the theoretical aspect, the stability analysis of specific missions and the inclusion of nonholonomic issues are of great interest.

REFERENCES

[1] G. Antonelli and S. Chiaverini, "Kinematic control of a platoon of autonomous vehicles," in *Proc. IEEE Int. Conf. Robot. Autom.*, Taipei, Taiwan, Sep. 2003, pp. 1464–1469.

[2] G. Antonelli and S. Chiaverini, "Obstacle avoidance for a platoon of autonomous underwater vehicles," in *Proc. 6th IFAC Conf. Manoeuvring and Control of Marine Craft*, Girona, Spain, Sep. 2003, pp. 143–148.

[3] R. C. Arkin, "Motor schema based mobile robot navigation," *Int. J. Robot. Res.*, vol. 8, no. 4, pp. 92–112, 1989.

[4] T. Balch and R. C. Arkin, "Behavior-based formation control for multirobot teams," *IEEE Trans. Robot. Autom.*, vol. 14, no. 6, pp. 926–939, Dec. 1998.

[5] T. D. Barfoot, C. M. Clark, S. M. Rock, and G. T. M. D'Eleuterio, "Kinematic path-planning for formations of mobile robots with a nonholonomic constraints," in *Proc. IEEE/RSJ Int. Conf. Intell. Robots Syst.*, Lausanne, Switzerland, Oct. 2002, pp. 2819–2824.

[6] B. E. Bishop, "On the use of redundant manipulator techniques for control of platoons of cooperating robotic vehicles," *IEEE Trans. Syst., Man, Cybern.*, vol. 33, no. 5, pp. 608–615, Sep. 2003.

[7] B. E. Bishop and D. J. Stilwell, "On the application of redundant manipulator techniques to the control of platoons of autonomous vehicles," in *Proc. IEEE Int. Conf. Control Appl.*, México City, México, Sep. 2001, pp. 823–828.

[8] R. A. Brooks, "A robust layered control system for a mobile robot," *IEEE J. Robot. Autom.*, vol. RA-2, no. 1, pp. 14–23, Feb. 1986.

[9] Y. Uny Cao, A. S. Fukunaga, and A. B. Kanhg, "Cooperative mobile robotics: Antecedents and directions," *Auton. Robots*, vol. 4, pp. 7–27, 1997.

[10] P. Chiacchio, S. Chiaverini, L. Sciavicco, and B. Siciliano, "Closed-loop inverse kinematics schemes for constrained redundant manipulators with task space augmentation and task priority strategy," *Int. J. Robot. Res.*, vol. 10, no. 4, pp. 410–425, 1991.

[11] S. Chiaverini, "Singularity-robust task-priority redundancy resolution for real-time kinematic control of robot manipulators," *IEEE Trans. Robot. Autom.*, vol. 13, no. 3, pp. 398–410, Jun. 1997.

[12] J. P. Desai, V. Kumar, and J. P. Ostrowski, "Control of changes in formation for a team of mobile robots," in *Proc. IEEE Int. Conf. Robot. Autom.*, Detroit, MI, May 1999, pp. 1556–1561.

[13] I.-A. F. Ihle, J. Jouffroy, and T. I. Fossen, "Formation control of marine surface craft using Lagrange multipliers," in *Proc. 44th IEEE Conf. Decision, Control/8th Eur. Control Conf.*, Sevilla, Spain, Dec. 2005, p. 8.

[14] A. A. Maciejewski and C. A. Klein, "Obstacle avoidance for kinematically redundant manipulators in dynamically varying environments," *Int. J. Robot. Res.*, vol. 4, no. 3, pp. 109–117, 1985.

[15] Y. Nakamura, H. Hanafusa, and T. Yoshikawa, "Task-priority based redundancy control of robot manipulators," *Int. J. Robot. Res.*, vol. 6, no. 2, pp. 3–15, 1987.

[16] L. E. Parker, "Designing control laws for cooperative agent teams," in *Proc. IEEE Int. Conf. Robot. Autom.*, Atlanta, GA, May 1993, vol. 3, pp. 582–587.

[17] S. T. Pledgie, Y. Hao, A. M. Ferreira, S. K. Agrawal, and R. Murphey, "Groups of unmanned vehicles: Differential flatness, trajectory planning, and control," in *Proc. IEEE Int. Confe. Robot. Autom.*, Washington, DC, May 2002, pp. 3461–3466.

[18] B. Siciliano, "Kinematic control of redundant robot manipulators: A tutorial," *J. Intell. Robot. Syst.*, vol. 3, pp. 201–212, 1990.

[19] D. J. Stilwell, B. E. Bishop, and C. A. Sylvester, "Redundant manipulator techniques for partially decentralized path planning and control of a platoon of autonomous vehicles," *IEEE Trans. Syst., Man, Cybern.*, vol. 35, no. 4, pp. 842–848, Aug. 2005.

[20] C. W. Wampler, II, "Manipulator inverse kinematic solutions based on vector formulations and damped least-squares methods," *IEEE Trans. Syst., Man, Cybern.*, vol. SMC-16, no. 1, pp. 93–101, Jan. 1986.

[21] H. Yamaguchi and J. W. Burdick, "Asymptotic stabilization of multiple nonholonomic mobile robots forming group formations," in *Proc. IEEE Int. Conf. Robot. Autom.*, Leuven, Belgium, May 1998, pp. 3573–3580.

Visual Tracking in Cluttered Environments Using the Visual Probabilistic Data Association Filter

Cheng-Ming Huang, David Liu, and Li-Chen Fu

***Abstract*—Visual tracking in cluttered environments is attractive and challenging. This paper establishes a probabilistic framework, called the visual probabilistic data-association filter (VPDAF), to deal with this problem. The algorithm is based on the probabilistic data-association method for estimating a true target from a cluster of measurements. There are two other key concepts which are involved in VPDAF. First, the sensor data are visual, similar to the target in the image space, which is a crucial property that should not be ignored in target estimation. Second, the traditional probabilistic data-association filter for the underlying application is vulnerable to stationary disturbances in image space, mainly due to some annoying background scenes which are rather similar to the target. Intuitively, such persistent noises should be separated out and cleared away from the continuous measurement data for seeking successful target detection. The proposed VPDAF framework, which incorporates template matching, can achieve the goal of reliable realtime visual tracking. To demonstrate the superiority of the system performance, extensive yet challenging experiments have been conducted.**

***Index Terms*—Data association, probabilistic data-association filter (PDAF), visual tracking.**

I. INTRODUCTION

In the practical implementation of visual tracking in the real world, it is hard to get perfect image data from the target. Disturbances due to mechanical vibration and optical projection will seriously interfere with the true target's detection or estimation in the image space. Mechanical noise also usually arises from undesirable shaking of the camera, whereas optical interference may come from similar objects, cluttered background, occlusions, lighting changes, etc. Under such circumstances, it is hard to distinguish the true target from the fake ones through the target-detection process. To cope with this difficulty, apparently we have to incorporate a more sophisticated tracking methodology into our visual tracking system. Instead of using only the best measurement among the perceived ones and discarding the rest, an alternative approach is to consider multiple measurements at the same time, using the probabilistic data-association filter (PDAF) [1].

The PDAF algorithm has been extensively applied to radar systems. But in the last decade, its applications gradually included tracking with imaging sensors or some other visual aspects. Bar-Shalom *et al.* [2]–[4], [15] have put a lot of effort into how to use probabilistic data association (PDA) algorithms to track single or multiple objects through image-domain reasoning, and PDAF has been modified in various ways in order to adapt to different task environments. Besides the above, PDAF based on a nonlinear motion model for establishing correspondences between two consecutive frames was shown in [5]. Furthermore, Hager *et al.* [6], [16] reported many experiments on visual-object tracking performed using the PDAF idea with different

This paper was recommended for publication by Associate Editor D. Sun and Editor L. Parker upon evaluation of the reviewers' comments. Manuscript received February 4, 2006; revised June 26, 2006. This work was supported by the National Science Council of Taiwan under NSC-95-2752-E-002-007-PAE. This paper was presented in part at the IEEE/RSJ International Conference on Intelligent Robots and Systems, Maui, HI, November 2001.

C. M. Huang and D. Liu are with the Department of Electrical Engineering, National Taiwan University, Taipei 106, Taiwan, R.O.C.

L. C. Fu is with the Department of Electrical Engineering and Department of Computer Science and Information Engineering, National Taiwan University, Taipei 106, Taiwan, R.O.C. (e-mail: lichen@ntu.edu.tw).

Digital Object Identifier 10.1109/TRO.2006.882960

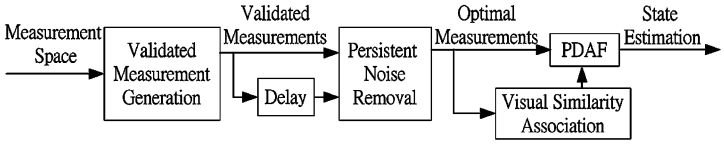

Fig. 1. Overall architecture of the VPDAF.

modalities for image analysis. They mainly resolved the overlapping problem with modified joint PDAFs given the tasks of tracking multiple objects. Finally, the contours connecting the edge points were used to derive the shape-based PDAF for shape tracking [7].

However, most of the above research views the visual measurements as the sensed light spots on the radar screen. The locations of multiple validated measurements are input to the PDAF-like estimator, which then outputs a prediction of the target's position. The most significant characteristics, such as visual similarity or image intensity, are unfortunately not taken into consideration [15]. The validated measurements, which are close to the prediction but exhibit low visual similarity, will deteriorate the tracking result, since they fail to originate from the true target of interest. Some other research [4], [6], [7] has used visual measurements to interpret estimations with more logical hypothetical events. They apparently can decrease the false alarm rate, but the visual similarity property of each event is not taken as a weighting when the associated probability is to be evaluated. Last but not least, the traditional PDAF has been challenged [8], in that its performance is degraded when a large number of nearly stationary false clutter appears. Many scene elements, including a static background and other objects [18], may also produce visually similar noise, and degrade the target estimate and tracking. One natural remedy is to separate this noise from the validated measurements according to the different motion trajectories.

In this paper, we describe a reliable visual-tracking system that employs the visual PDAF (VPDAF) algorithm. Fig. 1 illustrates the architecture of VPDAF. Conceptually, an image frame is just like a space concealing much of the target's information from the real world that is projected onto the imaging sensor domain. Through target detection within a reasonable region around the prediction, we can filter the data after image processing and then output validated measurements which should be refined again, since stationary noise can degrade the final estimate, as mentioned above. The traditional PDAF is then engaged and augmented with visual similarity. The associated probability of each measurement is evaluated in order to validate the sensed data and generate the final state estimate.

This paper is organized as follows. Section II describes the observational densities of measurements' features, which cooperate with the traditional PDAF to evolve into the augmented PDAF. Section III presents the modified PDAF, excluding the augmentation in Section II, with the persistent-noise-removal unit and a MATLAB simulation for comparison purposes. In Section IV, various applications of the proposed VPDAF, i.e., the PDAF with the visual-similarity evaluation unit and persistent-noise-removal mechanism, as shown in Fig. 1, are demonstrated with extensive experiments and discussed. Finally, conclusions are drawn in Section V.

II. Augmented PDAF With Observational Density

The measurements $\mathbf{Z}(k) = \{\mathbf{z}_i(k), i = 1, \ldots, m_k\}$, returned by the target-detection process and filtered through the validated measurement generation process, will process a visual similarity score as the basis for the so-called observational density. Let m_k be the number of current validated measurements. The state estimate should depend on both the feature-likelihood ratio and the position distribution of the validated measurements [15], i.e., the updated state can be estimated by

$$\hat{\mathbf{x}}(k|k) = E\{\mathbf{x}(k)|\mathbf{I}_{\mathbf{z}}(k), \mathbf{Z}^k\} \tag{1}$$

where $\mathbf{I}_{\mathbf{Z}}(k)$ is the observational density distribution of the validated measurement set $\mathbf{Z}(k)$, and $\mathbf{Z}^k = \{\mathbf{Z}(j), j = 1, \ldots, k\}$ denotes the cumulative measurement history. Thus, from the total probability theory with respect to the event $\theta_i(k)$, of which measurement is target-originated, (1) can be rewritten as

$$\begin{aligned}\hat{\mathbf{x}}(k|k) &= \sum_{i=0}^{m_k} E\{\mathbf{x}(k)|\theta_i(k), \mathbf{I}_{\mathbf{z}}(k), \mathbf{Z}^k\} P\{\theta_i(k)|\mathbf{I}_{\mathbf{z}}(k), \mathbf{Z}^k\} \\ &= \sum_{i=0}^{m_k} \hat{\mathbf{x}}_i(k|k)\beta_i(k)\end{aligned} \tag{2}$$

where $\beta_i(k) \triangleq P\{\theta_i(k)|\mathbf{I}_{\mathbf{z}}(k), \mathbf{Z}^k\}, i = 0, 1, \ldots, m_k$ is the mixed probability of each measurement associated with the final estimate. Notice that $\beta_0(k)$ is associated with the event $\theta_0(k)$, of which none of the measurements are target-originated. Applying Bayes' theory, we obtain the following expression:

$$\beta_i(k) = P\{\mathbf{I}_{\mathbf{z}}(k) \mid \theta_i(k), \mathbf{Z}^k\} P\{\theta_i(k) \mid \mathbf{Z}^k\}. \tag{3}$$

The joint probability of the observational density conditioned on the event $\theta_i(k)$ indicates the quality of this event arising from the ith measurement $\mathbf{z}_i(k)$. Since there is only one interesting target in this tracking scenario, no more than one measurement can originate from the same target. When the event $\theta_i(k)$ is considered, the associated probability of the measurement $\mathbf{z}_i(k)$ that is assumed to be the correct one should also be proportional to its observational density. Hence, the first factor in (3) can be broken down as follows [15]:

$$\begin{aligned}&P\{\mathbf{I}_{\mathbf{z}}(k)|\theta_i(k), \mathbf{Z}^k\} \\ &\quad = P\{\mathbf{I}_{\mathbf{z}_i}(k)|\mathbf{z}_i(k), \theta_i(k), \mathbf{Z}^{k-1}\} \\ &\qquad \cdot \prod_{j\neq i, j=1}^{m_k} P\left\{\mathbf{I}_{\mathbf{z}_j}(k)|\mathbf{z}_j(k), \theta_j(k), \mathbf{Z}^{k-1}\right\} \\ &\quad = \frac{I_i(\mathbf{z}_i)}{I_0(\mathbf{z}_i)} \prod_{j=1}^{m_k} I_0(\mathbf{z}_j), i = 1, \ldots, m_k\end{aligned} \tag{4}$$

where $I_i(\mathbf{z}_i)$ is the observational density of the validated measurement $\mathbf{z}_i(k)$, and $I_0(\mathbf{z}_j)$ is the probability distribution function of the observational density for measurement j when it is assumed to be false. We denote (4) as the conditional observational density of the measurement $\mathbf{z}_i(k)$. Obviously, there is no observational density for the case with $i = 0$, since there are no measurements taken under that circumstance.

The second factor $P\{\theta_i(k)|\mathbf{Z}^k\}$ in (3) corresponds to the probability of the event $\theta_i(k)$, given the validated measurements. This term can be referred to as the original PDA estimation [1]. Finally, the data-associated probabilities should be normalized so as to satisfy $\mathrm{d}\sum_{i=0}^{m_k} \beta_i(k) = 1$.

III. Persistent Noise Removal

The original PDAF assumes a single *a priori* probability density function (pdf) for the false measurements [1], [12]. The authors in [9] established a new false-measurement model for multiple *a priori* pdfs. Although these models might work for radar or sonar tracking purposes, they are not useful for visible light object tracking. Data-distortion factors, such as partial occlusion and the presence of clutter, can lead to false alarms whose distributions are unknown *a priori* [10], [11], but which usually occur at almost the same position in consecutive frames with little variance. Hence, false measurements should be modeled with nearly stationary, clustered, but unknown pdfs.

The above conclusion is based on experimental observations showing that even though the augmented PDAF has been strengthened by incorporating image features, the tracking system is still vulnerable to persistent interference. Neither the original PDAF nor augmented PDAF can filter out these persistent disturbances, such as a static background or other moving objects [18], with image features similar to the target. In other words, measurements that involve nearly stationary clutter will seriously contaminate the state estimate. In this section, we will propose a new method that can improve the PDAF's ability to deal with nearly stationary and clustered clutter. This approach also does not require *a priori* knowledge about the distribution of the clutter [8].

Define the measurement variation as $D_i(k) = \min_j \|\mathbf{z}_i(k) - \mathbf{z}_j(k-1)\|, i = 1, \ldots, m_k, j = 1, \ldots, m_{k-1}$. These measurement variations $D_i(k)$ can be separated into two clusters in one-dimensional (1-D) space. The cluster with the smaller value of $D_i(k)$, originating from current-previous measurement pairs $\mathbf{z}_i(k)$ and $\mathbf{z}_j(k-1)$, should have measurements that are all nearly stationary. Also, the cluster with the larger value of $D_i(k)$ may come from either nonstationary–nonstationary pairs or nonstationary–stationary pairs. The truly target-originated measurement will belong to the latter cluster, due to the motion of the target. We can use hypothesis testing or statistical pattern-recognition techniques to find the decision boundary for the two clusters, and the discriminant function will be very helpful in deriving the modified association probabilities.

By converting the distribution of $D_i(k)$ into a histogram, we can reduce the hypothesis-testing problem to a 1-D thresholding problem. We adopt the iterative threshold selection [13] for finding the threshold through the following steps.

Step 1) Assume no knowledge about whether a measurement is stationary (Class A) or nonstationary (Class B). Randomly select one measurement $\mathbf{z}_i(k), 1 \le i \le m_k$, assign it to Class A, and assign the rest to Class B. Moreover, set the initial threshold $\mathrm{Thr}^0 = 0$.

Step 2) At iteration t, compute

$$\mu_{\mathrm{st}}^t = \frac{\sum_{i \in \text{Class A}} D_i(k)}{\text{number of Class A measurements}}$$
$$\mu_{n\mathrm{st}}^t = \frac{\sum_{i \in \text{Class B}} D_i(k)}{\text{number of Class B measurements}}$$

and $\mathrm{Thr}^{t+1} = (\mu_{\mathrm{st}}^t + \mu_{n\mathrm{st}}^t)/2$.

Step 3) Use Thr^{t+1} as the threshold for classification

$$\begin{cases} \text{if } D_i(k) < \mathrm{Thr}^{t+1}, & i \in \text{Class A} \\ \text{else}, & i \in \text{Class B}. \end{cases}$$

Step 4) If the thresholds of two consecutive iterations are different, i.e., $\mathrm{Thr}^{t+1} \neq \mathrm{Thr}^t$, return to step 2. Otherwise, the optimal threshold Thr is found, and all measurements have been classified as belonging to either Class A or B.

After the optimal threshold is obtained, the measurements that belong to Class A are filtered out. The augmented PDAF mentioned in Section II is then performed on the reduced set of measurements.

A question arises if such a decision boundary or threshold does not exist or is ill-defined, e.g., when the target moves slowly or stays still, and is taken almost as the stationary false measurements. To prevent this, the noise-removal process will be turned off if the estimated displacement from PDAF is small or close to the optimal threshold Thr. In this case, all of the measurements will belong to the same class (nonstationary), and the proposed modified PDAF then returns back to the standard PDAF. This fact suggests that the proposed modified PDAF will work at least as well as the traditional standard PDAF in a pure nonstationary cluttered environment, and may perform better in a more complicated cluttered environment, i.e., with both nonstationary and stationary measurements.

To demonstrate the effectiveness of this proposed noise-removal process, simulations are done in the following. The plant equation, discretized with time interval $T = 1$ s, is

$$\mathbf{x}(k+1) = \mathbf{F}\mathbf{x}(k) \tag{5}$$

where the state is $\mathbf{x}(k) = [x(k) \;\; \dot{x}(k) \;\; y(k) \;\; \dot{y}(k)]^T$ and

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

The initial state is $\mathbf{x}(0) = [200\,\mathrm{m} \;\; 0 \;\; 10^4\,\mathrm{m} \;\; -15\,\mathrm{m/s}]^T$. The measurement model is determined using

$$\mathbf{z}(k) = \mathbf{H}\mathbf{x}(k) + \mathbf{w}(k) \tag{6}$$

where

$$E\{\mathbf{w}(k)\} = 0,\; E\left\{\mathbf{w}(j)\mathbf{w}(k)^T\right\} = \delta_{jk} \begin{bmatrix} 200\,\mathrm{m}^2 & 0 \\ 0 & 200\,\mathrm{m}^2 \end{bmatrix}, \delta_{jk}$$

is the Kronecker delta, and

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

At first, initialization was done in a clean environment with no false measurements, and false measurements are introduced into the system at $k = 10$. There are 200 measurements generated per unit time, including one target-originated measurement, 130 nonstationary false measurements, and 69 stationary false measurements. The nonstationary false measurements are generated at each time step independently and uniformly distributed within the whole 2-D space. Also, we also generate persistent stationary false measurements distributed near the moving trajectory of the actual target. The stationary false measurements are generated with uniform distribution, clustered in a confined region, and containing small perturbation at every sampling instant. The small perturbation is simulated using Gaussian noise with zero mean and identity covariance. One snapshot of all the measurements generation at $k = 200$ s is plotted in Fig. 2.

Fig. 2(a) and (b) shows the estimated trajectories of the original PDAF [1] and modified PDAF over the entire duration (0–200 s) in simulation time. All of the measurements have to be verified to see whether they are located within the validation region [1], whose size is

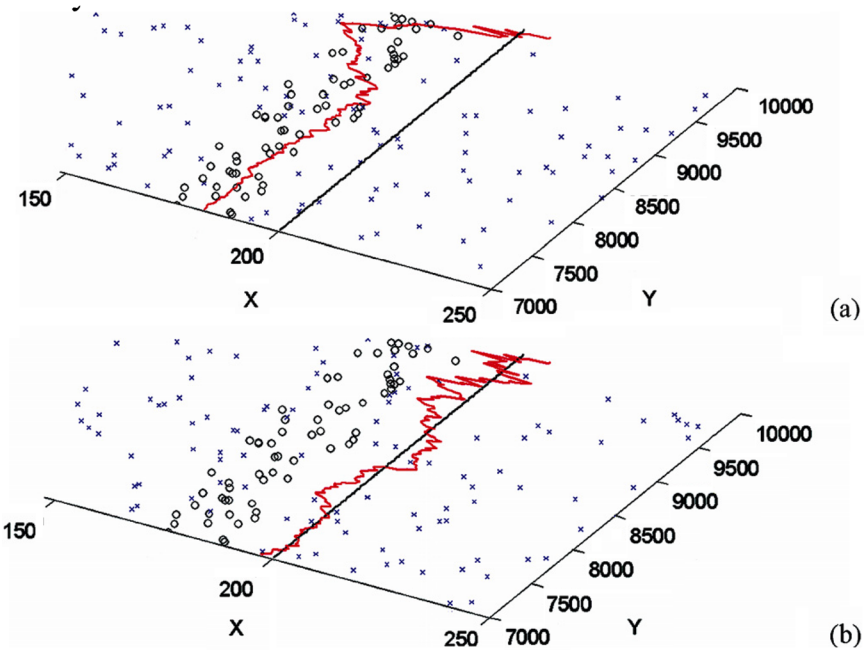

Fig. 2. The "×" and "◯" marks denote the nonstationary false measurements and persistent stationary noise at $k = 200$, respectively. Solid straight lines show the true trajectories over the entire duration ($k = 0$ to 200) of the target moving along $X = 200$ from $Y = 10\,000$ down to $Y = 7000$. The oscillating curve indicates the filtered trajectory over the entire duration. (a) Simulation result of the traditional standard PDAF (without removing stationary noise). (b) Simulation result of the modified PDAF with persistent noise removal. (Color version available online at http://ieeexplore.org.)

Fig. 3. Finger tracking with Kalman filter. The drawn square shows the best-matched and estimated position.

Fig. 4. Finger tracking with VPDAF. The squares denote the validated measurements used in VPDAF. The scenario is the same as in Fig. 3.

affected by the chosen threshold value $\gamma = 16$. In Fig. 2(a), the original PDAF-filtered trajectory is biased away from the real trajectory. However, the solid curve which indicates the modified PDAF-filtered trajectory is now close to the real trajectory in Fig. 2(b).

IV. Experiments

In the following experiments, we adopt the sum of absolute differences (SAD) template matching as the target-detection process. Template matching is implemented with the help of the Winner–Update (WinUp) algorithm [14] and the Intel MMX instructions to speed up the process. On the other hand, since the appearance and posture of the target will change over time, the template must be updated. We apply the Kalman-filter update strategy [17] to update the target template.

A. Kalman Filter versus VPDAF

First, we track fingers in front of stationary books, with the latter being treated as a similar and cluttered background. Two algorithms are considered for comparison purposes: the Kalman filter and VPDAF. The same SAD matching process is used for the two algorithms, but the Kalman filter only adopts the best (most similar) measurement for prediction. Fig. 3 shows that the Kalman filter fails to track the fingers. The Kalman filter misidentifies the books as the fingers, and gradually loses track of them in the background. On the other hand, VPDAF tracks the fingers successfully, as shown in Fig. 4. With the proposed VPDAF, a weighted average of the target estimates results, instead of relying on some particular measurement alone. This has a side effect in that the estimated target position may be at a small tolerable distance from the real target, but satisfactory tracking can still be achieved.

In this scenario, both the augmented PDAF mentioned in Section II and the modified PDAF with stationary noise removal in Section III will track the target as well as VPDAF. Since we can see that most of the validated measurements in Fig. 4 are originated from the hand, the disturbance from the poor measurement's visual similarity augmentation and the background noise on tracking is not evident. We will preset the improvements of VPDAF in the next two experiments.

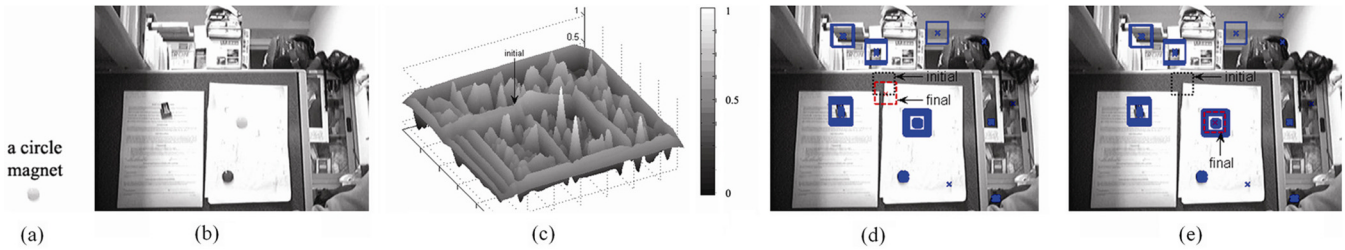

Fig. 5. Comparison of traditional PDAF and VPDAF estimation in one image frame. (a) Template. (b) Original image. (c) Normalized SAD score distribution. (d) PDAF without weight augmentation. (e) VPDAF. (Color version available online at http://ieeexplore.org.)

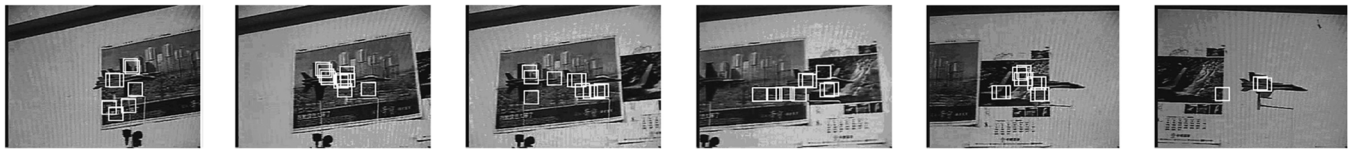

Fig. 6. Tracking an airplane model with VPDAF. The stationary airplane model acts as the "similar" but "stationary" false measurement.

Fig. 7. Tracking a pedestrian with VPDAF.

B. PDAF versus VPDAF

In this experimental comparison, we will focus on the target estimate from one single image frame. Through exhaustive SAD computation in the image space, the visual similarity scores are normalized, and the observational distribution is shown in Fig. 5(c). The summit of this distribution is the target, namely, a round magnet. In Fig. 5(d) and (e), the "$\times$" marks denote positions with observational density values greater than 0.75, and the rectangles represent those within the validation region, i.e., the validated measurements considered by PDAF or VPDAF. In order to show the improvement achieved by the augmentation which evaluates visual similarity in VPDAF as mentioned in Section II, we purposely set the predicted target position before performing the measurement update $\hat{\mathbf{x}}_0(k|k)$ near the false measurements. As shown in Fig. 5(d), since the associated probabilities of PDAF [1] only rely on the location, it is misled by the validated measurements which stay close to the prediction, but have low visual similarity. However, the VPDAF, which enhances PDAF with observational density, successfully performs an accurate update, as shown in Fig. 5(e).

Furthermore, some of the false measurements, originated from the background, with large visual similarity can still perturb the tracking result. The stationary noise removal will be introduced to improve the estimate in the next experiment.

C. Tracking Airplane Model

Instead of the fixed camera used in the above two scenarios, here the camera is mounted on a stepping-motor platform with two degrees of freedom (DOFs). In order to use the measurements' positions for the aforementioned classification purpose, we have to transform their image coordinates (u_i, v_i) into a well-defined spherical camera platform coordinate. By counting the shifting steps of the camera platform and using perspective projection, we can derive the azimuthal angle and polar angle of each measurement in this coordinate system as

$$\begin{bmatrix} \text{azimuth}_i \\ \text{polar}_i \end{bmatrix} = \begin{bmatrix} \tan^{-1}\frac{u_i}{f} & \tan^{-1}\frac{v_i}{f} \end{bmatrix}^T, \quad i = 1, \ldots, m_k \quad (7)$$

where f is the camera constant. Replace the states of (5) as

$$\mathbf{x}(k) = \begin{bmatrix} \text{azimuth}(k) & \Delta\text{azimuth}(k) & \text{polar}(k) & \Delta\text{polar}(k) \end{bmatrix}^T. \quad (8)$$

Then, substitute this result into VPDAF to fulfill the designed functions. The final output of VPDAF is the predicted position of the target in the camera spherical coordinate. We also command the stepping motors to turn to the desired prediction. Hence, the surveillance range can be expanded while realtime visual tracking is realized.

Here we consider a more complicated situation to test the functions of VPDAF. The tracking target is a moving airplane model, but there is also a stationary airplane model in the environment. The moving trajectory of the airplane passes through the cluttered background. In Fig. 6, we can see that many false measurements are generated for the stationary airplane model and background. These false measurements, which have the same observational density value as the actual target, are unlike the noise in Fig. 5. They can be eliminated by the persistent noise removal of VPDAF based on the motion situation. Only the similar measurements originating from the interested target are preserved for evaluation of the association probabilities.

D. Pedestrian Tracking

A visual tracker with the same moving camera platform and VPDAF is applied to track a pedestrian. We perform this experiment to verify the robustness of VPDAF in an outdoor environment. As shown in Fig. 7, the VPDAF succeeds in the outdoor tracking experiment, no matter how the light changes affect the captured images. Especially from frames (d) to (f), occlusion occurs on the tracked pedestrian so

that the target of interest disappears for a while, but tracking is maintained in this situation. The reason of success is explained as follows.

It is well known that the target detection process used, SAD template matching, is weak when it comes to tracking with occlusion. When occlusion happens, the reference template can not find similar measurements in the sensed image frame before the template is updated. However, this case is also considered in VPDAF by assuming event $\theta_0(k)$ in (4). The corresponding association probability $\beta_0(k)$ dominates the estimate. The updated state $\hat{\mathbf{x}}_0(k|k)$ in (2) increases the effect on the final weighted average $\hat{\mathbf{x}}(k|k)$ at the same time. The result is that the estimation only depends on the object motion model. Once the target of interest appears around this estimate again, the SAD template matching detects candidates, and the functions of VPDAF are recovered.

V. Conclusion

The VPDAF proposed in this paper is an extension of the traditional PDAF [1], which is designed mainly for target tracking from a cluster of measured data. It constructs the data-associated relationships of the measurements based on their positions. In the field of visual tracking, the measured data obtained through target detection in the image domain include not only position but also visual similarity information. Thus, in this paper, the significant visual characteristics are considered, while augmented association probabilities are derived. On the other hand, nontarget objects or some patterns in the background are major disturbances for visual target tracking. Fortunately, these persistent noises can be eliminated by classifying their motions. Through persistent noise removal, VPDAF generates modified data-association probabilities and outputs the final suboptimal target estimate.

VPDAF with template matching applied to target detection has been investigated through several experiments. The visual tracking system is constructed on either a fixed or 2-DOF camera platform. Reliable target-tracking performance has been achieved, even in a highly cluttered environment or one subjected to occasional occlusions. Although VPDAF takes a bit more time to evaluate the measurement variations and to filter out the stationary noise than the traditional PDAF, it still can be processed in about 30 ms for every image frame. Hence, real-time visual tracking is indeed achieved in our work.

References

[1] Y. Bar-Shalom and T. E. Fortmann, *Tracking and Data Association*. New York: Academic, 1988.

[2] H. M. Shertukde and Y. Bar-Shalom, "Tracking of crossing targets with imaging sensors," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 27, no. 4, pp. 582–592, Jul. 1991.

[3] A. Kumar, Y. Bar-Shalom, and E. Oron, "Precision tracking based on segmentation with optimal layering for imaging sensors," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 17, no. 2, pp. 182–188, Feb. 1995.

[4] T. Kirubarajan, Y. Bar-Shalom, and K. R. Pattipati, "Multiassignment for tracking a large number of overlapping objects and application to fibroblast cells," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 37, no. 1, pp. 2–21, Jan. 2001.

[5] Y. S. Yao and R. Chellappa, "Dynamic feature point tracking in an image sequence," in *Proc. 12th IAPR Int. Conf. Comput. Vis. Image Process.*, 1994, vol. 1, pp. 654–657.

[6] C. Rasmussen and G. D. Hager, "Probabilistic data association methods for tracking complex visual objects," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 23, no. 7, pp. 560–576, Jul. 2001.

[7] J. C. Nascimento and J. S. Marques, "Robust shape tracking in the presence of cluttered background," *IEEE Trans. Multimedia*, vol. 6, no. 6, pp. 852–861, Dec. 2004.

[8] D. Liu and L. C. Fu, "Target tracking in an environment of nearly stationary and biased clutter," in *Proc. IEEE/RSJ Int. Conf. Intell. Robots Syst.*, 2001, vol. 3, pp. 1358–1363.

[9] S. B. Colegrove and S. J. Davey, "PDAF with multiple clutter regions and target models," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 39, no. 1, pp. 110–124, Jan. 2003.

[10] M. Boshra and B. Bhanu, "Predicting performance of object recognition," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 22, no. 9, pp. 956–969, Sep. 2000.

[11] K. B. Sarachik, "The effect of Gaussian error in object recognition," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 19, no. 4, pp. 289–301, Apr. 1997.

[12] J. Chen, H. Leung, T. Lo, J. Litva, and M. Blanchette, "A modified probabilistic data association in a real clutter environment," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 32, no. 1, pp. 300–313, Jan. 1996.

[13] T. W. Ridler and S. Calvard, "Picture thresholding using an iterative selection method," *IEEE Trans. Syst., Man. Cybern.*, vol. SMC-8, pp. 630–632, 1978.

[14] Y. S. Chen, Y. P. Hung, and C. S. Fuh, "A fast block matching algorithm based on the winner-update strategy," in *Proc. 4th Asian Conf. Comput. Vis.*, 2000, vol. 2, pp. 977–982.

[15] Y. Bar-Shalom and X. R. Li, *Multitarget-Multisensor Tracking: Principles and Techniques*. Storrs, CT: Yaakov Bar-Shalom, 1995.

[16] G. Gennari and G. D. Hager, "Probabilistic data association methods in visual tracking of groups," in *Proc. Int. Conf. Comput. Vis. Pattern Recog.*, 2004, vol. 2, pp. II-876–II-881.

[17] C. Haworth, A. M. Peacock, and D. Renshaw, "Performance of reference block updating techniques when tracking with the block matching algorithm," in *Proc. IEEE Int. Conf. Image Process.*, 2001, vol. 1, pp. 365–368.

[18] A. Cavallaro, O. Steiger, and T. Ebrahimi, "Tracking video objects in cluttered background," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 15, no. 4, pp. 575–584, Apr. 2005.

Modeling and Control of a Small Autonomous Aircraft Having Two Tilting Rotors

Farid Kendoul, Isabelle Fantoni, and Rogelio Lozano

***Abstract*—This paper presents recent work concerning a small tiltrotor aircraft with a reduced number of rotors. The design consists of two propellers which can tilt laterally and longitudinally. A model of the full birotor dynamics is provided, and a controller based on the backstepping procedure is synthesized for the purposes of stabilization and trajectory tracking. The proposed control strategy has been tested in simulation.**

***Index Terms*—Backstepping methodology, helicopter dynamics modeling, tiltrotor rotorcraft, trajectory tracking.**

I. Introduction

Since the beginning of the 20th century, many research efforts have been done to create effective flying machines with improved performance and capabilities. The tiltrotor aircraft configuration [1] has the potential to revolutionize air transportation by providing an economical combination of vertical takeoff and landing (VTOL) capability with efficient high-speed cruise flight. Indeed, the Bell Eagle Eye unmanned

Manuscript received October 2, 2005; revised February 14, 2006. This paper was recommended for publication by Associate Editor G. Sukhatme and Editor F. Park upon evaluation of the reviewers' comments. This work was supported in part by the ONERA (French Aeronautics and Space Research Centre), in part by the DGA (French Arms Procurement Agency of the Ministry of Defence), and in part by the French Picardie Region Council. This paper was presented in part at the 44th IEEE Conference on Decision and Control/European Control Conference, Seville, Spain, December 2005. Color versions of Figs. 1–4 are available online at http://ieeexplore.org.

The authors are with Laboratoire Heudiasyc, UMR CNRS 6599, Université de Technologie de Compiègne, BP 20529–60205 Compiègne cedex, France (e-mail: fkendoul@hds.utc.fr; ifantoni@hds.utc.fr; rlozano@hds.utc.fr).

Digital Object Identifier 10.1109/TRO.2006.882956