Computer Vision and Image Understanding 116 (2012) 1076–1094

Contents lists available at SciVerse ScienceDirect

Computer Vision and Image Understanding

journal homepage: www.elsevier.com/locate/cviu

Fuzzy spatial constraints and ranked partitioned sampling approach for multiple object tracking ☆

Nicolas Widynski [a,b,*], Séverine Dubuisson [a], Isabelle Bloch [b]

[a] UPMC, CNRS LIP6, Paris, France
[b] Télécom ParisTech, CNRS LTCI, Paris, France

ARTICLE INFO

Article history:
Received 3 August 2011
Accepted 17 July 2012
Available online 25 July 2012

Keywords:
Multiple object tracking
Particle filter
Fuzzy spatial constraints
Partitioned Sampling

ABSTRACT

While particle filters are now widely used for object tracking in videos, the case of multiple object tracking still raises a number of issues. Among them, a first, and very important, problem concerns the exponential increase of the number of particles with the number of objects to be tracked, that can make some practical applications intractable. To achieve good tracking performances, we propose to use a Partitioned Sampling method in the estimation process with an additional feature about the ordering sequence in which the objects are processed. We call it *Ranked Partitioned Sampling*, where the optimal order in which objects should be processed and tracked is estimated jointly with the object state. Another essential point concerns the modeling of possible interactions between objects. As another contribution, we propose to represent these interactions within a formal framework relying on fuzzy sets theory. This allows us to easily model spatial constraints between objects, in a general and formal way. The association of these two contributions was tested on typical videos exhibiting difficult situations such as partial or total occlusions, and appearance or disappearance of objects. We show the benefit of using conjointly these two contributions, in comparison to classical approaches, through multiple object tracking and articulated object tracking experiments on real video sequences. The results show that our approach provides less tracking errors than those obtained with the classical Partitioned Sampling method, without the need for increasing the number of particles.

1. Introduction

Since the 1990's, the particle filter has been widely used in the mono-object tracking community, thanks to its natural dispositions for tracking purposes, its reliability to deal with non-linear systems, and its easiness of implementation. However, extending this methodology to multiple object tracking is not straightforward. Usually estimating more objects needs substantially more particles, the association problem between measures and objects has to be solved, and interactions between objects should be modeled.

The adaptation of particle filters to track several objects has been extensively addressed in the literature, in many different ways. In [1], the authors propose a Jump Markov System (JMS) that models and jointly estimates the object states, the number of objects to track, and the association hypotheses between measures and objects. In [2], the particle filter integrates interactions between objects and measures using a Joint Probabilistic Data Association Filter (JPDAF), that provides an optimal data solution in the Bayesian framework. In [3] the distribution of association hypotheses is computed using a Gibbs sampler. The authors in [4,5] use a Joint Multitarget Probability Density (JMPD) to estimate the number of objects and their states. In [6], the authors model the filtering distribution as a mixture model to handle multiple objects, and use an Adaboost procedure to detect objects leaving and entering the images. Two multi-object algorithms are proposed in [7], namely the Sequential Sampling Particle Filter (SSPF), which individually generates objects using a factorization of the importance weights, and the Independent Partition Particle Filter (IPPF), which considers that the associations between objects and measures are independent over the individual objects.

A major issue with the importance sampling, used in particular in particle filters, is that it suffers from the problem of the curse of dimensionality [8,9]. This means that the particle filter requires a number of particles that increases exponentially with

☆ This paper has been recommended for acceptance by Y. Aloimonos.

* Corresponding author. Address: University of Montreal, Department of Computer Science and Operations Research (DIRO), Montreal, Canada.

E-mail addresses: Nicolas.Widynski@gmail.com (N. Widynski), Severine.Dubuisson@lip6.fr (S. Dubuisson), Isabelle.Bloch@telecom-paristech.fr (I. Bloch).

1077-3142/$ - see front matter © 2012 Elsevier Inc. All rights reserved.
http://dx.doi.org/10.1016/j.cviu.2012.07.002

the number of objects, making the use of a particle filter for multiple object tracking intractable as soon as the number of objects is greater than three. Therefore, the authors in [10,11] proposed a particle filter that avoids this additional cost using a Partitioned Sampling strategy. In [10], they consider an exclusion principle to handle the data association problem (i.e. specifying that a measurement may be associated with at most one object). The principle of Partitioned Sampling is to partition the state space, by considering one element of the partition per object. The objects are processed in a prefixed order, that we call *scenario*, and particles that are the most likely to fit with the real state of the object are selected using a weighted sampling approach. The considered order matters since it can lead to unsuitable behaviors of the filter, such as loosing tracks, for example when the first considered object is hidden. In fact, as pointed in [12], the use of a prefixed order in the scenario for the joint state estimation can cause a lot of problems, caused by a particle cloud impoverishment effect. To avoid this, in [12], the filtering distribution is modeled as a mixture law, in which each component designates a specific order of processing of the objects, that is estimated using a Partitioned Sampling technique. This idea was also used in [13] to merge different features of an object. However, the number of particles used for each component is fixed, that can limit the performances of the filter if the order of processing is not optimal [12].

Handling several objects in a particle filter raises another type of problem. It is often necessary to model possible interactions between objects, in order to estimate jointly object states. Most approaches make specific hypotheses (see for example [14]), that are directly related to the application, making the model not generalizable. Generally speaking, spatial relations constitute an important type of structural information, useful in scene description and interpretation, as acknowledged in various domains. Although most relations have a clear intuitive meaning, expressing them mathematically is not obvious because they may be vague or imprecise. This is for instance the case for relative directions such as *to the right of*. Modeling such relations in a fuzzy set framework is then appropriate [15] since it allows considering them as a matter of degree, which can be tuned according to the context. For instance saying that two objects are *far* from each other depends on the application context, on the objects, and on the image resolution. Several definitions of fuzzy models of spatial relations have been proposed (see e.g. [16] for a review), and used in structural object recognition and image interpretation [17–19], data clustering [20], or single object tracking [21]. This type of information, to our knowledge, has however not been used so far in a multiple object tracking procedure (spatial constraints have only been used in specific ways, in a non fuzzy formalism [14]), and we show here that it is interesting to do so.

In this paper, that extends a preliminary work in [22], we propose to integrate fuzzy spatial constraints into the particle filter framework for multiple object tracking and to jointly estimate object states and their optimal processing order. We call this approach *Ranked Partitioned Sampling*. This allows us to consider the whole set of possible orders and to automatically prune irrelevant scenarios. We consider here that the number M of objects is known (but can vary over time), and we handle automatically possible hidden parts of objects by other ones.

This article is organized as follows: The particle filter for multiple object tracking is first presented in Section 2. The Partitioned Sampling procedure proposed in [10,11] is described in Section 3. Section 4 presents the fuzzy spatial constraint framework and its introduction in the particle filter framework, as our first contribution. The Ranked Partitioned Sampling, which constitutes the second main contribution of this paper, is then proposed in Section 5, and some considerations about the visibility of objects are given in Section 6. We show experimental results in Section 7, before concluding in Section 8.

2. Particle filter for multiple object tracking

2.1. Particle filter

Let us consider a classical filtering problem and denote by $\mathbf{x}_t \in \mathscr{X}$ the hidden state at time t of an object to be tracked, and by $\mathbf{y}_t \in \mathscr{Y}$ the measurement state extracted from the image sequence. The system describing the temporal evolution of $\mathbf{x}_t$ and the measurement equation are defined as follows:

$$\mathbf{x}_t = f_t(\mathbf{x}_{t-1}, \mathbf{v}_t) \tag{1}$$

$$\mathbf{y}_t = h_t(\mathbf{x}_t, \mathbf{w}_t) \tag{2}$$

where f_t models the non-linear temporal evolution of $\mathbf{x}_t$, h_t is the non-linear measurement equation, and $\mathbf{v}_t$ and $\mathbf{w}_t$ are independent white noises. The non-linear Bayesian tracking consists in estimating the *posterior* filtering density function $p(\mathbf{x}_t|\mathbf{y}_{1:t})$ (where $\mathbf{y}_{1:t}$ denotes the series of measures from time 1 to time t), expressed by:

$$p(\mathbf{x}_t|\mathbf{y}_{1:t}) = \frac{p(\mathbf{y}_t|\mathbf{x}_t)\ p(\mathbf{x}_t|\mathbf{y}_{1:t-1})}{\int_{\mathscr{X}} p(\mathbf{y}_t|\mathbf{x}'_t)\ p(\mathbf{x}'_t|\mathbf{y}_{1:t-1}) d\mathbf{x}'_t} \tag{3}$$

with $p(\mathbf{x}_t|\mathbf{y}_{1:t-1})$ the prediction (or *prior*) density function defined as:

$$p(\mathbf{x}_t|\mathbf{y}_{1:t-1}) = \int_{\mathscr{X}} p(\mathbf{x}_t|\mathbf{x}_{t-1})\ p(\mathbf{x}_{t-1}|\mathbf{y}_{1:t-1}) d\mathbf{x}_{t-1} \tag{4}$$

When the filtering density function cannot be computed in a closed form, i.e. when the system is non-linear and non-Gaussian, particle filters are used to approximate it by a weighted sum of N Dirac masses $\delta_{\mathbf{x}_t^{(n)}}(d\mathbf{x}_t)$ centered on hypothetic state realizations $\left\{\mathbf{x}_t^{(n)}\right\}_{n=1}^{N}$ of the state $\mathbf{x}_t$, also called particles [23]. Then, the filtering distribution $\mathbb{P}(d\mathbf{x}_t|\mathbf{y}_{1:t})$ is recursively approximated by the empiric distribution $P_N(d\mathbf{x}_t|\mathbf{y}_{1:t}) = \sum_{n=1}^{N} w_t^{(n)} \delta_{\mathbf{x}_t^{(n)}}(d\mathbf{x}_t)$, where $\mathbf{x}_t^{(n)}$ is the nth particle and $w_t^{(n)}$ its weight. If an approximation of $\mathbb{P}(d\mathbf{x}_{t-1}|\mathbf{y}_{1:t-1})$ is known, the process is divided into three main steps:

1. The diffusion step consists in estimating $p(\mathbf{x}_t|\mathbf{y}_{1:t-1})$ by propagating the particle swarm $\left\{\mathbf{x}_{t-1}^{(n)}, w_{t-1}^{(n)}\right\}_{n=1}^{N}$ using an importance function $q\left(\mathbf{x}_t|\mathbf{x}_{0:t-1}^{(n)}, \mathbf{y}_t\right)$.
2. The update step then computes new particle weights using the new observation $\mathbf{y}_t$, as:
$$w_t^{(n)} \propto w_{t-1}^{(n)} \frac{p(\mathbf{y}_t|\mathbf{x}_t^{(n)}) p\left(\mathbf{x}_t^{(n)}|\mathbf{x}_{t-1}^{(n)}\right)}{q(\mathbf{x}_t|\mathbf{x}_{0:t-1}^{(n)}, \mathbf{y}_t)}, \text{ such that } \sum_{i=1}^{N} w_t^{(n)} = 1.$$
3. Resampling techniques are employed to avoid particle degeneracy problems, leading for instance to the classical Sequential Importance Resampling (SIR) filter [23].

2.2. Multiple object tracking using particle filter

When dealing with multiple objects, the previous model has to be adapted. The first proposed approach might be the one in [24], and consists in simply applying the SIR filter [23] to $\mathbf{x}_t$ defined as a concatenation of several objects states $\mathbf{x}_t = (\mathbf{x}_t^1, \ldots, \mathbf{x}_t^M)$, with $\mathbf{x}_t^i \in \mathscr{X}^i$ the unknown state of the ith object and M the fixed number of objects. The same process as the one described in Section 2.1 can then be used. However, a problem involved in multiple object tracking is that the likelihood cannot be directly computed. Indeed, the presence of several objects in the images induces several