Available online at www.sciencedirect.com
ScienceDirect
Pattern Recognition 40 (2007) 920–930
PATTERN RECOGNITION
THE JOURNAL OF THE PATTERN RECOGNITION SOCIETY
www.elsevier.com/locate/pr

Affine and projective active contour models

Dipti Prasad Mukherjee[a,*], Scott T. Acton[b]

[a]Electronics and Communication Sciences Unit, Indian Statistical Institute, Kolkata, India
[b]Electrical and Computer Engineering, University of Virginia, Charlottesville, VA, USA

Received 14 December 2005; received in revised form 14 March 2006; accepted 7 June 2006

Abstract

We investigate the evolution of active contours in terms of progressive modification of an initial contour following the chosen Lie group of object-to-image transformations. Because of non-fronto-parallel viewing of an object or due to relative motion between the camera and the object, the resultant image may undergo affine or projective object-to-image transformations. In a recent paper we have shown that in the case of object tracking, frame-to-frame deformations of an initial curve obtained through Euler–Lagrange descent equations of a curve functional can be used to enact a desired Lie group of plane transformations [A.-R. Mansouri, D.P. Mukherjee, S.T. Acton, Constraining active contour evolution via Lie groups of transformation, IEEE Trans. Image Process. 13 (2004) 853–863]. In this work, we propose an energy functional that encodes the Lie group transformation parameters, which in turn guide shape distortion due to oblique viewing. Additional constraints, such as transformation smoothness, are imposed on the active contour by modifying the energy functional. The functional is minimized using numerical schemes similar to the conjugate gradient technique, and the convergence properties are discussed. The success of the technique for affine and projective scenes is demonstrated with both synthetic and real image examples and compared with the related approaches.

Keywords: Tracking in video; Active contours; Affine transformation; Projective transformation

1. Introduction

Many video analysis applications involve the delineation of an object boundary throughout a sequence of images. Active contours [1,2] provide a convenient framework for this tracking process. In this context, the evolution of contour is performed minimizing an energy functional that attempts to maximize the image gradient magnitude coincident with the contour position. An initial contour close to the object of interest can be constructed and iteratively modified based on the energy minimization process of gradient descent. As such, in this paper, we refer to the method of Ref. [1] as the *gradient-based active contour*.

For tracking an object in a temporal sequence of images, the object contour detected within the tth frame can serve as the initial contour for the same object within frame $t+1$ [3–5]. A prediction engine such as the Kalman filter can be used to predict the center of the initial contour in a subsequent frame [6,7], or to guide the movement of each active contour point [8]. Alternatively, geometric contours based on level set formulations can be used for tracking [9–11].

Given the parametric contour $\psi(s)$, $0 \leqslant s \leqslant 1$, which moves within the domain Ω of image I, the energy functional to be minimized can be written as

$$E = \int_{\Omega} F(s, \psi(s), \dot{\psi}(s), \ddot{\psi}(s), \ldots)\, \mathrm{d}s. \tag{1}$$

The minimization process should satisfy the curve evolution equation

$$\frac{\mathrm{d}\psi(s)}{\mathrm{d}t} = -\frac{\delta E(\psi(s))}{\delta\psi}, \tag{2}$$

for $t \geqslant 0$ and given $\psi(s)|_{t=0} = \psi_0$ as the initial contour. Standard variational calculus techniques are used to minimize this energy functional. In this paper, the image gradient-based energy functional is coupled with the shape constraints

* Corresponding author. Tel.: +91 33 2575 2913; fax: +91 33 2577 3035.
E-mail addresses: dipti@isical.ac.in (D.P. Mukherjee), acton@virginia.edu (S.T. Acton).

0031-3203/$30.00
doi:10.1016/j.patcog.2006.06.015

for an object, assuming that the object may be viewed from a non-fronto-parallel viewpoint.

Essentially, the traditional active contour model does not use specific shape information, apart from smoothness constraints [1]. If the shape is known exactly, the shape may be encoded directly into the energy functional so that the contour is biased toward certain shapes [12–14]. But, in a general case, the shape of the boundary may not be known a priori. Alternatively, the shape may change due to the viewpoint and viewing distance, rendering the fixed-shape approaches of Refs. [12–14] ineffective. A more recent work [15] preserves the shape of the target object by first finding the best move at each point in the discretized active contour using the standard gradient-based active contour of Ref. [1] and then computes a global Lie group planar transformation that approximates the local best moves of individual contour points.

Given the active contour at time t and $t+1$ (where the contour at $t+1$ is obtained using the method in Ref. [1] with the contour at t), an affine or projective transformation can be computed using the corresponding points. Since solving a 2D affine transformation requires three-point correspondences (2D projective transformation requires four-point correspondences) between the contours at t and $t+1$ and since more than the required number of correspondences are always available between the contours at t and $t+1$, the affine or projective transformation is solved as an over-determined case. These computed affine or projective transformation matrices are then used to transform individual points on contour at t, and the total deformation (that is total shift along the x and y direction for all the contour points) is derived. Finally, the mean deformation (total deformation divided by the number of contour points) gives the movement of each contour point. It is proved in Ref. [15] that such mean deformation actually converges to the target in the image frame at $(t+1)$. In contrast, the novelty of the approach developed in this paper is found in the transformation parameters that are embedded directly in the energy functional, as opposed to finding the best transformation parameters that mimic the gradient-based active contour movement. Therefore, in the proposed approach the objective is to find the 2D affine or projective transformation parameters from the minimization of active contour energy functional, which is not necessarily in agreement with the best aggregate movement of individual contour points.

For applications where a rigid or a semi-rigid object needs to be tracked in a spatiotemporal image sequence and where the image of the object shape undergoes transformations due to relative motion between the camera and the object, the curve evolution problem can be considered as generation of curves in the orbits of a chosen Lie group [16] of transformation. In this paper, we construct the energy functional such that the active contour evolution follows the chosen Lie group of object-to-image transformations. The two most common object-to-image transformations are planar affine and projective transformations having six and eight degrees of freedom, respectively. While parallel lines in the object remain parallel even after affine object-to-image transformations, the parallel lines appear to meet at vanishing point(s) after projective object-to-image transformation [17]. Usually, for viewing conditions in which the viewer or the camera is so close to the object that the distance between the camera and the object is *significantly* less compared to the scale of the object along the viewing direction, the object-to-image transformation is taken as a projective model.

In Ref. [18], affine motion models are applied to control points of a deformable B-spline contour for which the subsequent positions are determined based on maximization of image gradient along the contour. The motion parameters are further refined using a Kalman filter model. In another variation on this theme, tracking is achieved by combining region and contour-based deformable models [19]. In this case, affine and homographic motion models are applied on deformable regions. The detection of deformable region in subsequent frames is based on a normalized correlation criterion measuring the differences between the gray levels of the template region (subtracting the mean gray levels inside the region) in consecutive images. Subsequent to this process, the contour of the detected region is further refined using maximization of gradient energy. The approach present here, in contrast, is efficient in terms of computational complexity. Further, our approach incorporates smooth changes in affine and projective parameters via constraints imposed on the determinant of the shearing matrix.

In the next section, we present the construction and evolution of an active contour that follows affine object-to-image transformation. The corresponding result is shown in Section 3. The proposed algorithm is compared with related approaches in Section 3.1. We discuss convergence issues in Section 3.2. The extension of the model for the projective case is demonstrated in Section 4.

2. Active contour evolution under affine transformation

With the assumption that the target contour is an affine mapping of the initial contour, the energy functional (1) can be suitably modified to incorporate the affine constraints. The 2D affine transformation from pixel position (x, y) to (x', y') is given by

$$\begin{bmatrix} x' \\ y' \end{bmatrix} = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} + \begin{bmatrix} t_x \\ t_y \end{bmatrix}, \tag{3}$$

where the four parameters a_{11}, a_{12}, a_{21} and a_{22} control the rotation, scaling and shearing (unequal scaling along x- and y-axes) of the initial point (x, y) to the transformed point (x', y'). The vector $[t_x t_y]^{\mathrm{T}}$ represents translation along x- and y-axes, respectively. These six parameters $\theta = \{a_{ij}; t_i\}$ represent six degrees of freedom due to planar affine transformation. A major characteristic of affine transformation is that even though the aspect ratio of the shape may change due to oblique viewing, the parallel lines will still remain