Spatio-temporal target-measure association using an adaptive geometrical approach

Abir El Abed, Séverine Dubuisson *, Dominique Béréziant

Laboratoire d’Informatique de Paris 6, Université Pierre et Marie Curie, 4 place Jussieu, 75005 Paris, France

A R T I C L E   I N F O
Article history:
Received 30 July 2010
Available online 29 November 2011
Communicated by S. Sarkar

Keywords:
Data association
Target tracking
State and measure spaces

A B S T R A C T
Data association is of crucial importance to improve target tracking performance in many complex visual environments (non-linear dynamics, occlusions, etc.). Usually, association effectiveness is based on prior information and observation category. However, association becomes difficult if targets are similar. Problems also arise in cases of missing data, complex motions or deformations over time. To remedy, we propose a new method for data association, that uses the evolution of the dynamic model of targets. The main idea is to measure an adaptive geometric accuracy between possible trajectories of targets, by only using positions as information, that constitutes its main advantage.

© 2011 Elsevier B.V. All rights reserved.

1. Introduction
Traditionally, multiple target tracking consists in estimating states of moving targets from their associated available measures to get a track (Vermaak et al., 2005). The main difficulty is to determine over time the correct assignments of given measures to a target whose model is generally unknown. The literature contains some classical approaches to solve the problem of data association: we distinguish deterministic approaches from probabilistic ones.

Deterministic approaches determine the best associations by using a score function (Vermaak et al., 2005). The simplest method might be the nearest-neighbor standard filter (NNSF) (Rong and Bar-Shalom, 1996) that selects the closest valid measure to a predicted target and uses it for its state estimation. The proximity is generally given by the Mahalanobis distance. This filter has been used in many simple target tracking problems (Song et al., 2005; Konstantinova et al., 2003). But since it does not take into account the possibility of incorrect associations, its performance might be poor in some cases (especially for multi-target tracking, for example when two measures are equidistant from the same target or if target tracks cross) inducing association errors and degrading performance over time.

Probabilistic approaches are based on posterior probability estimation and make an association decision using a probability error (Rasmussen and Hager, 2001). Probabilistic data association filter (PDAF) (Cox, 1993; Singer et al., 1974) is an extension of the Kalman Filter (Bar-Shalom and Fortman, 1988) and an approximation of the optimal Bayesian filter. It considers two assumptions: (i) a measure is associated with a track or is a false alarm, and (ii) a track generates at most one measure. Here, a probability of association between the target and each measure is computed and used to weight the measures for the track update. Because of its simplicity, PDAF was used in target tracking problems (Jaward et al., 2006; Yilmaz et al., 2006), but its formulation has some limitations. First, it assumes that all measures come from the track being updated, that is not true in case of dense target conditions. Secondly, PDAF gives poor performances if the signal to noise ratio is small: in such cases, for example, the “good” measure can be located outside the validation region, and the filter cannot estimate the state of the target. Joint probabilistic data association filter (JPDAF) (Fortmann et al., 1983; Bar-Shalom and Li, 1995) is an extension of PDAF to the multiple target case (this number is known and constant): it works by running one PDAF per target, under the same hypothesis than PDAF, enforces an exclusion principle that prevents two or more trackers from latching onto the same target by calculating target-measure association probabilities jointly. JPDAF computes all possible joint events and their probabilities. Then, the algorithm determines to which of the marginal events this joint event contributes. JPDAF uses a weighted sum for all measures near the predicted state, each weight corresponding to the posterior probability for a measure to come from a target. JPDAF is widely used for target tracking because it deals with complex cases (Rasmussen and Hager, 2000; Chen et al., 2001) and provides an optimal solution in the Bayesian framework. However, the number of possible hypothesis increases exponentially with the number of targets, requiring prohibitive computation times, and tracks cannot be deleted or created. Multiple hypothesis tracking (MHT) was proposed by Reid (1979) for multiple target tracking. The number of targets is unknown, and varies over time. Reid’s algorithm defines
a systematic way in which multiple data association hypotheses are made and evaluated for the problem of multiple targets dealing with false alarms and cluttered background. When new data are available, each hypothesis is expanded into a set of new hypotheses by considering all association possibilities, that are propagated, implying the computation of each possible hypothesis, updated for each new available measure. To avoid an exponential increasing of the complexity of the algorithms, MHT eliminates less probable hypotheses, merges similar ones and keeps most probable. MHT is a recursive algorithm and easy to implement that consists in creating a tree: each hypothesis gives a path in this tree. This approach is very efficient because it considers all possible data association hypotheses, adds, removes or merges tracks, and in cases of conflicts, waits for more measures to decide for data association (Vermaak et al., 2005; Gelgon et al., 2005), and is then less sensible to false alarms. However, this exhaustive enumeration of hypotheses makes this approach very slow and costly. Probabilistic multiple hypothesis tracker (PMHT) (Gauvrit et al., 1997) supposes a target generates zero or several measures, and, contrary to the previous approaches, makes the posterior estimation of the state once all the measures are available, making it inappropriate for online tracking applications.

Usually, association effectiveness is based on prior information and available measures. In cases of lack of prior information, insufficient knowledge about the dynamics, large interval of time between two available observations, deformable or non distinguishable target (i.e. same color or same shapes), the association becomes difficult. Likewise, when only dealing with target positions, and if a measure is equidistant from several targets, all target association probabilities get close values, also making difficult the data association. No association method can efficiently handle all the cases previously illustrated.

In this paper, we propose a novel method for data association based on a geometric criteria, called energetic association filter (EAF). The idea is to define a term $E$, inspired by the evolution of the target dynamical model, from which, we deduce the probability of association between a given target and all available measures. The main advantages are it requires few parameters and prior knowledges: only information about target positions in the image plan are used. Besides, it handles the problem of association if a measure falls within the validation region of several targets and is equidistant from them. The outline of this paper is as follows. In Section 2, we introduce the notations used throughout the paper. In Section 3, we expose our approach for data association (EAF), derive its geometrical representation and its mathematical formulation. The proposed method is then tested on several sequences in Section 5. Finally, concluding remarks and perspective works are given in Section 6.

2. Definitions and notations

In this paper, $k = 1, \ldots, K$ is the discrete time, $l = 1, \ldots, L$ is the index associated to a measure and $m = 1, \ldots, M$ the index associated to a target. The set of targets at time $k$ is given by $\{x^0_1, \ldots, x^0_M\} \subseteq \mathcal{X}$, the state space. We mainly deal with predictions, because we focus on a missing data problem without knowledge about the real dynamical model: we then work into an optimal filtering framework, requiring prediction and correction steps: the prediction of target $x^0_m$ is denoted $\hat{x}^m_k$. At each time $k$, a sensor delivers zero or one observation containing at least one measure, giving the set $\{y^1_l, \ldots, y^c_l\} \subseteq \mathcal{Y}$, the measure space. Observations are supposed to be i.i.d. samples of a Gaussian distribution. A target is characterized by its own dynamic model. We then dispose $M$ models one for each target, plus one model for false alarms, and then have $M + 1$ tracks. In this paper, we also make the following assumptions (Bar-Shalom and Fortman, 1988):

1. A measure $y^c_l$ is associated to a track or considered as false alarm (false-track);
2. A track contains at most one measure at time $k$.

The global complexity for data association is $O(L \times (M + 1))$. To reduce the computation time, some algorithms introduce a validation region (Bar-Shalom and Fortman, 1988), that only contains measures with strong association probabilities, and exclude the others. This produces an ellipsoid (Cox, 1993) whose center is the predicted measure, and size and direction of axes are respectively given by the eigenvalues and by their associated eigenvectors of the covariance matrix. Then $y^c_l$ belongs to the validation region $V^c_m$ of predicted target $\hat{x}^m_k$ if:

$$[y^c_l - \hat{y}^c_m]^T (\Sigma^m_k)^{-1} [y^c_l - \hat{y}^c_m] \leq \tau^2$$

where $y^c_m$ is the predicted measure, defined by the projection of the predicted state $\hat{x}^m_k$ into the measure space (see Eq. (3)), $(y^c_l - \hat{y}^c_m)$ is the estimation error between the predicted measure $\hat{y}^c_m$ and the current measure $y^c_l$, and $\Sigma^m_k$ its covariance matrix (Therrien, 1989).

3. Energy association filter (EAF)

Generally, an effective data association method is based on measures. A sensor provides various measures such as shapes, colors, positions, directions, etc. Furthermore, if the measure is limited to the position, falls inside the validation region of several targets and is equidistant from them, an association ambiguity will occur if we use NNSF or JPDAF. Moreover, targets can be quite similar: even if information about the color or shape is available, the association task is difficult, even impossible in case of complex dynamics.

In this section, we describe our contribution for data association restricted to one category of measures: positions. Furthermore, it only needs few prior information concerning targets: exclusively the two anterior predicted positions are used as input for our algorithm. We first explain the concept of our approach before giving its mathematical formulation.

3.1. Concept

Prediction is made using a dynamic model which parameters are generally known or learned from a training sequence to represent plausible motions, such as constant velocity or damped oscillations (Blake and Isard, 1998; North et al., 2000). For complex dynamics, such as non-constant velocities or non-periodic oscillations, the choice of the parameters for the dynamic model is difficult. Furthermore, the learning step becomes impossible in cases of missing data (dynamics between two successive measures are unknown). For these reasons, the parameters of our dynamic model are set in an adaptive and automated way once measures are available (El Abed et al., 2006).

Predicted state $\hat{x}^m_{k-1}$ of target $m$ and its corresponding measure (i.e. after projection into the measure space) $\bar{y}^m_{k-1}$ at instant $k + 1$ are given by:

$$\hat{x}^m_{k-1} = f_{k-1}(\hat{x}^m_k) + V_{k-1}$$
$$\bar{y}^m_{k-1} = h(\hat{x}^m_{k-1}) + W_{k-1}$$

where $f_{k-1}$ models the dynamics of the state at time $t$ and $h$ is the projection of the state into the measure space. $V_{k-1}$ and $W_{k-1}$ are respectively the state and measure noises. Consider now an
available observation at instant $k$, containing one measure $y_k^m$, equivalent from several predicted measures. We suggest to determine the association probability at instant $k$ by predicting at instant $k + 1$ the measure associated with predicted target $x_{k+1}^m$ in two different ways.

Prediction “without association” (#1). The new state is predicted into the state space without taking into account any available measure (i.e. no measure comes from the target: we use Eq. (2)). The prediction of the state $\langle \hat{x}_{k+1}^m \rangle$ and its projection into the measure space $\langle \hat{y}_{k+1}^m \rangle$ are then given by:

$$\langle \hat{x}_{k+1}^m \rangle = f_{k+1}(\hat{x}_k^m) + V_{k+1}$$

$$\langle \hat{y}_{k+1}^m \rangle = h(\langle \hat{x}_{k+1}^m \rangle) + W_{k+1}$$  \hspace{1cm} (4)

Prediction “after association” (#2). The new measure $\langle \hat{y}_{k+1}^m \rangle$ is predicted into the measure space by associating the measure $y_k^m$ to the previous predicted measure (i.e. by taking into account the available measure). This prediction is given into the measure space by:

$$\langle \hat{y}_{k+1}^m \rangle = f_{k+1}(y_k^m) + V_{k+1}$$  \hspace{1cm} (5)

Our goal is to analyze the effect of an adaptive parametrization of the dynamic model of the state on its future prediction. The comparison of predictions “without association” and “after association” gives the influence of the available measure on the predicted track. In Fig. 1(b), we show an example of the predicted measure obtained with these two models at different times. At times $k - 1$ and $k - 2$, the same dynamic model (the initial one $f_{k+1}$) is used to predict the state of the target (see Eq. (4)). At time $k + 1$, two predictions are made: one without association, and one after association of measure $y_k^m$ with the target it comes from. During the interval $[k, k + 1]$, we can see the parametrization of the dynamic model (i.e. use of the current measure) significantly changes the predicted product. Our approach then consists in comparing these two trajectories to determine the measure-target association probability. In next subsections, we derive the mathematical model of EAF.

3.2. Data association algorithm

We define a term inspired from the evolution of the target’s dynamic model, described in terms of displacements in the state space. This term provides the measure-target association probability, and requires the predicted measures at times $k - 1$ and $k - 2$. We suppose only one sensor observes the dynamic scene, and can deliver at most one observation at each time, containing at least one measure. Each measure can be associated with a target, or be a false alarm. The score for the association of predicted measure $y_k^m$ with measure $y_k^c$ is given by a linear combination of three terms:

$$E(y_k^m, y_k^c) = \frac{1}{3} \sum_{i=1}^{3} a_i E(y_k^m, y_k^c)$$

$$x_i$$ is a weighting factor introduced to emphasize the relative importance attached to each term $E$ and is given by:

$$x_i = \frac{1}{\sum_{m=1}^{M} E(y_k^m, y_k^c)}$$

Each term deals with one category of movement into the observation space. First term $E^1$ only considers linear translations cases and corresponds to the Mahalanobis distance: it is inappropriate in cases of complex dynamics (non linear displacements, oscillatory motions, non-constant velocities, etc.). To remedy, we add a second term $E^2$ which measures the geometric accuracy between two dynamic models (i.e. “without association” and “after association”), and indicates their closeness. Finally, to distinguish some cases, such as occlusions between targets, we add a third term $E^3$, indicating the proximity of evolution. The global term $E$ is only computed for a measure that falls within several validation regions (measures not included in any validation region of targets are considered as clutters). In our case, the validation region corresponds to the one described in Section 2. We now detail subsequently the three parts of $E$.

- **Distance term $E^1$ (case of linear translations).**

$E^1$ corresponds to the Mahalanobis distance between the measure and its prediction at instant $k$, and is given by:

$$E^1(y_k^m, y_k^c) = \sqrt{(y_k^m - y_k^c)^T (\Sigma_k^m)^{-1} (y_k^m - y_k^c)}$$

where $y_k^m = h(\hat{x}_k^m) + W_k$, and $\Sigma_k^m$ is the covariance matrix. This distance leads to association errors if measures and predictions are equidistant, or in cases of crossed trajectories.

- **Evolution term $E^2$ (case of complex dynamics)**

To consider the case of complex dynamics, such as oscillatory motions or non-constant velocities, we add the evolution term $E^2(y_k^m, y_k^c)$. It introduces the notion of geometric accuracy between two sets of features whose evolutions are different. These dynamics are:
1. The current dynamic model, that predicts the measure \( (\hat{y}_m^{k+1})^{e_1} \) without considering the new available measure \( y_k \), i.e. without updating the parameters of the dynamics (Eq. (5));

2. The updated dynamic model, that considers the measure \( y_k \) is generated by \( m \)th target and updates the parameters of its dynamic model to predict the new measure \( (\hat{y}_m^{k+1})^{e_2} \) of the target (Eq. (6)).

\[ E^2(\hat{y}_m^{k}, y_k) \] measures the geometric closeness between the trajectories of two dynamic models. Our idea is to evaluate the parameters of the dynamic model in two cases: if the measure \( y_k \) comes from this target or not. We compute predicted measures \( (\hat{y}_m^{k+1})^{e_1} \) and \( (\hat{y}_m^{k+1})^{e_2} \), and then determine \( S_1 \), intersection surface between the two circumscribed circles of triangles \( \hat{y}_m^{k}, \hat{y}_m^{k+1}, \hat{y}_m^{k+2} \) and \( \hat{y}_m^{k-1}, \hat{y}_m^{k}, \hat{y}_m^{k+1} \), and \( S_2 \), intersection surface between the two circumscribed circles of triangles \( \hat{y}_m^{k}, \hat{y}_m^{k+1}, \hat{y}_m^{k+2} \) and \( \hat{y}_m^{k-1}, \hat{y}_m^{k}, \hat{y}_m^{k+1} \), see Fig. 2(a). \( E^2(\hat{y}_m^{k}, y_k) \) is given by:

\[ E^2(\hat{y}_m^{k}, y_k) = |S_1 - S_2| \tag{10} \]

By comparing these two sets we measure the ratio of similarity, defined by \( R_i = 1 - \frac{E^2(\hat{y}_m^{k}, y_k)}{E^2(\hat{y}_m^{k}, y_k)} \), between the predictions at \( k + 1 \) given by two different dynamic models for \( m \)th target. The similarity between dynamic models is then maximal if \( E^2(\hat{y}_m^{k}, y_k) \) is minimal (i.e. surfaces \( S_1 \) and \( S_2 \) are similar). Decreasing \( E^2(\hat{y}_m^{k}, y_k) \) means that prediction \( \hat{y}_m^{k} \) and measure \( y_k \) are close, so are the two predictions \( (\hat{y}_m^{k+1})^{e_1} \) and \( (\hat{y}_m^{k+1})^{e_2} \) and their associated dynamic models, showing \( y_k \) is generated by target \( x_k \). \( E^2(\hat{y}_m^{k}, y_k) \) is considered reliable if \( y_k \) and \( y_k \) are on the same side comparing to axis \( (\hat{y}_m^{k}, \hat{y}_m^{k+1}) \) (see left part of Fig. 2(b)). If \( y_k \) and \( y_k \) are diametrically opposite or if their positions are in opposite side comparing to the axis \( (\hat{y}_m^{k}, \hat{y}_m^{k+1}) \), both surfaces \( S_1 \) and \( S_2 \) are quite similar, which implies \( E^2(\hat{y}_m^{k}, y_k) \) to be close to zero (see right part of Fig. 2(b)). In such cases, corresponding to trajectory crosses, \( E^2(\hat{y}_m^{k}, y_k) \) is not sufficient to solve the association problem. We then incorporate a third term \( E^3(\hat{y}_m^{k}, y_k) \).

- **Evolution of proximity term** \( E^3 \)

The evolution of proximity term, \( E^3(\hat{y}_m^{k}, y_k) \), is the inverse of the surface \( S \) defined by the common area between triangles \( \hat{y}_m^{k-2}, \hat{y}_m^{k-1}, \hat{y}_m^{k} \) and \( \hat{y}_m^{k}, \hat{y}_m^{k+1}, \hat{y}_m^{k+2} \) (see Fig. 2(c), dashed area). This term evaluates the absolute accuracy between prediction \( \hat{y}_m^{k} \) and measure \( y_k \) at instant \( k \), and is given by:

\[ E^3(\hat{y}_m^{k}, y_k) = \frac{1}{S} \tag{11} \]

Increasing \( S \) means that \( \hat{y}_m^{k} \) and \( y_k \) are close at instant \( k \) (see Fig. 2(c)). In the right part of this figure, we now have two predictions at instant \( k \), \( (\hat{y}_m^{k+1})^{e_1} \) and \( (\hat{y}_m^{k+1})^{e_2} \), both equidistant from the

---

**Fig. 2.** Geometrical representation of \( E \) terms. (a) Intersection surfaces \( S_1 \) and \( S_2 \); (b) Differences between surfaces \( S_1 \) and \( S_2 \) extracted from the two dynamic models. (c) Left: intersection surface \( S \) (dashed area); right: case of two predicted measures \( (\hat{y}_m^{k+1})^{e_1} \) and \( (\hat{y}_m^{k+1})^{e_2} \) equidistant from \( y_k \).
observation \( y_l \). If we only use the distance to compute the proximity term, we get that both models have the same degree of similarity with the one defined by the dynamic model of points \( \{ y_{ml}^0, y_{ml}^1, y_{ml}^2 \} \). This leads to a contradiction with the reality, that can be explained by the fact that if they have both the same degree of similarity with the second dynamic model, their corresponding targets have then the same dynamics. For this reason, we have chosen to evaluate the similarity by measuring the intersection surface between triangles extracted from previous prediction, to add an information about the previous similarity of the models.

### 3.3 Measure-target posterior association probability

The posterior association probability between the available measure \( y_l \) and the predicted measure \( \hat{y}_m \) for target state \( x_m^0 \) is given by:

\[
\beta_{ml} = 1 - E(\hat{y}_m^0, y_l) \tag{12}
\]

The posterior association probability that target state \( x_m^0 \) is undetected is given by:

\[
\beta_{m0} = 1 - \frac{1}{L} \sum_{l=1}^{L} \beta_{ml} \tag{13}
\]

EA F algorithm is given in Algorithm 1.

#### Algorithm 1: Algorithm of Energetic Association Filter (EAF)

1. Input: \( \hat{y}_m^0, \ m = 1, . . . , M \) (offline preprocessing)
2. for \( k = 2 \) to \( K \) do
3.   foreach predicted target \( \hat{y}_m^0, \ m = 1, . . . , M \) do
4.     Compute predicted measure \( \hat{y}_m^0 \) of target:
5.     \[ \hat{y}_m^0 = h(\hat{x}_m^0) + W_k \]
6.     Compute state prediction “without association” (Eq. (4)):
7.     \[ (\hat{y}_{m1}^0)_{\text{#1}} = f_{k+1}(\hat{y}_m^0); V_{k+1} \]
8.     Compute predicted measures “without association” (Eq. (5)) and “after association” (Eq. (6)), respectively by:
9.     \[ (\hat{y}_{m1}^0)_{\text{#1}} = h(\hat{x}_{m1}^0) + W_{k+1}; \]
10. \[ (\hat{y}_{m1}^0)_{\text{#2}} = f_{k+1}(\hat{y}_m^0); V_{k+1} \]
11. foreach valid measure \( \hat{y}_l^0 \) (see Eq. (11)), \( l = 1, . . . , L \) do
12.     Compute distance term:
13. \[ E(\hat{y}_l^0, \hat{y}_m^0) = \sqrt{\left( \hat{y}_l^0 - \hat{y}_m^0 \right)^T \left( \hat{y}_m^0 - \hat{y}_m^0 \right)} \]
14.     Compute evolution term:
15. \[ E^2(\hat{y}_l^0, \hat{y}_m^0) = |S_l - S_m| \]
16.     Compute evolution of proximity term:
17. \[ E^0(\hat{y}_l^0, \hat{y}_m^0) = \frac{1}{2} \]
18.     Compute global score:
19. \[ E(\hat{y}_l^0, \hat{y}_m^0) = \frac{1}{3} \sum_{i=1}^{3} \alpha_i E_i(\hat{y}_l^0, \hat{y}_m^0) \] with \( \alpha_0 = \sum_{m=1}^{M} \frac{1}{E(\hat{y}_l^0, \hat{y}_m^0)} \)
20. Compute posterior association probability of data association between measures \( \hat{y}_l^0 \) and target \( x_m^0 \): \( \beta_{ml} = 1 - E(\hat{y}_l^0, \hat{y}_m^0) \)
21. \[ \beta_{m0} = \max_{l \neq m} \beta_{ml} \]
22. \[ \beta_{m0} \] is associated with predicted target \( \hat{y}_m^0 \)

We have described a novel approach for data association based on a geometric criteria to define the global score \( E(\hat{y}_l^0, \hat{y}_m^0) \) between an available measure \( y_l \) and the predicted measure \( \hat{y}_m^0 \) of target state \( x_m^0 \). The three components of \( E \) are extracted from geometrical representations (areas and distances) constructed with measures, previous states and prediction positions. Once this score is computed, we deduce the posterior association probability for all possible targets to measure associations. Finally, EAF associates measure \( l \) to target \( m \) if the posterior association probability \( \beta_{ml} \) is maximal. The purpose of choosing a geometrical definition for these terms refers to:

- Show the geometrical continuity of the system between predictions and previous states using two different dynamic models;
- Measure the similarity between predictions, at a particular time for the same target, using two different dynamic models, that logically must be quite same because they represent the same system.

### 4. Problem modeling for applications

For position prediction, we usually use a dynamic model describing the state process over time. For regular motions, the parameters of this model are often learned (North et al., 2000). For complex motions or online applications, the choice of parameters is difficult, especially in a missing data context. For this reason, an adaptive parametrization of the dynamic model is important when a new observation is available.

We consider a rectangular and planar observation region \( \mathcal{R} \subset \mathbb{R}^2 \). The state vector of target state \( x_m^0 \) contains coordinates in the image domain: \( x_m^0 = [r_m^0, \theta_m^0] \). Dynamical models \( f_k \) and \( f_k^{\text{ref}} \) (see respectively Eq. (4) and (6)) are cubic B-splines that extrapolate new predictions, fitted using previous predictions as control points (see Fig. 1(b)), allowing for example to deal with direction changes or non constant accelerations. \( x_m^0 \) is the predicted state of target \( m \) at instant \( k \).

Measures are also positions, available regularly or not, noisy and detected according to a probability \( p < 1 \). They can then also be false alarms, uniformly distributed in \( \mathcal{R} \). The observation model is given by:

\[
\hat{y}_l^0 = \begin{cases} h(\hat{x}_m^0) + W_k & \text{if measure } l \text{ comes from target } m \\ u_k & \text{otherwise} \end{cases} \tag{14}
\]

with \( W_k \) a Gaussian noise and \( u_k \) the random process of the false alarms. Because both predicted state and measure are positions, we have \( h(\hat{x}_m^0) = \hat{x}_m^0 \). We fixed \( \gamma = 5.99 \) as the threshold for the Mahalanobis distance (see Eq. 1) (Cox, 1993). We call \( \delta_k \) the time between two successive observations.

### 5. Experimental results

Algorithms have been implemented in Matlab on a PC with a 2.2 GHz Intel processor. We first validate our approach on a synthetic example in Section 5.1. Tests presented in Sections 5.2 and 5.3 concern the comparison of our filter (EAF, see Section 3) with PDAF and JPDAF (see Section 1) on real video sequences. For these tests, performance tools must be defined. The complexity of the data association problem depends on two parameters: (i) the false alarm probability \( p_f \), and (ii) the detection probability \( p_d \). Increasing the probability of false alarm \( p_f \), increasing the number of targets to track, or decreasing the probability of detection makes the association problem more complex, especially if targets move close from each other and if their trajectories cross. Performances are evaluated using the following terms:

1. The correct association probability (CAP):

\[
\text{CAP} = \frac{\# \text{correct associations}}{\# \text{possible associations}}
\]
2. The ratio of incorrect association probability to correct association probability (IAP/CAP):

$$\frac{\text{IAP}}{\text{CAP}} = \frac{\text{#incorrect associations}}{\text{#correct associations}}.$$ 

An association filter should be more accurate if IAP/CAP $\rightarrow 0$.

3. The run time.

For each test, we propose to make three experiments:

- **Experiment I.** Variation of the false alarm probability $p_f$ from 0.1 to 0.9.
- **Experiment II.** Variation of the detection probability $p$ from 0.1 (target missed) to 0.9 (detected target). In these experiments, all targets are supposed to have the same detection probability.
- **Experiment III.** Variation of $\delta_k$ between two observations. Increasing $\delta_k$ makes the data association more difficult, especially in case of erratic motions.

5.1. Synthetic test

We first consider the 1D synthetic example of Fig. 3(a), that explores the case of oscillatory motion with a constant phase. We have two targets $T_1 = x_1^k$ and $T_2 = x_2^k$ whose real dynamic models are known and defined by $x_1^k = \sin(k)$ and $x_2^k = \sin(2k) + \frac{1}{2}$. We consider a specific instant $k$, when the measure $y_k$ (full square in Fig. 3(a)), is equidistant from both predictions $\hat{x}_1^k$ and $\hat{x}_2^k$ of respectively targets $T_1$ and $T_2$ and falls into their validation regions. In such case, both targets are candidates to be associated with this measure. We compute the three terms for each target (see Table 1).
and obtain $E(y_1^k, y_k) < E(y_2^k, y_k)$: the observation is correctly associated with $T_1$.

5.2. Aperiodic oscillatory motion: “Tennis” sequence

In this test, we have $M = 2$ targets (the ball and the racquet), whose initial positions are known. Their real trajectories are given

Table 1
Terms computed for the two targets $T_1$ and $T_2$. The measure $y_k$ is associated with $T_1$ because $E(y_1^k, y_k)$ is minimal.

<table>
<thead>
<tr>
<th>$m$</th>
<th>$z_1E(y_1^k, y_k)$</th>
<th>$z_2E(y_2^k, y_k)$</th>
<th>$z_3E(y_3^k, y_k)$</th>
<th>$E(y_1^0, y_k)$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$y_1^k$</td>
<td>0.5</td>
<td>0.0001</td>
<td>0.4821</td>
<td>0.3274</td>
</tr>
<tr>
<td>$y_2^k$</td>
<td>0.5</td>
<td>0.9999</td>
<td>0.5179</td>
<td>0.6724</td>
</tr>
</tbody>
</table>

Fig. 4. “Tennis” sequence: comparative tests between EAF, PDAF et JPDAF. Left column, Experiment I; Right column, Experiment II; From top to bottom: CAP depending on $p_f$ and $p$, IAP/CAP ratio and run time in seconds, depending on $p_f$ and $p$.

Table 2
Comparative results of online performances for the three tested filters, depending on $d_k$ (interval between two observation acquisitions), with $p = 0.8$ and $p_f = 0.3$. Best values are written in bold font.

<table>
<thead>
<tr>
<th>$d_k$</th>
<th>0.1</th>
<th>0.2</th>
<th>0.5</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>EAF</td>
<td>JPD</td>
<td>PDA</td>
</tr>
<tr>
<td>CAP</td>
<td>0.92</td>
<td>0.8</td>
<td>0.7</td>
</tr>
<tr>
<td>IAP/CAP</td>
<td>0.99</td>
<td>0.25</td>
<td>0.429</td>
</tr>
<tr>
<td>Times (s.)</td>
<td>26</td>
<td>25.2</td>
<td>9</td>
</tr>
</tbody>
</table>
on Fig. 3(c). Their motion is complex, non linear and undergoes vertical and horizontal oscillations with several periods coupled with translation in two directions: it is difficult to learn the motion from a training set. Moreover, centers of the targets are very close in the region $[150,160] \times [50,70]$. For the experiments, $\mathcal{R} = [132,175] \times [20,180]$ is the observation region. The state noise is Gaussian, and its covariance matrix is $0.5^2 0 \quad 0.1^2$. Measures are available at regular times, and $\delta_t = 0.1$. The covariance matrix of the measure noise is $0.5^2 0 \quad 0.5^2$. We suppose each observation is made of four measures, from which at least two are false alarms. Fig. 3(c) shows the measures coming from the targets (crosses) and the false alarms (triangles) during all the test.

**Experiment I.** For this test, $p = 0.9$ for the ball, and $p = 0.7$ for the racquet: targets are not systematically detected. Left column of Fig. 4 shows performances of the three tested filters.

1. For $p_f = 0.1$. CAP > 0.8 for all approaches. JPDAF value is greater than PDAF value, because this filter uses all measures to update the state of targets. EAF value is the greatest one, almost equal to 1. This is due to the use of an adaptive dynamic model, that permits to deal with the aperiodic oscillatory movements of the targets to give a better data association probability.

2. For $p_f \rightarrow 0.9$. If the ambiguity about measure associations increases, performances of PDAF and JPDAF decrease faster than EAF’s one, that keeps good results even for high values of $p_f$ (CAP < 0.75, 1 for EAF, and < 0.4 for the two other filters).

3. About ratio IAP/CAP. For small $p_f$ values, IAP/CAP < 0.3. On the other hand, for greater values, IAP/CAP curves of PDAF and JPDAF increase very quickly: for example IAP/CAP > 1 for $p_f = 0.75$. However, EAF better detects false alarms: IAP/CAP increases slowly, and stays above 0.4 even for high values of $p_f$.

4. About run time. Due to its simplicity, PDAF is the less time consuming. EAF requires more computation time because it updates the parameters of the dynamic model for each available measure. This time is however acceptable, varying between 10 and 22 s for this experiment.

**Experiment II.** We suppose the probability for a target to be detected is the same for both targets ($p_f = 0.3$). The other parameters keep the same values than for Experiment I. Right column of Fig. 4 shows the performances of the three filters as a function of $p$.

1. For $p = 0.1$. CAP < 0.2 for PDAF and JPDAF, because these filters associate false alarms to targets. For EAF, we have CAP ≈ 0.8, showing our filter better deals with false alarms.

2. For $p \rightarrow 0.9$. The ambiguity decreases with the number of false alarms. When the maximal value is reached ($p = 0.9$), CAP value for JPDAF is greater than the one of PDAF. But EAF value it the greatest one, with CAP = 0.92: the performance of our filter is better even when targets are well detected.

3. About ratio IAP/CAP. For $p = 0.9$, IAP/CAP > 0.5: the minimal value for this ratio is obtained with our filter. When $p$ decreases, IAP/CAP increases quickly for PDAF and JPDAF, contrary to EAF. For a small value of $p$, IAP/CAP is always lower than 0.5. We obtain a better performance with our filter because it better takes into account the movement of targets, providing a better association probability.

4. About run time. PDAF is the fastest filter, because it requires less computations to associate measures to targets. EAF and JPDAF give similar computation times.

**Experiment III.** We test three values for $\delta_h$ (0.1, 0.2 and 0.5). For both targets, $p = 0.8$ and $p_f = 0.3$. Table 2 shows comparative

<table>
<thead>
<tr>
<th>$y_1^1$</th>
<th>$y_1^2$</th>
<th>$y_1^3$</th>
<th>$y_1^4$</th>
<th>$y_1^5$</th>
<th>$y_1^6$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$y_2^1$</td>
<td>0.07</td>
<td>0.47</td>
<td>0.26</td>
<td>0.44</td>
<td>0.49</td>
</tr>
<tr>
<td>$y_2^2$</td>
<td>0.01</td>
<td>0.02</td>
<td>0.01</td>
<td>0.14</td>
<td>0.47</td>
</tr>
<tr>
<td>$y_2^3$</td>
<td>0.01</td>
<td>0.45</td>
<td>0.01</td>
<td>0.09</td>
<td>0.2</td>
</tr>
<tr>
<td>$y_2^4$</td>
<td>0.03</td>
<td>0.31</td>
<td>0.01</td>
<td>0.09</td>
<td>0.2</td>
</tr>
<tr>
<td>$y_2^5$</td>
<td>0.23</td>
<td>0.06</td>
<td>0.01</td>
<td>0.01</td>
<td>0.01</td>
</tr>
<tr>
<td>$y_2^6$</td>
<td>0.03</td>
<td>0.01</td>
<td>0.01</td>
<td>0.01</td>
<td>0.15</td>
</tr>
<tr>
<td>$y_3^1$</td>
<td>0.12</td>
<td>0.03</td>
<td>0.01</td>
<td>0.01</td>
<td>0.15</td>
</tr>
<tr>
<td>$y_3^2$</td>
<td>0.12</td>
<td>0.08</td>
<td>0.02</td>
<td>0.06</td>
<td>0.05</td>
</tr>
<tr>
<td>$y_3^3$</td>
<td>0.17</td>
<td>0.11</td>
<td>0.02</td>
<td>0.06</td>
<td>0.05</td>
</tr>
<tr>
<td>$y_3^4$</td>
<td>0.21</td>
<td>0.02</td>
<td>0.02</td>
<td>0.09</td>
<td>0.05</td>
</tr>
<tr>
<td>$y_3^5$</td>
<td>0.07</td>
<td>0.02</td>
<td>0.03</td>
<td>0.01</td>
<td>0.24</td>
</tr>
<tr>
<td>$y_3^6$</td>
<td>0.39</td>
<td>0.15</td>
<td>0.28</td>
<td>0.01</td>
<td>0.18</td>
</tr>
<tr>
<td>$y_4^1$</td>
<td>0.22</td>
<td>0.06</td>
<td>0.18</td>
<td>0.16</td>
<td>0.16</td>
</tr>
<tr>
<td>$y_4^2$</td>
<td>0.18</td>
<td>0.06</td>
<td>0.12</td>
<td>0.09</td>
<td>0.06</td>
</tr>
<tr>
<td>$y_4^3$</td>
<td>0.74</td>
<td>0.93</td>
<td>0.93</td>
<td>0.07</td>
<td>0.92</td>
</tr>
<tr>
<td>$y_4^4$</td>
<td>0.33</td>
<td>0.4</td>
<td>0.36</td>
<td>0.06</td>
<td>0.35</td>
</tr>
<tr>
<td>$y_4^5$</td>
<td>0.33</td>
<td>0.4</td>
<td>0.36</td>
<td>0.36</td>
<td>0.07</td>
</tr>
<tr>
<td>$y_4^6$</td>
<td>0.16</td>
<td>0.14</td>
<td>0.09</td>
<td>0.11</td>
<td>0.10</td>
</tr>
<tr>
<td>$y_5^1$</td>
<td>0.01</td>
<td>0.01</td>
<td>0.01</td>
<td>0.02</td>
<td>0.01</td>
</tr>
<tr>
<td>$y_5^2$</td>
<td>0.23</td>
<td>0.2</td>
<td>0.32</td>
<td>0.75</td>
<td>0.11</td>
</tr>
<tr>
<td>$y_5^3$</td>
<td>0.13</td>
<td>0.12</td>
<td>0.14</td>
<td>0.29</td>
<td>0.08</td>
</tr>
</tbody>
</table>

Table 3

"Ant" sequence. Normalized evolution energies $x_i^E(y_i^t, y_i^t)$, $x_2^E(y_i^t, y_i^t)$, $x_1^E(y_i^t, y_i^t)$ and global energy $E(y_i^t, y_i^t)$ computation. Best values are written in bold font.
results of CAP, IAP/CAP and run times. EAF performs better: the increase of $d_k$ does not affect CAP, IAP/CAP and run time. The other filters quickly fail, mainly because of the aperiodic oscillatory motion of the two targets and their proximity into the region $[150,160] \times [60,80]$ (see Fig. 3(c)).

5.3. Interaction between targets and crossed trajectories, “Ant” sequence

This sequence shows six ants moving on a flat surface. Each ant has a non linear and erratic movement: translation while walking and deformations while moving its head, abdomen, or antennas (see trajectories in Fig. 3(e)). Ants frequently interact, sometimes four of them are close from each others and are quite similar even nondistinguishable, characterized by the same gray level distribution and shape: only their motion can be useful for the data association.

We first test the robustness of EAF. Images of Fig. 3(d) are observations at instants $k - 1, k - 1, k$ and $k + 1$, and respectively represent frames 10, 25, 35 and 45 (ants are labeled to show their real positions from a frame to another). There is a considerable interval of time between frames $k - 1$ and $k + 1$ ($d_k = 0.1$).

Table 3 gives the values of components $x_kE^1(y_k^{10}, y_k^1), x_kE^2(y_k^{10}, y_k^4)$ and $x_kE^3(y_k^{10}, y_k^4)$ computed between measures $y_k^i$ and the predicted measure $y_k^m$ of six predicted target state $x_m^i$: the first row represents the measure number and the first column the target number ($y_k^i$ really coming from measure $y_k^{10}$). If we associate measures to targets only using the Mahalanobis distance minimization, we associate $y_1^k$ to $y_1^m, y_2^k$ to $y_3^m, y_3^k$ to $y_2^m, y_4^k$ to $y_2^m$ and $y_5^k$ to $y_4^m$. This leads to two problems: only two measures are well associated, and two measures are associated to the same target, that is not possible under our assumptions. Global energies are also given in Table 3: we observe each measure is well associated with its corresponding target, that show the interest of our approach.

In this second part of our tests, the sensor, at instant $k$, provides an observation containing five measures (from targets or false alarm, see Eq. (14)): at least one ant is not detected (missing data).
6. Conclusion

This work proposes a new method for data association based on a geometric criteria. The developed approach can handle complex motions and highly non-linear systems, and deals with the lack of prior knowledge, by only using few parameters. The geometric illustration of energy components allows to measure the accuracy between two dynamic models and to define their degree of similarity. The three components are extracted from geometrical representations (areas and distances) constructed with measures, previous states and prediction positions. The purpose of choosing a geometrical definition for these terms refers measure the similarity between predictions, at a particular time for the same target, using two different dynamic models, that logically must be quite similar because they represent the same system.

As a perspective for this work, we suggest to integrate the energetic association filter within the classical particle filter to build a new framework for multiple target tracking. Moreover, since we consider erratic motions that cannot be learned from training sequences, we suggest to use an adaptive and automated way to set parameters of the dynamic model of the filter. Then, we could track targets under the restriction of the missing of prior information and especially if similar targets are evolving in the scene.

References


