Velocity-Guided Tracking of Deformable Contours in Three Dimensional Space

Reuven Zaritsky[1], Natan Peterfreund[2], and Nahum Shimkin[1]

[1] Department of Electrical Engineering, Technion – Israel Institute of Technology, Haifa 32000, Israel,
shimkin@ee.technion.ac.il
[2] Oak Ridge National Laboratory, Oak Ridge, TN 37831-6355, USA,
v4p@ornl.gov

Abstract. This paper presents a 3D active contour model for boundary tracking, motion analysis and position prediction of non-rigid objects, which applies stereo vision and velocity control to the class of deformable contour models, known as snakes. The proposed contour evolves in three dimensional space in reaction to a 3D potential function, which is derived by projecting the contour onto the 2D stereo images. The potential function is augmented by a velocity term, which is related to the three dimensional velocity field along the contour, and is used to guide the contour displacement between subsequent images. This leads to improved spatio-temporal tracking performance, which is demonstrated through experimental results with real and synthetic images. Good tracking performance is obtained with as little as one iteration per frame, which provides a considerable advantage for real time operation.

1 Introduction

Deformable contours have been established in the last decade as a major tool for low level vision functions, including edge detection, segmentation and tracking of non-rigid objects. Boundary detection and tracking based on deformable planar contours, known as snakes, were originally introduced by Kass, Witkins and Terzopoulos [10]. Energy-minimizing active contours are deformable contours that move under the influence of image-induced potential, subject to certain internal deformation constraints. The contour dynamics may be specified by the Euler-Lagrange equations of motion associated with the contour potential. Using the image gradient as the potential function, for example, results in edge-seeking deformation forces, leading the contour towards high contrast boundaries, thus performing boundary detection. Many variations and extension of the original snake model have been reported in recent literature, see [3,13,4,9] and references therein.

In this paper we present an active contour model in three dimensional space, which is intended to for temporal tracking of the contour of three dimensional objects in stereo image sequences.

D. Vernon (Ed.): ECCV 2000, LNCS 1842, pp. 251–266, 2000.

Contour tracking in a sequence of images may be performed by initializing the snake in each image using its last position in the previous frame. A significant improvement on this basic scheme may be obtained by incorporating inter-frame velocity estimates to improve the initial positioning. By starting the snake closer to the actual object boundary, we obtain improved convergence and reduced likelihood of trapping in local minima. Velocity information may be obtained by extrapolating the contour motion from previous images, using Kalman filtering or related methods; see, e.g., [18,4]. Another approach, more relevant to the present study, is to extract an estimate of local or global velocity from two adjacent images. In [1] the authors use correlation and rigid motion models to estimate the global motion of the object between the two frames. The work reported in [5] proposes an explicit calculation of the local velocity field along the detected contour, using optical flow methods, which is then used to predict the next position of the contour. A related idea [14] is to use a potential term related to the temporal difference between images, in order to focus the tracking contour on regions of temporal change, as an indicative of object motion against a stationary background.

Recently, and integrated spatio-temporal snake model was proposed for 2D contour tracking in combined spatio-velocity space [16,17]. The *velocity snake* model uses the optical flow constraints in order to propagate the contour in the direction of the local velocity field. Explicit computation of the optical flow is avoided by incorporating the optical flow constraint equation directly into the snake dynamics equations. The improved tracking performance of the model was demonstrated theoretically by treating the image sequence as continuous measurements along time, where it was shown that the proposed model converges with no tracking error to a boundary moving at a constant velocity [16]. This is in contrast to the basic snake model which is biased due to lack of image velocity input. The present paper generalizes these ideas to 3D contour tracking of shape and motion in stereo images.

Deformable contours are well suited for stereo matching, as they can perform simultaneously the two tasks of feature detection and correspondence. This was originally realized in [10], which proposes to use different 2D contours in each of the stereo images, coupled by a stereo disparity smoothness term in their respective potential functionals. A deformable spline model for stereo matching that evolves in 3D was proposed in [2], which employs an additive potential function based on the projections of the curve on each image. A similar idea will be used in the tracking scheme proposed here. In [6], the authors propose an affine epipolar geometry scheme for coupling pairs of active contours in stereo images to enhance stereo tracking of 3D objects.

The active contour model proposed here is a parameterized curve which evolves in three dimensional space under the influence of a three dimensional potential function. We use an additive potential which is derived by projecting the 3D contour onto the the two stereo images, where 2D gradient potentials are defined. An alternative multiplicative potential is also considered. The basic potential is augmented by a velocity term related to the optical flow in each of the

images. This term provides an additional force which approximately tracks the actual contour motion. The resulting dynamic equation of the tracking contour is realized in discrete time and space, with a single iteration step per frame.

It should be mentioned that the methods used in the paper are easily scalable to more than two cameras, as well as general geometric configurations.

The rest of the paper is organized as follows. In the next section we briefly describe the basic velocity snake model in two dimensions. The complete three dimensional model is presented in Section 3, followed by a discussion of some implementation considerations in Section 4. Experimental results are presented and discussed in Section 5, followed by some concluding remarks.

2 The Velocity Snake

In this section we present the basics of the two dimensional Velocity Snake proposed in [17]. Consider the closed contour $v(s,t) = (x(s,t)\,, y(s,t))$ for a parametric domain $s \in [0,1]$ and time $t \in [0,\infty)$. The snake Lagrangian, originally proposed by Terzopoulos et. al. [18], is given by:

$$L = \frac{1}{2}\int_0^1 \mu|v_t|^2 ds - \frac{1}{2}\int_0^1 \left(w_1|v_s|^2 + w_2|v_{ss}|^2\right) ds - \frac{1}{2}\int_0^1 P(v) ds \qquad (1)$$

The first term is the kinetic energy field, where $\mu(s)$ is the mass of the snake. The second term defines the internal deformation energy where $w_1(s)$ controls the tension and $w_2(s)$ controls the rigidity of the snake, respectively. The third term is the potential field energy of the contour, imposed by the image. The potential energy may be derived from a single (fixed) image, in which case the snake will converge to a static shape, or it may be derived from a temporal sequence of images, for tracking purposes. In the latter case, the potential becomes a function of time (c.f. [17]). In this paper, we consider for concreteness the common case of edge oriented potential:

$$P(x,y,t) = -c\,|\nabla\left[G_\sigma * I(x,y,t)\right]| \qquad (2)$$

This will cause the snake to be attracted to edges. The function $I(x,y,t)$ denotes the brightness of pixel (x,y) at time t in the image, and G_σ depicts a Gaussian filtering window with variance σ. The energy dissipation function which is used in conjunction with the Lagrangian (1) to describe the effects of nonconservative forces is defined here by:

$$D(v_t, v_t^b) = \frac{\gamma}{2}\int_0^1 \left\|L^T(v_t - v_t^b)\right\|^2 ds + \frac{\beta}{2}\int_0^1 \left\|\frac{\partial}{\partial s} v_t\right\|^2 ds \qquad (3)$$

where L is a real matrix and v^b and v_t^b denote the boundary position and velocity, respectively. The second term represents a smoothness constraint. Using the above Lagrangian (1) and the dissipation function (3), the Euler-Lagrange

equations of motion of the velocity snake are given by:

$$\mu v_{tt} + \gamma c(v_t, v_t^b) \; -\beta \frac{\partial}{\partial s}\left(\frac{\partial}{\partial s} v_t\right) - \frac{\partial}{\partial s}(w_1 v_s) + \frac{\partial^2}{\partial s^2}(w_2 v_{ss}) = -\nabla P(v(s,t)\,,\, t) \tag{4}$$

where: $c(v_t, v_t^b) = LL^T(v_t - v_t^b)$. Since the boundary velocity of the object is unknown in advance, we use instead the apparent velocity (optical flow) v_t^i of the image at the contour position.

A remarkable special case exists when $L = (\nabla I(v(s,t)))$ and v_t is replaced by the apparent velocity. In this case we obtain the integrated optical flow model [17]:

$$\mu v_{tt} + \gamma \nabla I(\nabla I^T v_t + I_t) \; -\beta \frac{\partial}{\partial s}\left(\frac{\partial}{\partial s} v_t\right) - \frac{\partial}{\partial s}(w_1 v_s) + \frac{\partial^2}{\partial s^2}(w_2 v_{ss}) = -\nabla P(v(s,t)\,,\, t) \tag{5}$$

This result is due to the optical flow constraint equation [9]:

$$\nabla I(x,t,t)^T \left(\frac{dx}{dt}\,,\, \frac{dy}{dt}\right)^T + I_t(x,y,t) = 0 \tag{6}$$

Compared to the general model (4), the model presented by (5) does not require an explicit estimation of the image velocity.

3 The 3D Velocity Snake Model

Given stereo image sequences, the present paper extends the two dimensional velocity snake to object tracking in three dimensions. The required 3D potential function is extracted at each point in time from the two stereo images. The velocity term serves to improve the quality of the 3D tracking. In the following, we briefly present the relevant 3D to 2D relations. The 3D space-velocity tracking model is then derived.

3.1 3D Projections

We assume here the following layout of the stereo cameras. The two cameras are parallel and positioned at the same height. The distance between the two cameras is b, and between the focus and the image plane is f. The focal point of each camera is positioned at the center of each image. The world coordinate axes is in the plane which includes the focal points of the two images and is at a distance of $b/2$ from each focal point.

Consider a point $P = (X, Y, Z)$ in world coordinates. In the left and right images the point will be projected onto $p_\ell = (x_\ell, y)$ and $p_r = (x_r, y)$ respectively. We have [11]

$$\begin{aligned} \frac{x_\ell}{f} &= \frac{X + b/2}{Z} \quad \frac{x_r}{f} = \frac{X - b/2}{Z} \\ \frac{y_1}{f} &= \frac{y_r}{f} = \frac{Y}{Z} \end{aligned} \tag{7}$$

For the velocity snake we will also require the equations linking the three dimensional velocity and the two dimensional image velocity. Differentiating equation (7) in time we obtain:

$$\begin{bmatrix} \dot{x}_\ell \\ \dot{y} \end{bmatrix} = H_\ell \begin{bmatrix} \dot{X} \\ \dot{Y} \\ \dot{Z} \end{bmatrix}, \begin{bmatrix} \dot{x}_r \\ \dot{y} \end{bmatrix} = H_r \begin{bmatrix} \dot{X} \\ \dot{Y} \\ \dot{Z} \end{bmatrix} \tag{8}$$

where:

$$H_\ell = [3D \to 2D]_\ell \triangleq \begin{bmatrix} \frac{f}{Z} & 0 & \frac{-\left(X + \frac{b}{2}\right)}{Z^2} f \\ 0 & \frac{f}{Z} & -\frac{Y}{Z^2} f \end{bmatrix}$$

$$H_r = [3D \to 2D]_r \triangleq \begin{bmatrix} \frac{f}{Z} & 0 & \frac{-\left(X - \frac{b}{2}\right)}{Z^2} f \\ 0 & \frac{f}{Z} & -\frac{Y}{Z^2} f \end{bmatrix}$$

3.2 The 3D Tracking Model

We define the snake as a 3D parametric contour $V(s,t)$ that depicts the points of the contour in world coordinates at time t, $V(s,t) = [X(x,t),\, Y(s,t),\, Z(s,t)]$, where $s \in [0,1]$ and $t \in [0,\infty)$. We will define the Lagrangian energy of the 3D snake similar to the 2D case as:

$$L = \frac{1}{2}\int_0^1 \mu |V_t|^2 ds - \frac{1}{2}\int_0^1 \left(w_1 |V_s|^2 + w_2 |V_{ss}|^2\right) ds - \frac{1}{2}\int_0^1 P(V) ds \tag{9}$$

The first term of equation (9) is the 3D kinetic energy field (where $\mu(s)$ is the mass of the snake). The second term defines the 3D internal deformation energy where $w_1(s)$ controls the tension and $w_2(s)$ controls the rigidity of the snake. We note that in this model, we do not include the torsion term (which exists in three dimensions, but not in two dimensions) in order to simplify the resulting equations. Our experimental results have indicated that such term does not contribute to the tracking performance.

The third term in equation (9) is the potential energy field. The 3D potential energy is obtained by an additive combination of the stereo potential fields

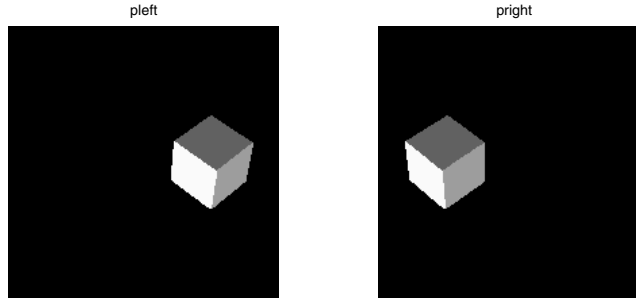

Fig. 1. Synthetic stereo-image-pair of a 3D cubic

evaluated at the projection of the 3D position onto the image planes:

$$P = \alpha^* P_\ell \left(\frac{X(s) + b/2}{Z(s)} f \,,\, \frac{Y(s)}{Z(s)} f \right) \;+\; \beta^* P_r \left(\frac{X(s) - b/2}{Z(s)} f \,,\, \frac{Y(s)}{Z(s)} f \right) \quad (10)$$

Here P_ℓ and P_r are the left and right 2D image potentials, respectively.

The desirable property of this potential function is that it has local minima at the combined local minima of the two component potentials, which in turn correspond to the local minima of the boundary edges in 3D space. While this may be the simplest combination with this property, other reasonable combinations exist. For example, an alternative potential to the one in (10) is the multiplicative potential

$$P = -\alpha \sqrt{P_\ell P_r} \quad (11)$$

(note that all potentials here are assumed negative). Let us briefly discuss and compare the properties of these two potentials. While both potential functions obtain local minima at the boundary-edges of the 3-D object, their behavior in the vicinity points may differ, depending on image content and clutter.

Figures 1–3 provide an illustration of these two potential functions. We use a synthetic stereo image pair of a 3-D cubic given in Figure 1. Figures 2 and 3 show the computed values of the potential functions (10) and (11) at Y=0, as functions of the 3D coordinates X and Z. Note that the absolute value of the potential is depicted, so that minima of the potential correspond to maxima in these figures; we shall refer in the following discussion to the magnitude of the potential. Both functions are seen to have four major local maxima, corresponding to the true locations of the four cubic edges intersecting $Y = 0$. Consider the additive potential (10) first. According to Figure 2, each of these maxima is formed by the intersection of two "ridges", each contributed by the the potential energy term of a single image. The "ridge" is formed along the epipole line intersecting the image at the object's edge, and due to the additive nature of this potential it leads to a considerable magnitude of the overall potential (and possible local maxima) even off the intersection with the other ridge. These spurious maxima disappear in the case of the multiplicative potential function (11), as illustrated in Figure 3, where the potential will be whenever one of its components is small

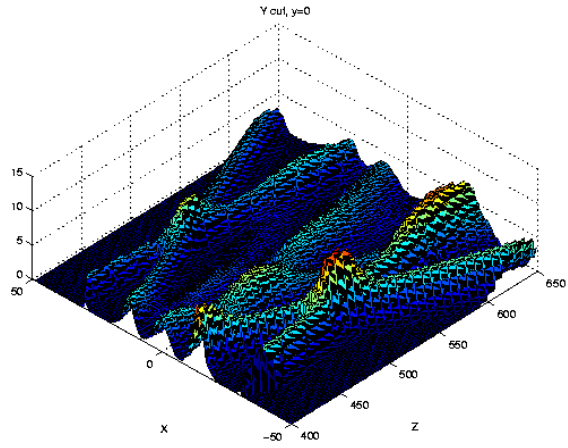

Fig. 2. The additive potential function (10) of the stereo-pair in Figure 1 along the coordinate Y=0, as a function of the 3-D coordinates X and Z. The magnitude (absolute value) of the potential is depicted.

enough. While this quenching of the potential off the intersection points seems to lead to a smoother function and possibly decrease the chances of entrapment in local extrema, the additive potential may posses some advantage when the deformable contour is far from the desired position as the information provided from one image can guide it to the correct position. Moreover, in the presence of partial occlusion in one image of the stereo-pair, the additive potential (10) provides partial support for tracking based on the contribution of a single image to the potential field, while the multiplicative potential field (11) is nulled. In the experiments performed within this work, where fairly close tracking was maintained, the two potentials yielded similar performance.

In analogy to the 2D case we propose the 3D dissipation function given by:

$$D(V_t, V_t^b) = \frac{\gamma}{2}\int_0^1 \left\| L_\ell^T (V_t - V_t^b) \right\|^2 ds + \frac{\gamma}{2}\int_0^1 \left\| L_r^T (V_t - V_t^b) \right\|^2 ds + \frac{\beta}{2}\int_0^1 \left\| \frac{\partial}{\partial s} V_t \right\|^2 ds \qquad (12)$$

where γ and β are positive scalars, V^b and V_t^b denote the 3D boundary position and velocity, respectively, and L_ℓ and L_r are real matrices which will later be defined as functions of the left and right image data. The third term represents a smoothness constraint. The first two terms provide the velocity tracking term in 3D space. In the following we show how these terms may be computed without explicit computation of V_t^b.

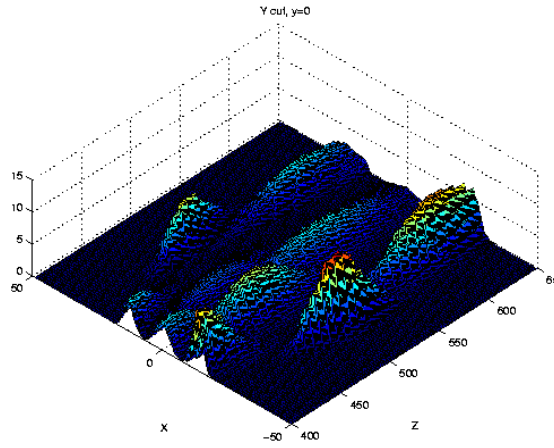

Fig. 3. The multiplicative potential function (11) of the stereo-pair in Figure 1 along the coordinate Y=0, as function of the 3-D coordinates X and Z.

Using the Lagrangian (9) and the dissipation function (12), the Euler-Lagrange equations of motion are given by:

$$\mu V_{tt} + \gamma \left(C_\ell(V_t, V_t^b) + C_r(V_t, V_t^b) \right) - \beta \frac{\partial}{\partial s} \left(\frac{\partial}{\partial s} V_t \right) \ldots$$
$$- \frac{\partial}{\partial s}(w_1 V_s) + \frac{\partial^2}{\partial s^2}(w_2 V_{ss}) = -\nabla P \left(V(s,t) \right), t) \tag{13}$$

where:

$$C_\ell(V_t, V_t^b) = L_\ell L_\ell^T (V_t - V_t^b),$$
$$C_r(V_t, V_t^b) = L_r L_r^T (V_t - V_t^b)$$

An explicit computation of the 3D object velocity V_t^b is obviously a demanding task. This may be circumvented by defining the following image dependent weight matrices $L_\ell^T = \nabla I_\ell^T H_\ell$ and $L_r^T = \nabla I_r^T H_r$ where H_ℓ and H are the projection matrices in (8) and I_ℓ and I_r denote the left and right images. This gives:

$$C_\ell(V_t, V_t^b) = H_\ell^T \nabla I_\ell (\nabla I_\ell^T H_\ell V_t - \nabla I_\ell^T H_\ell V_t^b)$$
$$C_r(V_t, V_t^b) = H_r^T \nabla I_r (\nabla I_r^T H_r V_t - \nabla I_r^T H_r V_t^b) \tag{14}$$

The term $\nabla I^T H V_t^b$ is equal to $\nabla I^T v_t^b$ where v_t^b is the apparent motion of the image boundary. Approximating v_t^b by the apparent motion of the image v_t^i at the snake position in each image, and then using the optical flow constraint equation (6), this term reduces to $-I_t$. C_ℓ and C_r then become:

$$C_\ell(V_t, V_t^b) = H_\ell^T \nabla I_\ell \left(\nabla I_\ell^T H_\ell V_t + (I_\ell)_t \right)$$
$$C_r(V_t, V_t^b) = H_r^T \nabla I_r \left(\nabla I_r^T H_r V_t + (I_r)_t \right) \tag{15}$$

The resulting Euler-Lagrange equation given by (13) and (15) is easily implemented by direct discretization.

3.3 Space-Time Discretization

We briefly describe in this subsection the discretization in space and time of the equations of motion of the 3D velocity snake. The continuous model is transformed into simple discrete-time equations with nonlinear image-based inputs through conventional finite difference techniques. Consider first the space discretization of the equations of motion as proposed above. Let $U = \lfloor U^1, \ldots, U^N \rfloor$ be the equidistant sampling vector of the contour $V(s)$ with a sampling distance of $h = 1/N$. Using a simple finite difference approximation for the space derivatives, and substituting (15) into (13), we obtain the discrete version of the equations of motion:

$$\begin{aligned} MU_{tt} + \gamma \left(H_\ell^T \nabla I_\ell \left(\nabla I_\ell^T H_\ell U_t + (I_\ell)_t \right) + H_r^T \nabla I_r \left(\nabla I_r^T H_r U_t + (I_r)_t \right) \right) \\ + \beta D D^T U_t + KU = -\nabla P(U) \end{aligned} \tag{16}$$

where K is the deformation matrix defined in [18], D is the "derivative" matrix defined in [17] and M is the mass matrix. For brevity we omit the details here. For comparison, a 3D tracking contour which employs the approach suggested by the original snake model (as outlined e.g. in [18]) does not have the velocity control and smoothing terms, and is given by

$$MU_{tt} + \gamma U_t + KU = -\nabla P(U) \tag{17}$$

The discretization in time of (16) is again performed by straightforward central difference approximation. The sampling interval T is taken as the video interframe interval.

4 Experimental Results

We demonstrate the performance of the proposed three dimensional velocity contour model by applying it to simulated and to real image sequences. We first show the tracking capability of the three dimensional velocity snake (16), applied a synthetic cube sequence, and compare the results to that of the three dimensional active contour (17) which doesn't employ the velocity control term. As was noted, the latter is a direct generalization of the Kalman Snake of Szeliski and Terzopoulos [18] to the three dimensional space. The tracking capability is then demonstrated on two real stereo-image sequences with rigid motion (a book) and with nonrigid one (a hand).

Prior to the calculation of the image gradients and derivatives, the image sequences were smoothed in space by a Gaussian filter with $\sigma = 2$, for the potential computation, and with $\sigma = 5$, for the velocity measurements. The sequence for the velocity measurements was also smoothed along time. Note

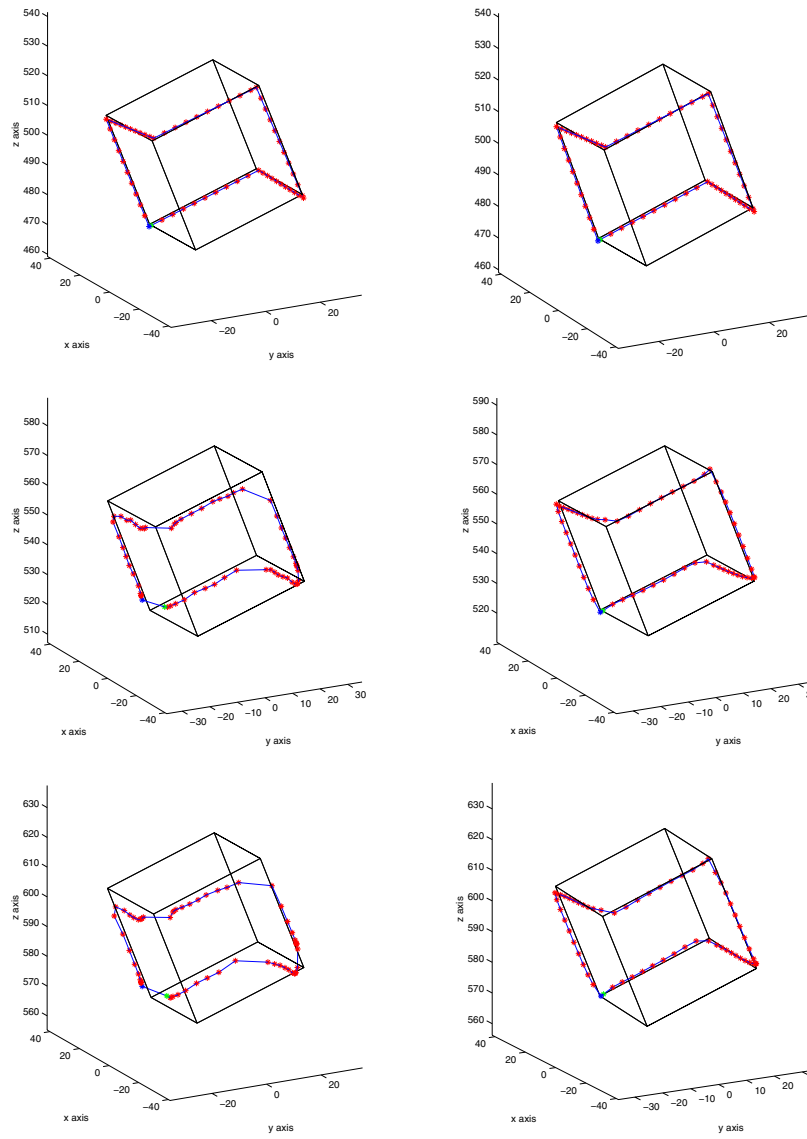

Fig. 4. Tracking results of the synthetic cube in Figure 1 with the regular snake (17) and the velocity snake (16), at the initial frames (top), and the 50th and 97th frames.

that the amount of smoothing was based on prior knowledge on the bounds on image motion. The spatial derivatives of the images were calculated by applying a simple 3×3 Prewitt operator, and the time derivative by simple subtraction of values of successive images, divided by the frame period.

The contour models (16) and (17) were chosen with the additive potential function (10) with $P = -\|\nabla I\|$ and with $\alpha = \beta = 0.5$. The results with the multiplicative potential model (11) were similar. The parameters of the contour models were chosen empirically based on multiple runs and on the resulting performance. This "pseudo learning" could be done automatically via parameter learning (e.g. the auto-regressive modeling and learning proposed in [4]).

Numerical Considerations: For the regular three dimensional snake (17), the contour position at each sampled image was computed based on 100 iterations. This number was chosen in order to allow the snake to stabilize on a solution without too much overhead in time. This heavy computation is in contrast to the reduced complexity provided by the velocity snake (16) which employs only one iteration per image sample.

In order to fully understand the tracking performance of the proposed 3D active contour, the tracking was performed on a synthetic stereo image sequence of a synthetic cube. We simulated a three dimensional cube $40 \times 40 \times 40$ moving in the positive Z direction (towards the camera) with a velocity of 25 units/sec. The cube was initially positioned at a distance of 500 units in the Z direction. The two dimensional stereo-image sequence was formed by using perspective projection with a camera focus of f=500 and with a grayscale of 64 colors. Each cube face was painted with a different color (grayscale level of 15, 40 and 63 respectively) with the background painted in black (grayscale level of 1) (see Figure 4). The image sequence was made up of 97 images, with a sampling interval along time taken as the inverse of frame rate T=1/25 sec.

Contour Parameters: We used contour models with a three dimensional spatial sampling distance of 4 units. The model (17) was used with $\mu = 1$, $w1 = 5$, $w2 = 0.1$ and $\gamma = 100$ and the three dimensional velocity snake with $\mu = 1$, $w1 = 5$, $w2 = 0.01$, $\gamma = 0.01$ and $\beta = 1000$. In addition, the potential was multiplied by a factor of 100 in order to allow the image to form the proper attraction force.

Tracking Results: The tracking results of the moving cube are shown in Figure 4. We show samples of the image sequence at the 50th and the 97th frames. The corresponding tracking results of the regular snake (17) and of the velocity snake (16) are shown below. It can be seen that both models successfully track the cube throughout the sequence. The tracking of the regular three dimensional snake shows two major flaws: Bunching of the sampling points towards the center of the cube's sides. This happens because there is no correspondence information between points in successive images. The second is that the contour lags after the synthetic cube. This is caused by the fact that the snake samples have no prior information on the velocity of the object being tracked instead they have to settle onto each new image. Both of the above flaws are solved by adding the velocity term into the three dimensional snake. This term gives an estimate of

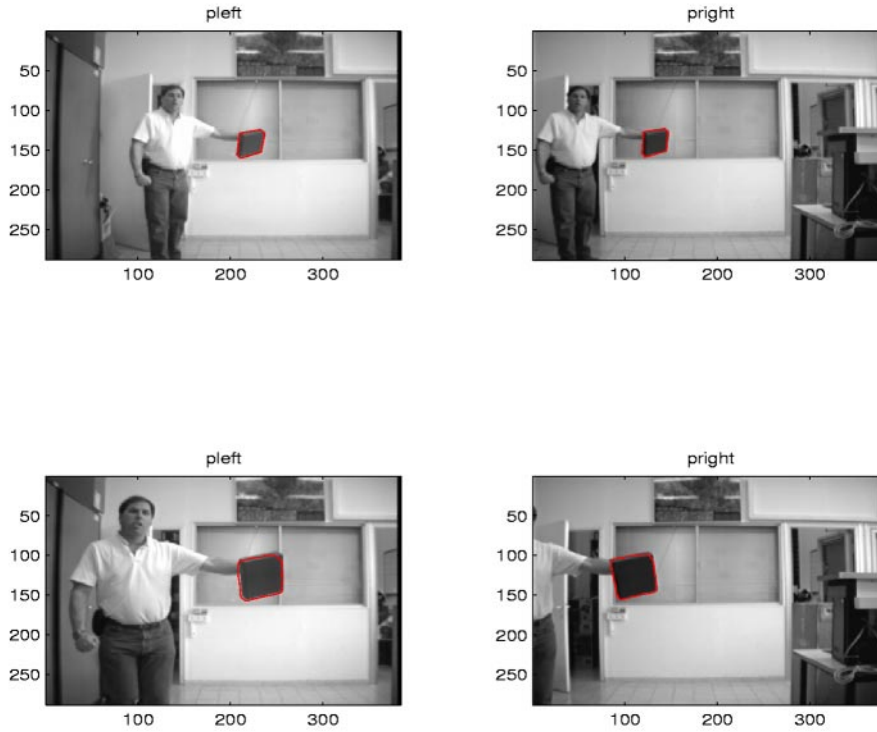

Fig. 5. Samples of a book tracking video sequence (1st frame and 101th frame) with the position of the three dimensional velocity snake projected onto them.

the objects velocity vector between subsequent frames which stops the bunching of the snake samples and the lagging behind the tracked contour. Note that using only one iteration per frame for the model (17) resulted with loss of tracking after few frames.

Next we present the results of tracking a book with the regular model (17) and with the velocity snake (16). The book had the dimensions of $24 \times 21 \times 6$ cm. and was moving towards the camera (-Z direction) from a distance of 3.1 meters to 2.3 meters (2.7 meters only for the regular snake). The image sequence was comprised of only 51 images for the regular three dimensional snake (we stopped after 51 frames because of poor tracking) and 101 images for the three dimensional velocity snake. The stereo images were captured using a stereo setup - two identical synchronized cameras positioned 58 cm. apart in parallel. The

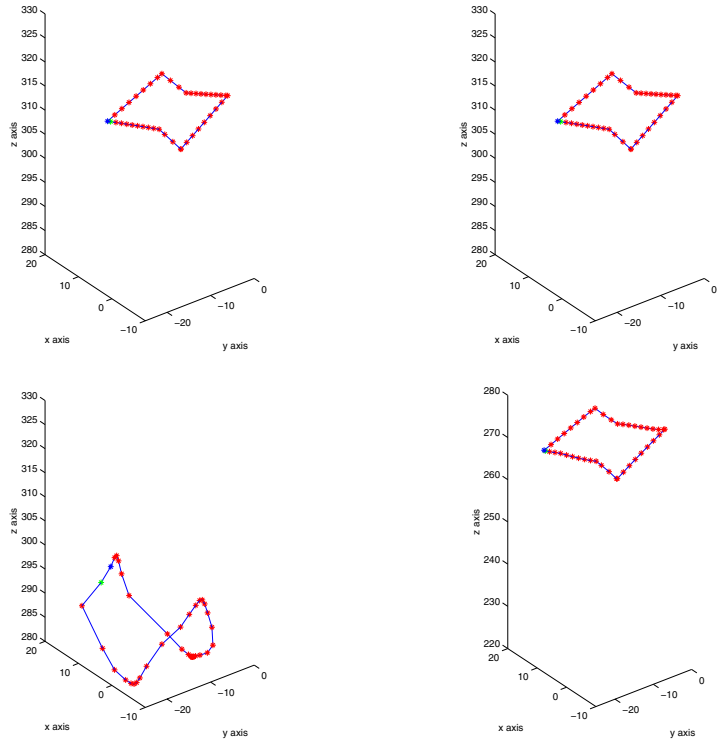

Fig. 6. The 3D snake contours corresponding to images in Figure 5.

camera focus was f=476. The images had a grayscale of 256 colors and were digitized on two Silicon Graphic workstations at a rate of 25 frames per second.

Contour Parameters: We used the contour models with a three dimensional sampling distance of 2.5 cm. The model (17) was used with $\mu = 1$, $w1 = 5$, $w2 = 0.01$ and $\gamma = 10$. For the three dimensional velocity snake, we used $\mu = 1$, $w1 = 1$, $w2 = 0.01$, $\gamma = 0.001$ and $\beta = 2000$. For the regular three dimensional snake, the potential was multiplied by a factor of 10, while a factor of 5 was used for the three dimensional velocity snake.

Tracking Results: The results of tracking of the book in 3D are shown in Figures 5 and 6. We show the results only up to the 51'th frame as the results of the model(17) were very poor following that frame. The velocity snake gave precise results along the 100 frames of the sequence.

In Figure 7 we show the results of tracking a moving hand with the velocity snake (16). The model (17) couldn't handle this sequence. In this sequence we wished to demonstrate the robustness of the three dimensional velocity snake in a long image sequence (about 14 seconds) which included changes in velocity of

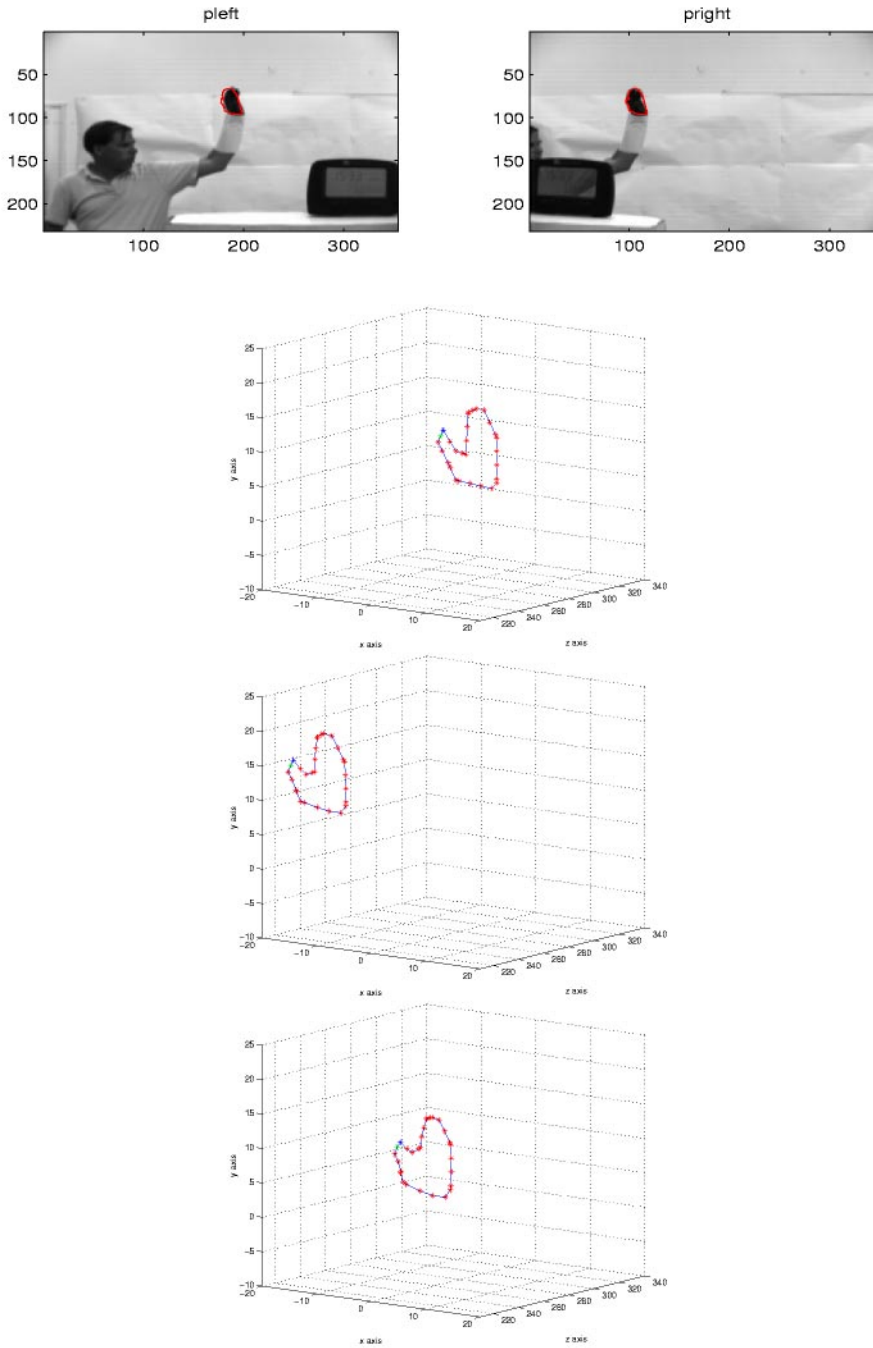

Fig. 7. Tracking results of a moving hand using the 3D velocity snake at the 1st, 100th and 225th frames.

the object being tracked (forward and backward motion) as well as a change in the object's shape (the hand turns sideways at the end of the sequence). The sequence was comprised of 345 images of a hand moving towards and away from the cameras from a distance of 3.4 meters to 2.4 meters. This sequence shows the robustness of the three dimensional velocity snake. The hand moved backwards and forwards as well as to the sides, with the three dimensional velocity snake keeping track throughout the sequence. There was a slight loss of track when the motion changed directions, but the three dimensional velocity snake quickly caught up to the object and continued tracking (the loss of track was caused by the fact that there is no acceleration prediction, only velocity prediction). This sequence also showed the three dimensional velocity snake's capability of tracking changes in shape of the three dimensional object.

5 Concluding Remarks

We have considered in this paper the problem of 3D tracking using stereo sequence data. The tracking scheme incorporated 3D active contours and optical flow velocity information was introduced to improve the tracking performance. A particular choice of the relevant parameters, which greatly reduces the computational requirements was presented. The experimental results which use this scheme indicate successful tracking of simulated and real scenes and clearly demonstrate the performance improvement associated with the velocity term.

The results of this paper clearly demonstrate the feasibility of of the proposed approach for real time 3D contour tracking. Additional work is required in order to optimize the computational load and improve numerical stability. The incorporation of similar ideas in the context of geometric snake models and related level set approaches is an interesting area for future research.

References

1. B. Bascle and R Deriche, "Energy-based methods for 2D curve tracking, reconstruction and refinement of 3D curves and applications", *SPIE*, Vol. 2031, pp. 282-293, 1993.
2. B. Bascle and R. Deriche, "Stereo matching, reconstruction and refinement of 3D curves using deformable contours", *IEEE, Int. Conf. on Computer Vision*, pp. 421-430, 1993.
3. A. Blake and A. Yuille (eds.), *Active Vision*, MIT Press, 1992.
4. A. Blake and M. Isard, *Active Contours*, Springer, 1998.
5. B. Caselles and B. Coll, "Snakes in movement," *SIAM Journal on Numerical Analysis,* Vol. 33, pp. 2445–2456, Dec. 1996.
6. T. Cham and R. Cipolla, "Stereo coupled active contours", *Proceedings of IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, pp. 1094-1099, 1997.
7. M. P. Do Carmo, *Differential Geometry of Curves and Surfaces*, Prentice Hall, 1976.

8. D.B. Gennery, "Visual tracking of known three-dimensional objects", *Int. J. of Computer Vision*, Vol. 7, No. 3, pp. 243-270, 1992
9. A.K. Jain, Y. Zhong and M. Dubuisson-Jolly, "Deformable template models: a review" Signal Processing **71**, 1998, pp. 109–129.
10. M. Kass, A. Witkin and D. Terzopoulos, "Snakes: active contour models", *Int. J. of Computer Vision*, Vol. 1, No. 4, pp. 321-331, 1987.
11. B.K.P. Horn, *Robot Vision*, MIT Press, 1986.
12. D.G.Lowe, "Robust model-based motion-tracking through the integration of search and estimation", *Int. J. of Computer Vision*, Vol. 8, No. 2, pp. 113-122, 1992
13. T. McInerney, and D. Terzopoulos, "Deformable models in medical image analysis: a survey," *Medical Image Analysis*, Vol. 1, No. 2, 91–108, 1996.
14. N. Paragios and R. Derichee, "A PDE based level set approach for detection and tracking of moving objects," *Proc. 6th Intern. Conf. on Computer Vision,* Bombay, India, Jan. 1998.
15. N. Peterfreund, "Robust tracking of position and velocity with Kalman snakes," *IEEE Trans. on PAMI*, 21(6), June 1999.
16. N. Peterfreund, "The velocity snake", *Proc. IEEE Non Rigid and Articulated Motion*, Puerto Ricoh, 1997.
17. N. Peterfreund, "The velocity snake: deformable contour for tracking in spatio-velocity space", *Computer Vision and Image Understanding,* **73**(3), 346–356, 1999.
18. D. Terzopoulos and R. Szeliski, "Tracking with Kalman Snakes", in A. Blake and A. Yuille (eds.), *Active Vision*, pp. 3–20, MIT Press, 1992.
19. Z. Zhang and O. D. Faugeras, "Three-Dimensional Motion Computation and Object Segmentation on a Long Sequence of Stereo Frames", *Intl. J. of Computer Vision*, Vol. 7, No. 3, pp. 211-241, 1992.

Acknowledgment

The research of the second author was funded by the Engineering Research Program of the Office of Basic Energy Sciences under contract DE-AC05-96OR22464 with Lockheed Martin Energy Research Corporation.