Signal Processing: *Image Communication* 36 (2015) 140–153

Contents lists available at ScienceDirect

Signal Processing: *Image Communication*

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

Iterative particle filter for visual tracking

Zhenhua Fan, Hongbing Ji*, Yongquan Zhang

School of Electronic Engineering, Xidian University, PO BOX 229, Xi'an 710071, PR China

ARTICLE INFO

Article history:
Received 13 January 2015
Received in revised form
1 July 2015
Accepted 1 July 2015
Available online 9 July 2015

Keywords:
Visual tracking
Particle filter
Iteration
Annealing
Sampling efficiency

ABSTRACT

Particle filter (PF) has been the subject of considerable attention in visual tracking. How to approach the true target state with computation cost as low as possible has always been an important issue. In this paper, a novel iterative PF (IPF) is proposed, which can converge to the true target state as close as possible by sampling the particles iteratively with the search scope contracted. The search scope is iteratively contracted around the centers determined by the previous converging results. Compared with annealed PF (APF) and PF, IPF can converge much closer to the true target state, thus improvement in sampling efficiency, clutter elimination, and tracking accuracy with the same computational burden.

1. Introduction

Particle filter (PF) has attracted a great deal of attention and found very wide application in the field of visual tracking [1–5]. The key of PF [6] is to represent the posterior probability density function (PDF) of target state distribution by a set of random samples with associated weights, and compute the expectation as the estimation of target state. In visual tracking, it is difficult to directly obtain the target position measurements, but there is much appearance information of the target besides its motion parameters (such as position, velocity, etc.). Thus, how to use this information becomes a complicated problem. As a consequence, compared with the point target tracking only using motion parameters, the visual tracking uses both the motion parameters and the features by some descriptors, i.e., using the feature measurements to update the target state in the Bayesian estimation. Therefore, the result of PF is the target state estimation with a given feature descriptor. Obviously, for different feature descriptors, the corresponding "true" states are slightly different. The aim of PF is to approach the true state as close as possible. Therefore, the visual tracking involves two main aspects: (a) Find a feature descriptor to make a compromise between robustness and discrimination which are generally constrained each other. (b) Find a filtering method, or searching strategy, to approach the true state with the computation cost as low as possible. In this paper, the latter is mainly addressed.

PF is confronted with three major problems: (a) The importance density function selection [7]. (b) The loss of particle diversity. (c) The computational burden.

1.1. Related works

Although the optimal importance density is hard to be obtained, Crisan proposed that the importance density closer to the true state is conducive to improve the tracking accuracy [8]. In the scenarios usually considered, the process noise covariance is smaller than that of the measurement noise, and the prior distribution is used as importance distribution for particle sampling. On the contrary, if the measurement noise covariance is smaller than that of the process noise, we can directly sample particles around the measurement to get a more precise result.

* Corresponding author. Tel./fax: +86 29 88201658.
E-mail address: hbji@xidian.edu.cn (H. Ji).

http://dx.doi.org/10.1016/j.image.2015.07.001

Besides, resampling [9] not only avoids the unnecessary computation cost of the particles with small weights, but also increases the percentage of effective particles to improve tracking accuracy by reproducing the particles with large weights.

In addition, marginalized PF (MPF) [10], also known as Rao-Blackwellized PF [11,12], separates the linear state from the nonlinear state, and combines PF with Kalman filter (KF) [13]. In MPF, the linear state variables are estimated by KF and the nonlinear ones are estimated by PF. Since only the nonlinear ones are randomly sampled, MPF reduces the space dimension and avoids blindly sampling in a higher dimensional space [14], which increases the number of effective particles and promotes tracking accuracy.

Recently, box particle filter [15,16] addresses the concept of box particle that can make sampling from coarse to fine, improving the sampling efficiency.

To summarize the improved PF algorithms above, the common point is increasing the sampling density close to the true state, and achieving high tracking accuracy.

Additionally, in visual tracking, Mean Shift (MS) [17] is also a simple and effective method. However, the traditional MS may be trapped in the local optimal solution due to the bandwidth limitation. To solve this problem, Shen and Li presented the multi-bandwidth MS [18] and multi-scale MS [19], respectively. Analogously, Neal proposed annealed importance sampling [20], which gradually sharpens the weight function surface, gently introduces the influence of the global peak and converges to the global optimal solution finally. As employing annealed importance sampling, annealed particle filter (APF) is widely used in visual tracking [21–23].

1.2. Contributions

Motivated by the success of both increasing the sampling density and simulated annealing, we propose a novel iterative PF (IPF), which can converge to the true target state as close as possible by sampling the particles iteratively with the search scope contracted. The search scope is iteratively contracted around the centers determined by the previous converging results. The main differences between APF and IPF are as follows. (a) The simulated annealing object is different. In IPF, the search scope is annealed, while in APF the flatness of the weight function is annealed. (b) The number of iterations. In APF, the number of annealing iterations is constant and set artificially, while in IPF the number of iterations is only determined by accuracy requirements, and not necessarily constant in each convergence procedure.

The main contributions of this paper are as follows. (a) *Accuracy*. The annealing iteration is introduced to PF. With particle number unchanged, contracting the search scope is equivalent to increasing the sampling density, thus effectively improving the tracking accuracy. (b) *Time efficiency*. By iteratively optimal sampling, IPF increases the sampling density near the true state so that the particles can be sufficiently used. Thus, it can meet the same precision requirements with fewer particles, an effective improvement of the time efficiency. (c) *Robustness*. The contraction of the search scope can exclude the clutter, eliminating the influence of the clutter in the fusion estimation. Accordingly, the final estimation results can reflect the true target state more closely.

The remainder of this paper is organized as follows. PF and Gaussian PF (GPF) are briefly described in Section 2. IPF is presented in detail in Section 3. In Section 4, the simulation results are shown and discussed. Finally, we draw conclusions in Section 5.

2. Backgrounds

2.1. Bayesian estimation and PF

Many nonlinear estimation problems can be handled by PF [6]. A general nonlinear dynamic state-space model can be formulated as

$$\boldsymbol{x}_k = f_k(\boldsymbol{x}_{k-1}) + \boldsymbol{\omega}_{k-1}, \tag{1}$$

$$\boldsymbol{z}_k = h_k(\boldsymbol{x}_k) + \boldsymbol{\upsilon}_k \tag{2}$$

where $\boldsymbol{x}_k$ is the state at time k, $\boldsymbol{z}_k$ is the measurement, $f_k(\cdot)$ is a nonlinear transition function, $h_k(\cdot)$ is a nonlinear measurements function, $\boldsymbol{\omega}_{k-1}$ is an independent identically distributed (i.i.d.) process noise sequence, $\boldsymbol{\omega}_{k-1} \sim \mathcal{N}(\boldsymbol{\omega}_{k-1}; \mathbf{0}, \mathbf{Q}_{k-1})$, and $\boldsymbol{\upsilon}_k$ is an i.i.d. measurement noise sequence, $\boldsymbol{\upsilon}_k \sim \mathcal{N}(\boldsymbol{\upsilon}_k; \mathbf{0}, \boldsymbol{R}_k)$.

The state sequence $\{\boldsymbol{x}_\ell, \ell = 0, \ldots, k\}$ is an unobserved Markov process with a prior PDF $p(\boldsymbol{x}_0)$ and transitional density $p(\boldsymbol{x}_k | \boldsymbol{x}_{k-1})$. Within the Bayesian framework, the posterior state PDF $p(\boldsymbol{x}_k | \boldsymbol{z}_{1:k})$ provides a complete description of the state up to time k, given the measurements $\boldsymbol{z}_{1:k}$. $p(\boldsymbol{x}_k | \boldsymbol{z}_{1:k})$ may be obtained recursively in two stages: prediction and update. Suppose that the PDF $p(\boldsymbol{x}_{k-1} | \boldsymbol{z}_{1:k-1})$ at time $k-1$, transitional density $p(\boldsymbol{x}_k | \boldsymbol{x}_{k-1})$ and likelihood function $p(\boldsymbol{z}_k | \boldsymbol{x}_k)$ are available, and then the prediction PDF and the measurement update PDF can be expressed, respectively, as

$$p(\boldsymbol{x}_k | \boldsymbol{z}_{1:k-1}) = \int p(\boldsymbol{x}_k | \boldsymbol{x}_{k-1}) p(\boldsymbol{x}_{k-1} | \boldsymbol{z}_{1:k-1}) d\boldsymbol{x}_{k-1}, \tag{3}$$

$$p(\boldsymbol{x}_k | \boldsymbol{z}_{1:k}) = \frac{p(\boldsymbol{z}_k | \boldsymbol{x}_k) p(\boldsymbol{x}_k | \boldsymbol{z}_{1:k-1})}{\int p(\boldsymbol{z}_k | \boldsymbol{x}_k) p(\boldsymbol{x}_k | \boldsymbol{z}_{1:k-1}) d\boldsymbol{x}_k}. \tag{4}$$

Through the recursive estimation of posterior PDF, we can obtain the complete analytical form of sequential estimation, and estimate the target state.

PF, also known as sequential Monte Carlo (SMC), is the most representative nonlinear filtering practical method within the Bayesian framework. PF recursively approximates the posterior state PDF $p(\boldsymbol{x}_k | \boldsymbol{z}_{1:k})$ as a set of random sampling particles with associated weights $\{\boldsymbol{x}_k^i, w_k^i\}_{i=1}^N$, and computes the expectation $\sum_i w_k^i \boldsymbol{x}_k^i$ as the estimation of target state.

The weights are chosen by the principle of importance sampling [6]. Although $p(\boldsymbol{x}_{0:k} | \boldsymbol{z}_{1:k})$ is difficult to draw samples, particles can be easily generated from a proposal $q(\boldsymbol{x}_{0:k} | \boldsymbol{z}_{1:k})$ called an importance density. Thus, $p(\boldsymbol{x}_{0:k} | \boldsymbol{z}_{1:k})$ can be approximated as

$$p(\boldsymbol{x}_{0:k} | \boldsymbol{z}_{1:k}) \approx \sum_i w_k^i \delta\left(\boldsymbol{x}_{0:k} - \boldsymbol{x}_{0:k}^i\right) \tag{5}$$