Real Time Pedestrian Tracking using Thermal Infrared Imagery

Jianfu Li
1. Key Laboratory for Optoelectronic Technology and System of the Education Ministry of China, Chongqing University, Chongqing 400030, China; 2. Department of Computer and Modern Education Technology, Chongqing Education College, Chongqing 400067, China
Email: lijianfu@yeah.net

Weiguo Gong
Key Laboratory for Optoelectronic Technology and System of the Education Ministry of China, Chongqing University, Chongqing 400030, China
Email: wggong@cqu.edu.cn

***Abstract*—In the study, a real time pedestrian tracking algortithm is presented using thermal infrared imagery. It makes use of the characteristics of pedestrian body regions in infrared images, which is based on a particle filter framework. The method constructs the regions of interest's (ROI) histogram representation in an intensity-distance projection space model, so as to hurdle the disadvantage of insufficient information when only intensity feature is considered. In addition, the tracking algorithm which embeds the above mentioned representation model in the particle filter framework and update the sample's representation model automatically. The experimental results are gained by using different infrared image sequences, which show that the proposed scheme achieves more robust and stable than the classical tracking method.**

***Keywords*—pedestrian tracking; infrared images; particle filter; intensity-distance projection; histogram representation**

I. INTRODUCTION

Pedestrian tracking plays a key role in many research fields, such as video surveillance, human-computer interaction, and automatic driving assistance systems nowadays [1-3]. However, the tracking problem becomes very difficult because of the pedestrian's appearance variability, especially in outdoor environments [4]. Compared with visible light images, infrared images have their unique advantages. The intensity of target object is mainly determined by its temperature and radiated heat and is independent of the current light conditions, so the tracking system can be applied indiscriminately day and night. Also the infrared images almost eliminate the influences of the color, texture, and illumination on the appearance variability of target object. Therefore, there is a latent development potential for pedestrian tracking in infrared images. In the early days, the infrared technology is mainly applied to the military fields for its high cost. Recently, the development of infrared technology and the drop of its cost have both make it possible to enable its use in many industrial and civil fields. Many researchers pay their attentions to the problems of robust pedestrian detection and tracking in infrared imagery [5-8].

In visible light images, pedestrian body have distinct discrimination features compared to non-pedestrian target object, such as different skin colors which can be easily used to construct the pedestrian representation model. In infrared images, the surrounding's non-pedestrian target objects' (such as cars, animals, light poles, and buildings) intensities are often similar to the pedestrian body's intensity, which makes it very fallibility to construct the pedestrian's representation model only depends on its intensity. Although infrared images enable better image segmentation, the performance of current tracking algorithms is poorer [9]. In reference [10], Darrel et al. present an interactive system for detecting and tracking several people. Their system detects and track people using the integration of the information provided by three models: a skin detector, a face detector and the disparity map. However, there is a main drawback for the system as it relies on a predefined color model. When the illumination conditions differ significantly from the training samples [11], there would be a degration on the tracking performance. In [12], a dynamic skin color model is used to solve this problem. In [13], Harville presents an interesting method to locate and tracking multiple people in stereo images. It first uses a sophisticated image analysis method [14] to build the environment model before the detection process. Then, the detection process is performed based on the merging information from two maps including an occupancy map and a height map. Human tracking is implemented using the Kalman filter combined with deformable templates. The main drawback of the system is that the detection heuristics employed can lead to incorrectly taking as people new objects in the scene which is similar to human beings. In reference [5], Yasuno et al propose a system for pedestrian detection and tracking which uses the fixed camera and requires

doi:10.4304/jcp.5.10.1606-1613

the human's head to be located in the center of field view. The system takes the former detected pedestrian's head as the template making use of the brighter property of the head in infrared images. Then it tracks the pedestrian's head by template match algorithm. In [8], F. L. Xu et al. adopt the combination of Kalman filter and mean shift algorithm to tracking human's head using the high brightness of human's head too. Therefore, the current algorithm for the tracking problem is mainly used for tracking parts of the human being, such as head or other parts. Also the tracking algorithms for infrared image sequences are based on the characteristic of intensity only and are applied for specific environments, so the application fields and performance of them are limited.

This paper addresses the problem of robust performance for tracking pedestrian from infrared image sequences. The proposed approach is based on a novel intensity-distance projection space, which hurdle the disadvantage of insufficient information represented by the current pedestrian representation model in infrared images. At first, we propose a method for constructing the robust pedestrian representation model which projects the pedestrian body's regions of interest (ROI) into the intensity-distance projection space. Then, we embed the pedestrian projection histograms in the particle filter framework and propagate the sample distributions over time.

This paper is organized as follows. Section 2 describes our pedestrian histogram representation model based on the intensity-distance projection space. In section 3, we propose our tracking algorithm which embeds the above mentioned representation model in the particle filter framework and update the sample's representation model automatically. In section 4 we successfully apply this algorithm to real infrared image sequences of pedestrians. Section 5 concludes the paper.

II. INTENSITY-DISTANCE PROJECTION SPACE BASED ON PEDESTRIAN REPRESENTATION MODEL

Pedestrian body in infrared images has the characteristic of high intensity, which can be used for tracking purpose. However, when only the intensity is considered, the algorithm's result is uncertainty because of the high intensity interference of other targets. Therefore we must find certain robust pedestrian representation model to correctly differentiate pedestrian from non-pedestrian target. When compared to non-pedestrian, the shape of different pedestrian body has more similarities. In infrared tracking domain, the variability of pedestrian movements in image sequences is small, and the pedestrian's shape can be taken almost as the same. So we can make use of shape feature to enhance the tracking performance. The paper presents an effective pedestrian representation model which constructs the model using the intensity-distance projection space. Furthermore, the method presented can be easily implemented and requires small computational time.

A rectangle window is used to represent the pedestrian target's region, and then the target's distance feature and intensity feature are analyzed respectively. Let m be the rectangle window's width, n be the length and I(x, y) be the pixel's intensity which is located in (x, y) in the current image. The pedestrian histogram representation model based on the intensity-distance projection space is constructed in the following steps.

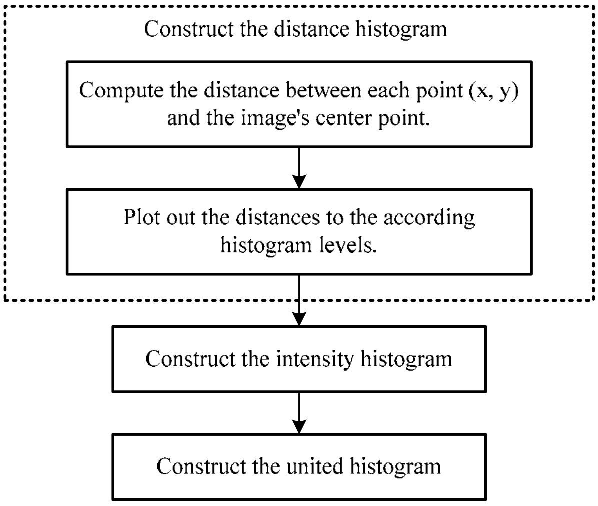

Figure 1. The pedestrian histogram representation model based on the intensity-distance projection space

In the construction of distance histogram, $d(x,y)=\sqrt{(x-x_0)^2+(y-y_0)^2}$ ((x0, y0) is the center point of the image) is used to compute the distance between each point (x, y) and the image's center point.

Plot out the distances according to histogram levels.

$$H_1(k)=\sum_{i=1}^{n}\sum_{j=1}^{m}\delta[d'(x,y)-k] \tag{1}$$

$$d'(x,y)=\begin{cases}1 & d(x,y)=0\\ floor(\frac{2zd(x,y)}{\sqrt{m^2+n^2}}) & others\end{cases} \tag{2}$$

where k = 1, 2, …z is the histogram's level and δ is the Kronecker delta function.

In the construction of the intensity histogram,to acquire the intensity histogram, we need to analyze the target's intensity feature and plot out different intensities to the according histogram.

$$H_2(f)=\sum_{i=1}^{n}\sum_{j=1}^{m}\delta[I'(x,y)-f] \tag{3}$$

$$I'(x,y)=\begin{cases}1 & I(x,y)=0\\ floor(\frac{lI(x,y)}{255}) & others\end{cases} \tag{4}$$

where l=1, 2 … l is the histogram's level and δ is the Kronecker delta function.

In the construction of the united histogram.We can obtain the final united histogram by fusing the upper two histograms. The united histogram is as the following.

$$H(k,f)=\sum_{x=1}^{n}\sum_{y=1}^{m}\delta[d'(x,y)-k]\delta[I'(x,y)-f] \quad (5)$$

After getting the final united histogram, we need to choose the histogram's levels z. The different value of z can affect the tracking result. In the study, we set z to 7 according to the experimental result.

III. Pedestrian Tracking in the Particle Filter Framework

In recent years there has been a flurry of interest in applying Particle Filtering (PF), also known as Condensation, Sequential Monte Carlo or Sequential Importance Sampling (SIS), to optimal estimation problems for non-linear non-Gaussian state-space models, such as computer vision problems [15-17]. Particle filter technique is a popular method in tracking domain for its excellent performance [18].It can solve the non-gaussian and non-linear problems in object tracking. Applications on parameterized or non-parameterized contour tracking [19-21], and human tracking [22, 23] have demonstrated its usefulness. The paper embeds the upper pedestrian representation model in the particle filter framework and implements adaptive tracking.

A. Overview of Particle Filtering

Particle filtering implements the recursive Bayesian filter using the non-parametric Monte Carlo approximation method, and can be applicable for any non-linear system which can be represented by the sate space model. The advantages of particle filtering are that the approximated probability density needs not to be the Gaussian distribution and it uses the non-parametric method to describe the target's state. Each particle represents the hypothetical state of the target in the sample set and the weight of the particle describes how the particle approximates the target's true state. During the evolution of the particles, the large weight particles are conserved, while the small weight particles are discarded. As the recursive state and measurement of the target update, the sample set trends to obey the distribution described by the Bayesian estimation. The new sample set can be regarded as to describe the a posteriori probability distribution. Therefore we can acquire the estimation of the current state of the target and track the target.

Let the vector X_t denotes the state of a tracked target while the vector Z_t describes all the observations $\{z_1,...,z_t\}$ up to time t. The basic idea of particle filtering is to approximate the target's a posteriori probability density by a weighted particle sample set $S=\{s^{(n)},\pi^{(n)} \mid n=1,...,N\}$ .Each particle sample describes one hypothetical state of the target , with a corresponding discrete weight π , where $\sum_{n=1}^{N}\pi^{(n)}=1$.

The weight of the particle is determined by the measurement equation and generally is set as $\pi^{(n)}=p(Z_t \mid X_t=S_t^{(n)})$. After computing all the particles' weights, we can get the expected value of the target's state.

$$E(S)=\sum_{n=1}^{N}\pi^{(n)}S^{(n)} \quad (6)$$

B. Intensity-distance Projection Space Based Particle Filtering

The state vector of the sampled particle can be set as $S=\{x,y,x',y',H_x,H_y\}$, where (x,y) represents the position of the target in the image , (x',y') is the speed of the target and (H_x,H_y) is the width and height. The state transition equation is defined by a second-order autoregressive model as follows.

$$S_t=2S_{t-1}-S_{t-2}+W_t \quad (7)$$

where W_t represents the noise vector corresponding to the state vector.

We need to determine the measurement equation when using the histogram to describe the target. The measurement equation is the distance between the feature histogram of the particle and the reference histogram of the target. The proposed tracker uses the Bhattacharyya distance to represent the similarity measurement function. The reference histogram is set as

$$q=\{q_u\}_{u=1,...,m}, \qquad \sum_{u=1}^{m}q_u=1 \quad (8)$$

and the *i*th particle's feature histogram is set as

$$p(i)=\{p_u(i)\}_{u=1,...,m}, \qquad \sum_{u=1}^{m}p_u(i)=1 \quad (9)$$

The Bhattacharyya coefficient which measures the similarity between the two histograms p and q is

$$\rho(i)=\rho\left[p(i),q\right]=\sum_{u=1}^{m}\sqrt{p_u(i)q_u} \quad (10)$$

The coefficient approaches to 1 when the two histograms are equal. We can define the measurement equation of the particle as

$$d(i)=\sqrt{1-\rho\left[p(i),q\right]} \quad (11)$$

The weight of the particle is set as a Gaussian function according to its measurement with the standard variation σ .

$$\pi(i)=\frac{1}{\sqrt{2\pi\sigma}}e^{-d^2(i)/2\sigma^2} \quad (12)$$

After acquiring the initial position of the target manually or automatically by background extraction, we can use the upper tracking algorithm to track the selected object. The tracking algorithm is shown in Table I.

TABLE I. THE TRACKING ALGORITHM

Initialization	Locate the target in the first image (manually or automatically using a detector); construct the referenced object's intensity-distance projection space based histogram representation using the formula (3).
State prediction	Compute each particle's accumulative weights: $c_t^0 = 0, c_t^n = c_t^{n-1} + \pi_t^n$.
	Generate a random number r with uniform distribution.
	Find the smallest j which satisfied $c_t^j > r$
	Predict the next state of the jth particle using the formula (5)
Measurement updating	Acquire the sample particle's histogram and compute the distance between the histogram of the particle and the reference histogram
	Compute the particles' weights using the formula (10) and normalize all the particles' weights.
State updating	Estimate the state of the current target: $\hat{X}_t = \sum_{n=1}^{N} \pi^{(n)} S^{(n)}$
Let t = t + 1 and go to step (2) until the last frame	

IV. EXPERIMENTAL RESULTS

The theory outlined above has been successfully tested on two different real infrared image sequences came from the OTCBVS Benchmark Database [24]. The two database image sequences are from OSU Color-Thermal Database [25] and taken at busy pathway intersections on the Ohio State University campus. Each sequence is taken by a still camera and the image size is 240x320. The cameras are mounted adjacent to each other on tripod at two locations approximately 3 stories above ground.

The experiments are divided into two groups. Firstly the optimal number of particles is acquired through experiments. We compare the performance of different number (10 to 200) of particles to that of the ideal number (unlimited in theory and 500 in reality) of particles and set the optimal number of particles as 60. Secondly, we separately use both the proposed algorithm and the intensity-based algorithm to track objects and compare the tracking results of the two algorithms. All the tests are run on a normal PC with an Intel Pentium 2.4GHz processor and the software MATLAB6.5. When the number of particles is set as 60, the tracking speed of the proposed algorithm approaches 5-10 fps/s. It would run faster if the system is implemented using the C language. Therefore, our proposed algorithm could tracking pedestrians accurately in real time.

A. The Relation Between Tracking Performance and Number of Particles

The number of particles affects the tracking performance in two aspects: the computation load and the loaction precisions. Like the traditional correlation tracking methods, the computation load of correlation tracking algorithm based on the particle filtering is depended mainly on the computation of correlation values. If there are N particles, then we need to compute the correlation values for N times. Suppose the size of search area is M1 x M2, then the ratio between the computation load (O1) of tracking algorithm based on the particle fitlering and the compuatation load (O2) based on the traditional tracking algorithm is $\frac{O1}{O2} \approx \frac{N}{M1xM2}$.

In general, if we add the number of particles, then we can track the target more accurately. However, when the number of particles is equal or bigger then one certain value (such as M1xM2), then the target state of some particles are possible the same. The computation load of those particles are wasted. So we try to acquire the optimal number of particles through several experiments. To add the step between targets in adjacent frames, we get one target every 5 frames in the original test image sequences. Figure 2-7 show the tracking error curves when the number of particles (N) is set 20, 30, 45, 50, 60 and 150. Table II summarizes the tracking error statistics for different number of particles. We can see that the tracking failed almostly when the number of particles is below 45. The tracking error is decreasing as the number is increasing. When the number of particles reaches 60, the mean tracking error is less than 0.7 pixel. Adding the number continuously can not improve the tracking performance distinctly. Therefore we set the optimal number of particles as 60 in the experiments.

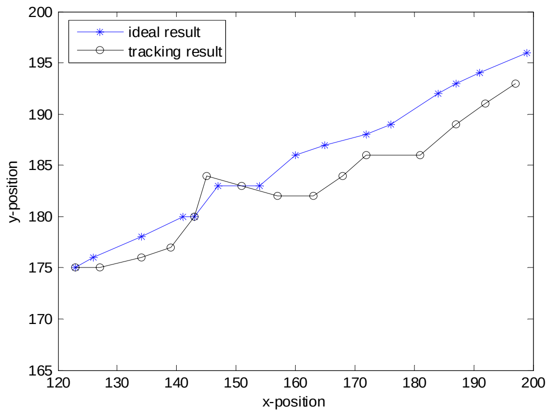

Figure 2. The tracking error cures (N=20)

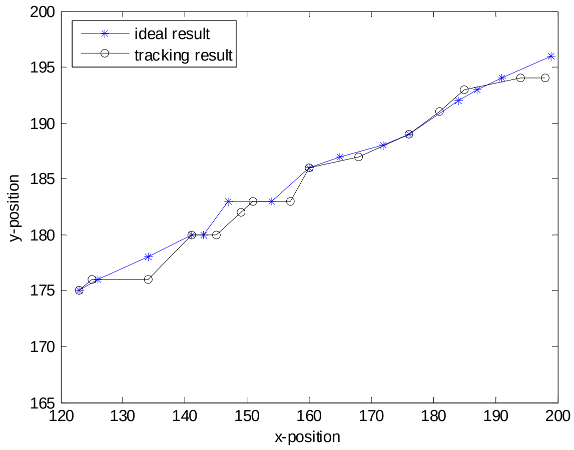

Figure 3. The tracking error cures (N=30).

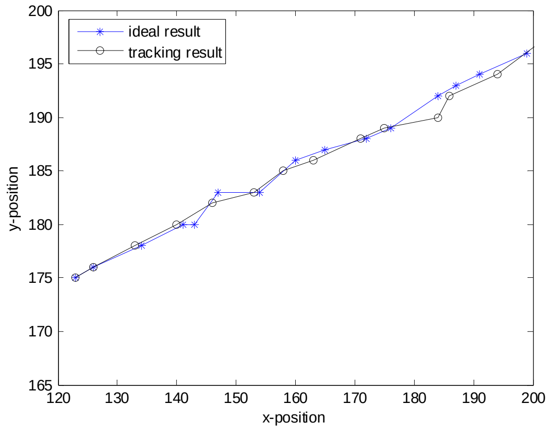

Figure 4. The tracking error cures (N=45).

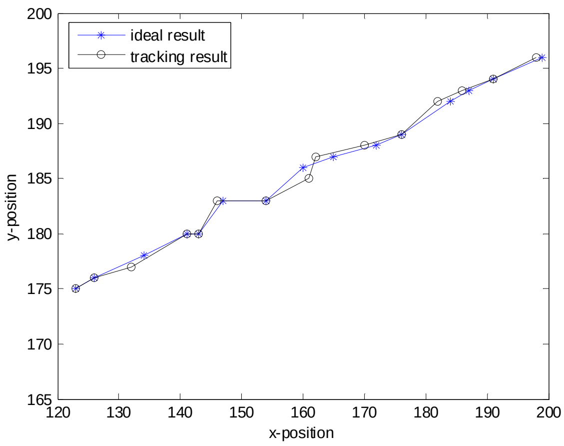

Figure 5. The tracking error cures (N=50).

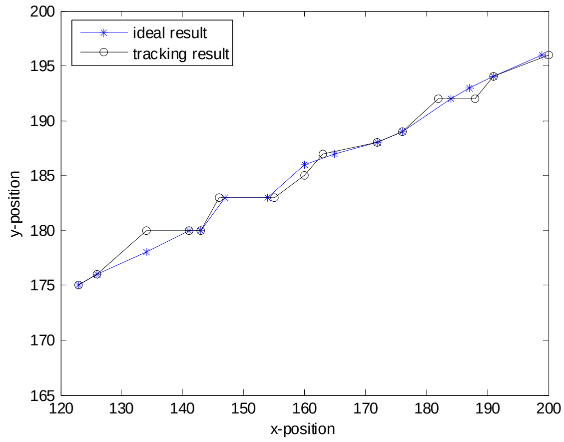

Figure 6. The tracking error cures (N=60).

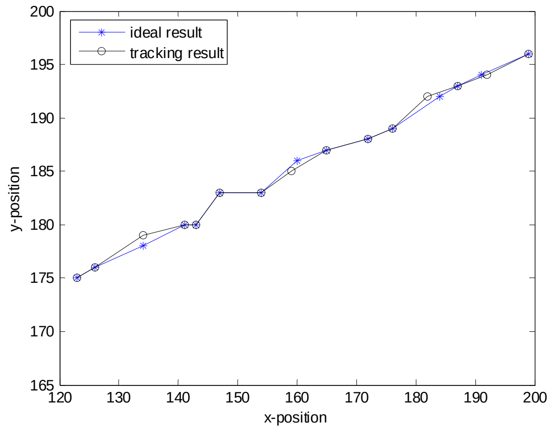

Figure 7. The tracking error cures (N=150).

B. The Comparisons Between Different Algorithms

We compare the tracking performance between our proposed algorithm and the traditional intensity-based algorithm. Figure 8 shows the tracking results from Frames No. 1, 10, 50, 250, 300, 400 in sequence #1 using our proposed algorithm (algorithm 1). Figure 9 shows the tracking result from Frames No. 1, 5, 10 in sequence #1 using the intensity-based algorithm (algorithm 2). From the two figures we can draw that our proposed algorithm keeps more accurate tracking result than the intensity-based algorithm does. The tracking rectangular window in algorithm 1 represents the target accurately both in the scale and the position, while the algorithm 2 can hardly track the target after frame 5. Figure 10 and 11 show the similar results using our proposed algorithm and the intensity-based algorithm.

The experimental tracking results show that our proposed intensity-distance projection space based pedestrian representation model considers both the appearance and shape features of the pedestrian and is superior in robust and feasibility to the intensity-based pedestrian representation model.

TABLE II. STATISTICS OF TRACKING ERROR FOR DIFFERENT NUMBER OF PARTICLES

Num / Err (pixel)	20	30	45	50	60	150
Biggest Distance	6.7	5.10	3.0	3.0	2.0	2.0
Mean Distance	3.4	2.28	1.4	0.9	0.7	0.4
Biggest horizontal location	4.0	5.0	3.0	3.0	2.0	2.0
Mean Horizontal location	1.8	2.0	1.2	0.9	0.5	0.3
Biggest Vertical location	6.0	3.0	2.0	1.0	2.0	1.0
Mean vertical location	2.6	0.6	0.4	0.1	0.3	0.1

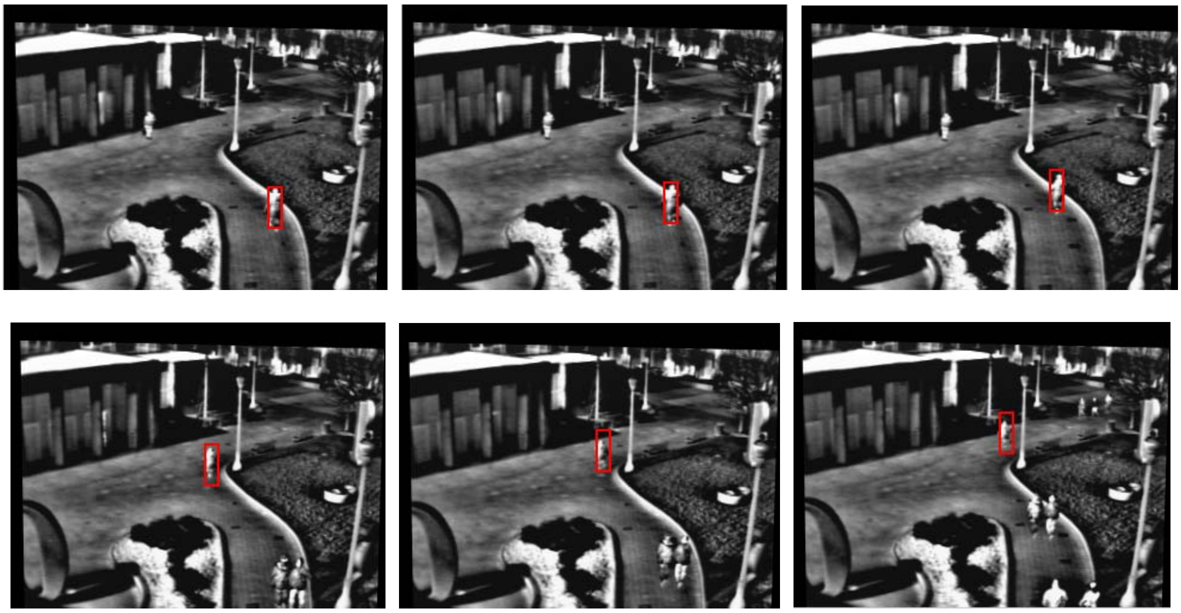

Figure 8. Tracking results for frames No. 1, 10, 50, 250, 300, 400 in sequence #1 using algorithm 1.

Figure 9. Tracking results for frames No. 1, 5, 10 in sequence #1 using algorithm 2.

Figure 10. Tracking results for frames No. 1, 10, 50, 150, 200, 250 in sequence #2 using algorithm 1.

Figure 11. Tracking results for frames No. 1, 10, 20 in sequence #2 using algorithm 2.

V. CONCLUSIONS

In this paper a novel and effective algorithm aimed at tracking of pedestrians in infrared image sequences has been discussed. The proposed algorithm projects the pedestrian body's regions of interest (ROI) into the intensity-distance projection space to construct the robust pedestrian histogram representation model firstly. Then the pedestrian tracking algorithm is designed under the particle filter framework and the sample distribution is propagated over time automatically. The proposed algorithm has been tested in the infrared image benchmark databases. Theoretical analysis and experimental results demonstrate that this approach is promising.

ACKNOWLEDGEMENT

This research was supported by the National High-Tech Research and Development Plan of China (863 Program) under Grant No. 2007AA01Z423, under Grant No. C10020060355 and the Natural Science Foundation Project of CQ CSTC under Grant No. 2008BB2199.

REFERENCES

[1] J.Zhou and J.Hoang, "Real time robust human detection and tracking system", in Proceedings of the 2nd Joint IEEE International Workshop on Object Tracking and Classification in and Beyond the Visible Spectrum,San Diego,Calif,USA,June 2005.

[2] D.C. McFarlane, and K.A.Latorella, "The scope and importance of human interruption in human-computer interaction design", Human-Computer Interaction, 17 (1), 2002, pp. 1-62.

[3] O. Tsimhoni and P. Green, "Time-sharing of a visual in-vehicle task while driving: effects of four key constructs", Proceedings of the International Symposium on Automobile Lighting, Darmstadt, Germany.

[4] D. Ross, J. Lim, and M.-H. Yang, "Adaptive probabilistic visual tracking with incremental subspace update", In Proceedings of the Eighth European Conference on Computer Vision, volume 2, pp. 470–482, 2004.

[5] M. Yasuno, N. Yasuda, M.Aoki, "Pedestrian detection and tracking in far infrared images", in Conference on Computer Vision and Pattern Recognition Workshop, 27-02 June 2004,pp. 125-131.

[6] C. Dai, Y. Zheng, X. Li, "Layered representation for pedestrian detection and tracking in infrared imagery", in 2nd Joint IEEE International Workshop on Object Tracking

and Classification in and Beyond the Visible Spectrum, San Diego, CA,USA,June 20,2005.
[7] M.G. Di Bono, G. Pieri and O. Salvetti, "Multimedia target tracking through feature detection and database retrieval", in Proceeding of the 22nd International Conference on Machine Learning-Workshop on Machine Learning Techniques for Processing Multimedia Content(ICML 2005), Bonn, Germany, 11th August 2005, pp. 19-22.
[8] F.L. Xu, X. Liu and K. Fujimura, "Pedestrian detection and tracking with night vision", IEEE Transaction on Intelligent Transportation Systems, 2005, 6(1):63-71.
[9] E. Goubet, J. Katz and F. Porikli, "Pedestrian tracking using thermal infrared imaging", Mitsubishi Electric Research Laboratories, Technical Report, TR2005-126.
[10] T. Darrell,G.Gordon, and J. Woodfill, "Integrated person tracking using stereo, color, and pattern detection", International Journal of Computer Vision, 2000, 37:175-185.
[11] B. Martinkauppi, M. Soriano and M. Pietikainen, "Detection of skin color under changing illumination: a comparative study", in 12th International Conference on Image Analysis and Processing, 2003, pp. 652-657.
[12] L. Sigal, S.Sclaroff and V. Athitsos, "Skin color-based video segmentation under time-varing illumination", IEEE Transactions on Pattern Analysis and Machine Intelligence, 2004, 26:862-877.
[13] M. Harvile, "Stereo person tracking with adaptive plan-view templates of height and occupancy statistics", Image and Vision Computing, 2004, 2:127-142.
[14] M. Harville, G. Gordon and J.Woodfill, "Foreground segmentation using adaptive mixture models in color and depth", in IEEE Workshop on Detection and Recognition of Events in Video, 2001,pp. 3-11.
[15] Y. Bresler, "Two-filter formula for discrete-time non-linear Bayesian smoothing", International Journal of Control, 1986, 43(2):629-641.
[16] S. Arulampalam, S. Maskell, N. Gordon and T. Clapp, "A tutorial on particle filters for on-line nonlinear/non-Gaussian Bayesian tracking", IEEE Transactions on Signal Processing, 2002, 50(2):174-188.
[17] A.M. Johansen and A. Doucet, "A note on auxiliary particle filters", Statistics and Probability Letters, 2008, 78(12):1498-1504
[18] D. Crisan and A. Doucet, "A survey of convergence results on particle _ltering for practitioners", IEEE Transactions on Signal Processing, 2002, 50(3): 736-746.
[19] P. Del Moral, "Feynman-Kac Formulae: Genealogical and Interacting Particle Systems with Applications", Series: Probability and Applications, Springer-Verlag, New York, 2004.
[20] P. Del Moral, A. Doucet and A. Jasra, "On adaptive resampling procedures for sequential Monte Carlo methods", Technical report. INRIA, 2008.
[21] W.R. Gilks, and C. Berzuini, "Following a moving target - Monte Carlo inference for dynamic Bayesian models", Journal of the Royal Statistical Society B, 2001, 63:127-146.
[22] C.Andrieu and A. Doucet, "Particle _ltering for partially observed Gaussian state space models", Journal of the Royal Statistical Society B, 2002, 64(4):827-836.
[23] A. Doucet, S. J. Godsill and C. Andrieu, "On sequential Monte Carlo sampling methods for Bayesian Filtering", Statistics and Computing, 2000, 10:197-208.
[24] OTCBVS Benchmark Dataset. http://www.cse.ohio-state.edu/otcbvs-bench>, accessed 08.09.
[25] Davis J, Sharma V. Background-Subtraction using Contour-based Fusion of Thermal and Visible Imagery. Computer Vision and Image Understanding 2007;106:162–182.