Pattern Recognition 42 (2009) 2372 – 2391

Contents lists available at ScienceDirect

Pattern Recognition

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

Variance reduction techniques in particle-based visual contour tracking

Daniel Ponsa[a,*], Antonio M. López[b]

[a]Centre de Visió per Computador, Universitat Autònoma de Barcelona, Edifici O, Bellaterra 08193, Spain
[b]Centre de Visió per Computador & Dept. Ciències de la Computació, Universitat Autònoma de Barcelona, Edifici O, Bellaterra 08193, Spain

ARTICLE INFO

Article history:
Received 25 June 2008
Received in revised form 30 March 2009
Accepted 4 April 2009

Keywords:
Contour tracking
Active shape models
Kalman filter
Particle filter
Importance sampling
Unscented particle filter
Rao-Blackwellization
Partitioned sampling

ABSTRACT

This paper presents a comparative study of three different strategies to improve the performance of particle filters, in the context of visual contour tracking: the unscented particle filter, the Rao-Blackwellized particle filter, and the partitioned sampling technique. The tracking problem analyzed is the joint estimation of the global and local transformation of the outline of a given target, represented following the active shape model approach. The main contributions of the paper are the novel adaptations of the considered techniques on this generic problem, and the quantitative assessment of their performance in extensive experimental work done.

1. Introduction

Visual contour tracking is an area of research that has received much attention by the computer vision community for many years. One essential reason for this to happen is that, in many application domains, the contour of an object is a very informative cue about its state or configuration. Proof of that is the application of contour tracking in areas like visual surveillance [1], traffic monitoring [2], medical diagnosis [3,4] and human–machine interaction [5,6], among others.

The tracking of contours has been posed mainly as a minimization or as an inference problem. Following the first perspective, the so-called *active contour* methods adapt iteratively an elastic curve to image edges, while imposing some constraints on it (e.g., smoothness and compactness). The *classical* snakes approach [7] performs that by minimizing an energy term associated to a parametric curve. Geodesic active contours [9], which generalize in most situations classical snakes [10], pose the problem from a geometric point of view. Targets are segmented using an implicit contour representation. A non-parametric surface is evolved according to image edges, being the tracked contour the zero level set of this surface. The main advantage of this *level set-based* approach is that topological changes of the original curve are naturally managed. Extensions of this work, where contours are defined in terms of the content of the region that they enclose conform the *active regions* methods [11–13].

An important disadvantage of minimization-based approaches is the possibility of converging into local minima and mistrack the target. This drawback can be treated in a principled way by posing contour tracking as an inference problem. Now the goal is estimating the posterior density of a contour given image observations. Minimization-based approaches can be interpreted as a way of determining the maximum a posteriori of this density, assuming implicitly its unimodality. Problems appear when this density is not unimodal, which can be eluded if the whole density is estimated. This paper studies contour tracking from this perspective.

Formally, given a parametric model of the contour to be tracked, the goal is estimating at each instant t the probability density function (PDF) of the model parameters $\mathbf{x}_t$ (i.e., the contour state), conditioned on the observations up to t (i.e., $\mathbf{y}_{1:t} = [\mathbf{y}_i]_{i=1}^{t}$). In many applications this PDF can be properly assumed Gaussian, and its parameters can be efficiently estimated by means of Kalman-based filters. However, in cluttered scenes, this Gaussian assumption is usually too rough, since the PDF presents in fact multiple modes. This happens when there is more than one model parameterization that fits tightly to image observations, due to the presence of the tracked shape and also of other distractors in the scene. In these cases, it seems reasonable to maintain more than one contour tracking hypothesis, and in that way assure to keep track of the one that effectively adjusts to the object of interest. A principled manner to perform that consists in representing $p(\mathbf{x}_t|\mathbf{y}_{1:t})$ by means of a population of *particles* (i.e., concrete

* Corresponding author. Tel.: +34 935812440; fax: +34 935811670.
E-mail addresses: daniel@cvc.uab.cat (D. Ponsa), antonio@cvc.uab.cat (A.M. López).

0031-3203/$ - see front matter
doi:10.1016/j.patcog.2009.04.007

$\mathbf{x}_t$ instances), distributed (ideally) according to this PDF. In that way, any arbitrary form of the filtering density can be properly managed, what results in a tracking performance more robust to clutter. Providing a proper particle-based representation of $p(\mathbf{x}_t|\mathbf{y}_{1:t})$ is the objective pursued by the so-called particle filters (PFs). Briefly, PFs are stochastic sampling methods that sequentially approximate $p(\mathbf{x}_t|\mathbf{y}_{1:t})$ by combining a particle-based representation of this density at the previous instant $t-1$ (i.e. $p(\mathbf{x}_{t-1}|\mathbf{y}_{1:t-1})$), and new collected observations $\mathbf{y}_t$. Due to that, they are commonly referred as sequential Monte Carlo methods and good reviews of their theoretical basis can be found in [14–16]. PFs were seminally applied to the problem of contour tracking by Isard and Blake [17,18], in a particular form that they termed as Condensation algorithm.

The Condensation algorithm is definitely the most popular form of PF applied in vision-based tracking applications. However, its computational cost (which depends on the amount of particles needed to represent $p(\mathbf{x}_t|\mathbf{y}_{1:t})$ properly) increases exponentially with the number of parameters of the target model used. That is, it suffers from the curse of dimensionality. This is a serious drawback in contour tracking problems. In general, targets being tracked present global transformations of their outline (e.g., translations, rotations, etc.), as well as simultaneous local shape deformations. Consequently, the dimension of the parametric model of the contour is rather big, what makes the cost of its robust tracking high. The good news is that, since the problem of Condensation with the state dimensionality is well known, different generic strategies have been proposed to counteract it. In this paper we analyze the performance of three of these strategies in the context of visual contour tracking: the unscented particle filter (UPF), the Rao-Blackwellized particle filter (RBPF), and the partitioned sampling (PS) technique. We contribute with their adaptation in the context of contour tracking using active shape model (ASM). Developed mainly by British research groups in Leeds, Oxford and Manchester [19–21], ASMs represent the outline of an object by means of a parametric model, whose representability is limited to a given space of transformations, whether generic (e.g., Euclidean or affine transformations of a basic shape) or specific (shape deformations *learned* from the statistic analysis of training data). Our study focuses on ASMs since exploit naturally the a priori knowledge on the feasible shapes that a target can take. As traditionally formulated, they do not consider topological changes of the contour. However, in [22] is shown that applying their same principles on implicit contour representations, a parametric model is obtained that can manage such cases. ASMs have been shown effective in many application domains [20], and thanks to their parametric nature, their use in inference-based contour tracking is direct. However, as it is shown in [23,24], strategies exist to consider also non-parametric contour representations inside this framework.

As will be stated in the respective sections, two of the three techniques studied in this paper (the UPF and the PS) have also been applied previously in the contour tracking problem by other authors. However, our proposals differ significantly from the ones in these previous works. On the one hand, we use a more complete contour model, accounting for global and local shape transformations. On the other, our model of the contour observation process is more rigorous and accurate, leading to a better interpretation of the evidence extracted from frames. Another contribution of this paper is an exhaustive study of the performance of the proposed algorithms. This has been done using synthetic sequences, distorted with different levels of noise. Using the knowledge of the parameters used to generate the sequences, the performance of each technique has been measured quantitatively. This has allowed us to rank proposed algorithms at each evaluated situation, and to identify their strengths and weaknesses. Algorithms have also been tested on real sequences, in the contexts of hand and pedestrian racking.

Table 1
List of acronyms.

AR	Auto-regressive
AR1	Auto-regressive process of first order
ASM	Active shape model
CBM	Constrained Brownian motion
IPPF	Independent partition particle filter
KF	Kalman filter
MCE	Mean contour error
OISD	Optimal importance sampling density
PDF	Probability density function
PF	Particle filter
PF-EIS	Particle filtering with efficient importance sampling
PS	Partitioned sampling
RB	Rao-Blackwellization
RBPF	Rao-Blackwellized particle filter
SIS	Sequential importance sampling
SISR	Sequential importance sampling with resampling
SNR	Signal-to-noise ratio
UKF	Unscented Kalman filter
UPF	Unscented particle filter

The remainder of this paper is organized as follows: Section 2 gives an overview of the used contour model representation (the ASM), and introduces the approach used to jointly account for the affine transformations and the local deformations of a given shape of interest. Then, Section 3 focuses on modeling the shape evolution along time, and Section 4 on how the shape model relates to observations in images. Section 5 formalizes the visual tracking of contours as a Bayesian inference problem, and presents the general solution to this problem given by the importance sampling technique, which has led to the so-called particle filtering. The main drawbacks of this approach are remarked, and three different strategies to deal with them are adapted in the following sections to the contour tracking problem: the UPF (Section 6), the RBPF (Section 7), and the PS (Section 8). A comparative study of the performance of these approaches is presented in Section 9, and final conclusions are provided in Section 10. A list of the abbreviations used in the paper is given in Table 1.

2. Contour representation

In many application domains the use of shape tracking algorithms is motivated by the need not only to localize a given target, but also to identify its specific pose or configuration. To fulfill that, a generative model of the target shape variability is required. Many authors have worked on developing representations of shape variability in many different ways. Refs. [19–21] review the major contributions on this field, and then focus on the description of a model-based approach to shape tracking, commonly denoted as the ASM approach. There exist different possibilities to represent parametrically the outline of an object. In the ASM formalism, the dominant approach is based on 2D contours modeled by B-spline parametric curves.

B-splines construct expressions of a 2D contour as a weighted sum of N_B basis functions. Contour point coordinates $\mathbf{r}(s)=[x(s)\ y(s)]^T$ of 2D shapes are obtained by an expression of the form

$$\begin{bmatrix} x(s) \\ y(s) \end{bmatrix} = \begin{bmatrix} \mathbf{B}(s)^T & \mathbf{0}_{N_B}^T \\ \mathbf{0}_{N_B}^T & \mathbf{B}(s)^T \end{bmatrix} \begin{bmatrix} \mathbf{q}^x \\ \mathbf{q}^y \end{bmatrix},$$

or more compactly $\mathbf{r}(s)=\mathbf{U}(s)\mathbf{q}$. We denote with $\mathbf{O}_{N_B}$ a column vector of N_B zero elements. $\mathbf{B}(s)$ is a vector maintaining N_B basis functions, which are curves composed of polynomials of degree d with finite support. They are C^{d-1} continuous, which means that the contour derivatives up to the $(d-1)$-th are smooth. $\mathbf{q}=[\mathbf{q}^x\ \mathbf{q}^y]^T$ is the vector of the N_{cp} *control points* that weights the basis functions to generate a desired curve. Thus, its dimensionality corresponds to $N_q=2N_{cp}$, being $N_{cp}=N_B$. The parameter s evaluates the linear combination of