ARTICLE IN PRESS

Neurocomputing ∎ (∎∎∎∎) ∎∎∎–∎∎∎

Contents lists available at ScienceDirect

Neurocomputing

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

Ordered over-relaxation based Langevin Monte Carlo sampling for visual tracking

Fasheng Wang [a,b], Peihua Li [a], Xucheng Li [c,d], Mingyu Lu [c,*]

[a] School of Information and Communication Engineering, Dalian University of Technology, No. 2, Linggong Road, Ganjingzi District, Dalian 116024, China
[b] Department of Computer Science, Dalian Neusoft University of Information, No. 8, Software Park Road, Ganjingzi District, Dalian 116023, China
[c] School of Information Science and Technology, Dalian Maritime University, No. 1, Linghai Road, Ganjingzi District, Dalian 116026, China
[d] Department of Software Engineering, Dalian Neusoft University of Information, No. 8, Software Park Road, Ganjingzi District, Dalian 116023, China

ARTICLE INFO

Article history:
Received 2 January 2016
Received in revised form
22 March 2016
Accepted 29 April 2016

Keywords:
Visual tracking
Bayesian filtering
Langevin Monte Carlo sampling
Ordered over-relaxation

ABSTRACT

Visual tracking is a fundamental research topic in computer vision community, which is of great importance in many application areas including augmented reality, traffic control, medical imaging and video editing. This paper presents an ordered over-relaxation Langevin Monte Carlo sampling (ORLMC) based tracking method within the Bayesian filtering framework, in which the traditional object state variable is augmented with an auxiliary momentum variable. At the proposal step, the proposal distribution is designed by simulation of the Hamiltonian dynamics. We first use the ordered over-relaxation method to draw the momentum variable which could suppress the random walk behavior in Gibbs sampling stage. Then, we leverage the gradient of the energy function of the posterior distribution to draw new samples with high acceptance ratio. The proposed tracking method could ensure that the tracker will not be trapped in local optimum of the state space. Experimental results show that the proposed tracking method successfully tracks the objects in different video sequences and outperforms several conventional methods.

© 2016 Elsevier B.V. All rights reserved.

1. Introduction

Visual tracking is an active research topic in computer vision community which has a wide range of potential applications including visual surveillance, medical research, sports analysis, human–machine interaction, and so on [1–6]. The main challenges for visual tracking are appearance variation induced by illumination changes, cluttered background and partial occlusions, and motion uncertainties induced by sudden dynamic change, low frame rate video and camera switching. In the past decades, visual tracking has been extensively studied by researchers and many tracking methods are proposed to tackle the challenges leading to a steady performance improvement [7–9]. These existing methods can be classified into two categories, that are deterministic methods and stochastic methods (sampling based methods). Among the stochastic methods, the particle filter (PF) and Markov Chain Monte Carlo (MCMC) method have shown their robustness for visual tracking [10–14]. PF is efficient in handling non-linearity and non-Gaussianity of system models. Isard and Blake [10] first applied PF to deal with visual tracking difficulties encountered by Kalman filters. Del Bimbo and Dini [15] proposed to use PF for visual tracking for long-term video sequences with first-order dynamic model and uncertainty adaptation. This method could continuously adapt the system model noise to balance uncertainty between the static and dynamic components of the state vector. Oron et al. [14] presented a locally orderless tracker (LOT) using PF framework which showed better performance in tracking both demormable and rigid targets.

MCMC is another popular sampling method for visual tracking. Khan et al. [12] proposed a MCMC based PF method (MCMC-PF) for tracking a variable number of objects. In the MCMC-PF method, the authors substituted MCMC posterior sampling for the importance sampling method aiming at suppressing the particle degeneracy and particle impoverishment problems. Cong et al. [16] proposed a MCMC based PF tracking method using the histogram of oriented gradients (HOG) based appearance model. It shows improved robustness in handling slight object occlusions. Pan and Schonfeld [17] extended the first-order Markov model commonly used in MCMC based visual tracking and introduced a high-order MCMC based tracking framework. Wang et al. [18] proposed a second-order MCMC based PF tracking method (2MCMC-PF). The 2MCMC-PF leverages a joint posterior probability density function based on x_t and x_{t-1} and shows better

* Corresponding author.
E-mail address: mingyu_lu@126.com (M. Lu).

http://dx.doi.org/10.1016/j.neucom.2016.04.063
0925-2312/© 2016 Elsevier B.V. All rights reserved.

Please cite this article as: F. Wang, et al., Ordered over-relaxation based Langevin Monte Carlo sampling for visual tracking, Neurocomputing (2016), http://dx.doi.org/10.1016/j.neucom.2016.04.063

ARTICLE IN PRESS

performance than traditional first order MCMC-PF. Kwon and Lee [19] proposed a robust tracker based on interactive MCMC sampling framework, that is visual tracking decomposition (VTD tracker). VTD is proven to be robust for appearance variation induced by occlusion and illumination change.

Although PF and MCMC have received great success in visual tracking, most of the trackers are based on smooth motion assumption which hypothesizes that the objects being tracked are moving with stable motion. However, many tracking scenarios contain abrupt motion difficulties induced by rapid motion, low frame rate video and camera switching. Abrupt motion implies that the object positions between two consecutive frames are changed suddenly. Traditional trackers often fail to tackle this thorny problem.

Recent studies have witnessed the progress in abrupt motion tracking. Most of the abrupt motion tracking methods are based on MCMC or PF and Bayesian filtering framework. Li et al. [20] proposed an abrupt motion tracking method for contour tracking based on particle swarm optimization and level set methods. However, the contour based representation depend highly on the accurate localization of the object boundary which is difficult to extract in low resolution images. A cascaded particle filter is used to track objects in low frame rate videos in [21]. But this method needs reliable prior observation model and off-line learning process. In addition, it cannot deal with the situations where objects are moving with large scale changes and rapid movements. Su et al. [13] proposed a saliency embedded particle filter (SPF) for abrupt motion tracking. During tracking, once the object is lost, the SPF could recover tracking by detecting the target region from salient regions. SPF could give robust tracking results for different kinds of abrupt motions. However, SPF cannot give satisfactory performance when camera switching occurs. Wang–Landau Monte Carlo sampling-based tracking scheme (WLMC) is proposed in [11,22]. The WLMC uses the density of states (DOS) term for the prior distribution of the Bayesian filtering. The DOS is adaptively updated via online learning during the sampling process. WLMC based tracking method has shown efficiency in tracking object with different kinds of abrupt motions. Zhou et al. [23,24] introduced an adaptive stochastic approximation Monte Carlo sampling (ASAMC) based tracking method for abrupt motion tracking. The ASAMC uses the DOS term to weight the samples which can avoid the local-trap problem encountered by traditional MCMC sampling methods. Both the WLMC and ASAMC trackers are based on DOS which needs first divide the sampling space (image observations) into several subspaces. But such division could decrease the efficiency of trackers when the frame size is large. Zhou et al. [25,26] proposed a nearest neighbor field (NNF) driven smoothing SAMC (SSAMC) tracking framework. The approximate NNF is used to reduce the sample space and the SSAMC sampling is adopted to explore the state space to localize the object. This method is efficient and robust for abrupt motion tracking. But again, SSAMC, SAMC and WLMC all need to visit the state space using the random walk method which often needs long iterations to reach the promising object state.

In this paper, we propose a novel tracking method based on the Bayesian filtering framework in order to solve the problems encountered by the traditional MCMC based tracking methods. The main contribution of our work is that the ordered over-relaxation based Langevin Monte Carlo (OR-LMC) sampling method is introduced into visual tracking. The OR-LMC sampling method leverages the gradient knowledge of the system model to design proposal distribution, while it uses the OR method to suppress the random walk behavior. We integrate the OR-LMC method into the Bayesian filtering based tracking framework accompanied by a color based appearance model, which consequently constructs a robust tracker to deal with various motion uncertainties. First, the LMC sampling method is introduced. LMC allows the sampler to efficiently search the state space through first-order gradients of the system model. Traditional MCMC sampling methods are mostly based on random walk in visiting the state space, which makes them exhaustive and inefficient in exploring the state space. The LMC could suppress the influence of random walk behavior because its implementation is based on simulating Hamiltonian dynamics [27]. In order to further improve the sampling efficiency and suppress the random walk behavior, we introduce the ordered over-relaxation (OR) method in the proposal step of the LMC method. Finally, the OR-LMC sampling method is integrated into the Bayesian tracking framework to formulate our proposed tracking algorithm.

This paper is organized as follows. In Section 2, we briefly introduce the Bayesian tracking framework. Afterward, in Section 3, we give the details of our proposed tracking method. Experimental results are shown in Section 4. The last section draws conclusion. The notations used in this paper are listed in Table 1.

Table 1
Notations used in this paper.

Notations	Description
x_t	Object state at time t
z_t	Observation at time t
x_t^x	x position of the object state x_t
x_t^y	y position of the object state x_t
x_t^s	Scale of the object state x_t
$p(\cdot)$	Probability distribution function
v_t	Process noise at time t
$f(\cdot)$	Object state transition function
$g(\cdot)$	Measurement function
$\hat{x}_t$	The best state sample at time t
ξ	Parameter for calculating the likelihood score
$D(\cdot)$	Bhattacharyya distance function
$hist_{ref}(\cdot)$	Histogram of the reference target model
$hist(\cdot)$	Histogram of an object region
σ	Bhattacharyya similarity coefficient
$H(\cdot)$	Total energy function
$U(\cdot)$	Potential energy function
$K(\cdot)$	Kinetic energy function
ν	Momentum term
M	Quantity matrix
P	Probability distribution
π	Arbitrary energy function
ε	Step size
k	Degree of over-relaxation
r	An integer generated from a distribution
s	The CDF value at ν_0
w	An integer generated from a distribution
α	Acceptance ratio
$Q(\cdot)$	Proposal density function
β	Acceptance ratio
w_g	Ground truth width
h_g	Ground truth height

2. Bayesian tracking framework

In computer vision community, object tracking is often formulated as Bayesian filtering problem. Thus, the goal of object tracking within this framework is to search the best object state x_t which could maximize the posterior probability density function $p(x_t|z_{1:t})$. A 3D vector $x_t = \left(x_t^x, x_t^y, x_t^s\right)$ is often used to represent the object state at time t, where the three components represent the x, y position and scale of the object, respectively. If the state of the object x_t at time t and the frame based observations $z_{1:t}$ up to time t are given, the posterior probability density $p(x_t|z_{1:t})$ can be

Please cite this article as: F. Wang, et al., Ordered over-relaxation based Langevin Monte Carlo sampling for visual tracking, Neurocomputing (2016), http://dx.doi.org/10.1016/j.neucom.2016.04.063

Download English Version:

https://daneshyari.com/en/article/4948095

Download Persian Version:

https://daneshyari.com/article/4948095

Daneshyari.com