TELKOMNIKA Indonesian Journal of Electrical Engineering
Vol.12, No.7, July 2014, pp. 5621 ~ 5628
DOI: 10.11591/telkomnika.v12i7.4152

Object Tracking Based on Multiple Features Adaptive Fusion

Jie Cao[1], Leilei Guo*[1,2], Jinhua Wang[3], Di Wu[3]
[1]College of Computer and Communication, Lanzhou University of Technology, Lanzhou 730050, China
[2]Technology & Research Center of Gansu Manufacturing Information Engineering, Lanzhou 730050, China
[3]College of Electrical and Information Engineering, Lanzhou Univ. of Tech, Lanzhou 730050, China
*Corresponding author, e-mail: 745541228@qq.com

Abstract

Multiple features fusion based tracking is one of the most active research in tracking literature, In this paper, a novel adaptive fusion strategy is proposed for multiple features fusion, based on two common used fusion rules: product rule and weighted sum rule. This strategy employs particle filtering technique, product rule and weighted sum rule are unified into an adaptive framework through defined features distance. In practice, the new fusion strategy shows more robustness than product fusion and weighted sun rule.

Keywords: o*bject tracking, particle filtering, features distance, multiple features fusion*

1. Introduction

Target tracking is one of the core technology in computer vision, it has widespread applications in human-computer interaction, surveillance, visual servoing and biomedical image analysis [1]. Tracking based on multi-feature fusion has therefore been an active research topic for over a decade, in order to make the tracking more robust and more stable, it fuses multiple features included color, edge and motion feature. The main strategies of feature fusion have product rule and weighted sum rule. The algorithms proposed by [2, 3] is the typical example of using product rule to multi-feature tracking. Further, the probability framework of a combination tracking algorithm proposed by [4], the algorithms based on product rule get reasonable analysis under this probability framework. On the other hand, weighted sum rule plays a very important role in multi-feature fusion tracking. The algorithm proposed by [5] which used weighted sum rule to approximate joint likelihood. Moreover, there are some fusion methods of multi-feature, such as the algorithm of online switching feature propsed by [6], the algorithm of online switching tracker proposed by [7]. Likewise, min max fusion rule proposed by [8] and democratic electoral fusion proposed by [9].

In order to combine product rule's advantage and weighted sum rule's advantage,there are two achieved ways at present. The first achieved way is that timely switching fusion strategy according to different scenarios. The algorithm proposed by [10] is the typical example of the first achieved way, this algorithm estimats second order moment of the weighted sample set and computing its Frobenius norm to denote how features are reliable, and then switch the two fusion rules in time. Obviously, if there are more samples, this algorithm's real-time need to test and verify. The second achieved way is that making product rule and weighted sum rule unified into an adaptive framework, which adjust the weights of product rule and the weights of weighted sum rule in the tracking results according to adaptive factor. The algorithm proposed by [11] is the typical example of the second achieved way,this algorithm defines a new feature uncertainty measurement method to adjust the relative contributions of different features.

Currently information entropy theory has been successfully applied to information fusion theory [12-14].This paper proposes an adaptive fusion strategy based on information entropy theory. our algorithm make product rule and weighted sum rule unified into an adaptive framework according to defined features distance. An extensive number of comparative

Received August 15, 2013; Revised March 20, 2014; Accepted April 9, 2014

experiments show that the proposed algorithm is more stable and robust than product rule and weighted sum rule in the object tracking.

In the rest of this paper, we explain the shortcomings and our algorithm in Section 2. Experimental results and analysis are reported in Section 3. We conclude this paper in Section 4.

2. Adaptive Fusion Strategy Based on Particle Filter

2.1. Adaptive Fusion Strategy

The model of product rule assumes that the feature is independent of each other. Namely, each feature generates independent observation, then the joint likelihood of n features can be expressed as:

$$p(z^1 \cdots z^n \mid x) = \prod_{i=1}^{n} p(z^i \mid x) \tag{1}$$

Among the above formulas: z^i is the observation of i-th feature which is independent of each other, x is the state of estimated target. Eq.(1) is simple but including rich information.

Another common fusion rule is weighted sum rule, it is a very effective tool for the complex of density estimation problem. its specific form as follows:

$$p(z^1 \cdots z^2 \mid x) = \sum_{i=1}^{n} \alpha^i p(z^i \mid x) \tag{2}$$

Among the above formulas: α^i shows the i-th feature corresponding to the weight of observation probability. the weights are normalized to ensure $\sum_{i=1}^{n} \alpha^i = 1$

This paper proposes an adaptive fusion strategy which combines the advantages of product rule and weighted sum rule, which can effectively solve the problem of difficulty to distinguish when the similar target close to the target. This algorithm is based on the fact: on the one hand, when the features support to each other, this illustrates that the features are influenced by a small degree of contamination, then using product rule can improve tracking accuracy. On the other hand, when the features don't support to each other, this illustrates that the features are influenced by a great degree of contamination, then using weighted sum rule can maintain the multi-modal of distribution, and suppress noise.

This paper will use the particle filter tracker. Therefore, we can obtain the sample's probability assignment. In order to convenience of illustration, two features are denoted C_1 and C_2, the particles set based on two features are $\left\{x_i, p(z^1 \mid x_i), p(z^2 \mid x_i)\right\}_{i=1}^{M}$, M is the number of particles, $p(z^1 \mid x_i)$ and $p(z^2 \mid x_i)$ is the weights of particles which obtained through defined likelihood model based two features.

Define 1. if $p(z^1 \mid x_i)$ and $p(z^2 \mid x_i)$ is the particle's weights of two features, then we build the model of based entropy as:

$$p(z^1 \mid x) \to \begin{Bmatrix} x_1 \cdots\cdots\cdots\cdots\cdots\cdots x_M \\ p(z^1 \mid x_1) \cdots p(z^1 \mid x_M) \end{Bmatrix} \qquad p(z^2 \mid x) \to \begin{Bmatrix} x_1 \cdots\cdots\cdots\cdots\cdots\cdots x_M \\ p(z^2 \mid x_1) \cdots p(z^2 \mid x_M) \end{Bmatrix} \tag{3}$$

The information distance of feature C_1 with feature C_2 can be defined as:

$$d(C_1, C_2) = \left| \sum_{j=1}^{M} [p(z^1 \mid x_i) + \alpha] \log_2^{\frac{[p(z^1 \mid x_i)+\alpha]}{[p(z^2 \mid x_i)+\alpha]}} \right| \tag{4}$$

Then the partial credible coefficient for C_1 can be defined as:

$$Sup(C_1)=\begin{cases}1\text{-}d(C_1,C_2) & if\ d(C_1,C_2)\le 1\\ 1/d(C_1,C_2) & if\ d(C_1,C_2)>1\end{cases} \tag{5}$$

The information distance of feature C_2 with feature C_1 can be defined as:

$$d(C_2,C_1)=\left|\sum_{j=1}^{M}[p(z^2\mid x_i)+\alpha]\log_2^{\frac{[p(z^2|x_i)+\alpha]}{[p(z^1|x_i)+\alpha]}}\right| \tag{6}$$

Then the partial credible coefficient for C_2 can be defined as:

$$Sup(C_1)=\begin{cases}1\text{-}d(C_2,C_1) & if\ d(C_2,C_1)\le 1\\ 1/d(C_2,C_1) & if\ d(C_2,C_1)>1\end{cases} \tag{7}$$

Define 2. The information distance between feature C_1 and feature C_2 can be defined as:

$$D(C_1,C_2)=\left|\sum_{j=1}^{M}[p(z^1\mid x_i)+\alpha]\log_2^{\frac{[p(z^1|x_i)+\alpha]}{[p(z^2|x_i)+\alpha]}}+\sum_{j=1}^{M}[p(z^2\mid x_i)+\alpha]\log_2^{\frac{[p(z^2|x_i)+\alpha]}{[p(z^1|x_i)+\alpha]}}\right| \tag{8}$$

Define 3. The credible coefficient between feature C_1 and feature C_2 can be defined as:

$$Sup(C_1,C_2)=\begin{cases}1\text{-}D(C_1,C_2) & if\ D(C_1,C_2)\le 1\\ 1/D(C_1,C_2) & if\ D(C_1,C_2)>1\end{cases} \tag{9}$$

Among the above formulas: in order to prevent the denominator is zero, α equals 0.0001.

$Sup(C_1,C_2)$ reflects the degree support of features. Namely, when $Sup(C_1,C_2)$ is relatively greater, which illustrates that features support each other, then the fusion result of product rule occupies an important position compared with weighted sum rule. On the other hand , when $Sup(C_1,C_2)$ is relatively smaller, which illustrates that features don't support each other, then the fusion result of weighted sum rule occupies an important position compared with product rule.

According to the simulation results of the experiments, thus we think product rule and weighted sum rule can be unified into an adaptive framework through the credible coefficient.

Define 4. Framework for adaptive multi-feature fusion can be defined as:

$$p(z^1\cdots z^n\mid x)=Sup(C_1\cdots C_n)\prod_{i=1}^{n}[p(z^i\mid x)+\alpha_i\cdot U(x)]+[1-Sup(C_1\cdots C_n)]\sum_{i=1}^{n}\alpha_i\cdot p(z^i\mid x) \tag{10}$$

In Equation (10), $\alpha_i=Sup(C_i)\Big/\sum_{i=1}^{n}Sup(C_i)$, considering that when the weights of particles (the value of feature's likelihood function)close to 0,because product rule make the another feature's contribution become small for these particles, we affiliate the uniform distribution which is directly proportional to this feature's support to every feature.

3. Tracking Algorithm Based on Proposed Fusion Strategy
3.1. Particle Filter
Particle filter is a filtering method based on Monte Carlo and recursive Bayesian estimation. In recent years, it has become an effective tool for target tracking under non-linear or non-Gaussian conditions [15, 16]. The detailed description refer paper [17], the particle filter's principle will not introduce in this section .

In the algorithm achieving process, we choice the ellipse to describe the target's state, namely $x=\{c_x,c_y,l_x,l_y,\theta\}$, c_x, c_y, l_x, l_y and θ are the center coordinates of ellipse, the long axis, the short axis and the deflection angle. Besides, in the particle filtering technique, we use the simplest and most commonly first-order linear system as the state transition model of the particle filter.

3.2. Extracted Features
Color is one of the main features for describing the target, researcher have made a lot of study for target's color feature. A color histogram proposed by [1] is used to describe the target's color feature, Its expression is:

$$h^c(u)=\sum_{u=1}^{B_c}\delta(I(x,y)-u) \tag{11}$$

$I(x,y)$ is the pixels of candidate district, B_c represents the long of color histogram.

Edge [18], as another efficient feature descriptor, can be used here to enhance the power of color feature.

Pixels $I(x,y)$ are evenly extracted within the ellipse, edge's strength G and direction angle α are defined as:

$$G(x,y)=\sqrt{I_x^2+I_y^2} \qquad \alpha(x,y)=\tan^{-1}(\frac{I_y}{I_x}) \tag{12}$$

The ellipse is divided into four parts according to two axes, for each part, direction angle is quantified B_e grade histograms, and then fuse the edge's strength information into each point, we can get the weighted gradient orientation histogram of each part. finally, the histograms of four parts are combined and normalized.

For the description of histogram, the Bhattacharyya coefficient is a popular similarity measure [19]. Considering discrete densities such as two histograms h_{mod} and h_{tar}, the coefficient is defined as:

$$\rho(h_{tar},h_{\mathrm{mod}})=\sqrt{1-\sum_{u=1}^{B}\sqrt{h_{tar}(u)h_{\mathrm{mod}}(u)}} \tag{13}$$

On this basis ,the observed likelihoods of two features can be defined as:

$$p(z^i\mid x)\propto\exp(-\lambda_i\rho_i^2(h_{tar},h_{\mathrm{mod}})) \tag{14}$$

$i\in\{1,2\}$, color is the first feature,edge is the second feature.

3.3. Proposed Particle Filter Tracking Algorithm
In section 3.2, we introduce how to extract features. Now, on the basis of them, the detailed processing of improved algorithm is given.

(1) Initialization: k=1, initialized particle sets $\{x_k^i, 1/N, i=1,\cdots,N\}$

(2) $k = 2, \cdots, N_f$

(a) Prediction: $\tilde{x}_k^i \sim p(x_k \mid x_{k-1}^{(i)}), i = 1, \cdots, N$

(b) Two features are empowered and normalized: Color feature: $\omega_1^{(i)} \propto p(z_k^1 \mid \tilde{x}_k^{(i)})$. Edge feature: $\omega_2^{(i)} \propto p(z_k^2 \mid \tilde{x}_k^{(i)})$

(c) According to Equation (10) fusing two features.

(d) Output: the target's state of k moment can be calculated by the weighted sum of the particles: $\tilde{x}_k = E[x_k \mid z_{1:k}] = \sum_{j=1}^{M} \tilde{\omega}_k^{(j)} \tilde{x}_k^{(j)}$

(e) According to the distribution of particles' weights deciding whether resample. If $(1 / \sum_{t=1}^{N} \omega_{t,j}^2) < N/2$, then $\{x_{k-1}^{(i)}, 1/M, i = 1, \cdots M\} \sim \{\tilde{x}_{k-1}^{(j)}, \tilde{\omega}_{k-1}^{(j)}, j = 1, \cdots, M\}$.Otherwise, don't deal.

4. Experimental Results and Analysis

First, we set experimental parameters as follows: the initial position of target is given manually, the number of particles is set to 200, the uniform distribution $U(x)$ equals $1/N$, direction histogram B equals 18,color histogram B equals 216. The value of λ are shown in Table 1.

Table 1. The Coefficients of Two Features

corresponding	color (λ_1)	edge (λ_2)
Video sequence of experiment 1	90	30
Video sequence of experiment 2	90	40

Simultaneously, in order to measure the tracking error, we define two measure mode.

$$AE = \left|\hat{x}_t - x_t\right|^2 \qquad RMSE = \frac{1}{T}\sum_{t=1}^{T}\left|\hat{x}_t - x_t\right|^2 \tag{15}$$

AE measures each frame error, *RMSE* measures all frames error.

4.1. Aircraft Video

Experiment 1 use the video of model aircraft, which length is 770 frames, The adaptive coefficients between product rule and weighted sun rule are shown in Figure 1, the credible coefficients of features are shown in Figure 2.

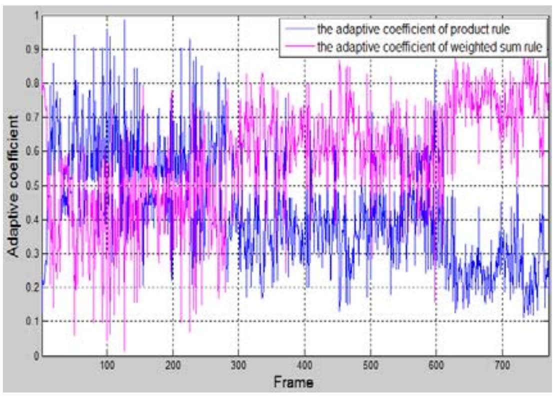

Figure 1. The Adaptive Coefficients of Two Fusion Rules

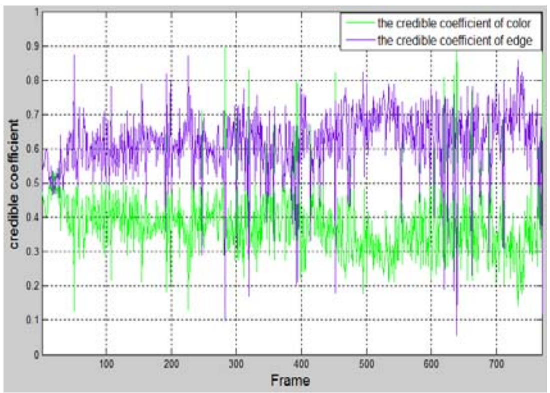

Figure 2. The Credible Coefficients of Features

From Figure 1, we can conclude that our algorithm can adaptively adjust the weights of product rule and weighted sum rule according to the change of environment. Before 300th frame, because the environment is relatively simple, product rule occupies an important position in the tracking. However, after 300th frame, because the environment is relatively complex, weighted sum rule occupies an important position in the tracking.

From Figure 2, we can clear observe that our algorithm can adjust the weights of two features according to the change of environment, Namely, our algorithm achieve the adaptive change of features' weights within the framework of a unified fusion. Figure 3 shows part of tracking results by the proposed algorithm, product rule and weighted sum rule.

Figure 3. Some Results on Experiment 1 by using :product rule (the first row), sum rule (the second row), and proposed method (the third row) (Frames:281,283,500,518,649,770)

From Figure 3. Because the change of illumination is small and the background is relatively simple from first frame to 500th frame, proposed algorithm, product rule and weighted sum rule can successfully track the target. However, as the target moves on, it is occluded by the tree, as shown in frame 500 and frame 518.These make it challenging for the tracking algorithms to follow the target. It can be seen that our algorithm overcomes these difficulties throughout the whole tracking process, whereas the product rule and weighted sum rule drifts away from the target during the tracking process and finally loses the target. This show the stability and robustness of proposed algorithm.

4.2. Speaker Tracking

Experiment 2 is the speaker tracking in smart meeting room, video length is 500 frames. Tracking difficulties are the rotation of head and the occlusion of others speaker.

First, based on equals (15), we give the error analysis. Besides, we only give the frames' tracking error between 382-th frame to 402-th frame when the target is occluded.

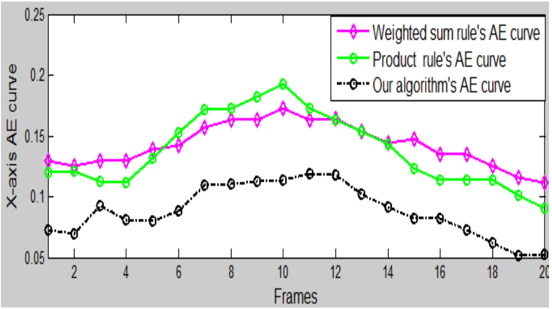

Figure 4. The X-axis Absolute Error

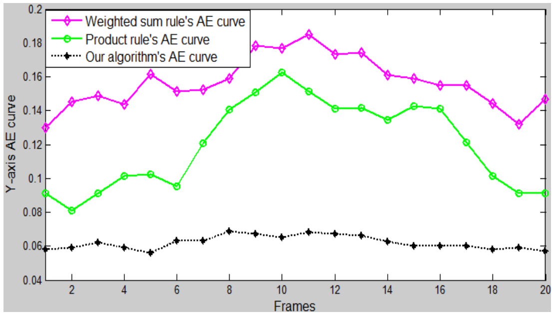

Figure 5. The Y-axis Absolute Error

From Figure 4 and Figure 5, we can conclude that product rule is sensitive to noise, so its AE is biggest in 392-th frame, moreover, weighted sum rule isn't sensitive to noise, the change of AE curve isn't obvious. The proposed algorithm combines the product rule's advantage and weighted sum rule's advantage, so its AE is smallest.

Table 2. Comparison of RMSE Value

Fusion rule	RMSE	
	X	Y
Weighted sum rule	0.2982	0.3129
Product rule	0.2936	0.2735
Our algorithm	0.2347	0.1970

From Table 2. Compared weighted sum rule, in the X-axis, our algorithm's error reduced 0.0635. in the Y-axis, our algorithm's error reduced 0.1159. Compared product rule, in the X-axis, our algorithm's error reduced 0.0589, in the Y-axis, our algorithm's error reduced 0.0765. So our algorithm has better accuracy. Figure 6 shows part of tracking results by the proposed algorithm, product rule and weighted sum rule.

Figure 6. Some Results on Experiment 1 by using: product rule (the first row), sum rule (the second row), and proposed method (the third row) (Frames:53,250,392,500)

From Figure 6, we can observe that when the target's head occludes at frame 392, the fusion results of product rule deviate the center of target. Besides, the fusion results of weighted sum rule exist that some particles begin to transpire. Our algorithm can correctly track the target. This show the effectiveness of proposed algorithm.

5. Conclusion

After analyzing the traditional fusion method including product rule and weighted sum rule,a novel fusion algorithm proposed by this paper which unify the traditional method into a adaptive fusion framework according to adaptive coefficient. In experiments, the new fusion strategy shows more robustness and stability than product fusion and weighted sun rule.

On the other hand, in this paper, we only consider the features of video, without considering the features of audio. Therefore, we will research the fusion strategy for fusing audio and video multi-feature in the next research.

Acknowledgements

This work was supported by the National Science and Technology Support Program (No. 2012BAF12B19) ,the National Natural Science Foundation of China (61263031). The Finance Department Foundation of Gansu Province, China(0914ZTB148).

References

[1] Zhiqinag Hou, Chongzhao Han.A survey of visual tracking. *Acta Automatica Sinica.* 2006; 32(4): 603-617.
[2] Shen CH, Hengel A, Dick A. *Probabilistic multiple cue integration for particle based tracking.* Proc. VIIth Digitl Image Computing: Techniques and Applications, Melbourne, Australia. 2003; I: 399-408.
[3] LI PH, Chaumentte F. *Image cues fus- ion for object tracking based on particle filter*. Proceedings of the 3rd International Workshop on Articulated Motion and Deformable Objects. Palma de Mallorca, Spain: Springer. 2004; 99-107.
[4] Leichter I, Lindenbaum M, Rivlin E. *A Probabilistic framework for combining tracking algorithms.* Proceeding of IEEE Conference on Computer Vision and Pattern Recognition, Washington, USA. 2004; (2): 445-451.
[5] Spengler M, Schiele B. Towards robust multi-cue intergration for visual tracking. *International Journal of Machine Vision and Applicat- ions.* 2003; (14): 50-58.
[6] Collins R, LIU Y. *On-line selection of discr- iminative tracking features.* Proceeding of Internation Conference on Computer Vision, Nice, France. 2003: 346-352.
[7] Toyama K, Hager G. Incremental focus of attention for robust vision-based tracking. International Journal Of Computer Vision. 1999; 35(1): 45-63.
[8] Kittler J, Hatef M, et al. On combining classifers. *IEEE Trans.on PAMI*. 1998; 20: 226-238.
[9] Triesh J, Malsburg C. Democratic integ- ration: selforganized integration of adaptive cue*. Neural Computation*. 2001; 13: 2049-2074.
[10] Xiaopin Zhong, Jianru Xue, Nanning Zheng, et al. An Adaptive Fusion Strategy Based Multiple Cue Tracking. *Journal of Electronics & Information Techonlogy*. 2007; 29(5): 1017- 1022.
[11] Xin Gu, Haitao Wang, Lingfeng Wang, et al. Fusing Multiple Features for Object Tracking Based on Uncertainty Measurement. *Acta Automatica Sinica.* 2011; 37(5): 550-559.
[12] Yong Deng, Dong Wang, Qi Li, et al. A new method to analyze evidence conflict. *Control Theory & Applications*. 2011; 28(16): 0839-0844.
[13] Ji-xiang Sun, Hui-min Shi, Hong-qiang Wang. The Theory Relative to Entropy in Information Fusion. *Chinese Journal of Computers.* 2003; 26(7): 796-801.
[14] Kullback Leibler divergence [EB/OL]. [2010- 11-09]. http://en.wikipedia.org/wiki/KULLBAC K_leibler_ divergence.
[15] Gordon NJ, Salmond DJ, Smith AFM. *Novel approach to nonlinear/non-gaussian bayesian state estimation.* IEE Proceedings on Radar and Signal Processing. 1993; 140(2): 107-113.
[16] Gao Y, Yuan GW, Zhou H, et al. Object Descriptor Combining Color and Detection*. TELKOMNIKA Indonesian Journal of Electrical Engineering*. 2013; 11(6): 3072-3079.
[17] Indah Agustien, Siradjuddin, M Rahmat Widyanto, T Basaruddin. Particle Filter with Gaussian Weighting for Human Tracking. *TELKOMNIKA Indonesian Journal of Electrical Engineering*. 2012; 10(4): 801- 806.
[18] Dalal N, Triggs B. *Histograms of oriented gradients for human detection.* Proceedingof the IEEE Computer Society Conference on Computer Vision and Pattern Recognition. San Diego, USA: IEEE. 2005; 886-893.
[19] Perez P, Vermaak J, Blake A. *Data fusion for visual tracking with particles*. Proceeding of the IEEE. 2004; 92(3): 495-51.