Expert Systems with Applications 41 (2014) 3944–3954

Contents lists available at ScienceDirect

Expert Systems with Applications

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

Review

Fight sample degeneracy and impoverishment in particle filters: A review of intelligent approaches

Tiancheng Li [a,*], Shudong Sun [a], Tariq Pervez Sattar [b], Juan Manuel Corchado [c]

[a] School of Mechatronics, Northwestern Polytechnical University, Xi'an 710072, China
[b] Center for Automated and Robotics NDT, London South Bank University, London SE1 0AA, UK
[c] Bioinformatic, Intelligent Systems and Educational Technology (BISITE) / Biomedical Research Institute of Salamanca (IBSAL), University of Salamanca, Plaza de la Merced S/N, 37008 Salamanca, Spain

ARTICLE INFO

Keywords:
Particle filter
Sequential Monte Carlo
Markov Chain Monte Carlo
Impoverishment
Artificial intelligence
Machine learning

ABSTRACT

During the last two decades there has been a growing interest in Particle Filtering (PF). However, PF suffers from two long-standing problems that are referred to as sample degeneracy and impoverishment. We are investigating methods that are particularly efficient at Particle Distribution Optimization (PDO) to fight sample degeneracy and impoverishment, with an emphasis on intelligence choices. These methods benefit from such methods as Markov Chain Monte Carlo methods, Mean-shift algorithms, artificial intelligence algorithms (e.g., Particle Swarm Optimization, Genetic Algorithm and Ant Colony Optimization), machine learning approaches (e.g., clustering, splitting and merging) and their hybrids, forming a coherent standpoint to enhance the particle filter. The working mechanism, interrelationship, pros and cons of these approaches are provided. In addition, approaches that are effective for dealing with high-dimensionality are reviewed. While improving the filter performance in terms of accuracy, robustness and convergence, it is noted that advanced techniques employed in PF often causes additional computational requirement that will in turn sacrifice improvement obtained in real life filtering. This fact, hidden in pure simulations, deserves the attention of the users and designers of new filters.

1. Introduction

The Sequential Monte Carlo (SMC) approach allows inference of full posterior distributions via Bayesian filtering in general nonlinear state-space models where the noises of the model can be non-Gaussian. There has been great interest in applying the SMC approach to deal with a wide variety of nonlinear filtering problems. This method is normally called the Particle Filter(ing) (PF) Del Moral, 1996, also referred to as Sequential imputations (Liu & Chen, 1995), the Monte Carlo filter (Kitagawa, 1996), the Condensation filter (Isard & Blake, 1998), and the survival of fittest and the likelihood weighting algorithm (Kanazawa, Koller, & Russel, 1995). To date, particle filters have been successfully applied in different areas including finance (Lopes & Tsay, 2011), parameter estimation (Andrieu, Doucet, Singh, & Tadic, 2004; Kantas, Doucet, Singh, & Maciejowski, 2009), geophysical systems (Van Leeuwen, 2009), wireless communication (Djuric et al., 2003), decision making (Caesarendra, Niu, & Yang, 2010; Kostanjčar, Jeren, & Cerovec, 2009), tracking and defense (Mihaylova et al., 2014; Ristic, Arulampalam, & Gordon, 2004; Yin, Zhang, Sun, & Gu, 2011), robotics (Thrun, 2002) and some nontrivial applications (Andrieu et al., 2004; Johansen, 2006). Additionally, a variety of strategies have been proposed to improve the performance of the particle filter in terms of accuracy, convergence, computational speed, etc. The staged survey of different years can be seen; examples include 2000 (Doucet, Godsill, & Andrieu, 2000), 2002 (Arulampalam, Maskell, Gordon, & Clapp, 2002), 2007 (Cappé, Godsill, & Moulines, 2007), 2009 (Doucet & Johansen, 2009), 2010 (Gustafsson, 2010), etc. However, PF continues to suffer from two notorious problems: sample degeneracy and impoverishment, which is arguably a long-standing topic in the community. A variety of methods have been investigated to fight these two problems in order to combat the weakness of the particle filter.

This study does not purport to give either a comprehensive review of the development of general particle filters or its special applications. Both are covered in the aforementioned survey papers. Our aim is to investigate a group of emerging 'intelligent' ways employed within PF that have benefited from a variety of intelligent and heuristic algorithms. These variations of techniques, acting in different ways to optimize the spatial distribution of particles namely Particle Distribution Optimization (PDO), are particularly effective in alleviating sample degeneracy and impoverishment, forming a very systematic standpoint that is both mathematically sound and practically efficient to enhance PF. In addition, approaches that are effective in dealing with

* Corresponding author. Tel.: +86 029 8849 4701.
E-mail addresses: t.c.li@mail.nwpu.edu.cn, lit3@lsbu.ac.uk (T. Li), sdsun@nwpu.edu.cn (S. Sun), sattartp@lsbu.ac.uk (T.P. Sattar), corchado@usal.es (J.M. Corchado).

0957-4174/$ - see front matter © 2013 Elsevier Ltd. All rights reserved.
http://dx.doi.org/10.1016/j.eswa.2013.12.031

high-dimensional filtering, another obstacle for the SMC, are reviewed. This study aims to coordinate these developments into a unifying framework, unveiling their pros and cons and thereby directing further improvements of existing schemes. This survey is specifically expected to serve as the first comprehensive coverage of artificial intelligence and machine learning techniques applied in PF.

The basic background of PF is presented in Section 2 with emphasis on its two fundamental difficulties: sample degeneracy and impoverishment. These two difficulties have motivated the development of a variety of PDO approaches, which are reviewed in the categories identified in Section 3. Further discussions on the PDO framework including the computational efficiency and high dimensionality challenge are given in Section 4. The conclusion is given in Section 5.

2. Sample degeneracy and impoverishment

Before we proceed, we provide a brief review of PF and define the notation. The primary notations used are summarized in Table 1.

Nonlinear filtering is a class of signal processing that widely exists in engineering, and is therefore a very broad research topic. The solution of the continuous time filtering problem can be represented as a ratio of two expectations of certain functions of the signal process. However, in practice, only the values of the observation corresponding to a discrete time partition are available; the continuous-time dynamic system has to be converted into a discrete-time simulation model, e.g., discrete Markov System, by discretely sampling the outputs through discretization. This paper is concerned with the problem of discrete filtering, which can be described in the State Space Model (SSM) that consists of two equations:

$$x_t = g_t(t, x_{t-1}, u_{t-1}) \quad \text{(state transition equation)} \tag{1}$$

$$y_t = h_t(t, x_t, v_t) \quad \text{(observation equation)} \tag{2}$$

The filtering problem recursively solving the marginal posterior density $p(x_t \mid Y_t)$ can be determined by the *recursive Bayesian estimation*, which has two steps:

(1) Prediction

$$p(x_t|Y_{t-1}) = \int_{\mathbb{R}^{n_x}} p(x_t|x_{t-1})p(x_{t-1}|Y_{t-1})dx_{t-1} \tag{3}$$

(2) Updating or correction

$$p(x_t|Y_t) = \frac{p(y_t|x_t)p(x_t|Y_{t-1})}{\int_{\mathbb{R}^{n_x}} p(y_t|x_t)p(x_t|Y_{t-1})dx_t} \tag{4}$$

In (3) and (4), the integration of often unknown and maybe high-dimensional functions is required, which can be computationally very difficult. This makes analytic optimal solutions such as the Kalman filter intractable. One flexible solution is the Monte Carlo approach, as the topic of this paper, which uses random number generation to compute integrals. That is, the integral, expressed as an expectation of $f(x)$ over the density $p(x)$, is approximated by a number of random variables $x^{(1)}, x^{(2)}, \ldots x^{(N)}$ (called samples or particles) that are drawn from the density $p(x)$ (if possible), then

$$\hat{f} = \int f(x)p(x)dx = \mathrm{E}_{p(x)}[f(x)] \approx \frac{1}{N}\sum_{s=1}^{N} f(x^{(i)}), \quad x^{(i)} \sim p(x) \tag{5}$$

This is an unbiased estimate and, provided the variance of $f(x)$ is finite, it has a variance which is proportional to $1/N$.

However, one limitation in applying Monte Carlo integration (5) in Bayesian inference (3) and (4) is that sampling directly from $p(x)$ is difficult, even impossible, if high density regions in $p(x)$ do not match up $f(x)$ with areas where it has a large magnitude. A convenient solution for this is the *Importance Sampling* (IS). Assuming the density $q(x)$ roughly approximates the density (of interest) $p(x)$, then

$$\hat{f} = \int f(x)\left(\frac{p(x)}{q(x)}\right)q(x)dx = \mathrm{E}_{q(x)}\left\{f(x)\left(\frac{p(x)}{q(x)}\right)\right\} \tag{6}$$

This forms the basis of Monte Carlo IS which uses the weighted sum of a set of samples from $q(x)$ to approximate (6):

$$\hat{f} \approx \frac{1}{S}\sum_{s=1}^{S} f(x^{(s)})\left(\frac{p(x^{(s)})}{q(x^{(s)})}\right) \tag{7}$$

An alternative formulation of IS is to use

$$\hat{f} \approx \hat{I} = \sum_{i=1}^{N} w^{(i)} f(x^{(i)}) \Bigg/ \sum_{i=1}^{N} w^{(i)}, w^{(i)} = \frac{p(x^{(i)})}{q(x^{(i)})} \tag{8}$$

with the variance given by

$$\mathrm{var}[\hat{f}] = (w^{(i)})^2 \left[\int \frac{(f(x)p(x))^2}{q(x)}dx - \mathrm{E}_p\{f(x)\}\right] \tag{9}$$

where $x^{(i)}$ is drawn from the proposal density $q(x)$. The variance is minimized to zero if $p(x) = q(x)$ (Doucet, de Freitas, & Gordon, 2001; Rasmussen & Ghahramani, 2003). There are many potential choices for $q(x)$ leading to various integration and optimization algorithms, as shown in the summary provided in Del Moral, Doucet, and Jasra (2006). In general, $q(x)$ should have a relatively heavy tail so that it is insensitive to the outliers.

In the importance sampling the estimator not only depends on the values of $p(x)$ but also on the entirely arbitrary choice of the proposal density $q(x)$. This results in heavy dependence on irrelevant information $q(x)$. This *sampling difficulty* seems inevitable as the density $p(x)$ of interest is generally always unknown; therefore, it is impossible to direct the sampling. For this reason, some advanced important sampling methods have been proposed, such as annealed importance sampling (Radford, 2001), Bayesian importance sampling (Rasmussen & Ghahramani, 2003), adaptive importance sampling (Liu & West, 2001), numerically accelerated importance sampling (Koopman, Lucas, & Scharth, 2011), and nonparametric importance sampling (Neddermeyer, 2011). Alternatively, sampling strategies such as rejection sampling (Gilks, Richardson, & Spiegelhalter, 1996), block sampling (Doucet, Briers, & Sénécal, 2006), Markov Chain Monte Carlo (MCMC) sampling (Gilks & Berzuini, 2001; Del Moral et al., 2006) and factored sampling (Banerjee & Burlina, 2010; Isard & Blake, 1998) have also been used, in addition to ad hoc strategies such as multiple stages of important sampling (Li, Ai, Yamashita, Lao, & Kawade, 2008). SMC samplers are specifically developed to sample sequentially from a sequence of probability distributions that allows the

Table 1
Primary notations.

x_t	The state of interest at time t
X_t	$X_t \triangleq (x_0, x_1, \ldots x_t)$, the history path of the state
y_t	Observation at time t
Y_t	$Y_t \triangleq (y_0, y_1, \ldots y_t)$, the history path of the observation
$g_t(\cdot)$	The state transition equation at time t
$h_t(\cdot)$	The observation equation at time t
u_t	Noise affecting the system dynamic equation $g_t(\cdot)$, at time t
v_t	Noise affecting the observation equation $h_t(\cdot)$, at time t
$x_t^{(i)}$	The state of particle i, at time t
$w_t^{(i)}$	The weight of particle i, at time t
N_t	The total number of particles at time t
$\delta_x(\cdot)$	The delta-Dirac mass located in x
$N(.:a,b)$	Gaussian density with mean a and covariance b,
$K_h(\cdot)$	A kernel function with bandwidth h