used to identify the article, but this is not machine recognition. The development of the more iconic face recognition technology can be divided into three stages [3].

The face recognition system can be classified from a variety of perspectives, such as classification according to the feature extraction method, classification according to the recognition method, classification according to the characteristics of the input signal, and so on. The well-known classification method is classified according to the recognition method. The mainstream faces recognition methods can basically be classified into three categories, namely: geometric feature-based methods, template-based methods, and model-based methods. The method based on geometric features is the earliest and most traditional method of face detection and recognition [4]. The essence is to form a feature vectored representing the face. It usually needs to be combined with other algorithms to have better results; template-based methods can be divided into Correlation matching method, eigenface method, linear discriminant analysis method, singular value decomposition method, neural network method, dynamic connection matching method, etc. Model-based methods include methods based on hidden Markov models, active shape models, and active appearance models.

2.2. Face recognition system based on deep neural network

The methods of using deep neural networks of face recognition can be divided into two categories. The core of the first type of method is big data, which recognizes the diversity of human faces through the accumulation of data. This type of method is suitable for large enterprises capable of big data collection. In the future, more and more comprehensive training data sets can be used to improve the accuracy of distinguishing face differences in different scenarios. The second type adopt the idea of synthesis, adding data sets through 3D models, etc, to synthesize different types of faces. This type of method has low operating cost and is more suitable for personal applications. The representative one is CNN-3DMM estimation.

CNN-3DMMC (3D morphable face models) is a three-dimensional variable face model based on convolutional neural networks. When applying 3D simulation in a real scene to increase the accuracy of face recognition, there is two types of problems. One is that the 3D simulation is unstable, which causes the 3D simulation of the same individual to be quite different; the other is that it is too generalized, which causes most of the synthesized pictures to be similar. Therefore, a robust 3D variable face
modelled generation method is studied, using a deep convolutional neural network method. Method of 3D faced reconstruction. They used a convolutional neural network (CNN) to adjust the face shape and texture parameters of the three-dimensional face model according to the input photos. This method can be used to generate a large number of labeled samples. There are two key points of this method: one is 3D reconstruction model training data acquisition; the other is 3D reconstruction model training [5].

In face recognition algorithms, face++ and others are service interfaces, not open-source codes, and the accuracy of CNN-3DMM evaluation depends on the construction of the deep neural network and the training results. Seetaface is known as a clear stream in the open sources to face recognition industry, and it has the characteristics of strong robustness and fast recognition speed. Therefore, this article will use seetaface algorithm to identify and track the target person.

3. Optimization of tracking algorithm

3.1. Target person tracking area selection based on face recognition

The selection of the tracking area in the first frame determines the degree of learning of target features. It is inevitable that background information will also be mixed into the selection, so it is necessary to circle as accurately as possible. In this paper, the tracking of the target person is based on face recognition. When the video image is obtained, real-time detection of whether there is a human face in the video, once the face similarity is greater than the set threshold, immediately according to the return to the current frame of the target person's face 5 feature pointed to coordinate (respectively the left eye coordinates, right Eye coordinates, nose coordinates, left mouth corner coordinates, right mouth corner coordinates). Based on the location of the target person's five senses, combined with the standard human body proportions, the tracking area is delineated through calculations.

Since it is considered that when the face of the target person is recognized in the picture, the body may be partially blocked by obstacles, the movement towards the person's head, and the complete body of the target person may not be fully captured in the picture. Combining the results of multiple experiments, the algorithm for the corners of the mouth selects a rectangular tracking area in the first frame where the target face is recognized. Compare the left eye and the left x coordinate and take the minimum value as X, compare the left eye and right eye and take the minimum value as Y, and take the maximum value of the difference between the eye distance and the x coordinate of the right eye to the left corner of the mouth. width, compare the difference between the right corner of the mouth and the right eye Y to compare the difference between the two coordinates and the difference between the right corner of the mouth and the left eye Y coordinated, take the maximum value and record it as height. The upper left vertex of the tracking area takes the left eye coordinate and moves it horizontally to the left by width, and then moves up to eight in the vertical direction [6]. The width of the rectangle is 3 times the width, and the height of the rectangle is 6 times the height.

3.2. The predictive effect of the Kalman filter

When tracking a moving target, the target may be blocked by obstacles. At this time, the tracking target does not appear on the screen, so no matter which feature matching method is used, the tracking target cannot be found.

In order to solve this problem, the function of correlation filtering can be used. The three major uses of filters are filtering, prediction, and smoothing. Considering the real-time nature of the algorithm, a simpler Kalman filtering method is selected. The overall idea is to use Kalman to predict a larger search area, and then use the reverse histogram to determine the tracking area within the search area [7].

The Kalman filter is generally used in the state estimation of linear systems, and the state of the target is assumed to be Gaussian. It is essentially a combination of prediction and correction. The prediction comes from the empirical model, which is derived from the modeling of the system by humans, that is, the movement towards the target of the next frame of image is predicted from the movement towards the target of the previous image. The other part is the measurement correction, that is, the correction of the target motioned model after the target state is measured on the current screen. Among them,
observation noise and measurement noise are quantitative descriptions of model and measurement uncertainty, respectively. Figure 1 shows the information flow diagram of the Kalman filter.

![Kalman filter information flow graph](image)

**Figure 1.** Kalman filter information flow graph

Normally, the state model is used in the prediction phase to predict the state variables at the next moment.

\[
\begin{align*}
\begin{bmatrix} x_k \\ p_k \\ \end{bmatrix} &= F_k \begin{bmatrix} x_{k-1} \\ p_{k-1} \\ \end{bmatrix} + B_k \begin{bmatrix} u_k \\ \end{bmatrix} \\
K &= \frac{P_k H_k}{P_k H_k + R_k}
\end{align*}
\]

K is the Kalman gain. R is the covariance matrix of the measurement noise. Bk is the observation matrix.

In the optimization algorithm, the prediction function of the Kalman filter is integrated into the traditional Cam Shift tracking algorithm, and the motion of the target object is taken as the research object, and the above prediction and correction process is carried out. Make the tracking process more oriented. In this paper, the motion state of the center point of the tracking area is taken as the object of prediction. In the first frame where the tracked person is determined to appear, its center point is obtained (the target person is framed by a rectangle), and its initial state is set to still. Predict the movement towards the object point before each subsequent frame of picture is running, and combine the position of the point in the current picture to obtain a predicted value of the point position in the next frame of picture. With the predicted point position as the center and the size of the tracking area in the previous frame as a reference, an area of interest that is slightly larger than the tracking area in the previous frame is established in the next frame [8]. Run the Cam Shift tracking algorithm only in the area of interest. After the tracking algorithm accurately determines the location of the target person in the area of interest, it returns the measured tracking area to correct the system.
4. Realization of Human Motion Tracking System Based on Facial Recognition

4.1. Introduction to the experimental system
The system processes real-time video, based on the stored face database, first calls the Seetaface library to do face recognition. Once the camera recognizes the target person, it immediately saves the first frame of the captured target person's image and returns to the facial features of the current image. Based on the facial features of the image, it is calculated according to the standard human body ratio. The target person is circled and tracked down the screen. Once the system switches to the tracking mode, the color histogram of the pure foreground image is first calculated. At the same time, the initial Kalman filter is used to predict the direction of movement away the target. Two coefficients are setting based on the size change of the tracking target, coefficient 1 is used to judge whether there is similar color interference, and coefficient 2 is used to judge whether occlusion occurs. Then use Kalman to set a larger search area based on the predicted motion direction, then use the Kalman Shift algorithm to determine a more accurate tracking area in the search area, and then return the tracking area to update the Kalman system state [9]. Once the coefficient 1 in a frame exceeds the threshold, the inter-frame difference algorithm is called to spend the period; once the coefficient 2 in a frame exceeds the threshold, the current frame only relies on Kalman to track the target.

4.2. Experimental designs.
The tracking algorithm of this work combines color features, Kalman prediction and inter-frame difference algorithm. Combine historical images and real-time images to determine the current situation, and then use the most suitable algorithm. In theory, compared with traditional algorithms, the optimized algorithm can effectively solve the problems of similar color interference and object occlusion, and to a certain extent can retrieve lost objects. This chapter designs comparative experiments for this optimization algorithm to verify its practicability.

4.3. Experimental results and data
Select consecutive frames, respectively read the actual location coordinates of the target person, Kalman predicted value coordinates, and the observed value coordinates using RCA-Dr. Since the area occupied by the target person has a certain range, the actual value coordinates are selected as the center of gravity coordinates of the target person. Select 20 consecutive frames of the video, and require that at least one tracking error has occurred in these 20 frames, and use this as a condition to filter.

5. Conclusion and Outlook
On the one hand, traditional tracking algorithms need to manually delineate the tracking target. After the tracking frame is delineated, the entire content of the frame is directly determined as the target. Manually delineating is inevitably not accurate enough. If the area is too small, the feature information is not clear enough, and the area is too large. It is easy to introduce background factors to cause interference; on the other hand, traditional algorithms for tracking moving objects often only capture a certain feature, usually only under certain conditions, the tracking effect is better and it is easy to be disturbed by noise. For example, the Kalman Shift algorithm is based on color features and is easily interfered by objects of similar colors; Kalman filter can predict the object's past movement, but once the movement is complicated and lost, it cannot be retrieved; the inter-frame difference algorithm is applicable to the condition of unchanged background. Once the background changes, it cannot be tracked accurately [10]. In addition, the short-term loss of the tracked target person is also one of the tracking problems.

For the optimization algorithm proposed to this paper, the following work can be improved and strengthened from the following aspects. First of all, the set of supervision parameters in this article only considers the background interference with similar colors and the short-term occlusion by obstacles in the foreground. Therefore, it is easy to make mistakes when the current scene occlusion is similar colors. The set of supervision parameters can be set up from other angles. Secondly, this article does
not verify the correctness of retrieving the target after the target is lost. It is to find the most probable area from the current picture, and you can judge the retrieved target again.

Acknowledgements
His study was funded by Humanities and Social Science Research Project of Hebei Education Department "Risk Analysis and Legal Countermeasures of Citizens' Privacy Protection under the Application of AI Facial Recognition Technology" (Grant Number: SZ2021078).

References
[1] Research on face recognition algorithm based on LBP [J]. Gu Wei. Office Automation. 2018(02)
[2] Overview of face recognition technology based on deep convolutional neural networks [J]. Jing Chenkai, Song Tao, Zhuang Lei, Liu Gang, Wang Le, Liu Kailun. Computer Applications and Software. 2018(01)
[3] Face recognition based on improved deep convolutional neural network [J]. Wu Yao, Qiu Weigen. Computer Engineering and Design. 2017(08)
[4] Application of adaptive deep convolutional neural network in face recognition[J]. Guo Xiaojie, Chen Liang, Shen Changqing, Liu Chengjian. Automation Technology and Application. 2017(07)
[5] Cross-age face recognition based on deep convolutional neural network [J]. Li Ya, Wang Guangrun, Wang Qing. Journal of Beijing University of Posts and Telecommunications. 2017(01)
[6] Feature weighted fusion face recognition method based on deep neural network [J]. Sun Jinguang, Meng Fanyu. Computer Application. 2016(02)
[7] A review of the application research of deep convolutional neural networks in computer vision [J]. Lu Hongtao, Zhang Qinchuan. Data Acquisition and Processing. 2016(01)
[8] Dynamic face recognition method and application based on deep neural network[J]. Chen Guoping, Du Shanshan. Microcomputer application. 2015(09)
[9] A review of multi-gesture face recognition[J]. Zou Guofeng, Fu Guixia, Li Haitao, Gao Mingliang, Wang Kejun. Pattern Recognition and Artificial Intelligence. 2015(07)
[10] Mobile object location and tracking algorithm based on particle filter[J]. Zhou Fan, Jiang Wei, Li Shuquan, Zhang Yuhong, Zeng Xue, Wu Yue. Journal of Software. 2013(09)