Integration of Fuzzy Spatial Information in Tracking Based on Particle Filtering

Nicolas Widynski, Séverine Dubuisson, *Member, IEEE*, and Isabelle Bloch, *Member, IEEE*

***Abstract*—In this paper, we propose a novel method to introduce spatial information in particle filters. This information may be expressed as spatial relations (orientation, distance, etc.), velocity, scaling, or shape information. Spatial information is modeled in a generic fuzzy-set framework. The fuzzy models are then introduced in the particle filter and automatically define transition and prior spatial distributions. We also propose an efficient importance distribution to produce relevant particles, which is dedicated to the proposed fuzzy framework. The fuzzy modeling provides flexibility both in the semantics of information and in the transitions from one instant to another one. This allows one to take into account situations where a tracked object changes its direction in a quite abrupt way and where poor prior information on dynamics is available, as demonstrated on synthetic data. As an illustration, two tests on real video sequences are performed in this paper. The first one concerns a classical tracking problem and shows that our approach efficiently tracks objects with complex and unknown dynamics, outperforming classical filtering techniques while using only a small number of particles. In the second experiment, we show the flexibility of our approach for modeling: Fuzzy shapes are modeled in a generic way and allow the tracking of objects with changing shape.**

***Index Terms*—Fuzzy-shape tracking, fuzzy spatial information, particle filter.**

I. Introduction

OBJECT TRACKING in video sequences has been widely addressed in the literature using a large variety of approaches [26]. Particle filters are more and more developed in this domain [9], [10]. Introducing high-level content, such as shape or spatial relations, and dealing with complex object dynamics are two main issues of tracking based on particle filtering which are addressed in this paper and for which we propose a unified framework.

While spatial relations are usually not explicitly used in particle-filter tracking, shape tracking for sequential Monte Carlo methods has been first introduced in [14], where the dynamics of the shape is learned using a second-order autoregressive process. More recently, the authors in [7] proposed to guide the particle-filter shape tracking using appearance models from a learning base and then distorted using affine transformations. A main drawback of these approaches is the use of a poor shape dynamics model that cannot handle local shape deformations. Hence, in [24], the authors introduced learned shapes represented by level sets and, for the tracking procedure, proposed new shapes by performing a gradient descent on the energy associated with the shape of the previous instant. However, this approach is computationally expensive since each particle is connected to a level set. Moreover, the size of the database must be important to handle most of the possible variations of the shape.

In practice, the classical use of a poor informative dynamics prior often leads to tracking errors. To avoid this problem, several approaches have been proposed in the literature. First, the optimal importance distribution guarantees the minimal weight variance, under assumptions of linear or partial linear Gaussian models [1], [10]. The common alternative choice consists in generating particles according to an efficient importance distribution. The Auxiliary Particle Filter (AUX-PF) [23] makes use of the current observation to select particles which may have the largest likelihood. When the observation model can be inverted, one can use the likelihood particle filter [2], which simulates particles according to the likelihood. Other importance distributions have also been proposed (see [8] for a review). On the other hand, a way to deal with dynamics involving strong discontinuities consists in using multiple dynamical models. For this purpose, the Interacting Multiple Model was introduced into the particle filter, the so-called Multiple Model Particle Filter (MMPF) [19]. When the process noise is high and despite an additional cost of the tracking algorithm, one can use the Auxiliary MMPF (AUX-MMPF) [17], which is an adaptation of the AUX-PF to the MMPF framework. However, none of these approaches makes use of high-level spatial information.

Our approach consists in defining fuzzy sets modeling this type of information and introducing them into the particle filter. The fuzzy sets can be fixed or learned and are generated using fuzzy fusion strategies. This representation is generic, since the probabilistic model is automatically deduced from the fuzzy framework and can handle different kinds of characteristics without using a restrictive mathematical modeling. Hence, we can, for example, easily introduce structural or fuzzy-shape information into the particle filter.

Technically, the fuzzy sets are integrated into the particle-filtering framework as fuzzy partitions on the parameter space of the unknown state. The proposed model then switches between the predefined fuzzy sets. Since it adds to the estimation

Manuscript received January 25, 2010; revised May 10, 2010; accepted July 9, 2010. Date of current version May 18, 2011. This work was supported in part by the Direction générale de l'Armement. This paper was recommended by Associate Editor J. Su.

N. Widynski is with the Laboratory of Computer Sciences (UPMC-LIP6), University Pierre and Marie Curie (Paris 6), 75005 Paris, France, and also with the Télécom ParisTech, Centre National de la Recherche Scientifique Laboratoire Traitement et Communication de l'Information (CNRS LTCI), 75013 Paris, France (e-mail: Nicolas.Widynski@telecom-paristech.fr).

S. Dubuisson is with the Laboratory of Computer Sciences (UPMC-LIP6), University Pierre and Marie Curie (Paris 6), 75005 Paris, France (e-mail: Severine.Dubuisson@lip6.fr).

I. Bloch is with the CNRS LTCI, Télécom ParisTech, 75013 Paris, France (e-mail: Isabelle.Bloch@telecom-paristech.fr).

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TSMCB.2010.2064767

a discrete random variable in the state space, which defines to which fuzzy set the parameter is related at a given time instant, the proposed model can be viewed as a switching dynamical model, as the ones proposed in [15] and [21] or the multiple-model approaches [17], [19]. However, the main difference between these methods and the proposed approach relies in the way the switching variable is used. In existing methods, the switch usually helps in deciding when to switch from one dynamical model to another one, while in the proposed approach, it is performed in the parameter space and allows switching from one fuzzy model to another one. Moreover, the fuzzy approach allows the automatic determination of the switch, which is also more gradual, leading to more flexibility.

We first give a brief description of particle filters for both state and parameter estimations in Section II. The fuzzy modeling is described in Section III, while its introduction in particle filters is detailed in Section IV. Results on tracking with simulated data, object tracking, and shape tracking are illustrated in Sections V–VII, respectively.

II. Bayesian Tracking and Parameter Modeling

Let us consider a classical filtering problem and denote by $\mathbf{x}_t \in \mathcal{X}$ the hidden state of a stochastic process at time t and by $\mathbf{y}_t \in \mathcal{Y}$ the measurement state. The temporal evolution of $\mathbf{x}_t$ and the measurement equation are given by

$$\begin{aligned} \mathbf{x}_t &= f_t(\mathbf{x}_{t-1}, \mathbf{v}_t) \\ \mathbf{y}_t &= h_t(\mathbf{x}_t, \mathbf{w}_t) \end{aligned} \tag{1}$$

where $\mathbf{v}_t$ and $\mathbf{w}_t$ are independent white noises. The nonlinear Bayesian tracking consists in estimating the posterior filtering distribution $p(\mathbf{x}_t|\mathbf{y}_{1:t})$ through a nonlinear transition function f_t and a nonlinear measurement function h_t.

There are different ways to model the unknown dynamics parameters $\mathbf{r}_t \in \mathcal{R}$ of the state $\mathbf{x}_t$, where $\mathcal{R}$ is the parameter space. A possible one is to decompose the dynamics vector parameter into R independent components $\mathbf{r}_t = (\mathbf{r}_t^1, \ldots, \mathbf{r}_t^R)$ and then to consider a random walk dynamics for each of them. This leads to the following system:

$$\begin{aligned} \mathbf{x}_t &= f_t(\mathbf{x}_{t-1}, \mathbf{r}_t, \mathbf{v}_t) \\ \mathbf{r}_t^r &= \mathbf{r}_{t-1}^r + \mathbf{s}_t^r \quad \forall r \in \{1, \ldots, R\} \\ \mathbf{y}_t &= h_t(\mathbf{x}_t, \mathbf{w}_t) \end{aligned} \tag{2}$$

where $\mathbf{s}_t = (\mathbf{s}_t^r)_{r=1,\ldots,R}$ are independent Gaussian noises. The independence and random walk dynamics assumptions are classical and allow the handling of most of the tracking problems. The resulting filtering distribution can be computed according to, first, a prediction step

$$\begin{aligned} p(\mathbf{x}_t, \mathbf{r}_t|\mathbf{y}_{1:t-1}) = \int_{\mathcal{X}} \int_{\mathcal{R}} & p(\mathbf{x}_t|\mathbf{x}_{t-1}, \mathbf{r}_t) p(\mathbf{r}_t|\mathbf{r}_{t-1}) \\ & \times p(\mathbf{x}_{t-1}, \mathbf{r}_{t-1}|\mathbf{y}_{1:t-1})\, d\mathbf{x}_{t-1}\, d\mathbf{r}_{t-1} \end{aligned} \tag{3}$$

and then a filtering step

$$p(\mathbf{x}_t, \mathbf{r}_t|\mathbf{y}_{1:t}) \propto p(\mathbf{y}_t|\mathbf{x}_t, \mathbf{r}_t)\, p(\mathbf{x}_t, \mathbf{r}_t|\mathbf{y}_{1:t-1}). \tag{4}$$

A closed form of the posterior distribution is available when the transition function of the state and the measurement equation are linear and their corresponding noise is Gaussian, yielding the Kalman filter. When not, particle filters, also known as sequential Monte Carlo methods, are used to approximate the posterior density function (*pdf*) by a weighted sum of N Dirac masses $\delta_{\mathbf{x}_t^{(n)}, \mathbf{r}_t^{(n)}}(d\mathbf{x}_t, d\mathbf{r}_t)$ centered on hypothetic state realizations $\{\mathbf{x}_t^{(n)}, \mathbf{r}_t^{(n)}\}$ of the state $(\mathbf{x}_t, \mathbf{r}_t)^{\mathrm{T}}$, also called particles. Then, the filtering distribution $\mathbb{P}(d\mathbf{x}_t, d\mathbf{r}_t|\mathbf{y}_{1:t})$ is recursively approximated by the empiric distribution $P_N(d\mathbf{x}_t, d\mathbf{r}_t|\mathbf{y}_{1:t}) = \sum_{n=1}^N w_t^{(n)} \delta_{\mathbf{x}_t^{(n)}, \mathbf{r}_t^{(n)}}(d\mathbf{x}_t, d\mathbf{r}_t)$, where $\{\mathbf{x}_t^{(n)}, \mathbf{r}_t^{(n)}\}$ is the nth particle and $w_t^{(n)}$ is its weight. If an approximation of $\mathbb{P}(d\mathbf{x}_{t-1}, d\mathbf{r}_{t-1}|\mathbf{y}_{1:t-1})$ is known, the process is divided into two main steps.

1) The diffusion step consists in estimating $p(\mathbf{x}_t, \mathbf{r}_t|\mathbf{y}_{1:t-1})$ by propagating the particle swarm $\{(\mathbf{x}_{t-1}^{(n)}, \mathbf{r}_{t-1}^{(n)}), w_{t-1}^{(n)}\}_{n=1}^N$ using an importance distribution $q(\mathbf{x}_t, \mathbf{r}_t|\mathbf{x}_{0:t-1}^{(n)}, \mathbf{r}_{0:t-1}^{(n)}, \mathbf{y}_t)$.
2) The update step then computes new particle weights using the new observation $\mathbf{y}_t$, as $w_t^{(n)} \propto w_{t-1}^{(n)} p(\mathbf{y}_t|\mathbf{x}_t^{(n)}, \mathbf{r}_t^{(n)}) p(\mathbf{x}_t^{(n)}, \mathbf{r}_t^{(n)}|\mathbf{x}_{t-1}^{(n)}, \mathbf{r}_{t-1}^{(n)}) / q(\mathbf{x}_t, \mathbf{r}_t|\mathbf{x}_{0:t-1}^{(n)}, \mathbf{r}_{0:t-1}^{(n)}, \mathbf{y}_t)$, such that $\sum_{i=1}^N w_t^{(n)} = 1$. Resampling techniques are then employed to avoid particle degeneracy problems.

The modeling of the dynamics parameters is of prime importance to successfully dealing with complex dynamics. Indeed, in object tracking, the true dynamics of the object is most of the time unknown and the tracking algorithm can fail. In this paper, we then make a weak assumption about dynamics parameters, i.e., we consider the prior pdf $p(\mathbf{r})$ as being uniform on the parameter space $\mathcal{R}$. We also introduce a new type of information, expressed as fuzzy sets. Considering parameters conditioned by fuzzy events, we can then restrain the parameter space under the considered fuzzy sets. This leads to an adaptive strategy, where the parameter sampling distribution only depends on which fuzzy set is considered. This modeling allows the definition of parameters in a generic way. For example, we applied our fuzzy information model to fuzzy spatial relations and fuzzy shapes. We also propose an efficient importance distribution dedicated to our model which is related to current observation and past realizations. This allows us to avoid local ambiguities and to deal with outliers while keeping hypotheses during time and new observation data to generate relevant samples according to the likelihood.

III. Fuzzy Information Modeling

The fuzzy-set theory allows the modeling of qualitative, vague, or imprecise knowledge often expressed in a linguistic form [12]. For example, when one wishes to make a poll, asking people how tall they are, it is easy to merge obtained answers with a chosen granularity: for instance, *Small*, *Medium*, and *Large*. Since these linguistic values are subjective and depend on origin, gender, and so on, it is very likely to think that the obtained classes, i.e., the fuzzy sets, own common values. Instead of considering probabilities, for instance, based on frequencies, fuzzy models allow the consideration of the set

of small persons as a fuzzy set, which better fits the idea of imprecision modeling.

Similarly, in image processing, and specifically in object tracking, it is possible to model several features or several behaviors using fuzzy representations. Hence, it is, for example, possible to describe an object volume by fuzzy sets, with a fixed granularity which changes the representation precision. In a similar way, the distance between positions of an object at two different instants may be characterized by the values *Small*, *Medium*, and *Large*, for which semantics are defined by fuzzy sets on $\mathbb{R}^+$. Here, again, the information to be modeled is deterministic but vague, and fuzzy models are then appropriate and very natural to express this type of imperfection. In the same way, we can introduce fuzzy-shape information, each fuzzy set corresponding to a posture for instance. In this section, we propose a general framework for modeling these types of information.

A. Fuzzy Information

We consider a collection of K fuzzy sets $\{\mathcal{A}_k\}_{k=1}^K$ defined by the membership functions $\{\mu_k\}_{k=1}^K$ such that $\forall k \in \{1,\dots,K\}$, $\mu_k : \mathcal{O} \to [0,1]$, $\mathcal{O} \subset \mathbb{R}^{n_\mathbf{o}}$, with $\mathcal{O}$ as the space of fuzzified elements and $n_\mathbf{o}$ as its dimension. We associate with each element $\mathbf{o} \in \mathcal{O}$ its degree of membership $\mu_k(\mathbf{o})$ to the fuzzy set $\mathcal{A}_k$. For example, if the set $\mathcal{A}_k$ defines impossible elements of a particular fuzzy concept, then $\mu_k(\mathbf{o})$ indicates to which degree $\mathbf{o}$ translates this impossibility. In the fuzzy-set terminology, the set of functions $\{\mu_k\}_{k=1}^K$ corresponds to a linguistic variable in which each function μ_k is a possible linguistic value of the variable. The granulometry (and the value of K) may be defined by the application. For instance, considering spatial orientations, linguistic values could be $\mathcal{A}_1 = East$, $\mathcal{A}_2 = North$, $\mathcal{A}_3 = West$, and $\mathcal{A}_4 = South$, defined by the membership functions $\{\mu_k\}_{k=1}^4$.

B. Fuzzy Functions

The fuzzy functions $\{\mu_k\}_{k=1}^K$ may be fixed, either by prior knowledge, hence linked to the application, or in a generic way, by learning. We now consider this last possibility. Let M^k be the number of observations owned by a class k that we denote $\{\mathbf{o}_m^k\}_{m=1}^{M^k}$, such that $\forall m \in \{1,\dots,M^k\}$, $\mathbf{o}_m^k \in \mathcal{O}$. The kth fuzzy function $\tilde{\mu}_k$ is then defined by a fuzzy combination of elements $\mathbf{o}_m^k$

$$\tilde{\mu}_k(\mathbf{o}) = \Phi^k\left(g_{\mathbf{o}_1^k}(\mathbf{o}),\dots,g_{\mathbf{o}_{M^k}^k}(\mathbf{o})\right) \tag{5}$$

with $\Phi^k : [0,1]^{M^k} \to [0,1]$ as a fuzzy fusion operator, detailed in the next section, and $g_{\mathbf{o}_m^k} : \mathcal{O} \to [0,1]$ as a function associated to $\mathbf{o}_m^k$

$$g_{\mathbf{o}_m^k}(\mathbf{o}) = \top\left[P_{\mathbf{o}_m^k}(\mathbf{o}), C_{\mathbf{o}_m^k}(\mathbf{o})\right] \tag{6}$$

with $\top$ as a t-norm (fuzzy conjunction) [12], $P_{\mathbf{o}_m^k} : \mathcal{O} \to [0,1]$ as a known and fixed fuzzy generic template, and $C_{\mathbf{o}_m^k} : \mathcal{O} \to [0,1]$ as an additional constraint on the shape of the final fuzzy function. The function $g_{\mathbf{o}_m^k}$ represents the fuzzification of a crisp observation $\mathbf{o}_m^k$, under a generic template $P_{\mathbf{o}_m^k}$ constrained by $C_{\mathbf{o}_m^k}$, which guarantees the global consistency of $g_{\mathbf{o}_m^k}$, with respect to $\mathbf{o}_m^k$.

These functions may be known *a priori* or learned, as we will see in the application described in Section VII. If there is no constraint, $C_{\mathbf{o}_m^k}$ corresponds to the neutral element of the t-norm, i.e., the value 1. The fuzzy template $P_{\mathbf{o}_m^k}$ can, for example, be set, if $\mathcal{O} \subset \mathbb{R}^L$, to a fuzzy ball of dimension L

$$P_{\mathbf{o}_m^k}(\mathbf{o}) = \begin{cases} 1, & \text{if } \|\mathbf{o}-\mathbf{o}_m^k\|_2 \in [0,b] \\ 1-\frac{\|\mathbf{o}-\mathbf{o}_m^k\|_2-b}{a}, & \text{if } \|\mathbf{o}-\mathbf{o}_m^k\|_2 \in [b;b+a] \\ 0, & \text{otherwise} \end{cases}$$

where $2(a+b)$ and $2b$ are the lengths of the support and the kernel, respectively. Finally, we define the final fuzzy function μ_k as a combination of functions $\tilde{\mu}_k$ and $\{\tilde{\mu}_h\}_{\setminus k} \triangleq \{\tilde{\mu}_h\}_{h\in\{1,\dots,K\}\setminus\{k\}}$ by a fuzzy fusion operator $\Psi^k : [0,1]^K \to [0,1]$

$$\mu_k(\mathbf{o}) = \Psi^k\left(\tilde{\mu}_k(\mathbf{o}), \{\tilde{\mu}_h(\mathbf{o})\}_{\setminus k}\right). \tag{7}$$

The function μ_k integrates both information of the inner class k, under an intraclass operator Φ^k, and information shared by all the classes, represented by the interclass operator Ψ^k.

It is then possible to define interactions between the fuzzy sets by calculating an intersection degree constrained by a minimum threshold. Let $R(k,h)$ be the interaction degree between the sets μ_k and μ_h

$$R(k,h) = \perp\left[\Xi^{kh}(\mu_k,\mu_h), \epsilon^{kh}\right] \tag{8}$$

with $\perp$ as a t-conorm (fuzzy disjunction) [12], $\Xi^{kh}(\mu_k,\mu_h)$ as the degree of intersection of the fuzzy sets μ_k and μ_h, and ϵ^{kh} as the minimal intersection threshold of these two sets, known and fixed. There are many operators computing an intersection degree of two fuzzy sets, and the reader may refer to [5] for a description of some of them.

C. Fusion Strategies

As we just noticed, the fuzzy sets $\{\mu_k\}_{k=1}^K$ are obtained by combining different fusion strategies, owing to the operators $\{\Phi^k\}_{k=1}^K$ and $\{\Psi^k\}_{k=1}^K$. The first one is an intraclass fusion operator, since it refers to a fusion of observations owned by only one class, whereas the second one makes a fusion of data conditioned by the other classes and then corresponds to an interclass fusion operator. It is hence possible to decline a nonexhaustive list of these fusion strategies (the reader may refer to [4] for a larger description of fusion operators).

Intraclass Strategies: Φ:

1) Disjunctive fusion using t-conorms: Each observation represents a partial vision of the possible world, and hence, their association defines the set. Then, $\mu_k(\mathbf{o})$ represents the degree to which $\mathbf{o}$ looks like at least one of these observations.
2) Conjunctive fusion using t-norms: It is of interest in at least two situations. When we have multiple observations of a unique phenomenon, each one marked by imprecision, we may want to dilate the corresponding fuzzy representations to include the truth and avoid conflicts. In this case, the conjunction just holds the common parts, given by all the observations, and hence allows the reduction of imprecision. Another situation occurs when

observations do not mean the possible values but the constraints (for instance, forbidden or impossible values). In this case, one may look for a conjunction of the complement of the constraints. With a conjunctive fusion, $\mu_k(\mathbf{o})$ measures the adequation between $\mathbf{o}$ and the set of the observations.

3) A compromised fusion using a sum or a mean: It allows a compensation between observations considered, for example, as an evaluation of a situation, in which each measure could be imperfect. It is then possible to average them to compensate the imperfection of each one. With an additive operator, one obtains a frequentist approach and reaches statistical estimation questions.

Interclass Strategies: Ψ*:*

1) Exclusive fusion: If an element has a membership shared by several classes, we consider that there is a semantic impossibility. This can be used for a binary state modeling, for which intermediate values do not exist.
2) Fusion by normalization: The membership value of each element to a class can be normalized by the sum of the membership values to all the classes.
3) None: It is the default behavior; each class is considered independently from each other.

These fusion strategies define a generic way to apply our model to several kinds of problems. It is then possible to combine these different strategies, leading to a particular representation, by defining those properties or their semantics. Examples will be given in the applications in Sections VI and VII.

IV. Spatial Information and Particle Filter

A. Introducing Spatial Information in a Probabilistic Context

The general idea behind the integration of fuzzy sets in a probabilistic context is to consider these sets as fuzzy events and to compute their probabilities under a particular probability measure $\mathbb{P}$. Let $\mathcal{A}$ be a fuzzy set defined by a measurable membership function μ_A. According to [28], the probability of a fuzzy event $\mathcal{A}$ is defined as

$$\mathbb{P}(\mathcal{A}) = \mathbb{E}[\mu_A] = \int_{\mathcal{O}} \mu_{\mathcal{A}}(\mathbf{o})\, d\mathbb{P}. \tag{9}$$

We then recall the expression of the conditional probability of a fuzzy event $\mathcal{A}$ of membership function $\mu_{\mathcal{A}}$ given a fuzzy event $\mathcal{B}$ of membership function $\mu_{\mathcal{B}}$

$$\begin{aligned}\mathbb{P}(\mathcal{A}|\mathcal{B}) &= \frac{\mathbb{P}(\mathcal{A}\mathcal{B})}{\mathbb{P}(\mathcal{A})} \\ &= \frac{\int_{\mathcal{O}} \mu_{\mathcal{A}}(\mathbf{o})\mu_{\mathcal{B}}(\mathbf{o})\, d\mathbb{P}}{\int_{\mathcal{O}} \mu_{\mathcal{A}}(\mathbf{o})\, d\mathbb{P}}\end{aligned} \tag{10}$$

where $\mathbb{P}(\mathcal{A}\mathcal{B})$ is the probability of the event $\mathcal{A}\mathcal{B}$ (product of $\mathcal{A}$ and $\mathcal{B}$) defined by the membership function $\mu_{\mathcal{A}}\mu_{\mathcal{B}}$.

B. Introducing Spatial Information in a Particle Filter

Let $\mathbf{r}$ be a continuous random variable representing the state parameter taking values in $\mathcal{R}$ and $d\mathbf{r}$ be a fuzzy event of infinitesimal support. The membership function $\mu_{d\mathbf{r}}(\mathbf{r})$ is defined as $\mu_{d\mathbf{r}}(\mathbf{r}) = 1$ if $\mathbf{r} \in [\mathbf{r} - (d\mathbf{r}/2), \mathbf{r} + (d\mathbf{r}/2)]$ and 0 otherwise. We consider a collection of K fuzzy sets $\{\mathcal{A}_k\}_{k=1}^K$ defined by the membership functions $\{\mu_k\}_{k=1}^K$. Our goal is to define distributions of interest on the dynamics parameter $\mathbf{r}$ given the considered set of fuzzy events $\{\mathcal{A}_k\}_{k=1}^K$.

We first define $\mathcal{Z}$ as the fuzzy set of all the possible events. Typically, it depends on the considered collection $\{\mathcal{A}_k\}_{k=1}^K$, as we will see in Section IV-E. We then define the parameter space as the crisp set of the possible values

$$\mathcal{R} = \{\mathbf{r} \in \Omega \text{ such that } \mu_{\mathcal{Z}}(\mathbf{r}) > 0\} \tag{11}$$

with Ω as the unconstrained parameter space. To introduce fuzzy information in a tracking based on particle filtering, we now make the assumption that we do not have specific prior information about the parameter pdf. We then model $p(\mathbf{r})$ as a uniform pdf and, accordingly, $\mathbb{P}(\mathbf{r} \in d\mathbf{r})$ as a uniform distribution. The probability of the event $\mathcal{Z}$ is then

$$\mathbb{P}(\mathcal{Z}) = \int_{\mathcal{R}} \mu_{\mathcal{Z}}(\mathbf{r})\, d\mathbf{r}. \tag{12}$$

We can then turn our model to be conditioned by the fuzzy set of the possible events $\mathcal{Z}$. This restricts the set of possible values of the parameters $\mathbf{r}$. Using (10), the conditional probability $\mathbb{P}(\mathbf{r} \in d\mathbf{r}|\mathcal{Z})$ is defined by

$$\begin{aligned}\mathbb{P}(\mathbf{r} \in d\mathbf{r}|\mathcal{Z}) &= \frac{\int_{\mathcal{R}} \mu_{d\mathbf{r}}(\mathbf{r}')\mu_{\mathcal{Z}}(\mathbf{r}')\, d\mathbf{r}'}{\int_{\mathcal{R}} \mu_{\mathcal{Z}}(\mathbf{r}')\, d\mathbf{r}'} \\ &= \frac{d\mathbf{r}\mu_{\mathcal{Z}}(\mathbf{r})}{\int_{\mathcal{R}} \mu_{\mathcal{Z}}(\mathbf{r}')\, d\mathbf{r}'}.\end{aligned} \tag{13}$$

Let $\mathbf{c}$ be a discrete random variable taking values in $\{1, \ldots, K\}$, which indicates the indices of the considered fuzzy set. We set the conditional distribution of $\mathbf{c}$ given the dynamics parameter $\mathbf{r}$ and $\mathcal{Z}$ as proportional to the ratio between the membership degree of $\mathbf{r}$ to the kth fuzzy set and its membership degree to the fuzzy set of the possible events $\mathcal{Z}$. The conditional distribution is finally obtained using a normalization term which corresponds to the ratio between the sum of membership degrees of the sets and that of $\mathcal{Z}$

$$\begin{aligned}\mathbb{P}(\mathbf{c} = k|\mathbf{r}, \mathcal{Z}) &= \frac{\mu_k(\mathbf{r})}{\mu_{\mathcal{Z}}(\mathbf{r})} \Big/ \frac{\sum_{h=1}^K \mu_h(\mathbf{r})}{\mu_{\mathcal{Z}}(\mathbf{r})} \\ &= \frac{\mu_k(\mathbf{r})}{\sum_{h=1}^K \mu_h(\mathbf{r})}.\end{aligned} \tag{14}$$

From (13) and (14), the conditional distribution of $\mathbf{c}$ given the fuzzy event $\mathcal{Z}$ is expressed as

$$\begin{aligned}\mathbb{P}(\mathbf{c} = k|\mathcal{Z}) &= \int_{\mathcal{R}} \mathbb{P}(\mathbf{c} = k|\mathbf{r}, \mathcal{Z})\mathbb{P}(\mathbf{r}|\mathcal{Z})\, d\mathbf{r} \\ &= \int_{\mathcal{R}} \frac{\mu_k(\mathbf{r})}{\sum_{h=1}^k \mu_h(\mathbf{r})} \frac{\mu_{\mathcal{Z}}(\mathbf{r})}{\int_{\mathcal{R}} \mu_{\mathcal{Z}}(\mathbf{r}')\, d\mathbf{r}'}\, d\mathbf{r}.\end{aligned} \tag{15}$$

We finally obtain the conditional distribution of $\mathbf{r}$ given $\mathbf{c}$ and $\mathcal{Z}$

$$\mathbb{P}(\mathbf{r} \in d\mathbf{r}|\mathbf{c} = k, \mathcal{Z}) = \frac{\mathbb{P}(\mathbf{c} = k|\mathbf{r}, \mathcal{Z})\mathbb{P}(\mathbf{r} \in d\mathbf{r}|\mathcal{Z})}{\mathbb{P}(\mathbf{c} = k|\mathcal{Z})} \propto \frac{d\mathbf{r}\mu_k(\mathbf{r})\mu_{\mathcal{Z}}(\mathbf{r})}{\sum_{h=1}^{K}\mu_h(\mathbf{r})} \tag{16}$$

where we have omitted the normalizing term for more clarity. Considering $\mathbf{c}_t$ as a Markovian process, with t representing time, it is then possible to define the transition from $\mathbf{c}_{t-1}$ to $\mathbf{c}_t$ by an intersection degree between fuzzy sets. Indeed, two close classes have close semantics, and we consider that the transition from a class to another one is gradual. This is the reason why the transition matrix is defined proportionally to the intersection degree (8)

$$\mathbb{P}(\mathbf{c}_t = k|\mathbf{c}_{t-1} = h, \mathcal{Z}) = \frac{R(k,h)}{\sum_{j=1}^{K} R(k,j)}. \tag{17}$$

Hence, the transition of the variables $\mathbf{c}_t$ can be automatically deduced from the expression of the fuzzy functions based on the idea that the probability of moving from class h to class k relies on the degree of intersection of their corresponding functions. We can note that $\mathbf{c}_t$ will always have a maximal probability to stay in the same state since $R(k,k) = \Xi^{kk}(\mu_k, \mu_k) \triangleq 1$ if the μ_k's are normalized (i.e., $\max_{\mathbf{r}} \mu_k(\mathbf{r}) = 1$).

C. Tracking Algorithm

Let $\mathbf{r}_t = (\mathbf{r}_t^1, \ldots, \mathbf{r}_t^R)$ be the parameter vector, all components being independent from each other. Each parameter $\mathbf{r}_t^r$ is modeled by a set of fuzzy functions $\{\mu_k^r\}_{k=1}^{K^r}$ and is hence defined on the associated set of the possible situations (11). We also consider the set of the fuzzy events $\mathcal{Z} = \{\mathcal{Z}^1, \ldots, \mathcal{Z}^R\}$ defined by fixed membership functions (as we will see in Section IV-E). Let $\mathbf{c}_t = (\mathbf{c}_t^1, \ldots, \mathbf{c}_t^R)$ be the vector indicating to which fuzzy set the parameters $(\mathbf{r}_t^1, \ldots, \mathbf{r}_t^R)$ are more likely to be associated with. Denoting by $k = (k^1, \ldots, k^R)$ and $h = (h^1, \ldots, h^R)$ two arbitrary realizations of the vector $\mathbf{c}$ and considering a classical independence assumption, we obtain the prior transition of $\mathbf{c}_t$

$$\mathbb{P}(\mathbf{c}_t = k|\mathbf{c}_{t-1} = h, \mathcal{Z}) = \prod_{r=1}^{R} \mathbb{P}\left(\mathbf{c}_t^r = k^r|\mathbf{c}_{t-1}^r = h^r, \mathcal{Z}^r\right) \tag{18}$$

with $\mathbb{P}(\mathbf{c}_t^r = k^r|\mathbf{c}_{t-1}^r = h^r, \mathcal{Z}^r)$ as the transition matrix (17). We model the vector of parameters such that $\mathbf{r}_t$ is conditionally independent of $\mathbf{r}_{t-1}$ given $\mathbf{c}_t$, which leads to

$$\mathbb{P}(\mathbf{r}_t \in d\mathbf{r}_t|\mathbf{c}_t = k, \mathbf{r}_{t-1}, \mathcal{Z}) = \prod_{r=1}^{R} \mathbb{P}\left(\mathbf{r}_t^r \in d\mathbf{r}_t^r|\mathbf{c}_t^r = k^r, \mathcal{Z}^r\right) \tag{19}$$

with $\mathbb{P}(d\mathbf{r}_t^r|\mathbf{c}_t^r = k^r, \mathcal{Z}^r)$ as the distribution introduced in (16). The filtering distribution at time t can be approximated using Algorithm 1.

Algorithm 1: Particle filter for fuzzy spatial information

Input: Past posterior empiric distribution
$P_N(d\mathbf{x}_{t-1}, d\mathbf{r}_{t-1}, d\mathbf{c}_{t-1}|\mathbf{y}_{1:t-1}, \mathcal{Z})$

Output: New posterior empiric distribution
$P_N(d\mathbf{x}_t, d\mathbf{r}_t, d\mathbf{c}_t|\mathbf{y}_{1:t}, \mathcal{Z})$

begin

- Particle propagation: **for** $n = 1, \ldots, N$ **do**
  - simulate $\tilde{\mathbf{c}}_t^{(n)} \sim q(\mathbf{c}_t|\mathbf{c}_{a_t:t-1}^{(n)}, \mathbf{r}_{t-1}^{(n)}, \mathbf{x}_{t-1}^{(n)}, \mathbf{y}_t, \mathcal{Z})$ (Eq. 20)
  - simulate $\tilde{\mathbf{r}}_t^{(n)} \sim \mathbb{P}(\mathbf{r}_t \in d\mathbf{r}_t|\tilde{\mathbf{c}}_t^{(n)}, \mathcal{Z})$ (Eq. 19)
  - simulate $\tilde{\mathbf{x}}_t^{(n)} \sim q(\mathbf{x}_t|\mathbf{x}_{t-1}^{(n)}, \tilde{\mathbf{r}}_t^{(n)})$
- Weight update: **for** $n = 1, \ldots, N$ **do**

$$\tilde{w}_t^{(n)} = w_{t-1}^{(n)} \frac{p(\mathbf{y}_t|\tilde{\mathbf{x}}_t^{(n)}, \tilde{\mathbf{r}}_t^{(n)}, \tilde{\mathbf{c}}_t^{(n)})}{q(\tilde{\mathbf{x}}_t^{(n)}|\mathbf{x}_{t-1}^{(n)}, \tilde{\mathbf{r}}_t^{(n)})} \times \frac{p(\tilde{\mathbf{x}}_t^{(n)}|\mathbf{x}_{t-1}^{(n)}, \tilde{\mathbf{r}}_t^{(n)})\mathbb{P}(\tilde{\mathbf{c}}_t^{(n)}|\mathbf{c}_{t-1}^{(n)}, \mathcal{Z})}{q(\tilde{\mathbf{c}}_t^{(n)}|\mathbf{c}_{a_t:t-1}^{(n)}, \mathbf{r}_{t-1}^{(n)}, \mathbf{x}_{t-1}^{(n)}, \mathbf{y}_t, \mathcal{Z})}$$

- Weight normalization: **for** $n = 1, \ldots, N$ **do**

$$w_t^{(n)} = \frac{\tilde{w}_t^{(n)}}{\sum_{m=1}^{N} \tilde{w}_t^{(m)}}$$

- Resampling, if necessary.

end

return $\sum_{n=1}^{N} w_t^{(n)} \delta_{\mathbf{x}_t^{(n)}, d\mathbf{r}_t^{(n)}, d\mathbf{c}_t^{(n)}}(d\mathbf{x}_t, d\mathbf{r}_t, d\mathbf{c}_t)$

Note that the measurement process may depend on indices $\mathbf{c}_t$ fuzzy sets; hence, it is possible to consider dedicated versions of the likelihood function, i.e., one by class.

D. Importance Distribution

One of the main issues of the proposed approach is to design the importance distribution based on the fuzzy intervals containing the dynamics parameters of the model. We then model it to be dependent both on past realizations of the vector $\mathbf{c}$ and on the new observation $\mathbf{y}_t$. Under an independence assumption, the importance distribution is given by

$$\begin{aligned} &q\left(\mathbf{c}_t|\mathbf{c}_{a_t:t-1}^{(n)}, \mathbf{r}_{t-1}^{(n)}, \mathbf{x}_{t-1}^{(n)}, \mathbf{y}_t, \mathcal{Z}\right) \\ &= \prod_{r=1}^{R} q\left(\mathbf{c}_t^r|\mathbf{c}_{a_t:t-1}^{r,(n)}, \mathbf{r}_{t-1}^{(n)}, \mathbf{x}_{t-1}^{(n)}, \mathbf{y}_t, \mathcal{Z}^r\right) \\ &= \prod_{r=1}^{R} \sum_{k=1}^{K^r} \delta_k\left(\mathbf{c}_t^r\right) q\left(\mathbf{c}_t^r = k|\mathbf{c}_{a_t:t-1}^{r,(n)}, \mathbf{r}_{t-1}^{(n)}, \mathbf{x}_{t-1}^{(n)}, \mathbf{y}_t, \mathcal{Z}^r\right) \end{aligned} \tag{20}$$

where a_t is set to $\triangleq t - 1 - m$ with m as a fixed constant value to deal with the $m + 1$ past realizations of $\mathbf{c}$.

The last term in (20) is a mixture distribution of the transition function defined in (15) and a distribution which extracts prior

information from the trajectory of $\mathbf{c}_{a_t:t-1}$ and the new observation $\mathbf{y}_t$ given by

$$q\left(\mathbf{c}_t^r = k|\mathbf{c}_{a_t:t-1}^{r,(n)}, \mathbf{r}_{t-1}^{(n)}, \mathbf{x}_{t-1}^{(n)}, \mathbf{y}_t, \mathcal{Z}^r\right) = \gamma\mathbb{P}\left(\mathbf{c}_t^r = k|\mathbf{c}_{t-1}^{r,(n)}, \mathcal{Z}^r\right) + (1-\gamma) \times \mathbb{P}\left(\mathbf{c}_t^r = k|\mathbf{c}_{a_t:t-1}^{r,(n)}, \mathbf{r}_{t-1}^{(n)}, \mathbf{x}_{t-1}^{(n)}, \mathbf{y}_t, \mathcal{Z}^r\right) \quad (21)$$

with $\gamma \in [0,1]$ as the weighting coefficient. The distribution $\mathbb{P}(\mathbf{c}_t^r = k|\mathbf{c}_{a_t:t-1}^{r,(n)}, \mathbf{r}_{t-1}^{(n)}, \mathbf{x}_{t-1}^{(n)}, \mathbf{y}_t, \mathcal{Z}^r)$ allows removing some local ambiguities by using a global information characterized by the trajectory $\mathbf{c}_{a_t:t-1}$, while maintaining some hypotheses during time.

Then, we consider prior probabilities $\pi_r^k = \mathbb{P}(\mathbf{c}_t^r = k|\mathcal{Z}^r)$ as a realization of the random vector $\Pi_r = (\Pi_r^1, \ldots, \Pi_r^{K^r})$ which takes values in $[0,1]^{K^r}$ such that $\sum_{k=1}^{K^r} \Pi_r^k = 1$. In order to simulate samples according to the distribution $\mathbb{P}(\mathbf{c}_t^r = k|\mathbf{c}_{a_t:t-1}^{r,(n)}, \mathbf{r}_{t-1}^{(n)}, \mathbf{x}_{t-1}^{r,(n)}, \mathbf{y}_t, \mathcal{Z}^r)$, probabilities $\pi_r = (\pi_r^1, \ldots, \pi_r^{K^r})$ are generated according to the importance distribution described in (22), as shown at the bottom of the page, where $\mathrm{nb}^j(l_{t_1:t_2}) = \sum_{i=t_1}^{t_2} \delta_j(l_i)$, $(t_1, t_2) \in \mathbb{N}^2$. The parameters $\{\alpha_k^r\}_{k=1}^{K^r}$ correspond to the degree of importance of the modes $\mathbf{c}_t^r$ and handle the global shape of the prior Dirichlet distribution $\mathcal{D}$. The multinomial distribution $\mathcal{M}$ uses the $m+1$ past realizations $\mathbf{c}_{a_t:t-1}^{r,(n)}$, which leads to a posterior Dirichlet distribution that integrates both trajectory of $\mathbf{c}_t$ and new prior probabilities of the modes. The distribution $q(\mathbf{c}_t^r = k|\mathbf{r}_{t-1}^{(n)}, \mathbf{x}_{t-1}^{(n)}, \mathbf{y}_t, \mathcal{Z}^r)$ in (22) can be written as the following mixture distribution:

$$q\left(\mathbf{c}_t^r = k|\mathbf{r}_{t-1}^{(n)}, \mathbf{x}_{t-1}^{(n)}, \mathbf{y}_t, \mathcal{Z}^r\right) = \beta^1\mathbb{P}\left(\mathbf{c}_t^r = k|\mathbf{r}_{t-1}^{r,(n)}, \mathcal{Z}^r\right) + \beta^2\mathbb{P}\left(\mathbf{c}_t^r = k|\mathbf{x}_{t-1}^{(n)}, \mathbf{r}_{t-1}^{(n)}, \mathbf{y}_t\right) + \beta^3\mathbb{P}\left(\mathbf{c}_t^r = k|\mathcal{Z}^r\right) \quad (23)$$

where $\sum_{k=1}^3 \beta^k = 1$, $\mathbb{P}(\mathbf{c}_t^r = k|\mathcal{Z}^r)$ is the prior distribution defined by (15), and $\mathbb{P}(\mathbf{c}_t^r = k|\mathbf{r}_{t-1}^{r,(n)}, \mathcal{Z}^r)$ is the posterior probability to consider at time t the kth fuzzy set representing a value of the rth fuzzy concept conditioned by the past realization of the parameter $\mathbf{r}_{t-1}^{r,(n)}$ (14). We propose to approximate the distribution $\mathbb{P}(\mathbf{c}_t^r = k|\mathbf{x}_{t-1}^{(n)}, \mathbf{r}_{t-1}^{(n)}, \mathbf{y}_t)$ by

$$\begin{aligned}\mathbb{P}\left(\mathbf{c}_t^r = k|\mathbf{x}_{t-1}^{(n)}, \mathbf{r}_{t-1}^{(n)}, \mathbf{y}_t\right) &\simeq \frac{p\left(\mathbf{y}_t|\mathbf{c}_t^r = k, \mathbf{r}_{t-1}^{(n)}, \mathbf{x}_{t-1}^{(n)}\right)}{\sum_{j=1}^{K^r} p\left(\mathbf{y}_t|\mathbf{c}_t^r = j, \mathbf{r}_{t-1}^{(n)}, \mathbf{x}_{t-1}^{(n)}\right)} \\ &= \frac{p\left(\mathbf{y}_t|g_t\left(\mathbf{x}_{t-1}^{(n)}, \bar{\mathbf{r}}_{t|t-1}^{r_k,(n)}, \mathbf{v}_t\right)\right)}{\sum_{j=1}^{K^r} p\left(\mathbf{y}_t|g_t\left(\mathbf{x}_{t-1}^{(n)}, \bar{\mathbf{r}}_{t|t-1}^{r_j,(n)}, \mathbf{v}_t\right)\right)}\end{aligned}$$

where

$$\bar{\mathbf{r}}_{t|t-1}^{r_k,(n)} = \left(\mathbf{r}_{t-1}^{1,(n)}, \ldots, \mathbf{r}_{t-1}^{r-1,(n)}, \bar{\mathbf{r}}_t^{r_k}, \mathbf{r}_{t-1}^{r+1,(n)}, \ldots, \mathbf{r}_{t-1}^{R,(n)}\right) \quad (24)$$

denotes a candidate value for parameter $\mathbf{r}_{t|t-1}$ and $\bar{\mathbf{r}}_t^{r_k}$ is a characteristic feature of the membership function $\mu_k^r(\mathbf{r})$, e.g., its mean value. Note that this approximation is reasonable since the probability for a class to stay in the same state is always maximal [according to (17)]. Although not considered in our experiments, a more robust strategy may be employed to propose a relevant candidate $\bar{\mathbf{r}}_{t|t-1}^{r_k,(n)}$. For instance, if the cardinality of $\mathbf{c}_t$ is quite small, one can generate a sample according to the following distribution:

$$\tilde{\mathbf{r}}_t^{r_k,(n)} \sim \frac{1}{U}\sum_{j_1=1}^{K^1} \cdots \sum_{j_{r-1}=1}^{K^{r-1}} \sum_{j_{r+1}=1}^{K^{r+1}} \cdots \sum_{j_R=1}^{K^R} \prod_{\forall h \neq r} \delta_{\bar{\mathbf{r}}_t^{h_{j_h}}}\left(\tilde{\mathbf{r}}_t^{h,(n)}\right) \times p\left(\mathbf{y}_t|g_t\left(\mathbf{x}_{t-1}^{(n)}, \tilde{\mathbf{r}}_t^{r_k,(n)}, \mathbf{v}_t\right), \mathcal{Z}^r\right)$$

with U as a normalization factor and where

$$\tilde{\mathbf{r}}_t^{r_k,(n)} = \left(\tilde{\mathbf{r}}_t^{1,(n)}, \ldots, \tilde{\mathbf{r}}_t^{r-1,(n)}, \bar{\mathbf{r}}_t^{r_k}, \tilde{\mathbf{r}}_t^{r+1,(n)}, \ldots, \tilde{\mathbf{r}}_t^{R,(n)}\right).$$

E. Fuzzy Spatial Information in Practice

Modeling the Spatial Relations: One way to easily integrate fuzzy spatial relations, such as fuzzy orientations $North, East, \ldots$, in the particle filter consists in considering the functions $\{\mu_k^r\}_{k=1}^{K^r}$, where K^r is the number of possible

$$\begin{aligned}&q\left(\pi_r|\mathbf{c}_{a_t:t-1}^{r,(n)}, \mathbf{r}_{t-1}^{(n)}, \mathbf{x}_{t-1}^{(n)}, \mathbf{y}_t, \mathcal{Z}^r\right) \\ &\propto q\left(\mathbf{c}_{a_t:t-1}^{r,(n)}|\pi_r\right)\, q\left(\pi_r|\mathbf{r}_{t-1}^{(n)}, \mathbf{x}_{t-1}^{(n)}, \mathbf{y}_t, \mathcal{Z}^r\right) \\ &= \mathcal{M}\left(\mathrm{nb}^1\left(\mathbf{c}_{a_t:t-1}^{r,(n)}\right), \ldots, \mathrm{nb}^{K^r}\left(\mathbf{c}_{a_t:t-1}^{r,(n)}\right)\right) \times \mathcal{D}\left(\pi_r\left\{1 + \alpha_k^r q\left(\mathbf{c}_t^r = k|\mathbf{r}_{t-1}^{(n)}, \mathbf{x}_{t-1}^{(n)}, \mathbf{y}_t, \mathcal{Z}^r\right)\right\}_{k=1,\ldots,K^r}\right) \\ &\propto \mathcal{D}\left(\pi_r\left\{1 + \alpha_k^r q\left(\mathbf{c}_t^r = k|\mathbf{r}_{t-1}^{(n)}, \mathbf{x}_{t-1}^{(n)}, \mathbf{y}_t, \mathcal{Z}^r\right) + \mathrm{nb}^k\left(\mathbf{c}_{a_t:t-1}^{r,(n)}\right)\right\}_{k=1,\ldots,K^r}\right) \quad (22)\end{aligned}$$

discrete values for $\mathbf{c}_t^r$ as a fuzzy partition of the bounded space $\mathcal{R}^r$, and then satisfying the following properties:

1) $\forall \mathbf{r}^r \in \mathcal{R}^r, \sum_{k=1}^{K^r} \mu_k^r(\mathbf{r}^r) = 1$;
2) $\forall (k,h) \in \{1,\ldots,K^r\}^2, \int_{\mathcal{R}^r} \mu_k^r(\mathbf{r}^r) d\mathbf{r}^r = \int_{\mathcal{R}^r} \mu_h^r(\mathbf{r}^r) d\mathbf{r}^r$.

Then, it is possible to model $\mathcal{Z}^r$ such that its membership function equals $\sum_{k=1} \mu_k^r = 1$. From these assumptions and considering trapezoidal fuzzy functions, it becomes easy to generate samples and to compute in a closed form (13), (15)–(17) as they simply reduce to

$$\mathbb{P}(\mathbf{r}^r \in d\mathbf{r}^r | \mathcal{Z}^r) = \frac{d\mathbf{r}^r}{|\mathcal{R}^r|} \tag{25}$$

$$\mathbb{P}(\mathbf{c}^r = k | \mathbf{r}^r, \mathcal{Z}^r) = \mu_k^r(\mathbf{r}^r) \tag{26}$$

$$\mathbb{P}(\mathbf{c}^r = k | \mathcal{Z}^r) = \frac{1}{K^r} \tag{27}$$

$$\mathbb{P}(\mathbf{r}^r \in d\mathbf{r}^r | \mathbf{c}^r = k, \mathcal{Z}^r) = \frac{d\mathbf{r}^r \mu_k^r(\mathbf{r}^r)}{|\mathcal{R}^r|/K^r}. \tag{28}$$

Fuzzy spatial information is then introduced using simple fuzzy templates of binary, triangular, or trapezoid shape.

It is also possible to define spatial relations by learning. For example, using a disjunctive intraclass strategy and a normalized interclass strategy systematically leads to (25) and (26).

Modeling General Spatial Information: When the constraints enunciated previously are too strong or we just want to model more general spatial information, there is no need to make additional assumptions. We then have to define a relevant fuzzy set of the possible values $\mathcal{Z}^r$. We present here a nonexhaustive list of strategies for defining $\mathcal{Z}^r$, leading to different properties of the set of the possible values with $\top$ as a t-norm and $\perp$ as a t-conorm [12].

1) Union: $\mathcal{Z}^r = \bigcup_{k=1}^{K^r} \mathcal{A}_k^r$

$$\Leftrightarrow \mu_{\mathcal{Z}^r}(\mathbf{r}^r) = \mathop{\perp}_{k=1}^{K^r} \mu_k^r(\mathbf{r}^r) \quad \forall \mathbf{r}^r \in \mathcal{R}^r.$$

2) Intersection: $\mathcal{Z}^r = \bigcap_{k=1}^{K^r} \mathcal{A}_k^r$

$$\Leftrightarrow \mu_{\mathcal{Z}^r}(\mathbf{r}^r) = \mathop{\top}_{k=1}^{K^r} \mu_k^r(\mathbf{r}^r) \quad \forall \mathbf{r}^r \in \mathcal{R}^r.$$

3) Product: $\mathcal{Z}^r = \prod_{k=1}^{K^r} \mathcal{A}_k^r$

$$\Leftrightarrow \mu_{\mathcal{Z}^r}(\mathbf{r}^r) = \prod_{k=1}^{K^r} \mu_k^r(\mathbf{r}^r) \quad \forall \mathbf{r}^r \in \mathcal{R}^r.$$

4) Sum: $\mathcal{Z}^r = \mathcal{A}_1^r \oplus \ldots \oplus \mathcal{A}_{K^r}^r$

$$\Leftrightarrow \mu_{\mathcal{Z}^r}(\mathbf{r}^r) = \sum_{k=1}^{K^r} \mu_k^r(\mathbf{r}^r) - \prod_{k=1}^{K^r} \mu_k^r(\mathbf{r}^r) \quad \forall \mathbf{r}^r \in \mathcal{R}^r.$$

The product and the sum are particular cases of $\top$ and $\perp$, respectively.

In a general case, i.e., when fuzzy sets are learned and characterized by complex shapes, it is often needed to approximate the distributions of interest, which is the subject of the following section.

F. Approximation

In several cases (probabilistic strategy, spatial relations, monodimensional space $\mathcal{R}^r$, etc.), the computing task or the simulation of the distributions defined in (13), (15)–(17) can be executed in an closed-form (or direct) way. In other cases, it is necessary to use an approximation method. Of course, this step is realized just once, when the fuzzy functions are learned. We propose here to approximate the prior distribution $\mathbb{P}(\mathbf{r}^r \in d\mathbf{r}^r | \mathcal{Z}^r)$ by an independent Metropolis–Hasting algorithm, but an accept–reject algorithm could also be suitable by modifying the proposal function. We refer to [18] for a comparison of convergence of these two algorithms. We set

$$F(\mathbf{r}^r) = \mu_{\mathcal{Z}}(\mathbf{r}^r) \quad \text{and} \quad G(\mathbf{r}^r) = \sum_{k=1}^{K^r} \sum_{m=1}^{M^k} P_{\mathbf{o}_m^k}(\mathbf{r}^r).$$

The functions F and G correspond to unnormalized versions of probabilistic density functions, the first one referring to $\mathbb{P}(\mathbf{r}^r \in d\mathbf{r}^r | \mathcal{Z}^r)$, whereas G could be obtained by an additive intraclass fusion, where neither an interclass strategy nor a constraint have been applied. Then, G has a support which includes that of F and a shape that easily enables the generation of samples. The approximation process can be done by Algorithm 2, where τ_{begin} corresponds to the period of burn-in and τ_{end} corresponds to the maximum number of iterations.

Algorithm 2: Independent Metropolis-Hasting algorithm

Input: Initial state $\mathbf{r}^{r,(0)}$
Output: Approximation of the distribution $\mathbb{P}(\mathbf{r}^r \in d\mathbf{r}^r | \mathcal{Z}^r)$
begin
  for $\tau = 0, \ldots, \tau_{end}$ **do**
    - Generate a candidate $\mathbf{r}^* \sim G(\mathbf{r}^r)$.
    - The candidate $\mathbf{r}^*$ is accepted with the probability :

$$\alpha = \begin{cases} \min\left\{1, \frac{F(\mathbf{r}^*)G(\mathbf{r}^{r,(\tau)})}{F(\mathbf{r}^{r,(\tau)})G(\mathbf{r}^*)}\right\} & \text{if } F(\mathbf{r}^{r,(\tau)})G(\mathbf{r}^*) > 0 \\ 1 & \text{otherwise.} \end{cases}$$

    - If accepted, $\mathbf{r}^{r,(\tau+1)} = \mathbf{r}^*$, otherwise $\mathbf{r}^{r,(\tau+1)} = \mathbf{r}^{r,(\tau)}$.
end
return $\{\mathbf{r}^{r,(i)}\,;\, Li = l(\tau_{end} - \tau_{begin}),\, l = 1,\ldots,L,\, i \in \mathbb{N}\}$

The approximation by an L-sample of the prior distribution defined in (13) is

$$\mathbb{P}(\mathbf{r}^r \in d\mathbf{r}^r | \mathcal{Z}^r) \approx \frac{1}{L} \sum_{l=1}^{L} \delta_{\mathbf{r}^{r,(l)}}(d\mathbf{r}^r). \tag{29}$$

It is then straightforward to deduce approximations of (15) and (16)

$$\mathbb{P}(\mathbf{c}^r = k | \mathcal{Z}^r) \approx \frac{1}{L} \sum_{l=1}^{L} \frac{\mu_k^r\left(\mathbf{r}^{r,(l)}\right)}{\sum_{h=1}^{K} \mu_h^r\left(\mathbf{r}^{r,(l)}\right)} \tag{30}$$

$$\mathbb{P}(\mathbf{r}^r \in d\mathbf{r}^r | \mathbf{c}^r = k, \mathcal{Z}^r) \approx \frac{1}{\mathbb{P}(\mathbf{c}^r = k | \mathcal{Z}^r)} \frac{1}{L} \times \sum_{l=1}^{L} \frac{\mu_k^r(\mathbf{r}^r) \delta_{\mathbf{r}^{r,(l)}}(d\mathbf{r}^r)}{\sum_{h=1}^{K} \mu_h^r(\mathbf{r}^r)} \tag{31}$$

while (14) remains unchanged.

V. Example: Tracking on Simulated Data

As an example, we consider a synthetic tracking problem, in which an object is defined by a state vector $\mathbf{x}_t = (\mathrm{x}_t, \mathrm{y}_t)$ and by a dynamics vector $\mathbf{r}_t = (\dot{\mathrm{x}}_t, \dot{\mathrm{y}}_t)$, for which we consider the following constant velocity dynamical system:

$$\mathbf{x}_t = g_t(\mathbf{x}_{t-1}, \mathbf{r}_t) = \begin{pmatrix} \mathrm{x}_{t-1} + \Delta t\, \dot{\mathrm{x}}_t + \mathbf{v}_t^1 \\ \mathrm{y}_{t-1} + \Delta t\, \dot{\mathrm{y}}_t + \mathbf{v}_t^2 \end{pmatrix} \tag{32}$$

where the hidden parameter $\mathbf{r}_t$ is estimated at time t using the proposed framework. We consider nine linguistic values for the couple $(\dot{\mathrm{x}}_t, \dot{\mathrm{y}}_t)^{\mathrm{T}}$, from *NorthWest* to *SouthEast*, modeling the relative position of the object at t with respect to its position at $t-1$. These parameters are designed in a flexible way by defining their membership to fuzzy sets

$$\mathbf{c}_t^1 \in \{NW, N, NE, W, Center, E, SW, S, SE\}.$$

A. Modeling the Fuzzy Functions

For fuzzy relations, we follow the model proposed in Section IV-E. It means that the distributions $\mathbb{P}(\mathbf{r}^r \in d\mathbf{r}^r | \mathcal{Z}^r)$ and $\mathbb{P}(\mathbf{c}^r = k | \mathcal{Z}^r)$ are uniform. We model the set of fuzzy functions corresponding to the velocity of the object as a fuzzy partition of $[-h_{\mathrm{x}}, h_{\mathrm{x}}] \times [-h_{\mathrm{y}}, h_{\mathrm{y}}]$, with h_{x} and h_{y} as the fixed horizon parameters. We then consider the following intersection degree between two classes k and h:

$$\Xi\left(\mu_k^r, \mu_h^r\right) = \frac{\int_{\mathcal{R}^r} \top\left[\mu_k^r(\mathbf{r}), \mu_h^r(\mathbf{r})\right] d\mathbf{r}}{\min\left[\int_{\mathcal{R}^r} \mu_k^r(\mathbf{r}) d\mathbf{r}, \int_{\mathcal{R}^r} \mu_h^r(\mathbf{r}) d\mathbf{r}\right]} \tag{33}$$

with $\top$ as a t-norm, for example, $\top = \min$.

B. Likelihood Function

We consider for this synthetic tracking problem a simple Gaussian observation model

$$p(\mathbf{y}_t | \mathbf{x}_t, \mathbf{r}_t) = \mathcal{N}\left(\mathbf{x}_t, \sigma_w^2 \mathbf{I}_{2\times 2}\right).$$

C. Experiments

The trajectory proposed in this first example has been artificially designed in order to present strong discontinuities in the dynamics, which allows the comparison of the proposed approach with different existing particle filters.

For the Sequential Importance Resampling (SIR) particle filter, the AUX-PF, and the proposed models, we use a standard constant velocity model. For the MMPF and the AUX-MMPF, we also consider clockwise- and counterclockwise-coordinated turn models that capture the maneuver dynamics. The reader may refer to [3] for additional explanations about these models. The position noise vector follows $\mathcal{N}(\mathbf{0}, 4^2 \times \mathbf{I}_{2\times 2})$, with $\mathbf{0} = (0\ 0)^{\mathrm{T}}$ and $\mathbf{I}_{2\times 2}$ as the 2×2 identity matrix. For the proposed models, the velocity horizons are fixed to $h_{\mathrm{x}} = h_{\mathrm{y}} = 40$. The shape of the fuzzy functions for the velocity parameter is triangular. For the parameters of the proposed model using an importance distribution, we used $\gamma = 0.3$ for (21) to give priority to the importance distribution, $m = 3$ for the length of the $\mathbf{c}$ trajectory, and $\beta^1 = \beta^3 = 0.1$ and $\beta^2 = 0.8$ to efficiently use the last measurement. The likelihood noise is set to $\sigma_w = 5$.

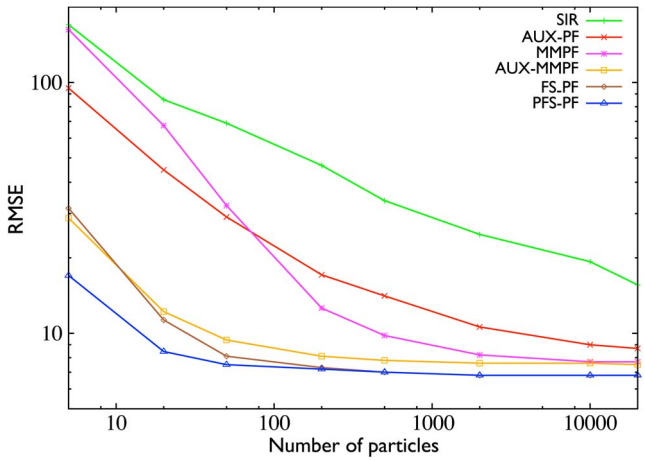

Fig. 1. Synthetic example: Root mean square tracking errors given by the different filters depending on the number N of particles. Both axes are logarithmic.

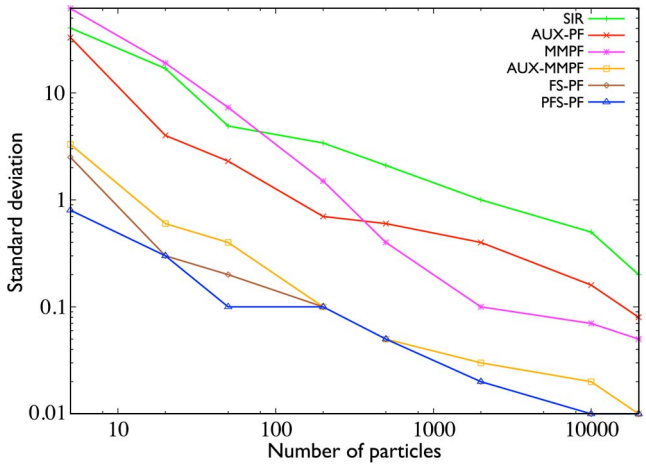

Fig. 2. Synthetic example: Standard deviations of estimators given by the different filters depending on the number N of particles. Both axes are logarithmic.

The shape parameters of the fuzzy functions and the length of the trajectory have been set experimentally but have shown reduced influence on the results, as could be observed in our tests. Similarly, parameters γ and β weakly depend on the application. The proposed methods are therefore robust to parameter settings, which do not require any specific optimization.

Fig. 1 shows root mean square errors (RMSEs), averaged over 100 simulations of the observation process, of the estimated positions obtained by the SIR filter, the AUX-PF, the MMPF, the AMMPF, and our approaches using the fuzzy framework without the importance distribution, called FS-PF (for Fuzzy Spatial Particle Filter), and with the importance distribution, called PFS-PF (for Proposal Fuzzy Spatial Particle Filter). Fig. 2 shows the standard deviations of estimators given by the different filters.

The limited results obtained by SIR and MMPF are due to discontinuities in the dynamics (Fig. 3), which are better handled when a new observation is used to generate particles. For the proposed models and the AUX-MMPF, we can appreciate the fast decrease of the RMSE and the standard deviation as a function of the number of particles. We can notice, for example,

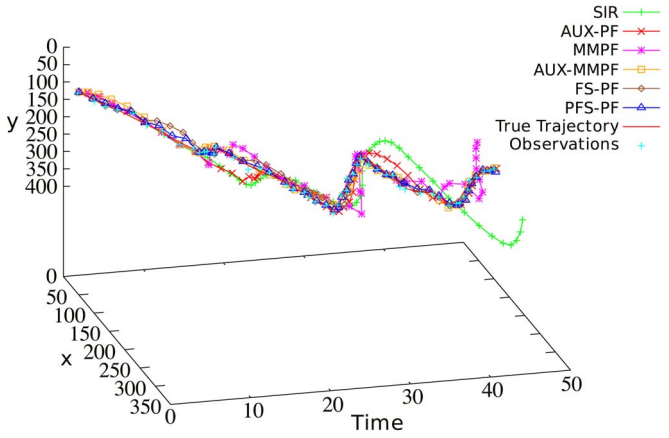

Fig. 3. Synthetic example: Estimated trajectories obtained by the different filters on a run with $N = 20$.

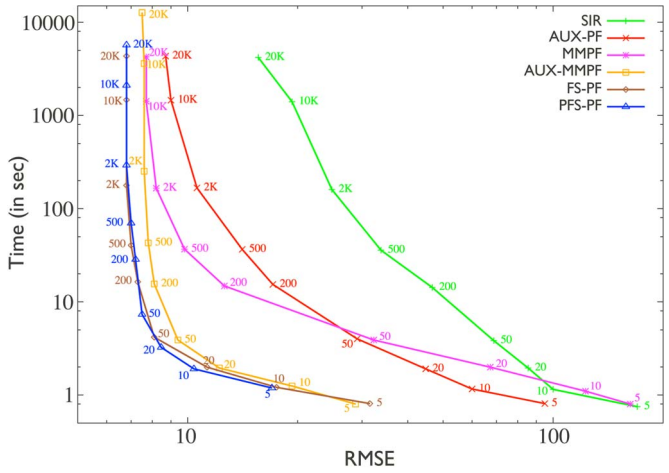

Fig. 4. Synthetic example: Computational costs (in seconds, over 100 runs) depending on the root mean square tracking errors obtained by the different filters. Both axes are logarithmic.

that results obtained by the FS-PF with $N = 20$ particles are better than those obtained by the SIR filter with $N = 2000$.

Fig. 4 shows the computational costs of the different filters given their RMSEs. We can first note that the PFS-PF has a computation time comparable with that of AUX-MMPF while the computation time of the FS-PF is equivalent to the SIR one. The advantage of the PFS-PF over the FS-PF is clear when using a small number of particles $N \leq 100$, since for the same error, both the processing time and the memory storage are lower. When using more particles, it is then more advantageous to use the FS-PF, since the two proposed methods quickly converge (Fig. 1).

In order to assess the significance of the observed differences, we performed a Welch's t-test [25], which is an adaptation of the Student's t-test in the case where the two samples may have unequal variances. With a risk of 0.05, this indicates that the error difference between the two proposed methods is no longer significant when the number of particles is greater than 2000, while the comparison with other methods remains always significant.

In the next two sections, we show that the observed behavior in the case of synthetic data is confirmed when tracking objects with complex and unknown dynamics in real video sequences.

VI. Application 1: Tracking of a Helicopter in a Video Sequence

We consider a single-object tracking problem in a video sequence. The sequence, in which we only consider the first 100 frames, shows a toy helicopter with complex dynamics.[1] The tracking task is not particularly simple in this sequence since the resolution is quite poor (this is a compressed sequence) and there are moving clouds, some defocusing, and instability (the sequence is shot with a camera by a person with shaky hands). The complex dynamics of the helicopter allows us to put in light the main contribution of our methods, which is to avoid strong assumptions on the dynamics and thus handle complex dynamics. We propose to estimate the ellipse surrounding the object over time, i.e., the joint unknown position and dynamics states $\mathbf{x}_t = (\mathrm{x}_t, \mathrm{y}_t)^{\mathrm{T}}$ and $\mathbf{r}_t = (\dot{\mathrm{x}}_t, \dot{\mathrm{y}}_t, \theta_t, \mathrm{w}_t, \mathrm{h}_t)^{\mathrm{T}}$, with θ_t as the object orientation, w_t as the width, and h_t as the height of the ellipse. For the proposed model, we consider the constant velocity dynamical system presented in (32).

We consider nine linguistic values for the couple $(\dot{\mathrm{x}}_t, \dot{\mathrm{y}}_t)^{\mathrm{T}}$, from *NorthWest* to *SouthEast*, five for the object orientation θ_t, from *North* to *South*, and nine for the couple of dimensions of the ellipse $(\mathrm{w}_t, \mathrm{h}_t)^{\mathrm{T}}$, from (*SmallWidth*, *SmallHeight*) to (*LargeWidth*, *LargeHeight*). These parameters are designed in a flexible way by defining their membership to fuzzy sets

$$\mathbf{c}_t^1 \in \{NW, N, NE, W, Center, E, SW, S, SE\}$$

$$\mathbf{c}_t^2 \in \{N, NE, E, SE, S\}$$

$$\mathbf{c}_t^3 \in \{(SW, SH), (MW, SH), (LW, SH), \ldots, (LW, LH)\}$$

$(S = small, M = medium, L = large, W = width, H = height)$.

A. Modeling the Fuzzy Functions

For fuzzy relations, we again follow the model proposed in Section IV-E. We model the set of fuzzy functions as a fuzzy partition of $[-h_{\mathrm{x}}, h_{\mathrm{x}}] \times [-h_{\mathrm{y}}, h_{\mathrm{y}}]$ for the velocity, with h_{x} and h_{y} as the fixed horizon parameters, of $[-h_{\mathrm{w}}, h_{\mathrm{w}}] \times [-h_{\mathrm{h}}, h_{\mathrm{h}}]$ for the dimension of the ellipse, with h_{w} and h_{h} as the fixed dimension parameters, and of $[-\pi/2, \pi/2]$ for the orientation. The intersection degree between two fuzzy relation sets is defined by (33) for $r = 1, 2, 3$.

B. Likelihood Functions

We design the likelihood function as the independent combination of a region-based component and an edge one and then

$$p(\mathbf{y}_t|\mathbf{x}_t, \mathbf{r}_t) = p\left(\mathbf{y}_t^R|\mathbf{x}_t, \mathbf{r}_t\right) p\left(\mathbf{y}_t^E|\mathbf{x}_t, \mathbf{r}_t\right)$$

where $p(\mathbf{y}_t^E|\mathbf{x}_t, \mathbf{r}_t)$ uses the gradient of the smoothed image such that

$$p\left(\mathbf{y}_t^E|\mathbf{x}_t, \mathbf{r}_t\right) \propto \exp\left(\frac{\lambda_1}{P_a} \sum_{p=1}^{P_{\mathbf{a}_t}} \nabla\left[I_t(a^p) \star \mathcal{N}(a^p, \sigma^2)\right]\right) \tag{34}$$

[1] http://www.youtube.com/watch?v=PQf404RumPE.

where ∇ and $\star$ are the gradient and the convolution operators, respectively, a^p is the pth contour point of the object ellipse of length P_a defined by $(\mathbf{x}_t, \mathbf{r}_t)^{\mathrm{T}}$, $I_t(a^p)$ is the grayscale value at point a^p at time t, $\lambda_1 \in \mathbb{R}^+$ is a fixed multiplicative constant value which aims at providing a good localization of high likelihood values, and $\mathcal{N}(a^p, \sigma^2)$ is a normal distribution of mean a^p and fixed variance σ^2.

The region-based likelihood is given by a distance measure between histograms. Let $\mathring{a}$ be the set of pixels covered by the ellipse of state $(\mathbf{x}_t, \mathbf{r}_t)^{\mathrm{T}}$. We denote by $h_o^\bullet = \{h_o^\bullet(u)\}_{u=1}^U$ the learned model histogram of the object, by $h_{BG}^\bullet = \{h_{BG}^\bullet(u)\}_{u=1}^U$ the learned model histogram of the background image, by $h_o(\mathbf{x}_t, \mathbf{r}_t) \triangleq h_o(\mathring{a}) = \{h_o(\mathring{a}; u)\}_{u=1}^U$ a candidate histogram of the object, and by $h_{BG}(\mathbf{x}_t, \mathbf{r}_t) \triangleq h_{BG}(\mathring{a}) = \{h_{BG}(\mathring{a}; u)\}_{u=1}^U$ a candidate histogram of the background, where u is the index of a bin of a histogram of length U. The likelihood function is a variant of that proposed in [22] and uses both object and background information to be robust to scale and orientation changes

$$p\left(\mathbf{y}_t^R | \mathbf{x}_t, \mathbf{r}_t\right) \propto \exp\left(-\lambda_2 \left(d^2\left(h_o^\bullet, h_o(\mathring{a})\right) - \max\left[d^2\left(h_{BG}^\bullet, h_o(\mathring{a})\right), d^2\left(h_o^\bullet, h_{BG}(\mathring{a})\right)\right]\right)\right) \quad (35)$$

with $\lambda_2 \in \mathbb{R}^+$ as a fixed multiplicative constant value. The distance $d[h^\bullet, h(\mathring{a})]$ is based on the Bhattacharyya coefficient

$$d\left[h^\bullet, h(\mathring{a})\right] = \left[1 - \sum_{u=1}^{U} \sqrt{h^\bullet(u) h(\mathring{a}; u)}\right]^{1/2}.$$

C. Experiments

The position noise vector follows $\mathcal{N}(\mathbf{0}, 4^2 \times \mathbf{I}_{2\times 2})$, with $\mathbf{0} = (0\ 0)^{\mathrm{T}}$ and $\mathbf{I}_{2\times 2}$ as the 2×2 identity matrix, the orientation noise vector follows $\mathcal{N}(0, 0.1)$, and the dimension noise vector follows $\mathcal{N}(\mathbf{0}, \mathbf{I}_{2\times 2})$. The velocity horizons are fixed to $h_{\mathrm{x}} = h_{\mathrm{y}} = 60$, and the dimensions of the ellipse are fixed to $h_{\mathrm{w}} = h_{\mathrm{h}} = 50$. The shape of the fuzzy functions for the velocity, orientation, and dimension of the ellipse parameters is triangular.

Following the choices made in [22], we use a hue, saturation, and value (HSV) color space decomposition, and for each pixel, we consider either the couple of hue and saturation if the saturation and the brightness are higher than the threshold values 0.1 and 0.2, respectively, or the brightness of the pixel otherwise. The resulting histogram is then composed of $U = U_h \times U_s + U_v$ bins. In our experiments, we used $U_h = U_s = U_v = 10$ and $\lambda_1 = \lambda_2 = 10$. In the same way as in Section V, parameters of the importance distribution have been set to $\gamma = 0.3$, $m = 3$, $\beta^1 = \beta^3 = 0.1$, and $\beta^2 = 0.8$.

We now focus on the implementation of the importance distribution used in the PFS-PF method. This is of prime interest since its complexity is related to the dimension of the vector of parameters and the number of fuzzy classes by parameter, which may make the method computationally expensive if it is not carefully designed. First, note that a straightforward optimization is possible by taking advantage of the resampling procedure performed at the previous time step, and hence, this allows us to compute (22) only for distinct particles. Moreover, for computation-time reasons, the importance distribution designed in this experiment has only been used for the velocity parameter. We also performed a small approximation in the likelihood used in the importance distribution, using only the first histogram distance term of the region-based likelihood in (35) and without using the edge likelihood (34). Of course, this approximation is only considered in the simulation step, and the likelihood computed in the weight update is the same for all methods.

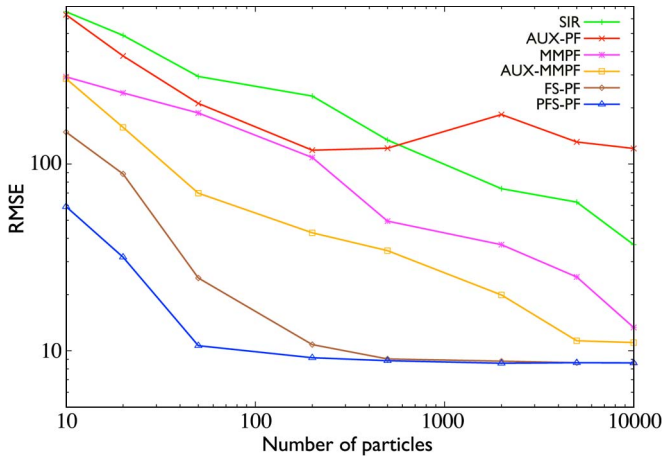

Fig. 5. Helicopter sequence: Root mean square tracking errors given by the different filters depending on the number N of particles. Both axes are logarithmic.

Fig. 5 shows the RMSEs averaged over 50 runs of the estimated helicopter positions obtained by the SIR filter, the AUX-PF, the MMPF, the AMMPF, and our approaches. Tracking results obtained by other filters, compared with those obtained by our approach, are much degraded, particularly for a small number of particles. This is mainly explained by the drastic changes of motion of the helicopter, often leading to unsuccessful tracking when the particle generation is not relevant. In particular, results of the AUX-PF are explained by the highly complex dynamics of the object: The mean state used in the simulation process does not provide sufficient information, and the resampled set of particles favors very local high responses of the likelihood [2]. FS-PF presents good results, since the model is flexible and allows dealing with strong changes of dynamics, which illustrates the relevance of our fuzzy framework. The PFS-PF shows the benefit of using the current observation in the simulation process, which improves the results, particularly when the number of particles is low.

Fig. 6 shows standard deviations of estimators given by the different filters. We can appreciate the robustness of the proposed approaches compared with the other ones. Fig. 7 shows estimated trajectories obtained on a typical run with $N = 50$ particles, and Fig. 8 shows captions of the associated tracking performed by the proposed approaches, the comparative models, and the ground truth with $N = 50$ particles. As we can see, SIR, AUX-PF, MMPF, and AUX-MMPF give tracking failures at about frames 30 and 75. This is due to the discontinuity of the dynamics when the cameraman is moving and the low number of particles used, whereas fuzzy models are more robust.

The robustness of the tracking process is the main challenge in this video sequence since the dynamics of the helicopter and the cameraman make this task really complex. This explains

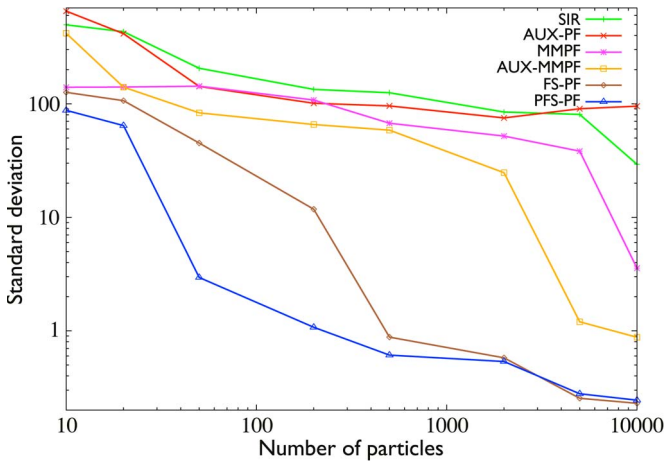

Fig. 6. Helicopter sequence: Standard deviation of estimators given by the different filters depending on the number N of particles. Both axes are logarithmic.

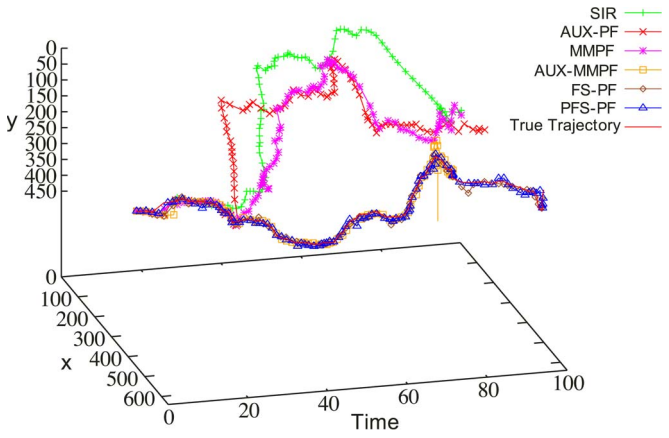

Fig. 7. Helicopter sequence: Estimated trajectories obtained by the different filters on a run with $N = 50$.

the high standard deviations observed in Fig. 6 for the SIR, AUX-PF, and MMPF models, even if the number of particles is large. Hence, RMSE results are strongly conditioned by failure or success of the tracking process. For the proposed approaches, the fuzzy partition used in this experiment describes a large searching space, which limits tracking failures since it allows the handling of strong dynamics discontinuities.

Fig. 9 shows the computational costs of the different filters given their RMSEs. Here, again, the proposed methods perform better in reasonable computation times. The benefit of the PFS-PS over the FS-PF is higher when using a number of particles lower than 300. This means that the PFS-FS method is interesting when a rate of more than 1 image/s should be achieved during the tracking process.

VII. Application 2: Tracking of the Lips in Face Image Sequences

To show the genericity of our model and the flexibility of our approach, we now present an application of shape tracking using the proposed fuzzy spatial information framework. The shape, shown here by the lips, is represented by a fuzzy set defined without a restrictive mathematical model, which allows the handling of deformable-shape models. Moreover, the fuzzy sets are learned in a generic way. The learning and test sets come from an annotated database freely available on the Internet[2] (Fig. 10).

The state vector $\mathbf{x}_t = (x_t, y_t)$ holds the mouth center coordinates at time t. The parameter vector $\mathbf{r}_t = (\dot{x}_t, \dot{y}_t, \theta_t, a_t)$ integrates the displacement vector $(\dot{x}_t, \dot{y}_t)$, the object orientation θ_t, and a vector of control points $a_t = (a_t^1, \ldots, a_t^P)$. We consider the constant velocity dynamics model defined in (32). Parameters of velocity, orientation, and lip shape are modeled in a flexible way by defining their possible values by fuzzy sets

$$\mathbf{c}_t^1 \in \{NW, N, NE, W, Center, E, SW, S, SE\}$$
$$\mathbf{c}_t^2 \in \{N, NE, E, SE, S\}$$
$$\mathbf{c}_t^3 \in \{Close, Open, Smile\}.$$

Hence, the vector of displacement can belong to nine fuzzy sets, from *North West* to *South East*, the object orientation, to five from *North* to *South*, and the control points of the fuzzy postures to three fuzzy sets *Close*, *Open*, and *Smile*.

A. Modeling the Fuzzy Functions

The velocity and the orientation of the object are modeled in the same way as in Section VI. The first one is characterized by a fuzzy partition of $[-h_{\mathrm{x}}, h_{\mathrm{x}}] \times [-h_{\mathrm{y}}, h_{\mathrm{y}}]$, with h_{x} and h_{y} as the fixed horizon parameters, whereas the second one is characterized by a partition of $[-\pi/2, \pi/2]$. The intersection degree is also similarly designed (33).

The set of fuzzy functions of lip shapes are obtained automatically by learning. Let $BD = \{(a^{(m)}, c^{(m)}), m = 1, \ldots, M\}$ be a database of elements $a^{(m)} \in \mathbb{R}^{2P}$ of P 2-D control points annotated by a label $c^{(m)} \in \{1, \ldots, K\}$. The learning algorithm of the fuzzy functions is summarized as follows.

1) database creation:
   a) collection of the M samples $\{(a^{(m)}, c^{(m)}), m = 1, \ldots, M\}$;
   b) rotation of all shapes performed to align their principal axes;
   c) distribution of the samples by classes.
2) creation of the fuzzy functions $\{g_{a^{(m)}}\}_{m=1}^{M}$ associated to the samples (6);
3) creation of the fuzzy functions of the classes $(\mu_k^3)_{k=1}^{K^3}$ (7);
4) approximation of the distribution of (13) with the method described in Section IV-F, if needed.

In the proposed application, we use a disjunctive intraclass fusion, with t-conorms ($\Phi = \max$), without interclass fusion strategy (hence, $\mu_k = \tilde{\mu}_k$). The particular set $\mathcal{Z}^3$ of the possible events has been defined using a union strategy (Section IV-E). This leads to classes with possibilistic semantics, which seems to be well adapted, among others, to learning applications: Since the database represents an actual observation, each element is possible (realizable).

We then introduce shape constraints into fuzzy functions $g_{a^{(m)}}$ in order to be able to define a fuzzy class which is consistent with the considered shapes. To define shape constraints

[2]http://personalpages.manchester.ac.uk/staff/timothy.f.cootes/data/talking_face/talking_face.html.

Fig. 8. Captions of surrounding ellipses for $N = 50$ particles at times 12, 18, 23, 28, 75, and 78. (Green) SIR, (red) AUX-PF, (purple) MMPF, (yellow) AUX-MMPF, (brown) FS-PF (proposed model without importance distribution), and (blue) PFS-PF (proposed model with importance distribution).

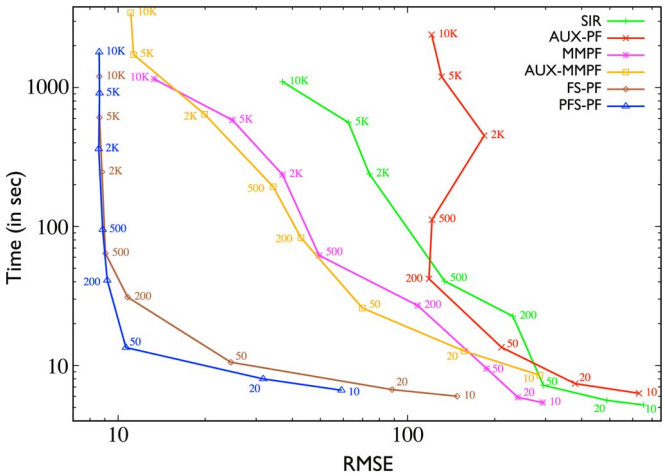

Fig. 9. Helicopter sequence: Computational costs (in seconds) depending on root mean square tracking errors obtained by the different filters. Both axes are logarithmic.

Fig. 10. Example of an annotated image from the database.

$C_{a^{(m)}}$, several strategies may be employed. For example, we can define them in a generic way by

$$C_{a^{(m)}}(o) = \Lambda\left(d\left(o, a^{(m)}\right)\right)$$

with $\Lambda : \mathbb{R}^+ \to [0,1]$ as a decreasing fuzzy function and $d : \mathbb{R}^{2P} \times \mathbb{R}^{2P} \to \mathbb{R}^+$ as a distance between two shapes. To guarantee a quasi-identical global shape to the sample $a^{(m)}$ while

Fig. 11. Lip simulations according to $\mathbb{P}(d\mathbf{r}^3|\mathbf{c}_t^3 = 3, \mathcal{Z}^3)$ with (a) a restrictive template support, (b) a large template support, and (c) a large template support with shape constraints. The red rectangle corresponds to the support of a control point ($P = 18$).

TABLE I
INTERSECTION DEGREE MATRIX WITH (a) A RESTRICTIVE TEMPLATE SUPPORT, (b) A LARGE TEMPLATE SUPPORT, AND (c) A LARGE TEMPLATE SUPPORT WITH SHAPE CONSTRAINTS. COLUMNS AND ROWS CORRESPOND TO FUZZY-SET INDICES

$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \text{(a)} \quad \begin{bmatrix} 1 & 0.11 & 0.01 \\ 0.11 & 1 & 0.07 \\ 0.01 & 0.07 & 1 \end{bmatrix} \text{(b)} \quad \begin{bmatrix} 1 & 0.20 & 0.02 \\ 0.20 & 1 & 0.16 \\ 0.02 & 0.16 & 1 \end{bmatrix} \text{(c)}$$

allowing scaling changes and local minor changes, we define the shape distance by

$$d(x,y) = \frac{1}{P}\sum_{p=1}^{P}\left|\widehat{(x^{p-1}, x^p, x^{p+1})} - \widehat{(y^{p-1}, y^p, y^{p+1})}\right|$$

with $\widehat{(a,b,c)}$ as the angle between the vectors $\vec{ab}$ and $\vec{ab}$. Other properties may be obtained using other kinds of distances. For example, a distance based on the histogram of gradient directions involves a response which is independent of the rotation, translation, and scaling of the object.

The use of a permissive fuzzy template $P_{a^{(m)}}$, for example, corresponding to a fuzzy ball, and strong constraints $C_{a^{(m)}}$ enables one to obtain a spread class with coherent shapes. Examples of lip simulations are shown in Fig. 11, using different template parameters and shape constraints and leading to different intersection degrees (Table I), which are directly connected to the transition between classes (17). We use the following intersection degree between classes k and h:

$$\Xi\left(\mu_k^3, \mu_h^3\right) = \frac{\sum_{l=1}^{L} \Upsilon\left[\mu_k^3\left(\mathbf{r}^{(l)}\right), \mu_h^3\left(\mathbf{r}^{(l)}\right)\right]}{\min\left[\sum_{l=1}^{L} \mathbb{1}^k\left(\mathbf{r}^{(l)}\right), \sum_{l=1}^{L} \mathbb{1}^h\left(\mathbf{r}^{(l)}\right)\right]}$$

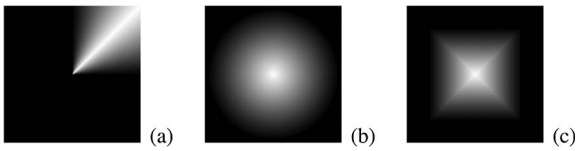

Fig. 12. Examples of fuzzy spatial relations (a) orientation *NorthEast* and position *Center* with (b) a fuzzy ball and (c) a fuzzy square.

with $\Upsilon[a,b]$ as a fuzzy operator such that $\Upsilon[a,b]=1$ if $a\times b\neq 0$ and 0 otherwise and $\mathbb{1}^k(x)=\Upsilon[\mu_k(x),1]$ as the indicator function of the class k. The intersection degree designed here is soft, since it emphasizes the intersection degree between two classes to make the transitions defined in (17) easier. Results in Table I show that the intersection degrees obtained with a large template support with shape constraints are higher than those obtained just with a large template support. Although this result may be unexpected, it can be explained by the denominator involved in the intersection-degree calculation. It means that the supports of fuzzy functions, although more restrictive, stay large when shapes are consistent with the fixed constraints and, more specifically, when membership areas are owned by several classes.

B. Likelihood Functions

Following the choices made in Section VI, we use a combined edge and region likelihood

$$p\left(\mathbf{y}_t|\mathbf{x}_t,\mathbf{r}_t,\mathbf{c}_t^3=k\right)=p\left(\mathbf{y}_t^R|\mathbf{x}_t,\mathbf{r}_t,\mathbf{c}_t^3=k\right)p\left(\mathbf{y}_t^C|\mathbf{x}_t,\mathbf{r}_t\right)$$

where $p(\mathbf{y}_t^C|\mathbf{x}_t,\mathbf{r}_t)$ corresponds to the edge-based likelihood defined in (34) computed along a B-spline interpolation of the contour points $(a_t^1,\ldots,a_t^P)$.

The region-based likelihood depends on the class indices $\mathbf{c}_t$. It is related to a distance between a mouth model histogram $h_k^\bullet=\{h_k^\bullet(u)\}_{u=1}^U$ of class k and a candidate histogram $h(\mathring{a}_t)=\{h(\mathring{a}_t;u)\}_{u=1}^U$ and to a distance between the candidate histogram and the image background histogram $h_0^\bullet=\{h_0^\bullet(u)\}_{u=1}^U$ [22]

$$p\left(\mathbf{y}_t^R|a_t,\mathbf{c}_t^3=k\right)\propto\exp\left(-\lambda_2\left(d^2\left[h_k^\bullet,h(\mathring{a}_t)\right]-d^2\left[h_0^\bullet,h(\mathring{a}_t)\right]\right)\right)$$

where the reference histograms $\{h_k^\bullet\}_{k=0}^3$ are created during the learning phase, for $k=1,2,3$, as the mean of the obtained histograms using fewer lips from the database and, for $k=0$, as the mean of the obtained histograms using fewer background images from the database.

C. Experiments

Fuzzy functions related to the lips are obtained using three classes of seven samples each. Tests are realized on the first 180 images. We have used a large template support with shape constraints [Fig. 11(c)] and with $P=18$ control points. The corresponding transition probabilities are described in Table I(c) and obtained without defining a minimal threshold between classes (8). Uniform fuzzy partitions of orientation and velocity are of triangular shape (spatial representation shown in Fig. 12), with a value of minimal intersection set to $\epsilon^{kh}=0.05$, allowing strong changes of parameter values.

TABLE II
LIP SEQUENCE: MEAN ESTIMATION ERROR OBTAINED DEPENDING ON THE NUMBER N OF PARTICLES

Model	Differences of areas (in %)					
	20	50	200	500	2000	5000
FS-PF	35.8	26.3	21.8	20.1	19.4	19.3

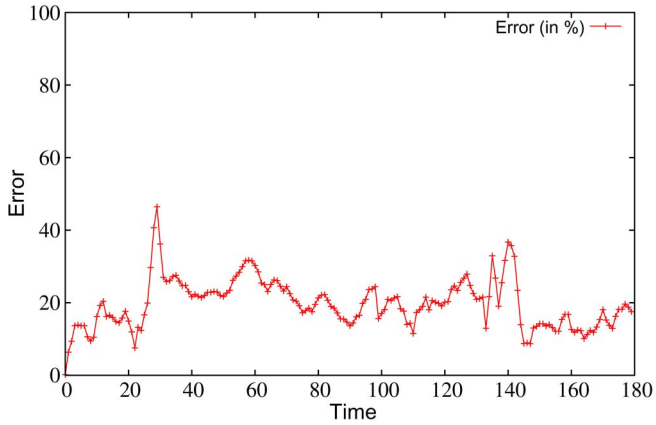

Fig. 13. Shape overlap error with $N=5000$ particles on 180 images. The error is computed as the difference areas between the result and the ground truth (in percentage).

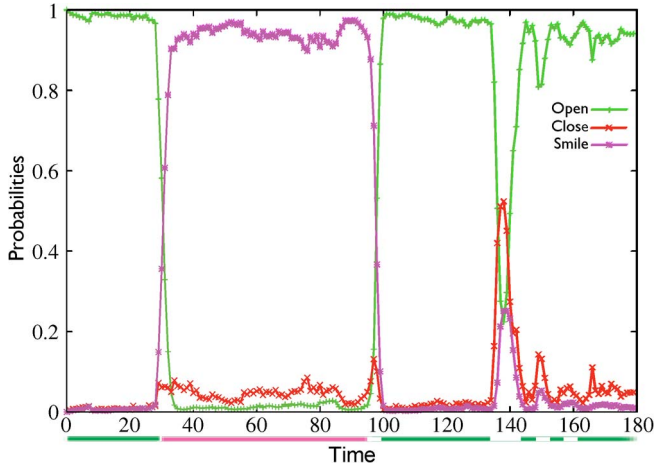

Fig. 14. Posterior probabilities of the classes with $N=5000$ particles and 180 images.

The position noise vector follows $\mathcal{N}(\mathbf{0},4\times\mathbf{I}_{2\times2})$, with $\mathbf{0}=(0\ 0)^{\mathrm{T}}$ and $\mathbf{I}_{2\times2}$ as the 2×2 identity matrix, and the velocity horizons were fixed to $h_{\mathrm{x}}=h_{\mathrm{y}}=20$. The shape of the fuzzy functions for the velocity and orientation parameters is triangular.

For the likelihood, we chose an HSV color space, with a histogram of length $U=U_T\times U_S\times U_V$, where $U_T=U_S=U_V=10$ is the number of bins by component. The multiplicative constants λ_1 and λ_2 are set to 20.

The importance distribution was not used in this section, although it was possible, since the purpose here is to illustrate the modeling potential of this approach. Then, considering the difficulty of the test sequence, dynamics parameters do not need an elaborated simulation process.

Table II presents mean errors of shape overlaps obtained using a growing number of particles, whereas Fig. 13 shows observed overlap shape errors during the complete sequence

Fig. 15. Mouth contour and class estimations at times (a) $t = 11$ (close), (b) $t = 66$ (smile), (c) $t = 140$ (open), and (d) $t = 146$ (close).

with $N = 5000$ particles. The degree of overlap corresponds to the ratio of the common points of a candidate shape CS and the true shape TS to the maximum of the two shape areas, i.e.,

$$\frac{\sum_{p\in\Omega_I} \mathbb{1}_{\mathrm{CS}}(p)\mathbb{1}_{\mathrm{TS}}(p)}{\max\left[\sum_{p\in\Omega_I} \mathbb{1}_{\mathrm{CS}}(p), \sum_{p\in\Omega_I} \mathbb{1}_{\mathrm{TS}}(p)\right]}$$

with Ω_I as the domain of image I and $\mathbb{1}_S(p)$ as the indicator function equal to 1 if the point p belongs to an object described by the shape S and equal to 0 otherwise.

The estimated marginal probabilities of the classes are shown in Fig. 14. The comparison with ground truth shows that our approach obtains, in majority, the good classes. Although we can observe a slight growing of the *open* value, obtained classes at steps $\sim$150 and $\sim$160 fail since this period corresponds to frames in which the volunteer opened his mouth just a little. The ground truth does not integrate the information of the opening degree of the mouth, which explains this result. Finally, the Monte Carlo expected value obtained in the test sequence is illustrated at several steps in Fig. 15.

VIII. Conclusion and Discussion

We have proposed in this paper an original approach that integrates fuzzy spatial information into particle filters. The proposed approach is generic and allows the handling of different kinds of information. To illustrate the method, we presented three applications, the first two ones only defining structural spatial information, known and fixed, and the other defining structural information and fuzzy learned shapes. We proposed a simple learning method for constrained fuzzy shapes, which can be adapted to any kind of shape, since the shape descriptors are basic. Then, the fuzzy model is introduced in the estimation of dynamics parameters in the particle filter. We also proposed an importance distribution which depends on the trajectory of fuzzy sets and new observations. The trajectory allows us to deal with outliers whereas observations attract particles toward high likelihood density. The synthetic example and the two proposed applications illustrate the flexibility of our model, since it can be easily extended to several kinds of spatial information. Tests realized on a complex dynamics sequence have shown that our approach outperforms reference particle filters, successfully dealing with complex dynamics, particularly for a small number of particles.

Nevertheless, the proposed framework has some limitations. First, the parameter space induced by the fuzzy set of possible values is, by definition, bounded, which may limit some applications. The second point concerns the importance distribution, which may be computationally expensive if not carefully designed, as its complexity is related to the dimension of the vector of parameters and the number of classes by parameter. We proposed a number of optimization and approximation steps to overcome this limit without degrading the results. However, if one has strict computational constraints and wants to reach the convergence performances, the use of our framework without the importance distribution may be more advised, as the complexity is then reduced to the SIR filter one. Finally, the proposed models consider, in essence, an independence assumption between dynamics parameters, which may not be suitable for all situations.

Integration of other kinds of fuzzy information in the proposed framework and multiobject aspects may be interesting orientations for future work. This may be made by modeling spatial relations between objects, such as distance or orientation concepts, or with more complex scene descriptors such as the object relation "between" proposed in [6].

References

[1] E. Arnaud and E. Mémin, "Partial linear Gaussian models for tracking in image sequences using Sequential Monte Carlo methods," *Int. J. Comput. Vis.*, vol. 74, no. 1, pp. 75–102, Aug. 2007.
[2] S. Arulampalam, S. Maskell, and N. Gordon, "A tutorial on particle filters for online nonlinear/non-Gaussian Bayesian tracking," *IEEE Trans. Signal Process.*, vol. 50, no. 2, pp. 174–188, Feb. 2002.
[3] S. M. Arulampalam, B. Ristic, N. Gordon, and T. Mansell, "Bearings-only tracking of manoeuvring targets using particle filters," *EURASIP J. Appl. Signal Process.*, vol. 2004, pp. 2351–2365, Jan. 2004.
[4] I. Bloch, "Information combination operators for data fusion: A comparative review with classification," *IEEE Trans. Syst., Man, Cybern.*, vol. 26, no. 1, pp. 52–67, Jan. 1996.
[5] I. Bloch, "Fuzzy spatial relationships for image processing and interpretation: A review," *Image Vis. Comput.*, vol. 23, no. 2, pp. 89–110, Feb. 2005.
[6] I. Bloch, O. Colliot, and R. Cesar, "On the ternary spatial relation between," *IEEE Trans. Syst., Man, Cybern. B, Cybern.*, vol. 36, no. 2, pp. 312–327, Apr. 2006.
[7] W. Y. Chang, C. S. Chen, and Y. D. Jian, "Visual tracking in high-dimensional state space by appearance-guided particle filtering," *IEEE Trans. Image Process.*, vol. 17, no. 7, p. 14, Jul. 2008.
[8] Z. Chen, "Bayesian filtering: From Kalman filters to particle filters, and beyond," Tech. Rep., McMaster Univ., Hamilton, ON, Canada, 2003.
[9] A. Doucet and A. M. Johansen, "A tutorial on particle filtering and smoothing: Fifteen years later," Tech. Rep., Dept. Statistics, Univ. British Columbia, Vancouver, BC, Canada, 2008.
[10] A. Doucet, S. Godsill, and C. Andrieu, "On sequential Monte Carlo sampling methods for Bayesian filtering," *Statist. Comput.*, vol. 10, no. 3, pp. 197–208, Jul. 2000.
[11] A. Doucet, N. De Freitas, and N. Gordon, *Sequential Monte Carlo Methods in Practice*. New York: Springer-Verlag, 2001.
[12] D. Dubois and H. Prade, *Fuzzy Sets and Systems: Theory and Applications*. New York: Academic, 1980.
[13] N. J. Gordon, D. J. Salmond, and A. F. Smith, "Novel approach to nonlinear/non-Gaussian Bayesian state estimation," *Proc. Inst. Elect. Eng.—F, Radar Signal Process.*, vol. 140, no. 2, pp. 107–113, Apr. 1993.

[14] M. Isard and A. Blake, "CONDENSATION—Conditional density propagation for visual tracking," *Int. J. Comput. Vis.*, vol. 29, no. 1, pp. 5–28, Aug. 1998.
[15] M. Isard and A. Blake, "A mixed-state condensation tracker with automatic model-switching," in *Proc. 6th Int. Conf. Comput. Vis.*, 1998, pp. 94–101.
[16] R. E. Kalman, "A new approach to linear filtering and prediction problems," *Trans. ASME, J. Basic Eng.*, vol. 82, pp. 35–45, 1960.
[17] R. Karlsson and N. Bergman, "Auxiliary particle filters for tracking a maneuvering target," in *Proc. 39th IEEE Conf. Decision Control*, 2000, vol. 4, pp. 3891–3895.
[18] J. S. Liu, "Metropolized independent sampling with comparisons to rejection sampling and importance sampling," *Statist. Comput.*, vol. 6, no. 2, pp. 113–119, Jun. 1996.
[19] S. McGinnity and G. W. Irwin, "Multiple model bootstrap filter for maneuvering target tracking," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 36, no. 3, pp. 1006–1012, Jul. 2000.
[20] K. Okuma, A. Taleghani, O. De Freitas, J. J. Little, and D. G. Lowe, "A boosted particle filter: Multitarget detection and tracking," in *Proc. ECCV*, 2004, pp. 28–39.
[21] V. Pavlovic, J. M. Rehg, and J. Maccormick, "Learning switching linear models of human motion," in *Proc. Neural Inf. Process. Syst.*, 2000, pp. 981–987.
[22] P. Pérez, C. Hue, J. Vermaak, and M. Gangnet, "Color-based probabilistic tracking," in *Proc. 7th Eur. Conf. Comput. Vis.—Part I*, 2002, pp. 661–675.
[23] M. Pitt and N. Shephard, "Filtering via simulation: Auxiliary particle filters," *J. Amer. Statist. Assoc.*, vol. 94, no. 446, pp. 590–599, Jun. 1999.
[24] Y. Rathi, N. Vaswani, and A. Tannenbaum, "A generic framework for tracking using particle filter with dynamic shape prior," *IEEE Trans. Image Process.*, vol. 16, no. 5, pp. 1370–1382, May 2007.
[25] B. L. Welch, "The generalization of 'Student's' problem when several different population variances are involved," *Biometrika*, vol. 34, no. 1/2, pp. 28–35, 1947.
[26] A. Yilmaz, O. Javed, and M. Shah, "Object tracking: A survey," *ACM Comput. Surv.*, vol. 38, no. 4, 2006, article 13.
[27] L. A. Zadeh, "The concept of a linguistic variable and its application to approximate reasoning," *Inf. Sci.*, vol. 8, no. 3, pp. 199–249, 1975.
[28] L. A. Zadeh, "Probability measures of fuzzy events," *J. Math. Anal. Appl.*, vol. 23, no. 2, pp. 421–427, 1968.

Nicolas Widynski was born in France in 1983. He received the Engineering degree from Ecole Pour l'Informatique et les Techniques Avancées, Le Kremlin-Bicêtre, France, and the Image Processing Master's degree (IMA) from University Pierre and Marie Curie (Paris 6), Paris, France, in 2007. He is currently working toward the Ph.D. degree at the Laboratory of Computer Sciences (UPMC-LIP6), University Pierre and Marie Curie (Paris 6).

He is also with Télécom ParisTech, Paris.

Séverine Dubuisson (M'06) was born in 1975. She received the Ph.D. degree in system control from the Compiègne University of Technology, Compiègne, France, in 2001.

Since 2002, she has been an Associate Professor with the Laboratory of Computer Sciences, University Pierre and Marie Curie (Paris 6), Paris, France. Her research interests include computer vision, probabilistic models for video sequence analysis, and tracking.

Isabelle Bloch (M'94) received the Bachelor's degree from the Ecole des Mines de Paris, Paris, France, in 1986, the Master's degree from the University Paris 12, Paris, in 1987, the Ph.D. degree from the Ecole Nationale Supérieure des Télécommunications (Telecom ParisTech), Paris, in 1990, and the Habilitation degree from the University Paris 5, Paris, in 1995.

She is currently a Professor with the Signal and Image Processing Department, Telecom ParisTech, in charge of the Image Processing and Understanding Group. Her research interests include 3-D image and object processing; computer vision; 3-D fuzzy logic; mathematical morphology; information fusion; fuzzy set theory; structural, graph-based, and knowledge-based object recognition; spatial reasoning; and medical imaging.