PDA Filter Based on the Winner-Update Strategy for Visual Tracking

David Liu[1] **Li-Chen Fu**[1,2] **Teng-Kai Kuo**[1] **Jun-Wen Yeh**[1]
Department of Electrical Engineering [1]
Department of Computer Science and Information Engineering [2]
National Taiwan University, Taipei, Taiwan, R.O.C.
Tel: (02) 23622209
Fax: (02) 23657887
E-mail: lichen@ccms.ntu.edu.tw

摘要

當影像複雜到無法正確分辨目標物與非目標物時，PDA提供一解決的辦法。以往PDA主要應用在雷達上，本論文將PDA應用在可見光之影像追蹤。由於使用可見光，目標物之形狀、顏色、大小等特徵皆提供了更充裕的資訊。我們將此些特徵與PDA做適當之結合，以提高追蹤成功率。其中使用的影像辨識方法為Winner-Update，此法又較傳統之SAD節省90%以上之運算量。

關鍵詞：PDA，Winner-Update，Kalman Filter，visual tracking。

Abstract

The PDA filter deals with tracking a target when there is uncertainty in the origin of the measurements. Its use has primarily been on radar tracking tasks. While the target information is provided by a visible-light camera, much more information is available, like the shape, color, and size. We propose a method which incorporates this additional information into the PDA filter. The Winner-Update strategy is utilized for target detection and provides the likelihood ratio for the PDA filter.

Keywords: PDA, Winner-Update, Kalman Filter, visual tracking.

1. Introduction

In a visual tracking problem the CCD camera provides the coordinates of each detected target. Measurements are affected by noise which is modeled as Gaussian. As the camera repeats capturing, the sequence of plots can be processed in a proper filter to smooth the measurement noise, thus providing target tracks. In the practical case, the problem is complicated by the false alarms owing to system noise and clutter. By clutter we refer to detections from nearby objects, weather, etc., that are generally random in location, intensity, and number. This may lead to several measurements in the "validation region" [1] of a single target. The information of the target of interest is called the "target-originated measurement" in the Probabilistic Data Association (PDA) context. The PDA Filter [1,2] is a technique to handle the false data difficulty where the measurements are processed with a probabilistic weighting within the state estimation procedure.

In the visible light target tracking case, as opposed to the radar tracking case, much more target information is available, like shape and color. Traditionally, target detection is handled by techniques like the Sum of Squared Difference (SSD) [3], or by motion-based recognition [4], etc. The Winner-Update Strategy (WinUp) [5], which is a SAD-like template matching strategy, can save 91.6% to 98.0% of the absolute operations needed by the full scan SAD algorithm, depending on the image sequence.

The objective of this paper is to combine WinUp with the PDA Filter. In section 2 and 3, we briefly review the essence of WinUp and PDA respectively. In section 4 we propose how to produce a likelihood ratio and incorporate it into the PDA filter. Section 5 discusses the control method. Section 6 presents the experiment results.

2. The Winner-Update Strategy

The Winner-Update Strategy is a special case of the branch-and-bound strategy. The basic idea is that one does not have to examine all pixels within a block to find out which block has minimum error, since most blocks have very large error even with very little pixels examined. This leads to a huge reduction in computation time. Define the partial sum of absolute difference (PSAD) of l pixels, $l = 1,2,\ldots,B^2$, as:

$$\begin{aligned}\mathrm{PSAD}^{l}_{(\mathrm{x,y})}(u,v) \equiv \sum_{m=0}^{l-1} | & I_t(x+i(m), y+j(m)) \\ & - I_{t-1}(x+u+i(m), y+v+j(m)) |\end{aligned}$$

where $\{(i(m), j(m)) \mid m = 0,\ldots, B^2-1\}$ is the index set of all the pixels in the block, and B is the block size. The index set determines the positions and the order of pixels in the matching block used for the accumulation of PSAD. Obviously, the following holds true:

$$\mathrm{PSAD}^1(u,v) \le \mathrm{PSAD}^2(u,v) \le \cdots$$
$$\le \mathrm{PSAD}^{B^2}(u,v) \equiv \mathrm{SAD}(u,v)$$

The last term, SAD (sum of absolute difference), is the most primitive template matching technique. It calculates the complete matching error, as opposed to the partial matching error PSAD. The Winner-Update Algorithm is summarized below:

The Winner-Update Algorithm

```
Given a template block at position $(x, y)$ in $I_t$
begin
  for each $(u,v)$ in the search range do
    begin (initialization)
      Calculate $\mathrm{PSAD}^1(u,v)$
      $\mathrm{PSAD}(u,v) := \mathrm{PSAD}^1(u,v)$
      $l(u,v) := 1$
    end
  select $(\hat{u},\hat{v})$ having minimum $\mathrm{PSAD}(\hat{u},\hat{v})$ to be the temporary winner
  while $l(\hat{u},\hat{v}) < B^2$ do
    begin
      $l(\hat{u},\hat{v}) := l(\hat{u},\hat{v}) + 1$
      calculate $\mathrm{PSAD}^{l(\hat{u},\hat{v})}(\hat{u},\hat{v})$
      $\mathrm{PSAD}(\hat{u},\hat{v}) := \mathrm{PSAD}^{l(\hat{u},\hat{v})}(\hat{u},\hat{v})$
      select $(\hat{u},\hat{v})$ having minimum $\mathrm{PSAD}(\hat{u},\hat{v})$ to be the new temporary winner
    end
  output $(\hat{u},\hat{v})$
end
```

3. The PDA Filter

The Probabilistic Data Association Filter (PDAF) [1,2] is used to handle the measurement origin uncertainty problem. It computes the posterior association probabilities for all current candidate measurements in a validation gate and uses them to form a weighted sum of innovations for updating the target's state in a suitably modified version of the Kalman Filter.

3.1 State Space Model

The dynamics of the object are modeled by the equation

$$\mathbf{x}(k+1) = \mathbf{F}\,\mathbf{x}(k) + \mathbf{w}(k)$$

where

$$\mathbf{x}(k) = \begin{bmatrix} x_i(k) \\ y_i(k) \\ v_{x_g}(k) \\ v_{y_g}(k) \\ a_{x_g}(k) \\ a_{y_g}(k) \end{bmatrix}$$

where i denotes the image frame, and g denotes the ground frame,

$$\mathbf{F} = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 & 0 & 1 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix}$$

and the process noise vector is taken to be white Gaussian, with

$$E\{\mathbf{w}(j)\mathbf{w}(k)'\} = \delta_{jk}\mathbf{Q}$$

The measurement system is modelled as follows. If the measurement originates from the object in track, then

$$\mathbf{z}(k) = \mathbf{H}\mathbf{x}(k) + \mathbf{v}(k)$$

where

$$\mathbf{z}(k) = \begin{bmatrix} zx_i(k) \\ zy_i(k) \\ zv_{x_g}(k) \\ zv_{y_g}(k) \end{bmatrix}, \quad \mathbf{H} = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 & 0 & 1 \end{bmatrix}$$

and the measurement noise is taken to be white Gaussian, with

$$E\{\mathbf{v}(j)\mathbf{v}(k)'\} = \delta_{jk}\mathbf{R}$$

3.2 The PDA Implementation

The PDA filter is implemented by first performing measurement validation and then state estimate update, as summarized below.

3.2.1 The Validation Region

The measurements are reduced to a set of validated measurements by defining the following validation region

$$\tilde{V}_{k+1}(\gamma) = \left\{\mathbf{z}_i : \mathbf{v}_i'(k+1)\,\mathbf{S}^{-1}(k+1)\,\mathbf{v}_i(k+1) \le \gamma\right\}$$

where $\mathbf{v}_i(k+1) = \mathbf{z}_i(k+1) - \hat{\mathbf{z}}(k+1\,|\,k)$ is the innovation. Each measurement $\mathbf{z}_i$ that lies within this region is considered validated. The threshold γ is obtained from tables of the chi-square distribution, since the weighted norm of the innovation that defines the validation region is chi-square distributed with number of degrees of freedom equal to the dimension of the measurement. The computation of the measurement prediction $\hat{\mathbf{z}}(k+1|k)$ as well the state estimate are shown next.

3.2.2 State Estimation

Suppose at time k there are a number of m_k validated measurements. The set of validated measurements at time k is denoted by

$$Z(k)=\{\,\mathbf{z}_i(k)\}_{i=1}^{m_k}$$

and the cumulative set of measurements up to time k is

$$Z^k=\{Z(j)\}_{j=1}^{k}$$

Define the events $\theta_i(k)=\{\,\mathbf{z}_i(k)$ is the target originated measurement $\}$, $i=1,\ldots,m_k$, and $\theta_0(k)=\{$ none of the measurements at time k is target-originated $\}$ with probability

$$\beta_i(k)=P\left\{\theta_i(k)|Z^k\right\},\;\; i=0,1,\ldots,m_k$$

The procedure that yields these probabilities is called PDA, and will be given in the next section. At the moment we assume they are known.

By the total probability theorem, the conditional mean of the state can be written as

$$\begin{aligned}\hat{\mathbf{x}}(k\,|\,k)&=E\left\{\mathbf{x}(k)\,|\,Z^k\right\}\\&=\sum_{i=0}^{m_k}E\left\{\mathbf{x}(k)\,|\,\theta_i(k),Z^k\right\}P\left\{\theta_i(k)\,|\,Z^k\right\}\\&=\sum_{i=0}^{m_k}\hat{\mathbf{x}}_i(k\,|\,k)\beta_i(k)\end{aligned}$$

where $\hat{\mathbf{x}}_i(k\,|\,k)$ is the updated state estimate conditioned on the event that the i^{th} validated measurement is correct. This is given by the standard Kalman Filter as

$$\hat{\mathbf{x}}_i(k\,|\,k)=\hat{\mathbf{x}}(k\,|\,k-1)+\mathbf{W}(k)\,\mathbf{v}_i(k)$$

where $\mathbf{v}_i(k)=\mathbf{z}_i(k)-\hat{\mathbf{z}}(k\,|\,k-1)$ is the corresponding innovation, and $\mathbf{W}(k)$ is the standard Kalman gain. The error covariance associated with the updated state estimate is defined as

$$\mathbf{P}(k\,|\,k)=\left\{\left[\mathbf{x}(k)-\hat{\mathbf{x}}(k\,|\,k)\right]\left[\mathbf{x}(k)-\hat{\mathbf{x}}(k\,|\,k)\right]'\middle|Z^k\right\}$$

and can be evaluated by

$$\begin{aligned}\mathbf{P}(k\,|\,k)&=\beta_0(k)\mathbf{P}(k\,|\,k-1)\\&\quad+[1-\beta_0(k)]\mathbf{P}^{\mathrm{c}}(k\,|\,k)+\tilde{\mathbf{P}}(k)\end{aligned}$$

where

$$\begin{aligned}\tilde{\mathbf{P}}(k)=\mathbf{W}(k)\Bigg(&\sum_{i=1}^{m_k}\beta_i(k)\,\mathbf{v}_i(k)\mathbf{v}_i'(k)\\&-\mathbf{v}\,(k)\,\mathbf{v}'(k)\,\Bigg)\,\mathbf{W}'(k)\end{aligned}$$

and

$$\mathbf{P}^{\mathrm{c}}(k\,|\,k)=\left(\mathbf{I}-\mathbf{W}(k)\mathbf{H}\right)\mathbf{P}(k\,|\,k-1).$$

Now that we have obtained $\hat{\mathbf{x}}_i(k\,|\,k)$, it remains to find out the values of $\beta_i(k)$.

3.2.3 The PDA Weights

The association probabilities $\beta_i(k)$ are developed in detail in [2] and require knowledge of the probability mass function of the number of false measurements (clutter). Assuming a Poisson density with parameter λV_k, the weights are given by

$$\beta_i(k)=\frac{e_i}{b+\sum_{i=1}^{m_k}e_i}\;,\;\; i=0,1,\ldots,m_k \tag{1}$$

$$\beta_0(k)=\frac{b}{b+\sum_{i=1}^{m_k}e_i} \tag{2}$$

where

$$e_i=\exp\left(-\frac{1}{2}\mathbf{v}_i'(k)\mathbf{S}^{-1}(k)\mathbf{v}_i(k)\right)$$

$$b=\left(2\pi/\gamma\right)^{n_z/2}\lambda V_k c_{n_z}\left(1-P_DP_G\right)/P_D$$

P_G is the probability that the target-originated measurement falls within the validation gate, and P_D is the probability that the correct measurement is detected, and c_{n_z} is the volume of the n_z dimensional unit hypersphere $\left(c_1=2\,,\,c_2=\pi\,,\,c_3=4\pi/3,\,etc.\right)$.

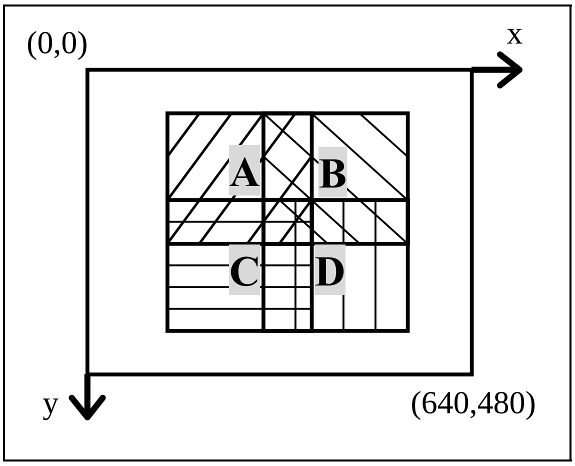

Fig. 1 Partition of the region of interest

4. PDAF with Winner-Update Strategy

To fully employ the image feature information of the target, we present the following method. The region of interest (ROI) is partitioned into four parts, call them part A, B, C, and D (see fig. 1). The WinUp algorithm is then performed in each of the four parts. Each will then get a similarity value, $\mathbf{SAD}_i$, with a smaller value indicating a higher likeliness. The similarity value is then converted into the likelihood ratio, defined by

$$\lambda_i = \frac{\sum_{i=1}^{4} \mathbf{SAD}_i}{\mathbf{SAD}_i + 1}, \quad i = 1,2,3,4 \cdot$$

We can see that following this definition, the region with smaller **SAD** will have a larger likelihood ratio λ. This likelihood ratio can be used to modify the standard PDA Filter's association probabilities in equation (1),(2), as follows [6]:

$$\beta_i(k) = \frac{e_i \exp(k\lambda_i)}{b + \sum_{i=1}^{m_k} e_i \exp(k\lambda_i)}, \quad i = 0,1,\ldots,m_k \qquad (3)$$

$$\beta_0(k) = \frac{b}{b + \sum_{i=1}^{m_k} e_i \exp(k\lambda_i)} \qquad (4)$$

where $k \geq 0$ is included to adjust the influence of the likelihood ratio on the standard PDA Filter's association probabilities. If $k = 0$, equation (3),(4) return back to their original form (eq.(1),(2)).

5. Motor Control

For the pan-tilt camera platform, the predicted target velocity in ground frame (v_{x_g}, v_{y_g}) and the predicted target position in image frame $(\hat{x}_i, \hat{y}_i)$ are utilized for motor control. The control input is obtained by

$$\begin{pmatrix} control\ input\ x \\ control\ input\ y \end{pmatrix} = f\begin{pmatrix} \hat{v}_{x_g} \cdot \Delta t \\ \hat{v}_{y_g} \cdot \Delta t \end{pmatrix} + \mathbf{u} + k_p \cdot \begin{pmatrix} image\ position\ error\ x \\ image\ position\ error\ y \end{pmatrix}$$

and

$$\mathbf{u} \leftarrow \mathbf{u} + k_I \cdot \begin{pmatrix} image\ position\ error\ x \\ image\ position\ error\ y \end{pmatrix}$$

where **u** is initialized to zero. The function f is obtained by coarse calibration. The f term and the k_p term constitute the basic control component based on the nominal system. As the nominal part would not represent the servomechanism exactly, additional enhancements to achieve effective control is required, and it is useful to consider an overall control law with the learning control [7] component **u** included.

6. Real Experiments and Results

The proposed method is demonstrated by a real indoor target tracking experiment. The target is an combat airplane model moving from left to right along a straight track, and the background is composed of one big poster with cluttered textures, and another stationary combat airplane model. These constitute the cluttered background and are used to compare the performance of the WinUp-PDAF with the standard Kalman Filter. The whole image is of size 640×480, and the block size B for template matching is 16×16 pixels. The search range for parts A, B, C, and D is 72×72, and the overlapping length of two parts is 16 pixels, so that the total search range is 128×128 (fig. 1). For comparison, the WinUp-PDAF tracking module is compared with the standard Kalman Filter module. Both modules are using WinUp as the target detection method, while the Kalman Filter only keeps track of the point with minimum SAD. The template matching block is first locked onto the centroid of the target, and thereafter it is kept not updated.

Smooth control input is the desired goal. Since the smoother the control input, the less oscillation occurs on the camera, and this yields a higher confidence in target detection.

The results are shown in fig. 2(a~e) and table 1. Tracking with WinUp-PDAF yields smoother control input. This in turn yields higher confidence in target detection, since fast switching control input, as shown in fig. 2.(c),(d), results in large oscillation and serious image blurring. Even worse, the system will go lost of track as in fig. 2(e).

7. Acknowledgement

The Winner-Update algorithm is implemented in C language by the authors of [5] and downloadable at ftp://smart.iis.sinica.edu.tw.

Table 1. Results of the experiments

	Variance of control input (Unit in motor steps)
Experiment 1 : Fig.2(a) WinUp-PDAF	100.61
Experiment 2 : Fig.2(b) WinUp-PDAF	115.49
Experiment 3 : Fig.2(c) WinUp-Kalman Filter	142.54
Experiment 4 : Fig.2(d) WinUp-Kalman Filter	173.22

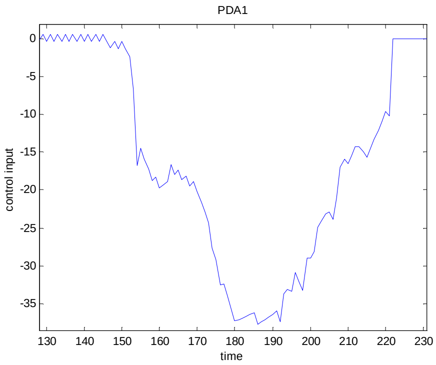

Fig. 2(a)

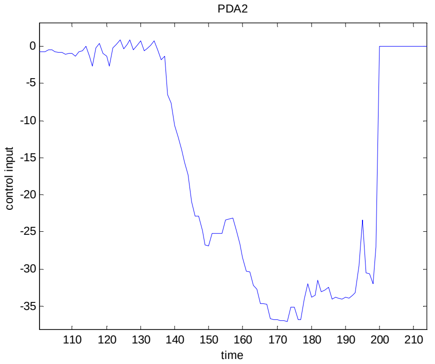

Fig. 2(b)

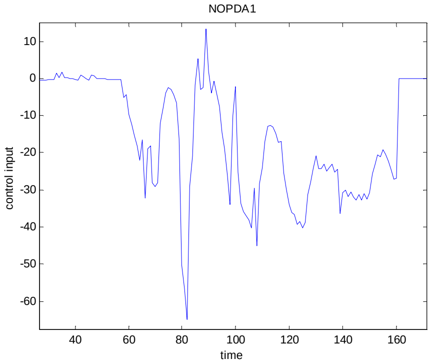

Fig. 2(c)

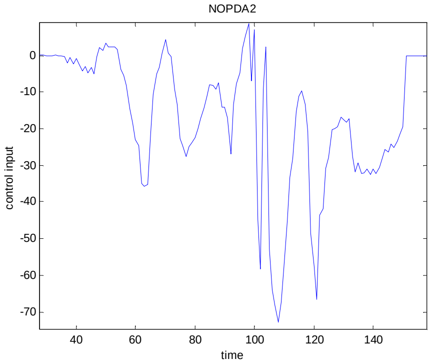

Fig. 2(d)

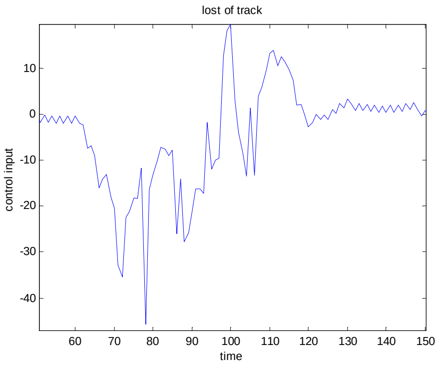

Fig. 2(e)

Fig. 2 (a)~(e) : The horizontal 'time' axis has unit length equal to one cycle, which is about 120ms. The vertical 'control input' has unit length equal to one step of the stepping motor, while 10,000 steps correspond to 2π radians. Fig.2(a&b) are WinUp with PDAF. Fig.2(c&d) are WinUp with standard Kalman Filter. Fig.2(e) demonstrates a case where large oscillation finally lead to lost of track.

8. References

[1] Y. Bar-Shalom and E. Tse, "Tracking in a Cluttered Environment with Probabilistic Data Association," *Automatica*, Vol. 11, pp. 451-460, 1975.

[2] Y. Bar-Shalom and T. E. Fortmann, *Tracking and Data Association*, Academic Press, 1988.

[3] N. P. Papanikolopoulos and P. K. Khosla, "Adaptive Robotic Visual Tracking: Theory and Experiments," *IEEE Trans. Automatic Control*, Vol. 38, no. 3, pp. 429-445, Mar. 1993.

[4] C. Cédra and M. Shah, "Motion-based recognition: a survey," *Image and Vision Computing*, Vol. 13, no. 2, pp. 129-155, Mar. 1995.

[5] Y. S. Chen, Y. P. Hung, and C. S. Fuh, "A Fast Block Matching Algorithm Based on the Winner-Update Strategy," in *Proceedings of the Fourth Asian Conference on Computer Vision*, Vol. 2, pp. 977-982, Jan. 2000.

[6] D. Lerro and Y. Bar-Shalom, "Automatic Track Formation with Target Amplitude Information," in Proceedings of Ocean Technologies and Opportunities in the Pacific for the 90's, pp. 1460-1467, 1991.

[7] T. J. Jang, C. H. Choi, and H. S. Ahn, "Iterative learning control in feedback systems," *Automatica*, Vol. 31, No. 2, pp. 243-245, 1995.