A Pedestrian Tracking Method Based on Adaboost and Feature Fusion

Xueying Huang
School of Automation Wuhan University of Technology Wuhan, China
Corresponding author e-mail: xueying_huang@yeah.net

Abstract. In view of the existing time-consuming and identity exchange problems in the pedestrian tracking process. A multi feature fusion tracking method based on adaboost detection is proposed. Firstly, the multi-scale wavelet transform is used to extract the foreground regions of the moving target, and the Haar-like feature is used in the foreground region to detect human heads. Then the particle filter tracking algorithm is introduced by combining color and gradient direction histogram to track the detected heads. Finally, the pedestrians are tracked through the association matrix. The experimental results show that the proposed method has a good tracking result when there are the cross motion and short occlusion between the pedestrians.

1. Introduction
Pedestrian tracking is a very important research topic in the field of computer vision. It is the basis and premise of the understanding and recognition of the behavior of moving target. It has a broad application prospect in outer space exploration, robot navigation, intelligent video surveillance and video analysis [1] [2]. However, there are many difficulties in the implementation of multi pedestrians tracking, such as the emergence of new pedestrians in video scene, the departure of the old pedestrian, the change background and the color of the target, and the mutual occlusion among the targets [3]. For this reason, many scholars have devoted to the study of multi-target tracking problems, and put forward some tracking algorithms [4-6]. There are Kalman filter tracking algorithm and particle filter algorithm based on color model. The Kalman filter tracking algorithm uses the estimated value of the previous moment and the observed values of the present time to update the estimation of the state variable and can track the moving target in real time, but it is aimed at the Gauss linear process, but has limits the in the effect of the non-Gauss or nonlinear process. The tracking loss will be lost when the pedestrian is blocked [7]. Particle filter algorithm based on color model has poor tracking effect under the condition of little difference between background and target color image [8]. Most of the methods used to deal with multiple tracking problems with feature data fusion, which is complicated and time-consuming, and is easy to have identity exchange in the process of tracking [9, 10].

Through the analysis of different tracking methods, a pedestrian tracking method based on adaboost and feature fusion in proposed in this paper. First, we use multi-scale wavelet to detect the moving target, and then extract the target through Haar-like feature by cascade classifier. On this basis, a particle filter tracking algorithm based on the fusion color and gradient direction histogram is introduced. Finally, the pedestrians are tracked through the association matrix.
2. Target detection based on wavelet transform

2.1. Moving target foreground segmentation

A foreground segmentation algorithm based on multi-scale wavelet transform and frame difference method is proposed in this paper. According to the excellent performance of the multi-scale wavelet transform in both the time domain and the space domain, the adjacent frame images are first performed in the RGB color space, and then the frequency extraction features of the multi-scale wavelet transform are used. Suppose at some points from adjacent frames \(f(t_{n-1}, x, y)\), \(f(t_n, x, y)\) are obtained respectively, and the difference image \(\text{Diff}(x, y)\) is obtained by pixel-by-pixel difference between the two frames.

\[
\text{Diff}_R(x, y) = |f_R(t_n, x, y) - f_R(t_{n-1}, x, y)|
\]

\[
\text{Diff}_G(x, y) = |f_G(t_n, x, y) - f_G(t_{n-1}, x, y)|
\]

\[
\text{Diff}_B(x, y) = |f_B(t_n, x, y) - f_B(t_{n-1}, x, y)|
\]

Where \(\text{Diff}_R\), \(\text{Diff}_G\) and \(\text{Diff}_B\) correspond to the three components of the differential image red, green and blue respectively, and \(|F|\) is the absolute value of \(F\). Based on the above adjacent frame difference, the foreground moving object region is segmented by wavelet transform. According to two dimensional image \(I(x, y)\) in the scale \(2^j\) and \(k\) direction of wavelet transform:

\[
W_{2^j}^k f(x, y) = I * \psi_{2^j}^k (x, y), k = 1, 2
\]

Then the wavelet function in the direction of \(x\) and \(y\) can be expressed as:

\[
\psi^1(x, y) = \frac{\partial \theta(x, y)}{\partial x}
\]

\[
\psi^2(x, y) = \frac{\partial \theta(x, y)}{\partial y}
\]

In the formula, \(\theta(x, y)\) is a smoothing filter function. From this, we can confirm that the wavelet transform of \(I(x, y)\) is filtered at different scales as:

\[
\nabla_{2^j} I(x, y) = (W_{2^j}^1 I(x, y), W_{2^j}^2 I(x, y)) = \frac{1}{2^{j+2}} \nabla I * \theta_{2^j}(x, y)
\]

According to this, the edge points of different scales can be determined. Since noise is sensitive to scale changes, it is impossible to suppress noise effectively by seeking local amplitude maxima. In order to overcome this effect effectively, the edge points of different scales are determined by seeking the method that the gradient amplitude is higher than a certain threshold value instead of seeking the local maximum amplitude.

\[
E = \sqrt{(I \otimes h)^2 + (I \otimes v)^2} \geq T
\]

Where \(h\) and \(v\) are filter operators in horizontal and vertical directions respectively. \(T\) is threshold and \(\otimes\) is convolution operator. According to the above analysis, the formula (5) can be changed to:
\[ E_{k,R,G,B} = \sqrt{\left(\Delta I_k \otimes s\right)^2 + \left(\Delta I_k \otimes t\right)^2} \geq T \]

It represents \( \text{DiffR} \), \( \text{DiffG} \) and \( \text{DiffB} \), that is, the gray difference between the three components of red, green and blue in the adjacent frames.

**2.2. Target human head detection**

On the basis of motion foreground objects based on inter frame multi-scale wavelet transform, first of all, we should use Haar-like features of a large number of picture samples to train the classifier, and get a cascade of Boosted classifier. Here we collect 7100 samples of human head pictures and more than 12000 negative samples. All positive samples are normalized to 12×12 pixels. After the adaboost cascade classifier based on the class Haar-like features is obtained, the classifier is applied to the region of interest in the input image (the same size as the training sample) for the target detection. If the head area is detected, the classifier outputs 1, or output 0. In order to detect the whole image, move the search window in the image to detect each location to determine possible human head targets.

**3. Tracking method based on feature fusion**

The traditional particle filter tracking algorithm based on color histogram can lead to the failure of tracking in complex environment. In this paper, a particle filter tracking algorithm which combines color histogram and gradient direction histogram features is proposed in this paper. The basic steps of the algorithm are as follows:

1. Initialization: in the current frame is considered a new target in the scene, the comprehensive histogram of the target area is calculated and \( N \{X_i^0\}_i=1 \) is generated by adding Gauss noise. Since this article considers the head target, the initial weight of each example is \( w_i^0 = 1/N = 1/30 \);

2. State transfer: when reading the next frame image, according to the two order linear autoregressive equation and the particle set \( \{X_i^k\}_i=1 \) of the last frame, the new particle set \( \{X_i^{k+1}\}_N=1 \). Two order linear autoregressive equation of the current moment is described as follows:

\[ X_i = 2X_{i-1} - X_{i-2} + \xi_i \]

Where \( X_i \) represents the state of particles at the time \( I \), and \( \xi_i \) is the variance of the random noise of Gauss distribution.

3. Weight calculation and updating: after the new particle set \( \{X_i^N\}_i=1 \) is obtained, the comprehensive histogram of each particle is calculated, and then the observation likelihood function \( p(Y_i|X) \) corresponding to each particle state is calculated. The particle weight updating formula is shown as follows:

\[ w_i^k = w_i^{k-1} \rho(Y_i | X_i^k) \]

4. Target state estimation output: Based on the weight of particles, the minimum mean square error estimation of \( K \) time targets can be calculated.

\[ X_k = \sum_{i=1}^{N} w_i^k X_i^k \]

5. Resampling: in order to avoid the fading of particle weight, if the weight value \( W_i^k \) is satisfied (9), a new particle set is resampled from the particle weight in the \( \{X_i^k\}_N=1 \), and the weight value of the particle is reassigned. \( W_i^k = 1/N \).
Finally, complete content and organizational editing before formatting. Please take note of the following items when proofreading spelling and grammar:

4. Experimental results and analysis
In order to verify the effectiveness of the detection based multi person recognition tracking method proposed in this chapter, we use a number of set of test sequences to test and compare with several existing related algorithms. The experimental running environment is Visual Studio 2010 combined with OpenCV 2.4.13, Intel (R) 2.8GHz CPU, 8G RAM.

The experimental results of multi- pedestrian tracking based on feature fusion are as Figure 2.

Through the comparison and analysis of the experiment, the improved adaboost detection algorithm and the particle filter tracking algorithm have good tracking effect in all kinds of scenes. The algorithm is obviously superior to the proposed multi person tracking algorithm, especially in complex situations such as occlusion, cross motion and so on. The algorithm achieves stable and reliable tracking of multiple head targets in complex scenes.
5. Conclusion
The traditional color feature based particle filter tracking algorithm does not consider the similarity between the target color and the background. This paper proposes an improved particle filter tracking algorithm based on the integrated histogram of color histogram and gradient direction histogram, which can track the target accurately in complex scenes.

References
[1] M. Zhao, D. H. Sun, Y. Tang and H.P. He. Head detection based on 21HT and circle existence model. Proceeding of the 10th World Congress on Intelligent Control and Automation, 2012: 4875-4880.
[2] D. Xie, L. Dang and R. Tong. Video based head detection and tracking surveillance system. Proceedings of IEEE International Conference on Fuzzy System and Knowledge Discovery, 2012: 2832-2836.
[3] J. Vermaak, A. Doucet and P. Pérez. Maintaining multimodality through mixture tracking. Proceedings of IEEE International Conference on Computer Vision, 2003: 1110-1116.
[4] I. Ali and M.N. Dailey. Multiple human tracking in high-density crowds. Image and Vision Computing, 2012, 30(12): 966-977.
[5] J. Hu, H. Zhang, H. Huang, J. Feng and G. Wang. An improved Kalman filter algorithm based on quaternion correlation in object tracking. Proceedings of IEEE International Conference on Fuzzy Systems and Knowledge Discovery, 2012: 1725-1729.
[6] C.C.R. Wang, J.J.J. Lien. AdaBoost learning for human detection based on histograms of oriented gradients [C]. Proceeding of Asian Conference on Computer Vision, 2007: 885-895.
[7] N. Dalal, B. Triggs. Histograms of oriented gradients for human detection [C]. Proceeding of IEEE Computer Society Conference on Computer Vision and Pattern Recognition, 2005, 1: 886-893.
[8] P. Viola, M. Jones. Rapid object detection using a boosted cascade of simple features [C]. Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition, 2001, 1: I-511-I-518.
[9] G.H. Park, H.H. Cho, M.R. Choi. A contrast enhancement method using dynamic range separate histogram equalization [J]. IEEE Transactions on Consumer Electronics, 2008, 54(4): 1981-1987.
[10] A.J. Lipton, H. Fujiyoshi, R.S. Patil. Moving target classification and tracking from real-time video [C]. Proceedings of IEEE Workshop on Applications of Computer Vision, 1998: 8-14.