Pattern Recognition 47 (2014) 194–207

Contents lists available at ScienceDirect

Pattern Recognition

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

Gravity optimised particle filter for hand tracking

Malik Morshidi, Tardi Tjahjadi *

School of Engineering, University of Warwick, Gibbet Hill Road, Coventry CV4 7AL, United Kingdom

ARTICLE INFO

Article history:
Received 14 September 2012
Received in revised form
31 March 2013
Accepted 27 June 2013
Available online 16 July 2013

Keywords:
Particle filter
Articulated hand tracking
Finger movement
Gravity
Convexity defects
CamShift

ABSTRACT

This paper presents a gravity optimised particle filter (GOPF) where the magnitude of the gravitational force for every particle is proportional to its weight. GOPF attracts nearby particles and replicates new particles as if moving the particles towards the peak of the likelihood distribution, improving the sampling efficiency. GOPF is incorporated into a technique for hand features tracking. A fast approach to hand features detection and labelling using convexity defects is also presented. Experimental results show that GOPF outperforms the standard particle filter and its variants, as well as state-of-the-art CamShift guided particle filter using a significantly reduced number of particles.

1. Introduction

Among human body parts, the hand is the most effective means of non-verbal communication and plays an effective role in human computer interaction such as gesture recognition [1,2], virtual reality [3], and computer games [4,5]. Hand gesture recognition involves hand tracking and the difficulties in developing a hand tracking system include: high dimensionality of the tracking; finger self-occlusion; high processing speed; and rapid hand motion [6]. Hitherto, an accurate and real-time hand tracking system is still far from realised. In this paper we propose an approach which balances the need for real-time requirement and acceptable accuracy. The approach incorporates an improved particle filter for hand tracking and detects hand features by computing the convexity defects of the hand silhouette contour.

The standard particle filter (PF) also known as conditional density propagation (i.e., condensation) algorithm (e.g., [7]) incorporates the use of complex motion models and is highly robust to clutter. However, it lacks the ability to run in real-time since the number of samples (or particles) is large in order to account for sudden movements of the object being tracked. A large number is also needed to overcome the samples impoverishment problem by populating some areas of the state-space that may be left empty due to prediction of the motion model that tends to cluster the samples in some area due to the predicted motion.

The ICondensation algorithm [8], an extension of PF, permits the combination of the original random sampling filter representation with the information available from alternative sensors in the form of an importance function. The importance function aids the sampling from prior method by generating and concentrating samples in areas of the state-space that contain most information about the posterior. This helps alleviate the samples degeneracy problem by avoiding to generate samples which have low weights to represent the posterior. Samples drawn from the importance function are systematically formulated in such a way that they do not change the probabilistic model of the tracker.

In [9] a local optimiser based on the stochastic meta-descent tracker is integrated after the standard particle propagation step. The new particles generated by the optimiser are combined with the original particles distribution which results in a smart particle filter that can track high-dimensional articulated structure with far fewer particles. In [10], after particles propagation the mean shift embedded particle filter (MSEPF) is applied to move the particles to the nearby local modes with high likelihood resulting in better posterior estimation. Using significantly fewer particles, MSEPF operates robustly and with lesser computational burden. In [11] CamShift Guided Particle Filter (CAMSGPF) employs a simplified version of CamShift which iterates at a smaller and fixed interval to reduce the computational burden. This is unlike CamShift that will iterate until convergence. The optimisation procedure in CAMSGPF takes into account the current observation which results in new proposal density of the current samples for which new samples are drawn.

Since the introduction of PF for object tracking, most of the works derived thereafter evolve around the use of weighted

* Corresponding author. Tel.: +44 24 76523126; fax: +44 24 76418922.
E-mail address: t.tjahjadi@warwick.ac.uk (T. Tjahjadi).

0031-3203/$ - see front matter © 2013 Elsevier Ltd. All rights reserved.
http://dx.doi.org/10.1016/j.patcog.2013.06.032

particles to represent the posterior density. In this paper we propose the gravity optimised particle filter (GOPF) that generates a new set of particles based on current observation and allows each weighted particle to have its own gravitational force (from an analogy with Newton's law of universal gravitation) that will attract nearby particles towards the peak of the particles distribution. The newly generated particles is combined with the particles generated in the same way as with a PF. The weights for the new particles depend on their position relative to the current observation to enable the combined particles to maintain the multiple hypotheses nature of the tracker.

There are generally two approaches to hand tracking: model-based and appearance-based. The model-based approach [12,13] is based on parametric models of the shape and kinematic structure of the 3-dimensional (3D) hand. Using the motion history and dynamics of the hand, this approach predicts the model hypotheses when a new observation is made by measuring the dissimilarity between the expected model hypotheses with the actual observation. However, seeking the optimised solution in a multi-dimensional model parameters space especially with the complexity of hand motion and self occlusion makes this approach unreliable for long video sequences. Furthermore, the computational requirements are high. Recent work in [12] addresses this problem using GPU-based software implementation and off-the-shelf Kinect sensor which demonstrates robust 3D articulated hand tracking in near real-time (15 Hz) over a long sequence.

The appearance-based approach [14,15] is based on appearance-specific 2D image mapping from a set of image features to a limited set of predefined poses. The approach is goal-oriented, where a small set of predefined poses needs to be recognised, making it impossible to estimate other poses. Variable length Markov model is combined with annealed PF in [16] to account for discontinuous changes during the observations. The framework operates by automatic switching between first-order Markov model and PF. The former is used for the case of previously observed events, whereas the latter operates for the case of unseen events. In [13], appearance-guided PF is used for high degree-of-freedom (DOF) tracking in image sequences. PF is extended by using state space vectors that act as attractors, and a probability propagation framework is derived to find the approximation for the maximum a posteriori solution. This solution avoids the drifting effect of PF.

In this paper we present a model-based framework for tracking hand features. The framework incorporates GOPF and a fast method for detecting hand features using convexity defects. The main advantage using the proposed hand features detection is due to its reasonable accuracy and fast processing. It also does not depend on offline learning from database of hand features or postures.

This paper is organised as follows: Section 2 provides an overview of the proposed hand tracking framework. Section 3 presents the details of GOPF. Section 4 provides detailed explanation of the measurement model. Finally, Sections 6 and 7, respectively, present the experimental results and conclusions.

2. Overview of the proposed hand tracking framework

In addition to the weighted particles generated in the same way as with a PF, the proposed hand tracking framework uses GOPF which utilises the gravitational concept to attract nearby particles towards the virtual likelihood particle, i.e., the current observation. The framework involves the combination of the standard PF and GOPF. At every time step, N particles of the $2N$ particles from the previous time step are resampled and propagated based on the PF framework. Using the locations of the N propagated particles as reference, a new set of m particles (where $m < N$) are replicated where some of the N particles close to the current observation should move due to the attraction effect (referred to as Algorithm 1). Another set of new n particles (where $m+n=N$) are randomly propagated within the expected direction of the next observation (referred to as Algorithm 2). The N, m, and n particles are combined to give the new total of $2N$ particles. This enhances the proposal density of the likelihood distribution while maintaining the original Bayesian framework of the weighted particles. As the proposal density improves, the need for a large number of particles reduces significantly. The framework also incorporates a model for hand features extraction with slight modification on the localisation and labelling of fingers as in [17]. Unlike in [17] where the dominant features are extracted using a k-cosine curvature, we use a faster approach using convexity defect [18]. Tracking the fingertips throughout an image sequence using only the hand features extraction algorithm might not work adequately, especially when the hand encounters partial or complete occlusion. Therefore, GOPF (referred to as Algorithm 3) is incorporated in the framework to boost the stability of the hand tracking when encountering noisy measurements and occlusions.

3. Gravity optimised particle filter (GOPF)

The ability of PF to handle multiple hyphotheses with non-linear motion and non-linear observations has made it the most widely used tracking technique [19,10,11,20]. Object tracking in the PF framework [21,7] is the process of sequentially estimating the state parameters x_t at time t. Given the set of observations history $Z_t = \{z_1, \dots, z_t\}$, the Bayesian formulation of PF is expressed as the computation of posterior probability, i.e.,

$$p(x_t|Z_t) = \eta p(z_t|x_t) \int p(x_t|x_{t-1}) p(x_{t-1}|Z_{t-1})\, dx_{t-1}, \tag{1}$$

where η is the normalisation constant, $p(z_t|x_t)$ is the likelihood model, $p(x_t|x_{t-1})$ is the motion model and $p(x_{t-1}|Z_{t-1})$ is the temporal prior. At each time-step t of the PF iteration, the posterior probability is estimated by assigning each particle $s_t^{(n)}$ with a weight $\pi_t^{(n)}$. The weighted particle set $\{(s_t^{(n)}, \pi_t^{(n)}),\ n = 1, \dots, N\}$ represents the hypothetical states of the conditional state-density $p(x_t|Z_t)$ (i.e., the posterior probability) of the object at time t. The best approximation of the state is determined by either the highest weighted particle, or the average of particles' weights. The particles set at the next time-step is resampled and propagated according to the motion model. Computing the posterior probability with an integral over all possible state values [21,7] in each iteration is computationally infeasible. To alleviate this problem, importance sampling or resampling is used to combine the prior knowledge of the object position and motion with any additional knowledge extracted from auxiliary sensors [8]. Since the dynamics of hand motion is non-linear, we adopt a general motion model of the Gaussian random walk [22]. The measurement in the current image is used to hypothesise the likelihood regions of the state space.

3.1. Theoretical basis of GOPF

Particles in a PF framework can be thought of as point masses of physical bodies in the theory of gravity. Each weighted particle is considered to have an additional property of gravity with a gravitational force proportional to its mass. Newton's law of universal gravitation [23] states that any point mass m_1 attracts every other single point mass m_2 by a force. The two point masses that are attracted towards each other can be viewed as time dependent in which the magnitude of the force at time t, i.e., F_t, becomes greater. This in turn affects the acceleration of a point