Particle Filtering Tracking Study of Automatic Extraction Tracking Range

Jian Cao¹, a, Da Wu¹, b
¹Air Force Engineering University of PLAXi’an, China
a cao_jian1972@163.com, b wwuda2003@163.com

Abstract. What needs to be solved is the problem of automatic tracking of pedestrians in a complex monitoring environment. In the actual monitoring environment, there are usually chaotic scenes, noise, light changes, and constant changes in human motion, in this context, the post-test probability and observation probability are non-Gaussian, nonlinear, so the framework of particle filtering is chosen to solve the pedestrian tracking problem. In target modeling, human motion is non-rigid body deformation, and color features for the target plane rotation, non-rigid deformation, partial masking and other situations are more robust, so in tracking pedestrians to choose color features. This paper proposes an overlay algorithm that automatically selects the maximum attribute area to determine the trace area. Finally, this paper uses color features to realize the automatic tracking of pedestrians under the theoretical framework of particle filtering.

1. Introduction

Color features are robust for target plane rotation, non-rigid deformation, partial masking, etc., and are suitable for deformation target tracking. Comaneci et al. [1] proposed a tracking algorithm based on Color Histogram, which uses homogeneity offset to track non-rigid targets and achieves good tracking results in some scenarios. However, if the target is obscured or the motion displacement of the adjacent two frames target is large, their algorithm may fail to track it because the homogeneity offset can only search for local state space. To overcome this problem, Perez et al. [2] and Numara et al. [3] embedded color features in particle filtering frames for tracking, resulting in more satisfactory tracking results in complex and even obscured situations. Birchfield et al. [4] further proposed the spatial color histograms tracking algorithm, which not only considers the color value of pixels but also considers the spatial relationship between pixels, which improves the performance of the tracking algorithm.

Particle filtering is a practical algorithm to solve Bayes probability, and particle filtering can be applied to any nonlinear system that can be represented by state space model, as well as nonlinear systems that cannot be represented by traditional Kalman filtering (KF), and the accuracy can force into optimal estimation. [5] Its core idea is to approximate the probability density function of state variables with discrete random sampling points (particles). When the number of sampling points is large enough, these particles can well approximate the post-test probability density function. Compared to extending Kalman filtering (EKF) algorithms and Unscented Kalman filtering (UKF) algorithms, particle filtering algorithms can handle arbitrary nonlinear functions and non-Gaussian distributions. Particle filtering methods are flexible, easy to implement and have parallel structures.

2. Tracking Issues Based on Bayes’s Estimates

The target tracking problem can be converted into a process of optimal estimation within the framework of Bayes’ theory, usually described by a state-space model, and this section uses a discrete moment state
shown in Fig. 1 to represent a dynamic system, with the aim of estimating the probability distribution of the implied target state based on observations up to the current moment. Where the system (implied) state, according to the system equation shown in the equation (1) evolves over time.

\[ x_{t+1} = f_t(x_t) + w_t \]  
\[ y_t = g_t(x_t) + v_t \]

Description of the observation equation shown in the equation (2).

\[ w_t \] and \[ v_t \] are noise vectors with a known probability distribution that is independent of each other, assuming that the probability density function of the initial state is known. It is also known as target state, real state, and so on. It is also known as observational information or observational information. In the state space model, the system state satisfies the first-order Markov process as shown in the equation (3).

\[ p(x_t) = p(x_0) \prod_{t=1}^{N} p(x_t|x_{t-1}) \]  
\[ p(Y_t|X_t) = \prod_{t=1}^{N} p(y_t|x_t) \]

When a given state vector \( X_t \) is given, the observations \( Y_t \) are independent of each other, as shown in equation (4).

Given the dynamic state-space model, the current tracking question is how to estimate the state variables \( x_t \) for each moment. The recursive Bayes estimate method provides a reliable solution to such problems by calculating post-probability density through recursive Bayes on reasoning at a given and recent observation. Recursive Bayes reasoning consists of two steps: bayes prediction steps (time updates) as shown in equation (5) and correction steps (observation updates) as shown in equation (6):

\[ p(x_t|Y_{t-1}) = \int p(x_t|x_{t-1}) p(x_{t-1}|Y_{t-1}) dx_{t-1} \]
\[ p(x_t|Y_t) = \frac{p(y_t|x_t)p(x_t|Y_{t-1})}{p(y_t|Y_{t-1})} \]

In recursive Bayes reasoning, according to different hypothetical conditions, there are also many solutions to \( p(x_t|Y_t) \) the calculation of the pair, when the dynamic system meets the linear Gauss hypothesis, the best solution can be obtained through the Kalman filter, commonly referred to as "best" and "optimal" estimation is based on the minimum mean square error as the criterion. When the dynamic system is nonlinear, the high-dimensional integral operation in recursive Bayes reasoning is usually a difficult problem to solve, and many approximate methods appear in order to relieve the computational pressure of nonlinear dynamic Bayes analysis.
Particle filtering (PF) is a filter algorithm based on Monte Carlo simulation, which gets rid of the constraint that random quantity must meet Gauss's distribution to solve the problem of nonlinear filtering.

The post-test probability distribution can be obtained approximately from a set of discrete sample sets. However, in general, it is multivariable and non-standard, so it is often difficult to sample particles directly from them. The importance sampling method is an effective solution. Importance sampling is a common sampling technique in Monte Carlo methods. In the sampling process, if you can't sample directly from the target probability distribution, you can sample from a probability distribution that is similar to the target probability distribution and easy to sample, which is the basic idea of importance sampling as shown in Fig.2. For example, the estimation function $E(f_t(x_t))$ expects that if sampling cannot be taken directly, it is expected to choose sampling $p(x_t|Y_t)$ from an approximate probability distribution $\pi(x_t|Y_t)$.

During the materiality sampling process, it can be reasoned. Fig.2 is sampling diagram.

Fig.3 is legend description of particle filter algorithm.

3. **Automatically Selected Tracking Area**

At present, a lot of research has been done based on PF tracking, but mainly focused on the a priori distribution of estimation, weight calculation and importance function of the study, and the selection of particles are mostly based on a specific range or manual selection, so it is difficult to automatically track pedestrians in intelligent monitoring system practical. In addition, the selection of human tracking areas affects the accuracy and complexity of the tracking algorithm. When modeling a target based on color features, rectangular or elliptical areas are usually selected as tracking areas, and reasonable color models are selected for the target area to determine the characteristics of moving targets. The contour shape of the human body is irregular, and the image shown in Fig.4 is processed by the previous motion target detection technique.
This section selects colored features in the rectangular area representing pedestrian characteristics as tracking features. How to automatically select the appropriate rectangular area is directly related to the accuracy and computational complexity of the tracking algorithm, such as Fig.5, Fig.6 and Fig.7 represent the three selected cases, Fig.5 indicates the selection of the human body external rectangular area, Fig.7 indicates the selection of the smaller rectangular area inside the human body, Fig.6 indicates that the rectangular area intersects the outer contour of the human body, is a situation between Fig.5 and Fig.7.
If the target area is selected according to the results of Fig.5, the whole human body is in the selected rectangular area, all color information of the human body will not be lost, but also bring a lot of information redundancy, with the advance of pedestrians, background changes, color characteristics in the target area is very unstable, and the complexity of the operation will increase. If the target area is selected according to the results of Fig.7, with the advance of pedestrians, the background changes, the color characteristics in the target area are very stable, the computational complexity is also very low, but the selected area has obvious locality, which can easily lead to pedestrian tracking errors. Fig.6 shows that the rectangular area intersects the outer contours of the human body, a condition between Fig.5 and Fig.7.

Therefore, the reasonable selection result is to take into account the stability characteristics of the region and the maximum regional characteristics, as shown in Fig.8 is the ideal result to meet the two conditions, the rectangular area is the human body's internal rectangle. However, due to the limitations of sports human body detection and extraction technology in practice, it is difficult to obtain complete human body information, as well as the diversity of human body shape, so it is difficult to fully extract to the internal rectangle. In some cases, the extracted ideal inner rectangle may still fall into obvious local characteristics. So, in practice, determining rectangular areas is usually not a problem of finding the optimal solution but of the sub-optimal solution. In general, the result is the torso.

3.1. SVM
Support Vector Machine (SVM), proposed by Vapnik, et al. in 1995, has relatively good performance indicators. This method is a machine learning method based on the theoretical basis of statistical learning. By learning the algorithm, SVM can automatically find out those support vectors that have better differentiation ability to classify, and the classifier constructed can maximize the interval between classes and classes, so it has better adaptability and high accuracy. This method only needs to determine the final classification result by the category of boundary samples for each domain.

3.2. Multiplayer Split
In the monitoring system, when the whole contour of pedestrians enters the monitoring area, after obtaining the human contours by combining background subtraction and differential method, the method of regional fusion is combined with the principle of morphological corrosion and expansion to ensure the integrity of the contours. Each frame of the video as a matrix, each part of the matrix is the pixel value of the image, because for a fixed camera, it obtains a fixed image size, so for each frame of the video, the matrix used to represent its row vector and column vector are also fixed.

For multiplayer segmentation, the coverage ratio of each column of the matrix is calculated first, and then the separation of people is made using SVM to mark the spatial position of each pedestrian.

4. Particle Filter Pedestrian Tracking Algorithm
Incorporating the particle filtering algorithm and automatic selection tracking area algorithm into the real-time tracking application of pedestrians, a particle filter automatic tracking algorithm based on color
characteristics is proposed, and active tracking of pedestrians is realized in the intelligent monitoring system.

In order to verify the effectiveness of the target tracking algorithm, the real-time tracking experiment is carried out by outdoor surveillance camera. The first sequence is the pedestrian movement from right to left, figure a pedestrian has just entered the monitoring area, the human body is not complete, do not choose the target area; The experimental results show that the method of automatically selecting the tracking area proposed in this part has the same effect as the method of manually selecting the area, and the real-time tracking of pedestrians can be realized in intelligent monitoring.

![enter the area](image1)

(b) track the process

Figure 9 Track sequence

5. Conclusion

The coverage ratio algorithm is proposed to automatically select the tracking area, and the real-time active tracking of pedestrians is realized based on color characteristics under the framework of particle filtering, which lays a good foundation for the behavior recognition of moving targets in the later stages.

Some problems remain for further study, such as the selection of particle number, the selection of important functions, and better solutions to particle degradation. Or there are multiple motion objects in the scene at the beginning, and these objects are connected from time to time, and are separated from time to time, or there is a large number of moving objects clustered in an area of the scene for a long time, in these cases, will bring difficulties to tracking, resulting in poor tracking of the system, there is no more effective solution.

References

[1] D. Comaniciu, V. Ramesh and P. Meer. Kernel-based object tracking. IEEE Transaction on Pattern Analysis and Machine Intelligence, 2017, 25(5): 564-577.
[2] P. Perez, C. Hue, J. Vermaak, et al. Color-based probabilistic tracking. In: Proceedings of 7th European Conference on Computer Vision. 2015: 661-675.
[3] K. Nummiaro, E. Koller-Meier and L. Van Gool. An adaptive color-based particle filter. Image and Vision Computing, 2018, 21(1): 99-110.
[4] S. T. Birchfield and S. Rangarajan. Spatiograms versus histograms for region-based tracking. In: Proceedings of IEEE Computer Society Conference on Computer Vision and Pattern Recognition, San Diego, CA, United States. 2019, 2: 1158-1163.

[5] D. Crisan and A. Doucet. A Survey of convergence results on particle filtering methods for practitioners. IEEE Trans. Speech and Audio Proc. 2020, 10(3): 173-185.