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

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

parallel after affine object-to-image transformation. Given that the object to be tracked is viewed under the affine transformation condition, an initial contour ψ_0 will be distorted following Eq. (3). Therefore, for a 2D image matrix I, the curve evolution should seek a local minimum in the gradient-based energy functional given by

$$E = -\int_{\Omega} |\nabla I(x', y')| \, \mathrm{d}s. \tag{4}$$

We can use a conjugate gradient method in $\Re^n$ (in this case $n=6$ for the six degrees of freedom of the affine transformation) to minimize the energy in Eq. (4). The minimization gives affine parameters $\theta = \{a_{ij}; t_i\}$ that provide the position of the curve in the current frame.

In an image sequence depicting the motion of a rigid or a semi-rigid body, the evolution of object contour delineating the object shape is expected to be gradual and smooth. For the tracking application discussed in this paper (see Section 3), an object is being tracked in consecutive frames at 30 frames per second. Therefore, changes of between-frame affine transformation parameters are minimal. Assuming the inter-frame shape transition lies in the orbit of the affine group, we may impose a smoothness constraint by precluding any sudden change in the determinant of the 2×2 matrix $\mathbf{a}$ of Eq. (3). It can be shown that the linear transformation $\mathbf{a}$ is a combination of isotropic scaling of the shape, rotation about the optic axis and an expansion along the eigenvector of a symmetric matrix decomposed from $\mathbf{a}$ [17]. A large change in the determinant of $\mathbf{a}$ between two consecutive video frames will correspond with either a change in the overall scaling of the shape or an expansion along the eigenvector or both. By restricting the rate of change of the determinant we are constraining the possibility of excessive between-frame shape distortion. Since translation vectors are not responsible for any shape distortion due to oblique viewing, they are ignored in this process.

Incorporating the determinant constraint in Eq. (4), the modified energy equation is given by

$$E = -\int_{\Omega} |\nabla I(x', y')| \, \mathrm{d}s + \lambda_1 \sum_{a_{ij}} \frac{\partial D^2}{\partial a_{ij}}, \tag{5}$$

where $D = (a_{11}a_{22} - a_{12}a_{21})$. The parameter λ_1 in Eq. (5) controls the relative weight between the edge strength and the determinant measure. The weight λ_1 can be selected empirically depending on the extent of shape distortion of the object being tracked. Also, the selection influences the stability and convergence of the minimization of Eq. (5), which will be discussed in Section 3.2.

We also capitalize upon a special property of affine transformation. From the definition of the affine transformation, a determinant $\det(\mathbf{a}) > 0$ indicates absence of reversal of orientation for the transformation, while $\det(\mathbf{a}) < 0$ indicates an orientation reversing transformation. In most physically meaningful applications like object tracking, it is unlikely that gradual changes in affine parameters lead to a transition from $\det(\mathbf{a}) > 0$ to $\det(\mathbf{a}) < 0$. For the smooth evolution of the initial contour, we constrain the evolution in such a way that the energy functional penalizes a change in the sign of the determinant. This constraint can be incorporated in the energy functional using a Dirac delta approximation. First, consider a Heaviside function $\mathrm{H}(z)$:

$$\mathrm{H}_{\varepsilon}(z) = \begin{cases} 1 & \text{if} \quad z > \varepsilon, \\ 0 & \text{if} \quad z < -\varepsilon, \\ \dfrac{1}{2}\left[1 + \dfrac{z}{\varepsilon} + \dfrac{1}{\pi}\sin\left(\dfrac{\pi z}{\varepsilon}\right)\right] & \text{if} \quad -\varepsilon \leqslant z \leqslant \varepsilon \end{cases} \tag{6}$$

and the corresponding Dirac delta approximation is given by

$$\rho_{\varepsilon}(z) = \mathrm{H}'_{\varepsilon}(z) = \begin{cases} 0 & \text{if} \quad |z| > \varepsilon, \\ \dfrac{1}{2\varepsilon}\left[1 + \cos\dfrac{\pi z}{\varepsilon}\right] & \text{if} \quad |z| \leqslant \varepsilon. \end{cases} \tag{7}$$

For $-\varepsilon \leqslant z < 0$, the value of $\rho_{\varepsilon}(z)$ is taken as $\delta_{\varepsilon}(z)$, which is the negative (additive inverse) of the $\rho_{\varepsilon}(z)$ value calculated from Eq. (7). The constraint on $\det(\mathbf{a})$ through the Dirac measure in Eq. (7) penalizes the situation in which there is a transition in the sign of $\det(\mathbf{a})$. Utilizing the Dirac measure prevents the situation where $\det(\mathbf{a}) = 0$. Consider the plot of Eq. (7) for $-1 \leqslant z \leqslant +1$ and $\varepsilon = 1$ as shown in Fig. 1. For $0 \leqslant D \leqslant 0.5$, $\rho_{\varepsilon}(D)$ has a value approximately between 1 and 0.5. In the case where D is close to zero, $\rho_{\varepsilon}(D)$ is close to 1. So, $\det(\mathbf{a})$ will be encouraged to remain non-zero during inter-frame tracking. Similarly, in the case where $-0.5 \leqslant D < 0$, $\delta_{\varepsilon}(D)$ varies approximately between -0.5 and -1. So, Eq. (7) prohibits a sign change in the determinant. The complete energy functional is given by

$$E = -\int_{\Omega} |\nabla I(x', y')| \, \mathrm{d}s + \lambda_1 \sum_{a_{ij}} \frac{\partial D^2}{\partial a_{ij}} + \lambda_2 \rho_{\varepsilon}(D). \tag{8}$$

As desired, $\rho_{\varepsilon}(D)$ (or $\delta_{\varepsilon}(D)$ when $-\varepsilon \leqslant D < 0$) has no effect on change in the active contour energy when $D > \varepsilon$ or when $D < -\varepsilon$. The weight λ_2 can be selected such that $\lambda_2 \gg \lambda_1$, as the determinant sign constraint is a hard constraint.

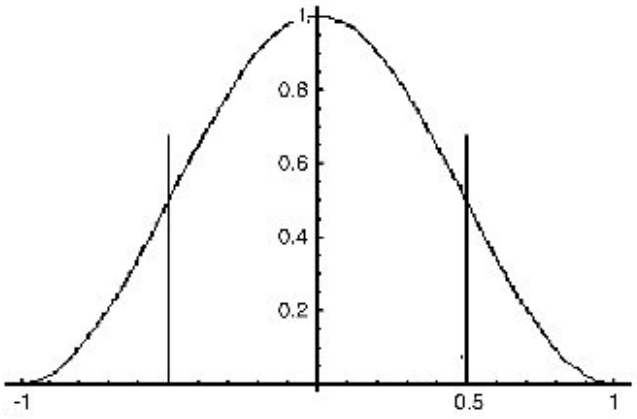

Fig. 1. Plot of Eq. (7) for $-1 \leqslant z \leqslant +1$. The parameter z is plotted along x-axis while $\rho_{\varepsilon}(z)$ is plotted along y-axis.

In the framework of a steepest descent method, iterative minimization of the function $\{f(z)|z \in \Re^n\}$, assuming $f(\cdot)$ is a strictly convex and differentiable function, involves computing $f(z_{i+1}) = \min\{f(z_i + \lambda h_i)|\lambda \geqslant 0\}$ such that $\{h \in \Re^n | \langle \nabla f(z_i), h\rangle < 0\}$. The gradient of function is denoted by ∇f and $\langle \cdot, \cdot \rangle$ denotes the Euclidean scalar product. For minimization of the energy functional in Eq. (8), the gradient vectors in $\Re^n | n = 6$ are evaluated using the partial derivatives of the energy functional, which gives the update direction of the affine variables:

$$\frac{\partial E}{\partial a_{11}} = -\int_\Omega x I'_x \, ds + \lambda_1 P_{a_{11}} + \lambda_2 Q_{a_{11}}, \tag{9.1}$$

where $P_{a_{11}} = 2a_{22}^2 - 2a_{21}a_{22} - 2a_{12}a_{22} + 4a_{11}a_{22} - 2a_{12}a_{21}$, $Q_{a_{11}} = -(\pi a_{22}/2\varepsilon^2)\sin \pi D/\varepsilon$ and

$$\frac{\partial \nabla I(x', y')}{\partial a_{11}} = \frac{\partial \nabla I(x', y')}{\partial x'}\frac{\partial x'}{\partial a_{11}} + \frac{\partial \nabla I(x', y')}{\partial y'}\frac{\partial y'}{\partial a_{11}} = x I''_x \quad \text{with } \frac{\partial \nabla I(x', y')}{\partial x'} = I''_x.$$

Similarly,

$$\frac{\partial E}{\partial a_{12}} = -\int_\Omega y I''_x \, ds + \lambda_1 P_{a_{12}} + \lambda_2 Q_{a_{12}}, \tag{9.2}$$

where $P_{a_{12}} = 2a_{21}^2 - 2a_{21}a_{22} - 2a_{11}a_{22} + 4a_{12}a_{21} - 2a_{11}a_{21}$, and $Q_{a_{12}} = (\pi a_{21}/2\varepsilon^2)\sin \pi D/\varepsilon$.

$$\frac{\partial E}{\partial t_x} = -\int_\Omega I''_x \, ds, \tag{9.3}$$

$$\frac{\partial E}{\partial a_{21}} = -\int_\Omega x I''_y \, ds + \lambda_1 P_{a_{21}} + \lambda_2 Q_{a_{21}}, \tag{9.4}$$

where $P_{a_{21}} = 2a_{12}^2 - 2a_{12}a_{22} - 2a_{11}a_{22} + 4a_{12}a_{21} - 2a_{11}a_{12}$, $Q_{a_{21}} = (\pi a_{12}/2\varepsilon^2)\sin \pi D/\varepsilon$ and $\partial \nabla I(x', y')/\partial y' = I''_y$.

$$\frac{\partial E}{\partial a_{22}} = -\int_\Omega y I''_y \, ds + \lambda_1 P_{a_{22}} + \lambda_2 Q_{a_{22}}, \tag{9.5}$$

where $P_{a_{22}} = 2a_{11}^2 - 2a_{11}a_{12} - 2a_{11}a_{21} + 4a_{11}a_{22} - 2a_{12}a_{21}$ and $Q_{a_{22}} = -(\pi a_{11}/2\varepsilon^2)\sin \pi D/\varepsilon$. Finally,

$$\frac{\partial E}{\partial t_y} = -\int_\Omega I''_y \, ds. \tag{9.6}$$

The process is initialized with a set of affine parameters $\theta^0 = \{a_{ij}^0; t_i^0\}$. Given n points (x_i, y_i) representing the initial contour and following minimization process of Eqs. (9.1)–(9.6), the affine parameters at iteration $(t+1)$ are updated from previous step as follows:

$$a_{11}^{t+1} \leftarrow a_{11}^t - \Delta a_{11}\left(\sum x_i^t I''_x + \lambda_1 P_{a_{11}} + \lambda_2 Q_{a_{11}}\right), \tag{10.1}$$

$$a_{12}^{t+1} \leftarrow a_{12}^t - \Delta a_{12}\left(\sum y_i^t I''_x + \lambda_1 P_{a_{12}} + \lambda_2 Q_{a_{12}}\right), \tag{10.2}$$

$$t_x^{t+1} \leftarrow t_x^t - \Delta t_x\left(\sum I''_x\right), \tag{10.3}$$

$$a_{21}^{t+1} \leftarrow a_{21}^t - \Delta a_{21}\left(\sum x_i^t I''_y + \lambda_1 P_{a_{21}} + \lambda_2 Q_{a_{21}}\right), \tag{10.4}$$

$$a_{22}^{t+1} \leftarrow a_{22}^t - \Delta a_{22}\left(\sum y_i^t I''_y + \lambda_1 P_{a_{22}} + \lambda_2 Q_{a_{22}}\right), \tag{10.5}$$

$$t_y^{t+1} \leftarrow t_y^t - \Delta t_y\left(\sum I''_y\right). \tag{10.6}$$

The parameters $\Delta\theta = \{\Delta a_{ij}, \Delta t_i\}$ are the time steps in the corresponding steepest descent directions. For our implementation, we have considered each of the coupled nonlinear parameter update equations as independent. As every update reduces the energy such that $E(\theta^{new}) \leqslant E(\theta^{old}) + \alpha\langle \nabla E, (\theta^{new} - \theta^{old})\rangle$ for a very small $\alpha, 0 < \alpha < 1, \alpha \approx 10^{-4}$, it can be shown that the value of each element of $\Delta\theta$ less or equal to $\frac{1}{2}$ facilitates stability [20]. For the image gradient portion of the update equations, assuming image intensity is normalized within [0, 1], the maximum possible values of $\sum x_i I''_x \approx \sum x_i I''_y < \sum x_i$, $\sum y_i I''_x \approx \sum y_i I''_y < \sum y_i$ and $\sum I''_x \approx \sum I''_y < N$ where N is the number of contour points. Considering a single time step for each $\Delta\theta$, the time steps are normalized and taken as: $\Delta a_{11} = 1/(2\sum x)$, $\Delta a_{12} = 1/(2\sum y)$, $\Delta t_x = 1/(2N)$, $\Delta a_{21} = 1/(2\sum x)$, $\Delta a_{22} = 1/(2\sum y)$, $\Delta t_y = 1/(2N)$.

In the next section we present results for the numerical implementation of this minimization process on both synthetic and real images.

3. Results of affine curve evolution

The numerical implementation of the minimization technique given by Eq. (10) depends on three factors. First, the initial contour must be sufficiently close to the desired boundary, and, second, the time step must be sufficiently small to guarantee convergence. The third important constraint for curve evolution involves the initial set of affine parameters $\theta^0 = \{a_{ij}^0; t_i^0\}$. For the results presented in this paper, we have selected an initial affine matrix that imposes no geometric distortion in the initial contour including no initial translation values. Therefore, the initial values are: $a_{11} = 1$, $a_{12} = 0$, $a_{21} = 0$, $a_{22} = 1$ and $t_x = 0$ and $t_y = 0$. Notice that this initial set of affine parameters force $x' = x$ and $y' = y$ following Eq. (3). So, our initial affine parameters do not enforce any form of distortion on the shape and allow the image gradient to be the driving force in delineating the boundary. This prescription is an ideal initial condition if we observe Eq. (8). Any deviation from this initial condition will bias the shape toward a particular deformation that may be desirable only in the case where we have a priori knowledge of the shape and the transformation.

Figs. 2 and 3 are synthetic examples comparing curve evolution due to the proposed affine model and to the standard parametric snake [1]. Fig. 2(a) shows the transformation of a circular contour into an oriented elliptical shape, while Fig. 2(b) is the same except that the target elliptical contour contains noisy and jagged edges. The blue contours are initial contours; the green contours are the final contours,

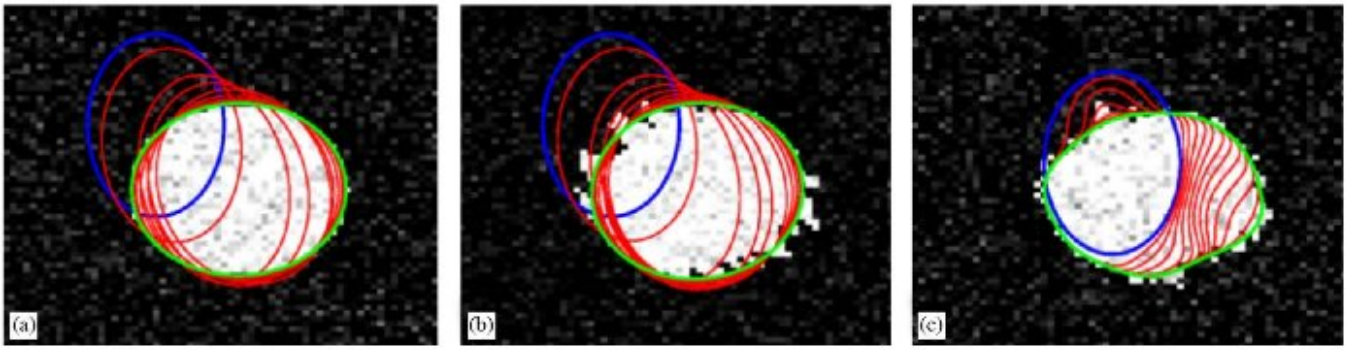

Fig. 2. (a) Initial contour (blue) is evolved and converges to oriented ellipse (green) following affine constraints. Red contours are intermediate steps. (b) Proposed affine curve evolution to detect ellipse with jagged edge. (c) Result of gradient-based active contour evolution.

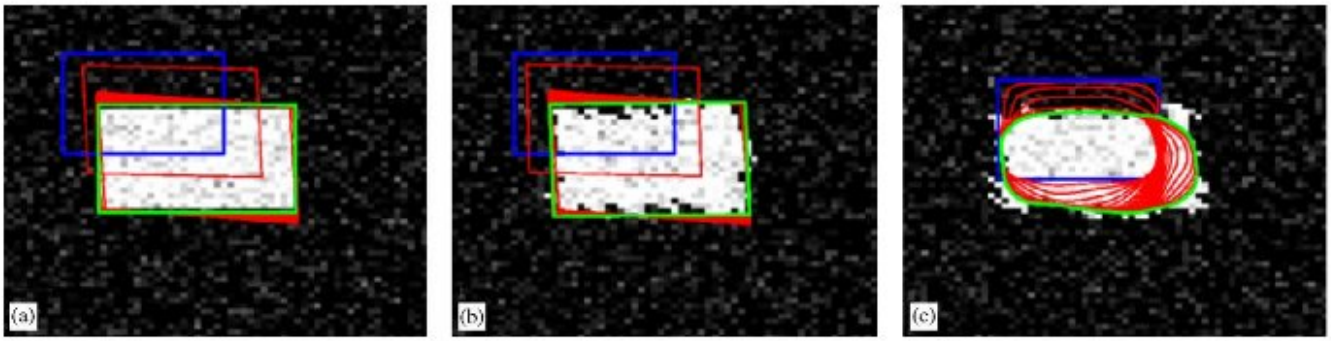

Fig. 3. (a) Proposed affine curve evolution for a rectangular shape. (b) Proposed affine curve evolution for noisy rectangular shape. (c) Result of gradient-based active contour evolution.

and the red contours are intermediate positions of the evolved curve. Fig. 2(c) shows the evolution of the gradient-based active contour algorithm following Ref. [1]. Fig. 3 presents the identical experiments for a rectangular initial contour in which unequal stretching along x- and y-axes captures the final contour. For both Figs. 2 and 3, the affine technique yields improved shape consistency compared to the gradient-based active contour.

Fig. 4 depicts a video sequence containing a bus in motion. From the affine active contour model, tracking results for the 1st, 5th, 10th and 15th frames are given in Figs. 4(a), (b), (c) and (d), respectively. Again, the green contours are the final results. There is an affine transformation of the rectangular bus top due to the relative motion with respect to the fixed camera, and the affine contour model is able to maintain shape consistency despite movement and shape transformation. The salient observation emerging from the results in Figs. 4(b)–(d) is the shape preservation in presence of occluding objects. The converged contour in Fig. 4(b) could lock onto edges from the lamppost in subsequent frames; however, the tracked contour in the 10th frame (Fig. 4(c)) shows that due to the affine parameter constraints, the shape of the rectangle is not distorted. Such shape consistency despite noise and irrelevant scene details is maintained throughout the sequence as shown in Fig. 4(d).

The corresponding result using the gradient vector flow (GVF) active contour is shown in Fig. 5 [2]. The GVF external forces are achieved via a vector diffusion operation of the gradient of the initial edge map. This diffusion extends the capture range of an object so that snakes can find objects that are distant from the initial snake position. Despite the fact that the initial contour is close to the desired boundary (see Fig. 5(a)), the active contour cannot delineate the rectangular bus top in the cluttered scene. The tracking results for the first, fourth and ninth frames are shown in Figs. 5(a), (b) and (c), respectively. Beyond the ninth frame, the contour has drifted from the object of interest. Due to the fact that each contour point is computed independently, as opposed to using the six parameters of the affine model, the GVF-based model is more expensive computationally. The energy functional for the active contour used in these results is [1]:

$$E(\psi(s)) = \frac{1}{2}\int_0^1 (\alpha\|\dot{\psi}(s)\|^2 + \beta\|\ddot{\psi}(s)\|^2)\,\mathrm{d}s + \gamma\int_0^1 E_{ext}(\psi(s))\,\mathrm{d}s. \quad (11)$$

The weights used for smoothness constraints are $\alpha=0.5$ and $\beta=1$, and the weight for the external force $E_{ext}(\cdot)$ is selected as $\gamma=4$ where the external force is derived from GVF [2].

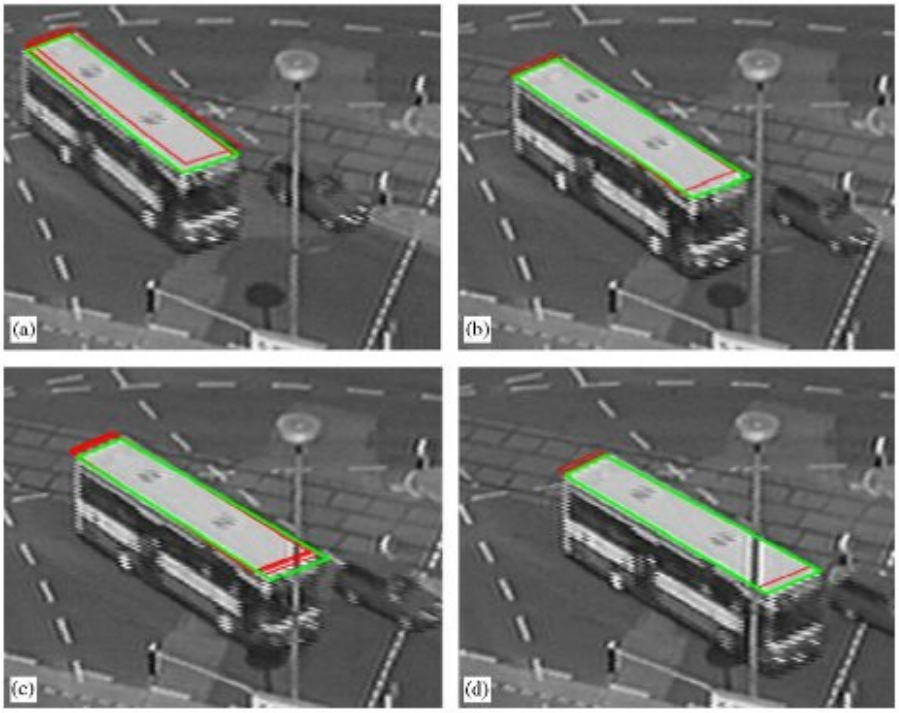

Fig. 4. (a)–(d) Red initial contour is evolved following affine model to track the moving bus top at 1st, 5th, 10th and 15th frame of the sequence. The tracking result is shown using green contour.

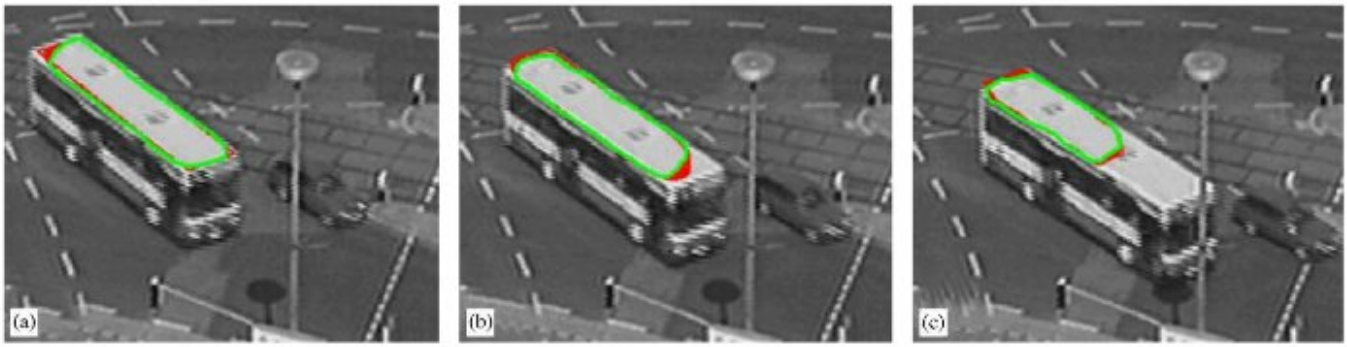

Fig. 5. (a)–(c) GVF-based active contour evolution after first, fourth and ninth frame.

Fig. 6 shows frames from a video sequence image where a white blood cell inside a venule of the mouse cremaster muscle is tracked using affine curve evolution. The elliptical shape of the white blood cell undergoes scaling and shearing as it flows through the microvasculature. In this case, the proposed affine technique is able to capture the cell in the entire sequence maintaining the shape consistency, in contrast to the GVF-based active contour, as shown in Fig. 7.

The results in Figs. 8 and 9 illustrate the importance of the restriction imposed on the determinant of the affine matrix. Figs. 8(a)–(e) give the output after affine contour evolution in every fourth frame of a sequence using Eq. (10) for non-zero values of λ_1 and λ_2, namely, $\lambda_1 = \lambda_2 = 0.1$. For Fig. 9(a)–(e), the initial contour is also evolved with affine model on same frames, but without any restriction imposed on the determinant of the affine matrix, i.e., we set $\lambda_1 = \lambda_2 = 0$ in Eq. (10). We can clearly see that the shape consistency and the delineation of blood cell contour are improved with the constraint, as shown in Fig. 8. This positive result is possible because arbitrary changes in the determinant of the affine matrix are not allowed, and the algorithm is less sensitive to minor variations in the image intensity.

3.1. Comparison with related approaches

The proposed curve evolution can be compared with the GVF-based active contour more objectively by considering

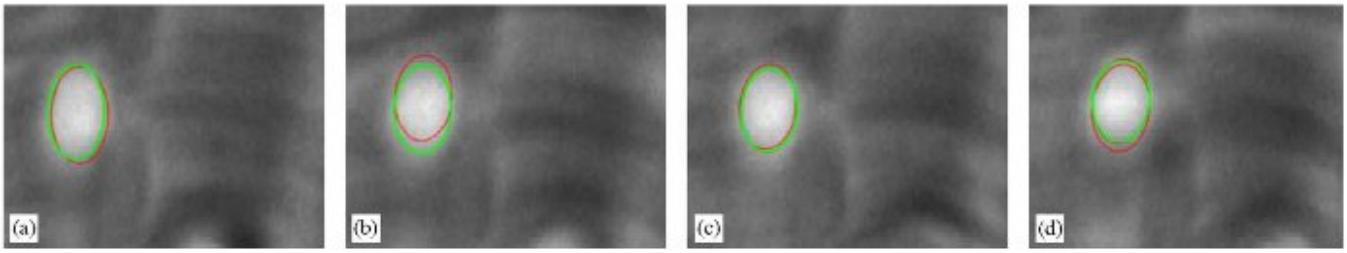

Fig. 6. (a)–(d) Tracking a white blood cell in 1st, 5th, 10th and 15th frame of the video sequence using affine curve evolution.

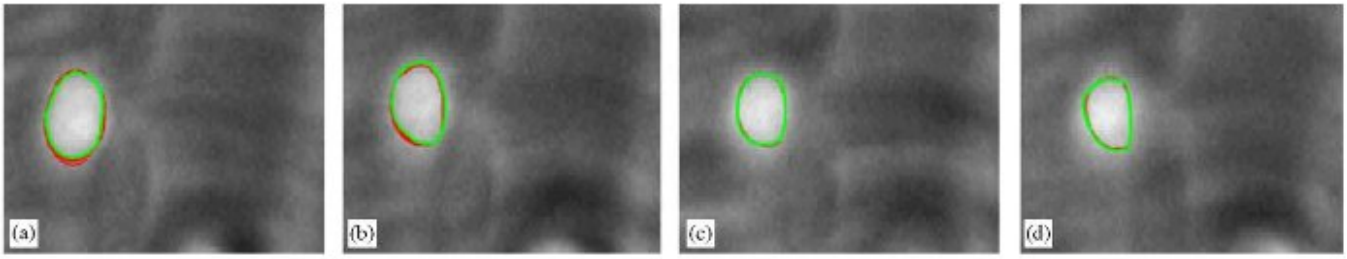

Fig. 7. (a)–(d) Tracking a white blood cell in 1st, 5th, 10th and 15th frame of the video sequence using GVF-based active contour evolution.

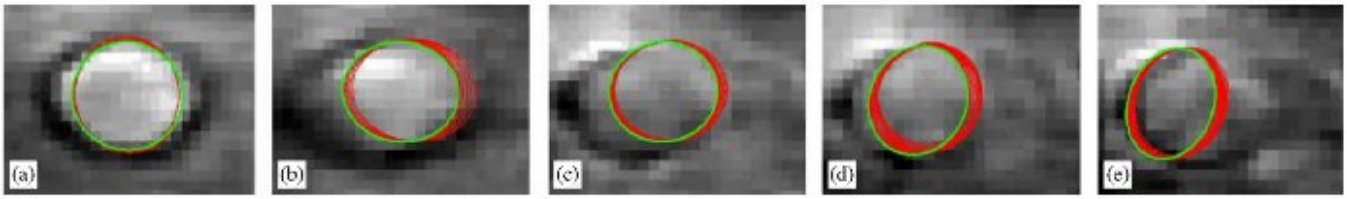

Fig. 8. (a)–(e) Blood cell sequence where cell boundary is captured using Eq. (10).

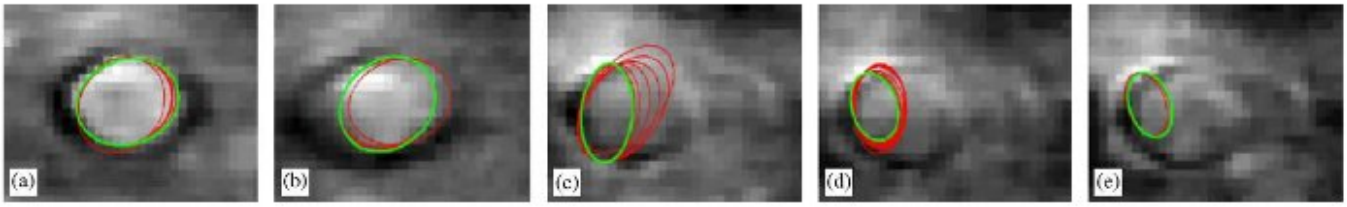

Fig. 9. (a)–(e) Blood cell sequence where cell boundary is captured using Eq. (10) for the case of $\lambda_1 = \lambda_2 = 0$.

the time and space complexity. Here, we also consider the convexity of the methods.

In the gradient-based active contour, minimizing energy equation (11) is typically implemented with the following implicit scheme:

$$\mathbf{x}^{t+1} = (\Delta t\,\mathbf{I} + \mathbf{A})^{-1}(\mathbf{f}_x + \mathbf{x}^t), \tag{12.1}$$

$$\mathbf{y}^{t+1} = (\Delta t\,\mathbf{I} + \mathbf{A})^{-1}(\mathbf{f}_y + \mathbf{y}^t), \tag{12.2}$$

with time step Δt and identity matrix $\mathbf{I}$. The set of active contour points on $\psi(s)$ is represented by two vectors, $\mathbf{x}$ and $\mathbf{y}$. $\mathbf{f}_x$ and $\mathbf{f}_y$ are either vectors containing the x and y directional derivatives of the image intensity, respectively, constituting the external image gradient force E_{ext} of Eq. (11) or the force field derived from GVF [2]. The matrix $\mathbf{A}$ is a penta-diagonal banded matrix of dimension $n \times n$, where n stands for the number of points in the active contour [1]. The traditional active contour evolution requires inversion of a matrix of large dimension (of the same size as number of points on the active contour) in each iteration in which resampling of the contour is implemented. Also, n equations need to be updated for n new positions of the contour points. In the proposed method we need an update for only six parameters representing the affine transformation. The processing time for curve evolution using both traditional and the proposed approach is shown in Table 1. Both implementations utilize Matlab 7 executed on a Pentium 4 (1.7 GHz) CPU. These anecdotal results show that the proposed affine snake model

Table 1
Comparison of curve evolution time using the proposed affine technique and the GVF-based active contour

Figure	Proposed technique (s)	GVF (s)
Bus sequence	7.1	11.1
Cell sequence	3.1	8.3
Tower sequence	6.4	7.4

incurs less computational expense than incurred by the traditional gradient snake model in actual implementation.

The approach presented in Ref. [15] has two distinct steps. In each iteration, the contour evolution is first implemented using the algorithm presented in Ref. [1]. This contour evolution is followed by estimation of the appropriate Lie group parameters to model the between-frame evolution. Since the 2D affine or projective transformation has six or eight degrees of freedom and the number of contour points are significantly greater in number, the transformation parameters are solved as an over-determined case using a least-squares solution. After the transformation parameters are estimated, the next contour positions are calculated by applying the Lie group transformation. The approach in Ref. [15] has a computational complexity similar to that of the combination of Ref. [1] and the numerical solution of the over-determined linear least-squares solution followed by a Lie group transformation. Therefore, the current approach that calculates Lie group transformation parameters directly through image gradient magnitude maxima of the affine transformed contour is appealing in a computational sense.

Results obtained using proposed approach and that of Ref. [15] are comparable except cases as seen in Fig. 4(c) or (d) where a lamppost is occluding the bus rooftop. In Ref. [15] since the contour will be mimicking gradient-based active contour of Ref. [1], the final deformation of the contour will be distracted by such an occlusion. In contrast, the approach given in this paper can successfully track the rooftop rectangular shape even in the presence of occlusion.

3.2. Convexity analysis

The convexity of the discrete active contour energy is important as emphasized in Refs. [21,22]. Under the assumption that a large number of contour points are being evolved in the active contour to estimate an object boundary, in Ref. [21], the convexity condition is derived by defining a ribbon region close to the desirable boundary. In Ref. [22], two important observations are made that are important in context of convexity analysis of the proposed technique. It is shown that the parameters of the model should be chosen in a way such that the solution (that is the object boundary) we seek is an attractor in the topology of the solution space. A singular point x is an attractor if the energy is strictly convex in the neighborhood of x, that is, if x is a local minimum of the energy functional. The energy functional presented here that seeks the maxima of the image gradient magnitude map satisfies the properties of an attractor, not globally, albeit locally.

For our proposed model, the energy functional in Eq. (8) is not a globally convex functional. A standard method to test the convexity of the function is through evaluating the principal minors of the Hessian of the function. For an $n \times n$ matrix there are n principal minors, and for the function to be strictly convex, each of the principal minors should be positive. Expressing Eq. (8) as $E = E_{affine} + \lambda_1 E_1 + \lambda_2 E_2$, and taking $D = (a_{11}a_{22} - a_{12}a_{21})$, we find that principal minors of E_1 and E_2 do not satisfy the convexity condition. For the E_{affine} to be convex, the principal minor $M(H(E_{affine})) > 0$ or the minimum eigenvalue of $H(E_{affine})$ should be strictly greater than zero where $H(E_{affine})$ is the Hessian of E_{affine}. Following Ref. [21], it can be shown that the minimum eigenvalue of $H(E_{affine})$ is positive when the initial contour is within a ribbon-like area centered on the desired boundary of the object. The desired object boundary represents a local minimum in the energy space E_{affine} as given in Eq. (8). The thickness of the ribbon is equal to the size of the kernel used to calculate the image gradients. Therefore, the convergence of the proposed scheme is guaranteed in the case where the spatial distance between the initial and the desired contour within a frame is less or equal to the mask size used to calculate I_x and I_y. Also, since E_1 and E_2 do not contribute toward convexity, it is desirable to reduce the weights λ_1 and λ_2 of Eq. (8) in evolving the affine curve.

In the next section we explore the extension of the proposed technique to the projective case.

4. Extension to projective curve evolution

As discussed in the previous sections, the affine active contour accommodates rotation, scaling and shear with different scaling along x- and y-axes. However, due to *non-fronto-parallel* imaging *very close* (compared to the scale of the object along viewing direction) to the object, the shearing effect on the shape may be pronounced in such a way that even the parallelism constraint of the object is not maintained. Viz., the parallel lines in the object appear as intersecting lines at a vanishing point. For such 2D object-to-image mappings, the projective transformation model used is

$$\begin{bmatrix} x_p \\ y_p \\ 1 \end{bmatrix} \cong \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & 1 \end{bmatrix} \begin{bmatrix} x \\ y \\ 1 \end{bmatrix}, \tag{13}$$

where the coordinates of points (x_p, y_p) and (x, y) are expressed in the homogeneous coordinates $(x_p, y_p, 1)$ and $(x, y, 1)$ [17]. The transformation in Eq. (13) contains eight unknowns $\{a_{ij}\}$ where the equality in Eq. (13) holds up to a scale factor and the spatial coordinates are expressed in

homogeneous space. The corresponding Euclidean coordinates are given by

$$x_p = (a_{11}x + a_{12}y + a_{13})/(a_{31}x + a_{32}y + 1) \tag{14.1}$$

and

$$y_p = (a_{21}x + a_{22}y + a_{23})/(a_{31}x + a_{32}y + 1). \tag{14.2}$$

Beyond the six degrees of freedom for the affine model in Eq. (3), the additional parameters $\{a_{31}, a_{32}\}$ encode the direction of the vanishing point.

Following the energy model of the affine contour given in Eq. (4), the external force for a projective active contour deformation can be modeled as

$$E = \int_{\Omega} |\nabla I(x_p, y_p)|\,\mathrm{d}s + \lambda_1 \sum_{a_{ij}} \frac{\partial D^2}{\partial a_{ij}} + \lambda_2 \delta_{\varepsilon}(D), \tag{15}$$

where regularization on determinant terms are added following derivations in Eqs. (5) and (8). In this case, the determinant of the upper left 2×2 matrix of the projective transformation matrix (responsible for shearing, scaling and rotation) is used to calculate D.

Minimizing E of Eq. (15) using conjugate gradient methods as in Section 3, the corresponding parameter update equations are

$$\frac{\partial E}{\partial a_{11}} = -\int_{\Omega} (x/A) I_x''\,\mathrm{d}s + \lambda_1 P_{a_{11}} + \lambda_2 Q_{a_{11}}, \tag{16.1}$$

with $A = (a_{31}x + a_{32}y + 1)$ and

$$\frac{\partial \nabla I(x_p, y_p)}{\partial a_{11}} = \frac{\partial \nabla I(x_p, y_p)}{\partial x_p}\frac{\partial x_p}{\partial a_{11}} + \frac{\partial \nabla I(x_p, y_p)}{\partial y_p}\frac{\partial y_p}{\partial a_{11}} = \frac{x}{A} I_x'' \quad \text{and} \quad \frac{\partial \nabla I(x_p, y_p)}{\partial x_p} = I_x''.$$

The notation for determinant regularization remains the same as that of Eq. (8). Similarly,

$$\frac{\partial E}{\partial a_{12}} = -\int_{\Omega} (y/A) I_x''\,\mathrm{d}s + \lambda_1 P_{a_{12}} + \lambda_2 Q_{a_{12}}, \tag{16.2}$$

$$\frac{\partial E}{\partial a_{13}} = -\int_{\Omega} (1/A) I_x''\,\mathrm{d}s, \tag{16.3}$$

$$\frac{\partial E}{\partial a_{21}} = -\int_{\Omega} (x/A) I_y''\,\mathrm{d}s + \lambda_1 P_{a_{21}} + \lambda_2 Q_{a_{21}}, \tag{16.4}$$

$$\frac{\partial E}{\partial a_{22}} = -\int_{\Omega} (y/A) I_y''\,\mathrm{d}s + \lambda_1 P_{a_{22}} + \lambda_2 Q_{a_{22}}, \tag{16.5}$$

$$\frac{\partial E}{\partial a_{23}} = -\int_{\Omega} (1/A) I_y''\,\mathrm{d}s, \tag{16.6}$$

$$\frac{\partial E}{\partial a_{31}} = -\int_{\Omega} \left(\frac{\partial x_p}{\partial a_{31}} I_x'' + \frac{\partial y_p}{\partial a_{31}} I_y''\right)\mathrm{d}s, \tag{16.7}$$

$$\frac{\partial E}{\partial a_{32}} = -\int_{\Omega} \left(\frac{\partial x_p}{\partial a_{32}} I_x'' + \frac{\partial y_p}{\partial a_{32}} I_y''\right)\mathrm{d}s. \tag{16.8}$$

We now provide results of projective curve evolution both with synthetic and real scenes. Akin to the affine case, projective curve evolution requires the specification of eight initial projective parameters $\{a_{ij}\}$. However, this selection is critical especially for the parameters $\{a_{31}, a_{32}\}$ that determine the vanishing direction. An improper selection of $\{a_{31}, a_{32}\}$ may be a valid projective transformation but may generate visually unacceptable curve positions in curve evolution. Illegal initial parameters are avoided by employing the gradient-based active contour. For the initial curve ζ, a curve ζ' is evolved with a single gradient descent step using Eq. (11). A projective transformation as in Eq. (13) is solved between ζ and ζ'. The equation is solved as an overdetermined system as we have more than four-point correspondences (a minimum of four-point correspondences are required to solve eight unknowns of Eq. (13)). After only one iteration of evolution of gradient-based active contour (which transforms ζ to ζ' where ζ' has marginal shape deformation with respect to ζ) and reparametrization of ζ', the number of contour points are identical for both ζ and ζ'. Therefore, the correspondence between ζ and ζ' is solved from the ordered lists of contour points of ζ and ζ'. Note that evaluation of the projective transformation parameters $\{a_{ij}\}$ between ζ and ζ' is solved only once for the initialization and is not required in subsequent iterations in which case the proposed model of Eq. (16) takes care of the update of $\{a_{ij}\}$. In case of examples involving projective transformation, since the deformation is more prominent compared to that of the affine case, this first iteration of active contour-based initialization is preferred over the identity transform used to calculate initial parameters in Section 3.

Fig. 10 contains a synthetic image used to demonstrate the projective curve evolution. In Fig. 10(a), an initial contour in the shape of a trapezoid is evolved to capture the rectangular object. This is a case of projective transformation where lines that are not parallel (the left and right sides of the trapezoid) are transformed to a parallel line pair in the rectangle. For initialization, the curve is evolved with a single gradient descent move (using Eq. (11)), and the projective transformation between the initial and the evolved curve[1] (Fig. 10(b)) is solved. The blue contour close to the initial red contour (Fig. 10(b)) is then generated using the projective transformation, and the coefficients are used as initialization for projective contour evolution. Successful delineation of the rectangular shape is demonstrated in Fig. 10(c). The corresponding result using gradient-based active contour evolution is shown in Fig. 10(d).

Fig. 11 shows frames from a video sequence in which a large tower is viewed aerially from a close distance such that projective distortion occurs in the face of the tower.

[1] The deformed curve where shape coherency is not maintained.

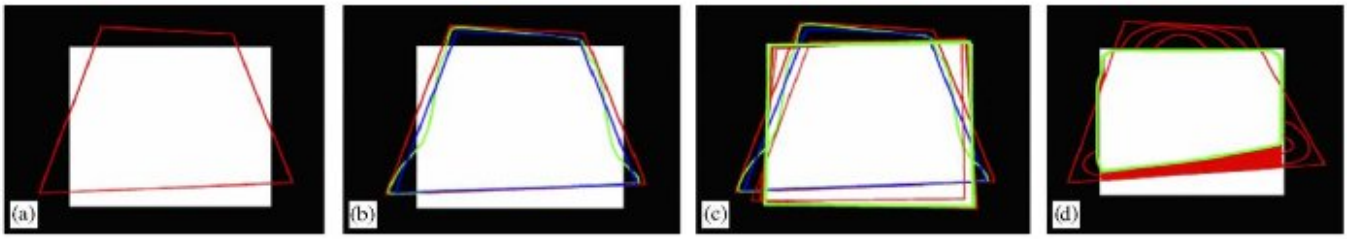

Fig. 10. (a) Initial contour in the shape of a trapezoid for detecting rectangular object boundary. (b) Projective contour initialization using single iteration of gradient-based active contour evolution. (c) Successful detection of the rectangular boundary after projective curve evolution. (d) Result of gradient-based active contour with initialization as in (a).

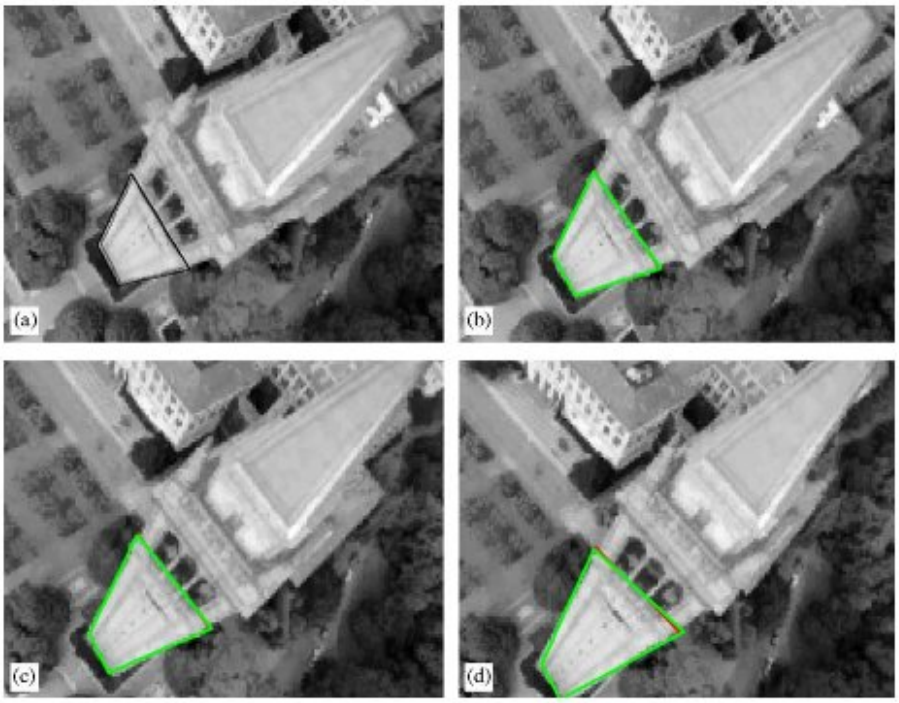

Fig. 11. (a) Initialization for the projective tower sequence. (b)–(d) Tracking of the tower face for (every third) frames using projective active contour model.

Fig. 12. (a)–(c) Tracking of frames as in Figs. 11(b)–(d) using GVF-based active contour with similar initialization as in Fig. 11(a).

The initial contour is shown in Fig. 11(a). In Fig. 11(b)–(d), we have shown that it is possible to track the rectangular face of the tower, which appears to be trapezoidal due to perspective, using projective active contour. The corresponding tracking task using the GVF active contour results in failure as shown in Fig. 12(a)–(c).

5. Summary

We put forth an active contour evolution process that is governed by the constraints of affine and projective object-to-image transformations. The motivation is to accommodate tracking problems where objects are being viewed from non-fronto-parallel vantage points and from close viewing distances. The process maintains shape coherency of the object and is robust to shape details and intensity variations, which in the case of the gradient-based active contour may deform the active contour in an undesirable manner. The efficacy of the method is demonstrated with a series of synthetic and real image sequences. Compared to related approaches, the modicum of parameters in the affine/projective models leads to computational efficiency. The convexity analysis shows that, similar to conventional active contours, initialization of the contour needs to be near the desired boundary.

Current work on the affine and projective active contours is focused on a combination of local (at each active contour point) and global (transformation parameters) constraints for contour evolution. We are also considering the extension of these models to 3D object tracking.

Acknowledgments

The authors wish to thank the reviewers for the meticulous critique that has led to an improved paper.

References

[1] M. Kass, A. Witkin, D. Terzopoulos, Snakes: active contour models, Int. J. Comput. Vis. 1 (4) (1988) 321–331.
[2] C. Xu, J.L. Prince, Generalized gradient vector flow external forces for active contours, Signal Process. 71 (1998) 131–139.
[3] F. Leymarie, M.D. Levine, Tracking deformable objects in the plane using an active contour model, IEEE Trans. Pattern Anal. Mach. Intell. 15 (1993) 617–634.
[4] S. Logbregt, M.A. Viergever, A discrete dynamic contour model, IEEE Trans. Med. Imaging 14 (1995) 12–14.
[5] C. Vieren, F. Cabestaing, J.-G. Postaire, Catching moving objects with snakes for motion tracking, Pattern Recognition Lett. 16 (1995) 679–685.
[6] S.T. Acton, K. Wethmar, K. Ley, Automatic tracking of leukocytes in vivo, Microvasc. Res. 63 (2002) 139–148.
[7] A. Blake, M. Isard, Active Contours: The Application of Techniques from Graphics, Vision, Control Theory and Statistics to Visual Tracking of Shapes in Motion, Springer, Berlin, 1998.
[8] D. Terzopolous, R. Szeliski, Tracking with Kalman snakes, in: A. Blake, A. Yuille (Eds.), Active Vision, MIT Press, Cambridge, MA, 1992, pp. 3–20.
[9] R. Deriche, N. Paragios, Geodesic active contours and level sets for the detection and tracking of moving objects, IEEE Trans. Pattern Anal. Mach. Intell. 22 (2000) 266–280.
[10] A.-R. Mansouri, Region tracking via level set PDEs without motion computation, IEEE Trans. Pattern Anal. Mach. Intell. 24 (2002) 947–961.
[11] D.P. Mukherjee, N. Ray, S.T. Acton, Level set analysis for leukocyte detection and tracking, IEEE Trans. Image Process. 13 (4) (2004) 1–11.
[12] A.A. Amini, T.E. Weymouth, R.C. Jain, Using dynamic programming for solving variational problems in vision, IEEE Trans. Pattern Anal. Mach. Intell. 12 (9) (1990) 855–867.
[13] K.F. Lai, R.T. Chin, Deformable contours: modeling and extraction, in: Internatinal Conference on Visual Pattern Recognition, 1994, pp. 601–608.
[14] N. Ray, S.T. Acton, K. Ley, Tracking leukocytes in vivo with shape and size constrained active contours, IEEE Trans. Med. Imaging 21 (10) (2002) (special issue on Image Analysis in Drug Discovery and Clinical Trials).
[15] 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.
[16] M. Nakahara, Geometry, Topology and Physics, Institute of Physics Publishing, Philadelphia, 1992.
[17] D.P. Mukherjee, A. Zisserman, M. Brady, Shape from symmetry—detecting and exploiting symmetry in affine images, Philos. Trans. R. Soc. London Ser. A 351 (1995) 77–106.
[18] B. Bascle, P. Bouthemy, R. Deriche, F. Meyer, Tracking complex primitives in an image sequence, in: Proceedings of the 12th ICPR, 1994, pp. 426–431.
[19] B. Bascle, R. Deriche, Region tracking through image sequences, in: Proceedings of ICCV, 1995, pp. 302–307.
[20] W.H. Press, et al., Numerical Recipes in C, Cambridge University Press, UK, 1992.
[21] C.A. Davatzikos, J.L. Prince, Convexity analysis of active contour algorithms, Image Vis. Comput. 17 (1999) 27–36.
[22] G.A. Giraldi, A.F. Oliveira, Convexity analysis of snake models based on hamiltonian formulation, Technical Report, Universidade Federal do Rio de Janeiro, Dep. Eng. Sistemas e Computação, ⟨http://www.cos.ufrj.br/relatorios/reltec99/⟩, 1999.

About the Author—DIPTI PRASAD MUKHERJEE, Ph.D., is the Professor in the Electronics and Communication Sciences Unit of the Indian Statistical Institute, Kolkata. He has published three books and more than 70 peer-reviewed papers. He is the recipient of pre-doctoral fellowship to University of Oxford, UK and UNESCO-CIMPA fellowships to INRIA, France and to ICTP, Italy and held visiting faculty positions at Oklahoma State University, University of Virginia, USA and University of Alberta, Canada. He is the senior member of IEEE and Computer Society of India and had served in the editorial board of IEEE Signal Processing Letters.

About the Author—SCOTT T. ACTON received the M.S. degree in Electrical and Computer Engineering and the Ph.D. degree in Electrical and Computer Engineering from the University of Texas at Austin in 1990 and 1993, respectively, where he was a Microelectronics and Computer Development Fellow. He received the B.S. degree in Electrical Engineering from Virginia Tech, Blacksburg in 1988 as a Virginia Scholar and Marshall Hahn Fellow. He is a 1984 graduate of state football champion Oakton High School.
Dr. Acton has worked in industry for AT&T, Oakton, VA, the MITRE Corporation, McLean, VA and Motorola, Inc., Phoenix, AZ and in academia for Oklahoma State University, Stillwater. Currently, he is a Professor at the University of Virginia (U.VA), where he is a member of the Charles L. Brown Department of Electrical and Computer Engineering and the Department of Biomedical Engineering. At U.VA, he was named the Outstanding New Teacher in 2002 and was elected a Faculty Fellow in 2003. In 1997, he was named the Eta Kappa Nu Outstanding Young Electrical Engineer—a national award that has been given annually since 1936. His research interests include anisotropic diffusion, analysis of circuit board images, active contours, biomedical segmentation problems, biomedical tracking problems and military image analysis.