Robust Tracking of Position and Velocity With Kalman Snakes

Natan Peterfreund

Abstract—A new Kalman-filter based active contour model is proposed for tracking of nonrigid objects in combined spatio-velocity space. The model employs measurements of gradient-based image potential and of optical-flow along the contour as system measurements. In order to improve robustness to image clutter and to occlusions an optical-flow based detection mechanism is proposed. The method detects and rejects spurious measurements which are not consistent with previous estimation of image motion.

Index Terms—Active contour, Kalman snakes, optical-flow, robust tracking.

1 Introduction

In a variety of applications of image technology, such as medical image analysis [1], human motion modeling, and visual servoing, it is desirable to track the boundaries of nonrigid objects and to analyze their motion. Considerable work has been done during the past few years in boundary tracking and motion analysis of nonrigid objects in the context of snake models [6], [9], [15], [16], [21]. These techniques are based on spatial measurements of gray levels or on the position of image features [5], [21]. Recently, a new snake model was proposed, namely the *velocity snake*, which is an active contour for real time tracking in combined spatio-velocity space [20]. The proposed work presents a stochastic version of the velocity snake and a robust Kalman filtering approach for it.

Boundary tracking with deformable planar contours, known as snakes, was originally introduced by Terzopoulos et al. (e.g., [21], [15]). Snakes are deformable contours that move under the influence of image-intensity "forces," subject to certain internal deformation constraints. In segmentation and boundary tracking problems, these forces relate to the gradient of image intensity. Considerable work has been done to overcome the numerical problems associated with the solution of the equations of motion and to improve robustness to image clutter and occlusions. Curwan and Blake [9] proposed a B-spline representation of active contours, Dubuisson et al. [10] employed polygonal representation in vehicle tracking problems, and Metaxas and Terzopoulos [19] proposed a deformable superquadric model for modeling of shape and motion of 3D nonrigid objects. Other active contour methods for matching and for reconstruction of nonrigid objects, were proposed by Bascle and Deriche [6], in the case of 3D modeling, and by Cootes et al. [8]. All of the above methods are based on measurements of image intensity or on the positions of image features.

Real-time visual tracking of rigid and nonrigid objects has been the subject of intensive research during the last few years. Tracking of 3D objects via model-based techniques can be found in the works of Gennery [11] and Lowe [17]. Both methods use temporal measurements of features and a Kalman filtering estimation approach [12]. Contour tracking of objects, using Kalman filter estimation [12], can be found in the original Kalman snake model of Terzopoulos and Szeliski [21] and in the B-spline active contour of Blake and Isard [5]. The latter, which uses measurements of a basic tracker to optimize the performance of the tracking scheme, was recently extended to obtain the optimal Bayes estimator [14]. The result is obtained through estimation of the probability distribution function of position. This method, however, is computationally expensive. The most related work to the velocity snake model is the tracking method proposed in [6], [7], wherein the algorithm searches for the best model-based transformation of shape between successive images. The resulting contour is then relaxed to nearby boundary through edge-based energy minimization.

In this work, we present a robust Kalman filtering approach for the velocity snake model. The proposed method employs optical-flow measurements and a robust Kalman filtering method (e.g., [4]) to detect and reject measurements which belong to other objects, thus resulting in a tracking scheme which is robust to partial occlusions and to image clutter. In the following, we present the basics of the Velocity Snake originally proposed in [20].

Consider the closed contour $v(s, t) = (x(s, t), y(s, t))$ for some spatial parametric domain $s \in [0, 1]$ and time t. Let $v_s \stackrel{\Delta}{=} \frac{\partial v}{\partial s}$ and $v_{ss} \stackrel{\Delta}{=} \frac{\partial^2 v}{\partial s^2}$. The Lagrangian energy of the snake is given by [21]

$$\frac{1}{2}\int_0^1 \left(w_1\|v_s\|^2 + w_2\|v_{ss}\|^2\right)ds + \frac{1}{2}\int_0^1 \mu\|v_t\|^2 ds + \int_0^1 P(v,t)ds, \quad (1.1)$$

where $P(v, t)$ is the potential field energy of the contour. Given the image sample at time t, $I(\mathbf{x}, t)$, where $\mathbf{x} = (x, y)$ denote the spatial coordinates, typical potential field energy for boundary segmentation, satisfies $P = -\|\nabla I(\mathbf{x}, t)\|$, where $\nabla I = (I_x, I_y)$ denotes the spatial gradient of $I(\mathbf{x}, \cdot)$. Let v_t^i denote the apparent velocity of the image at the contour position. The energy dissipation function which is used to dampen the Langrangian energy (1.1) is given by [20]

$$D\left(v_t, v_t^i\right) = \frac{\gamma}{2}\int_0^1 \left\|L^T\left(v_t - v_t^i\right)\right\|^2 ds + \frac{\beta}{2}\int_0^1 \left\|\frac{\partial}{\partial s}v_t\right\|^2 ds \quad \gamma > 0, \beta > 0, \quad (1.2)$$

where L is a real matrix. The second term represents a smoothness constraint. Using (1.1) and (1.2), the Euler-Lagrange equations of motion of the *velocity snake* are given by [20]

$$\mu v_{tt} + \gamma C\left(v_t, v_t^i\right) - \beta\frac{\partial}{\partial s}\left(\frac{\partial}{\partial s}v_t\right) - \frac{\partial}{\partial s}\left(w_1 v_s\right) + \frac{\partial^2}{\partial s^2}\left(w_2 v_{ss}\right) = -\nabla P\left(v(s,t),t\right), \quad (1.3)$$

where the velocity control term $C(v_t, v_t^i)$ satisfies

$$C\left(v_t, v_t^i\right) = LL^T\left(v_t - v_t^i\right). \quad (1.4)$$

Using basic results from feedback control theory [22], we showed in [20] that the contour will converge to a boundary moving at a constant velocity if the initial contour is sufficiently close to the boundary of the moving object and if we have the velocity of the object at the contour position (which could be computed via optical flow techniques). This result is due to the combined velocity and state control terms (1.4) and $-\nabla P$. The stability and the convergence rate of the system improve as the absolute value of eigenvalues of L increase [20]. However, large gain control will be required for an increased sampling-rate along time when solving the system equations. Also, it may increase the influence of measurement noise. In [20], we showed that the lack of motion control in the original snake model [21] causes a bias in the contour position. This could lead to serious tracking problems even in the case of noncluttered environments [20].

Various techniques for estimation of the apparent velocity v_t^i at the contour position had been proposed in [20]. These methods are

• *The author is with the Center for Engineering Systems Advanced Research (CESAR), Oak Ridge National Laboratory, P.O. Box 2008, Oak Ridge, TN 37831-6355. E-mail: v4p@mars.epm.ornl.gov.*

Manuscript received 4 June 1997; revised 1 Dec. 1998. Recommended for acceptance by R. Szeliski.
For information on obtaining reprints of this article, please send e-mail to: tpami@computer.org, and reference IEEECS Log Number 107719.

based on the optical flow constraint equation [13] and on integration of measurements along time [2].

In [20], we showed that, in the special case where $L = \nabla I(v(s, t))$, the velocity control (1.4) becomes

$$C\left(v_t, v_t^i\right) = \nabla I\left(\nabla I^T v_t + I_t\right). \tag{1.5}$$

This result is due to the optical-flow constraint equation [13]. In this model, there is no need to estimate the image velocity since the optical flow term provides a measure of the error in velocity estimation. Compared to the model in (1.3) and (1.4), however, this scheme is more sensitive to measurement noise and to numerical approximation as, instead of the velocity error, it has a projected version on the direction of ∇I. In the following, we shall refer to the system with the control (1.4) as the Batch-Mode model and to the one with (1.5) as the Real-time one.

The discretization of the velocity snake model in space is based on equidistant sampling of $v(s, \cdot)$ along s, with $u = [u_1, \ldots, u_M]$ and $u_i = (x_i, y_i)$ denoting the snake points, and on finite difference approximation of partial derivatives in space [20], [21].

2 Kalman Snake Model in Spatio-Velocity Space

In this section, we propose a Kalman filtering approach to the deterministic, least-squares based, velocity snake model (1.3) [20]. We propose two types of estimation models: Batch Mode model, in which the velocity is estimated independently of the contour dynamics and treated as an input to the tracking model, and Real Time Mode, which uses measurements of image velocity and spatial intensities to adjust the expected position and the covariance of the contour's parameters.

2.1 Batch-Mode Kalman Filter Model

Consider the discrete model of the velocity snake (1.3) with the control (1.4). Let $\xi = [x_1, x_2 \ldots x_M]^T$ and $\eta = [y_1, y_2 \ldots y_M]^T$ denote the vector of sampling points corresponding to u, with $u_i = [x_i, y_i]$, $V = [\xi^T, \eta^T]^T$, and $\dot{V} = [\dot{\xi}^T, \dot{\eta}^T]^T$, where $\dot{V}$ denote the derivative of V with respect to t. We denote the Gaussian probability density with mean vector h and covariance matrix Q by $N(h, Q)$. Using finite difference approximation of partial derivatives along space [16], the first order form of the velocity snake model (1.3), with the velocity control (1.4) and $L = I_2$, is given by

$$\frac{d}{dt}\begin{bmatrix} V \\ \dot{V} \end{bmatrix} = \begin{bmatrix} O_{2M\times 2M} & I_{2M} \\ -\frac{1}{\mu}\Sigma & -\frac{1}{\mu}\left(\beta\Psi\Psi^T + \gamma I\right) \end{bmatrix}\begin{bmatrix} V \\ \dot{V} \end{bmatrix} - \frac{1}{\mu}\begin{bmatrix} O_{2M\times 2M} \\ I_{2M} \end{bmatrix} u^p + \frac{\gamma}{\mu}\begin{bmatrix} O_{2M\times 2M} \\ I_{2M} \end{bmatrix} U_t^i, \tag{2.1}$$

where $O_{N\times M}$ is an $N \times M$ zero matrix, I_{2M} is a $2M \times 2M$ Identity matrix,

$$\Sigma = \begin{bmatrix} K & O_{M\times M} \\ O_{M\times M} & K \end{bmatrix} \text{ and } \Psi = \begin{bmatrix} D & O_{M\times M} \\ O_{M\times M} & D \end{bmatrix}.$$

Here, the matrix K consists of the deformation constraints imposed by w_1 and w_2 in (1.3) [21], [16], and D is the discrete-derivative matrix [20]. The input vector of apparent velocity is denoted by $U_i^t[\mathbf{x}_t^i, \mathbf{y}_t^i]$ where $\mathbf{x}_t^i$ and $\mathbf{y}_t^i$ are the components of the apparent velocity on u in the x and y directions, respectively. We assume that the apparent velocity of the image at the contour obeys the differential equation

$$\frac{d}{dt}U_t^i = -\alpha U_t^i + q \quad \alpha > 0, \quad q \sim N(0, Q). \tag{2.2}$$

This model is characterized by an exponential decaying of measurements along time and, thus, allows the system to adapt to changes in the apparent velocity. Note that the changes in the projected motion may be due to motion in 3D or to nonrigidity property of the object. As an alternative to the proposed nonrigid motion model, Bascle and Deriche [7] proposed estimating the best match of shape between successive frames under a given rigid transformation. This method would break, however, in the presence of nonrigidity, image clutter and occlusions. Let $\Pi = [\Pi_x, \Pi_y]$, where Π_x and Π_y are diagonal matrices with the diagonal elements given by $I_x(u)$ and $I_y(u)$, respectively, and $\partial P = [P_x^T(u), P_y^T(u)]^T$. The measurement vector is given by

$$\begin{aligned} \partial P &= u^p \\ z &= \begin{bmatrix} I_t \\ 0_{2M} \end{bmatrix} = \begin{bmatrix} -\Pi \\ \Psi \end{bmatrix} U_t^i + \begin{bmatrix} w_1 \\ w_2 \end{bmatrix} \quad w_i \sim N(0, W_i). \end{aligned} \tag{2.3}$$

Here, 0_{2M} is a $2M$ dimensional vector of zeros. The first component in z corresponds to the optical flow constraint equation [13]. The second term measures the smoothness of the apparent velocity along the contour line. Similar modeling of the measurement vector z, in the context of optical flow estimation, can be found in [18]. The estimation $\hat{U}_t^i$ of the apparent velocity U_t^i is given by the solution of the continuous Kalman filter (e.g., [12]). This result, along with $u^p = \partial P$, is used as input control to the velocity model (2.1). As will be shown, this model further allows for detection and rejection of measurements which belong to image clutter.

Next, we present a 3D model-based estimation approach for image velocity. In the presence of detection mechanism of spurious measurements, this method allows for estimation of boundary position under occlusion. The latter subject is further elaborated in the next section. In the following, we consider the problem of tracking of shallow objects.

Under the assumption of a planar surface object, the projected velocity satisfies (e.g., [3])

$$\begin{aligned} w_x &= a_1 + a_2 x + a_3 y + a_7 x^2 + a_8 xy \\ w_y &= a_4 + a_5 x + a_6 y + a_7 xy + a_8 y^2, \end{aligned} \tag{2.4}$$

where the scalars $a_1, \ldots, a_8$ are functions of the translation and angular velocities of the object in 3D space. Let $\psi_{u_t^i} = [a_1 \ldots a_8]^T$. Similar to image velocity model (2.2), we define the velocity parameter model as

$$\frac{d}{dt}\psi_{u_t^i} = -\alpha\psi_{u_t^i} + r \quad \alpha > 0, \quad r \sim N(0, R). \tag{2.5}$$

Here, the covariance matrix R is nondiagonal due to the dependence between the components of $\psi_{u_t^i}$ (e.g., [3]). The measurement vector corresponding to (2.3) is given by

$$\begin{aligned} \partial P &= u^p \\ z = I_t &= H\psi_{u_t^i} + w \quad w \sim N(0, W), \end{aligned} \tag{2.6}$$

where

$$H = -\Pi\begin{bmatrix} 1_M & \xi & \eta & 0_M & 0_M & 0_M & \xi^2 & \xi\cdot\eta \\ 0_M & 0_M & 0_M & 1_M & \xi & \eta & \xi\cdot\eta & \eta^2 \end{bmatrix}.$$

Here, 1_M is an M dimensional vector of ones, $\xi_i\xi_i \triangleq [\xi^2]_i$ and $\xi_i\eta_i \triangleq [\xi\eta]_i$. In this model, we do not have a measure of velocity smoothness since the velocity model is already smooth. The state estimation of (2.5) (and (2.1)) is performed by the Kalman Filtering method, similar to state estimation in the original model (2.1) and (2.2).

The Kalman snake model (2.1) is also used for prediction of position [12]. In this case, the input control is given by $u^p \equiv 0$ and

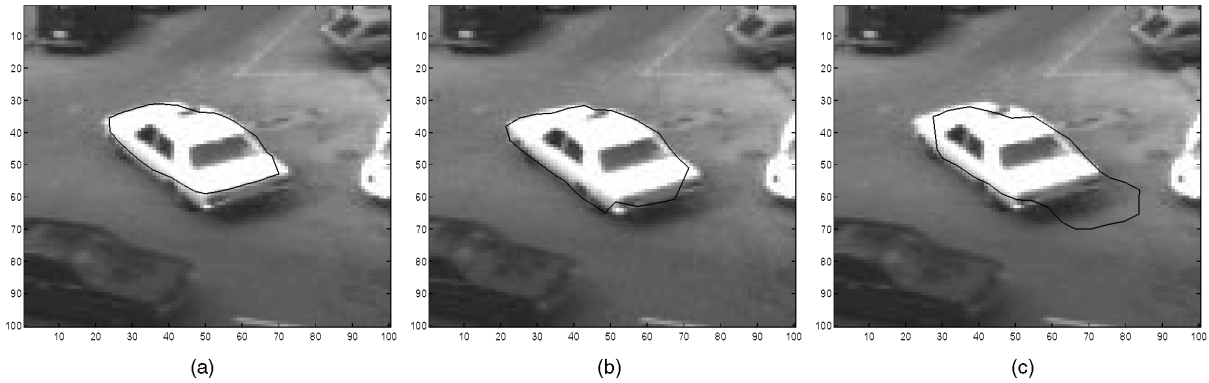

Fig. 1. Tracking results of the moving car with the velocity snake (after 19 frames) [20]: (a) Batch-Mode model, (b) Real-Time model, (c) original Kalman snake model [21].

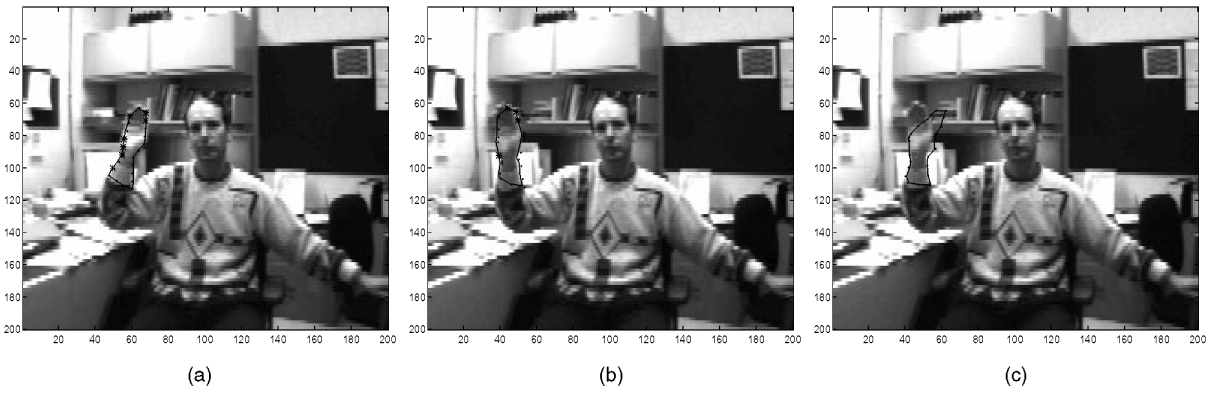

Fig. 2. Tracking results of the waving hand with Real-Time Kalman snake: (a) frame 3, (b) frame 40, (c) original Kalman snake model [21].

$\hat{\psi}_{u_t^i} = \hat{\psi}_{u_\tau^i}$, where τ defines the latest estimation using system measurements (2.6) (in the prediction case, we assume $\alpha = 0$).

2.2 Real Time Kalman Snake Model

Consider the Kalman filtering approach to the optical-flow constraint model (1.3) with velocity control (1.5). In the proposed model, the optical-flow constraint equation [13] and image-gradient ∇P are treated as system measurements. Similar to (2.1), the system model is given by

$$\frac{d}{dt}\begin{bmatrix} V \\ \dot{V} \end{bmatrix} = \begin{bmatrix} O_{2M\times 2M} & I_{2M} \\ -\frac{1}{\mu}\Sigma & -\frac{\beta}{\mu}\Psi\Psi^T \end{bmatrix}\begin{bmatrix} V \\ \dot{V} \end{bmatrix} - \frac{1}{\mu}\begin{bmatrix} O_{2M\times 2M} \\ I_{2M} \end{bmatrix} u^p + \begin{bmatrix} 0_{2M} \\ q \end{bmatrix} \quad q \sim N(0, Q). \tag{2.7}$$

The model assumes random acceleration with covariance Q, which accounts for changes in image velocity. This term, which increases the covariance matrix of estimation by Qdt each time interval dt [12], has direct influence on the amount of smoothing of measurements along time [11]. As the eigenvalues of Q become larger, old measurements are given relatively low weight in the adjustment of state. This allows the system to adapt to changes in the target velocity. The entries of Q could be obtained from error analysis of motion measurements relative to the proposed model of motion (e.g., [17]), or from manual tuning to obtain proper convergence, in the presence of a given class of objects and motions [19]. The measurement vector satisfies

$$z = \begin{bmatrix} \partial P(V,t) \\ I_t \end{bmatrix} = \begin{bmatrix} u^p \\ -\Pi\ \dot{V} \end{bmatrix} + \begin{bmatrix} 0_{2M} \\ w \end{bmatrix} \quad w \sim N(0, W). \tag{2.8}$$

Note that the velocity measurements result from first order approximation of the intensity preserving equation $I(V,\ t + dt) = I(V - \dot{V}dt, t)$. The above estimation problem is solved by the Extended Kalman filter [12]. The state estimation of (2.7) is then given by

$$\frac{d}{dt}\begin{bmatrix} \hat{V} \\ \hat{\dot{V}} \end{bmatrix} = \begin{bmatrix} O_{2M\times 2M} & I_{2M} \\ -\frac{1}{\mu}\Sigma & -\frac{\beta}{\mu}\Psi\Psi^T \end{bmatrix}\begin{bmatrix} \hat{V} \\ \hat{\dot{V}} \end{bmatrix} - \frac{1}{\mu}\begin{bmatrix} O_{2M\times 2M} \\ I_{2M} \end{bmatrix}\partial P(\hat{V}, t) + K(\hat{V}, t)\left(I_t(\hat{V}, t) + \Pi\ \hat{\dot{V}}\right), \tag{2.9}$$

where, for

$$H = \left[O_{M\times 2M}, -\Pi\right] \text{ and } L = \begin{bmatrix} L_1 & L_{12} \\ L_{12}^T & L_2 \end{bmatrix}, \tag{2.10}$$

the covariance matrix of $[\hat{V}^T \hat{\dot{V}}^T]^T$, we have the Kalman gain matrix

$$K(\hat{V}, t) = LH^T W^{-1} = -\begin{bmatrix} L_{12} \\ L_2 \end{bmatrix}\Pi^2 W^{-1}. \tag{2.11}$$

The covariance matrix L is given by the solution of the continuous matrix Riccati equation (e.g., [12]). Note that this equation is also function of the nonlinearity in $\partial P(\hat{V}, t)$ [12]. As in the Batch mode model, measurements of the optical-flow constraint whose error exceed certain threshold and the corresponding components in ∂P, are considered to be spurious and, thus, are being ignored. In the absence of measurements, the position of snake points could be predicted based on the solution of the system model with $u^p = 0$ and $K(\hat{V}, t) = 0$. We note, however, that, since the apparent velocity of a 3D object varies with position, good predictions could be obtained only in the case of a 3D translation motion which is parallel to image plane.

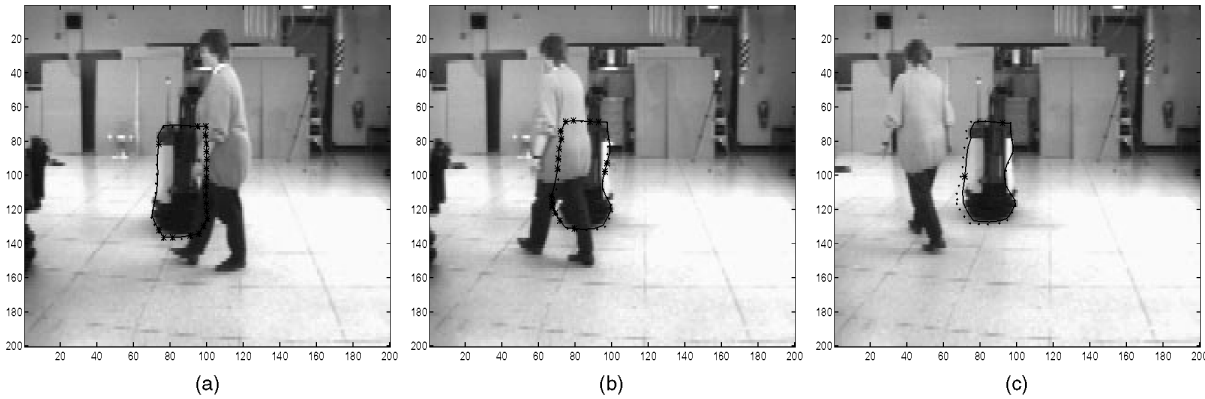

Fig. 3. Tracking results with Batch-mode Kalman snake and with the affine velocity model (2.4).

3 Robust Tracking in a Cluttered Environment

The random estimation approach proposed in previous section provides a probabilistic framework for detection of spurious measurements. A measurement is treated as noise or spurious if the error in the measurement equation exceeds certain threshold. Usually, this threshold corresponds to low probability of measurements values. In this section, we use this method to detect and reject spurious measurements of velocity and potential field. This capability does not exist in existing active contour models (e.g., [21], [5], and [9]).

In the following, we limit the discussion to the case of the Real-time model. The results can be directly extended to the Batch-mode model. Consider the system model (2.7) and (2.8). We assume that the elements of $I_t = [I_{t1} \dots I_{tm}]$ are not correlated, i.e., $W = \sigma I_M$ for some $\sigma > 0$.

Let $\mathbf{V} = [V^T, \dot{V}^T]^T$ denote the system state, $\hat{\mathbf{V}} = [\hat{V}^T, \hat{\dot{V}}^T]^T$ the estimated state, and L the corresponding covariance matrix. The velocity measurements are, by (2.8),

$$z_i = I_{ti} = -\Pi_i \dot{V} + w_i \quad w_i \sim N\left(0, \sigma^2\right) \quad i = 1 \dots M, \tag{3.1}$$

where Π_i is the ith row of Π. Using (2.7), the discrete approximation of the prediction of state is

$$\hat{\mathbf{V}}^- = \left(I_{4M} + \Delta t \begin{bmatrix} O_{2M\times 2M} & I_{2M} \\ -\frac{1}{\mu}\Sigma & -\frac{\beta}{\mu}\Psi\Psi^T \end{bmatrix} \right) \hat{\mathbf{V}} - \frac{\Delta t}{\mu} \begin{bmatrix} O_{2M\times 2M} \\ I_{2M} \end{bmatrix} \partial P\left(\hat{V}\right), \tag{3.2}$$

where Δt denote the sampling interval. The covariance matrix of $\hat{\mathbf{V}}^-$, L^-, is computed in a similar manner [12]. The prediction of the velocity measurements are, by (3.1),

$$\hat{z}_i^- = I_{ti} = -\Pi_i \hat{\dot{V}}^- \quad i = 1 \dots M$$
$$S_i = \Pi_i L_{22}^- \Pi_i^T + \sigma^2, \tag{3.3}$$

where L_{22}^- denote the covariance of $\hat{\dot{\mathbf{V}}}^-$. It can be shown that both $\hat{z}_i^-$ and S_i are only functions of the measurements of velocity and potential field at the ith sample $u_i = (x_i, y_i)$. No other components of ∂P and I_t along the contour affect these functions. Using this property, each measurement z_i can thus be examined separately. We assume that $P(z_i) = N(\hat{z}_i^-, S_i)$ for $i = 1, \dots, M$. The validation region Ω_i of each measurement z_i is defined as (e.g., [4])

$$\Omega_i(\gamma_i) \triangleq \left\{ z_i : \left[z_i - \hat{z}_i^-\right]^T S_i^{-1} \left[z_i - \hat{z}_i^-\right] \le \gamma_i \right\}, \tag{3.4}$$

where the value of γ_i relates to probability of detection P_D^i according to $P_D^i = P(z_i \in \Omega_i(\gamma_i))$. Within the proposed detection approach, all measurements associated with the ith point of the snake, u_i, are rejected if the error in velocity measurement z_i exceeds the validation region Ω_i. Note that, since $\partial P(u_i)$ uses the same gradient information of image intensity, it is rejected and replaced by zero value, which is the optimal estimator under zero mean assumption.

4 Experimental Results

We demonstrated the performance of the proposed contour model by applying it to real image sequences with both rigid (car, mobile robot) and nonrigid (waving hands) objects. The initial contour lines were generated manually, forming a rough polygonal approximation to the object's boundary. This task could be done automatically via ATR -based methods.

Prior to computation of image gradients and velocities, the image sequences were smoothed, both in space and time, by a Gaussian filter. In the proposed examples, we used a fixed Gaussian filter with $\sigma = 2$, both in space and time. The purpose of this prefiltering is to obtain reliable optical flow measurements [3]. The spatial derivatives of the images were calculated by applying a simple 3×3 Sobel operator. We note that the smoothing operation, in general, causes bias in the optical-flow measurements. In our model, however, the detection mechanism is used to detect and reject measurements with large spurious bias which may appear in the presence of large boundary clutter.

The discretization in time of the differential equations was based on the implicit differencing approximation $X_{n+1} = X_n + \dot{X}_{n+1}\Delta t$, where X_n and $\dot{X}_n$ are the state and the time-derivative of state at time step n, respectively. This scheme was found to derive the most stable results. In order to understand the complexity of both algorithms, we counted the number of multiplications [11]. It follows that the Batch-Mode model contains about $BM_{op} = 28M^2 + 16M^3$ operations and the Real-Time about $RT_{op} = 60M^2$. For M = 15 (15 snake points),we have $BM_{op} = 60{,}300$ and $RT_{op} = 13{,}500$.

The parameters of the snake model γ and μ were initialized using the scheme proposed in [20], which defines the limits on these parameters under which the system remains stable. These parameters and the ones of the noise components were then adjusted experimentally and were set to $\gamma = \mu = 2$ and $\beta = 30$. The covariance matrix Q was set to a diagonal form with variance of 0.2 (pixels) to account for changes in velocity between successive frames. The resulting system was found to perform well in various of tracking scenarios. This task of parameter adjustment could be also done using the learning method proposed in [5].

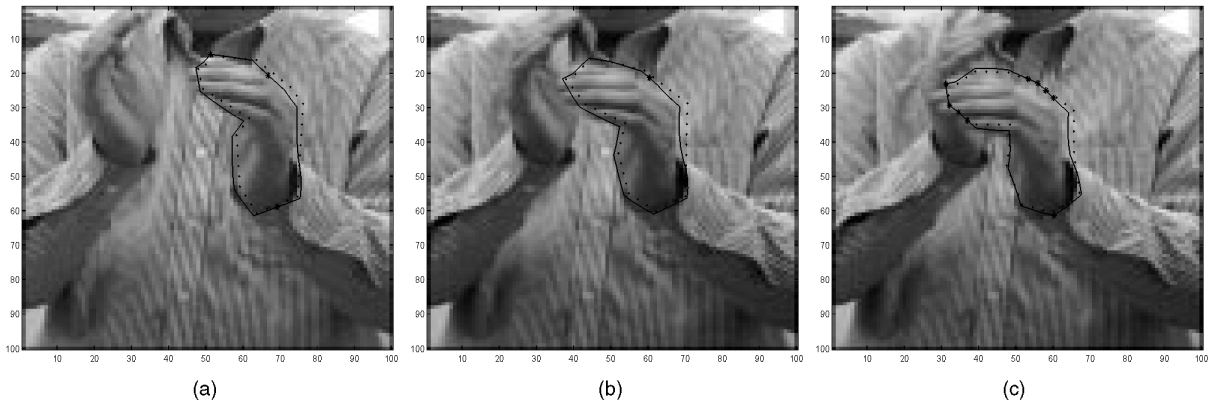

Fig. 4. Tracking results of a moving hand with Real-Time Kalman snake model.

The tracking algorithm of the Real-Time model is composed of the following steps. The Batch-Mode model was defined in a similar manner.

1) $i = 0$; Initiate (manually) contour position at image(i).
2) Initiate contour velocity based on image(i) and image(i + 1) using optical-flow measurements [20].
3) Perform prediction of velocity measurements $\hat{z}_i$ and S_i ((3.2) and (3.3)).
4) Reject spurious measurements outside the validation region Ω_i ((3.4)).
5) Update the snake state using (2.9), excluding the rejected measurements.
6) $i = i + 1$; Sample a new measurement vector (2.8) based on image(i) and image(i + 1); Goto 3.

The tracking schemes were tested on the following image sequences which were sampled at video rate: a moving car (rigid body in noncluttered environment; 20 images), mobile robot (rigid body under occlusions; 90 images), waving hand (cluttered environment; 60 images), and nonrigid hand (nonrigid object; 30 images). The results are presented in Figs. 1, 2, 3, and 4. In Fig. 1a, we present the results of tracking the rigid car with the velocity snake model (1.3) and (1.4) [20]. The results with the real-time control (1.5) and the ones with the original snake model [21] are given in Figs. 1b and 1c, respectively [20]. The latter result illustrates the effect of the bias in the tracking scheme of the original model [20]. In Figs. 2a and 2b, we present the results of the Real-Time Kalman snake with detection threshold $\gamma_i = 1$, after 3 and 40 images, respectively. In these figures "*" denotes snake points which were detected as spurious and "·" denotes the contour points in previous frame. It can be seen that most of the points that were detected to be spurious belong to regions with edges of image clutter. The tracking result with the original Kalman snake [21] is given in Fig. 2c. The poor results are due to the lack of detection mechanism of spurious measurements and due to the bias in tracking which increase the influence of nearby clutter edges. We note, however, that the proposed detection mechanism fails in the presence of small velocity measurements I_{ti} which are in the order of measurement noise. In this case, the spurious edges affect the tracking scheme as they do to the original snake model. In Figs. 3a, 3b, and 3c, we show the results of tracking the mobile robot under occlusion. In this case, we used the Batch-Mode model with the structured velocity (2.4). It can be seen from Figs. 3a and 3b that nearly all occluded points were detected as spurious. Furthermore, it is shown in Fig. 3b that the right side of the contour detected the reappearing edge and converged back to it. In a couple of places, due to measurement noise, the tracking scheme did not detect the occlusion in some places (Fig. 3b). The original Kalman snake could not cope with occlusions. In this case, the tracking contour switched to the occluding object. Finally, in Figs. 4a, 4b, and 4c, we show the tracking results of the Real-Time model in the presence of nonrigid motion of a hand. It is shown that the contour could track an object, which is characterized by varying motion along the object boundary. It is shown in Fig. 4c that edges within the moving object, if they exist in the vicinity of the object's boundary, could attract the contour during tracking. As was noted, however, edges which do not belong to the object (and which do not have the same motion properties) do not have such influences.

5 Concluding Remarks and Future Work

Using basic results in robust Kalman filtering, we proposed in this work a new Kalman filtering approach for the velocity snake, which uses image gradient and optical flow measurements along the contour, as system measurements. According to the proposed detection approach, velocity measurements which are not consistent with previous estimation of motion and the corresponding edge-based potential field measurements are rejected during the update step of the snake state. The proposed method was tested in tracking problems of both rigid and nonrigid objects, in the presence of occlusions and image clutter. It was shown that the velocity detection method allows detection of spurious measurements and occluded boundary points, thus improving the robustness of the tracking scheme.

In the work, we proposed two types of tracking methods: a Batch-Mode model, which uses estimates of image velocity at the contour position as an input control, and a Real-Time mode, which uses optical-flow measurements, along with measurements of image-gradient, as the system measurements. The former was found to derive better tracking results. The latter scheme, however, does not need separate estimation of image velocity and has similar results in case of low clutter background. Under a model-based estimation of motion, the Batch-Mode model demonstrated good tracking results of the object's contour under high percentage of boundary occlusion. The quality of prediction could be further improved by applying more sophisticated models of image motion [19], [5]. This modeling may also prevent the active contour from converging to other contour edges which belong to the moving object, as we observed in the case of the nonrigid hand. The present form of the Real-Time model cannot cope with occlusions, as the projected velocity of the object varies with position.

In order to obtain reliable measurements of optical-flow and potential field, the image sequences were smoothed by applying a low-pass filter, both along space and time [3], [20], resulting in an effective motion which is in the order of one pixel [3]. Here, the term effective motion refers to the ratio between motion amplitude and the spatial correlation length between image pixels. According to this scheme, larger velocities will be needed for lower bandwidth

prefiltering. Note that coarse-to-fine approximation of motion could be included within the proposed scheme, as is traditionally done in optical-flow estimation. Alternatively, one could predict the position $\hat{\mathbf{V}}^-$ of the estimated boundary and then sample the optical-flow measurements which relate to the error in velocity estimation, i.e., $\tilde{I}_t = -\Pi_i(\hat{\hat{V}} - \hat{\hat{V}}^-) + w$, where $\tilde{I}_t = I(\hat{V}^-, t+1) - I(\hat{V}, t)$. In order for the error in velocity estimation to be in the order of one pixel, a Gaussian prefiltering should be applied with covariance matrix proportional to L^-, the covariance of $\hat{\hat{V}}^-$. Due to the smoothing operation, the image-clutter, if exists, may cause for bias in the optical-flow measurements. This bias, which increases with image clutter, cause for error in the estimation of velocity. In the experiments we made, we found that the system could cope with clutter edges that are less or are in the order of the boundary characteristics. Larger biased measurements were determined to be spurious by an appropriate setting of detection threshold. The smoothing effect could be further improved if we instead used a directional-based measurements of velocity and image potential field. This topic is the subject of current research.

ACKNOWLEDGMENTS

This research was performed at the Center for Engineering Systems Advanced Research, Oak Ridge National Laboratory. Funding provided by the Engineering Research Program of the U.S. Office of Basic Energy Sciences under contracts DE-AC05-96OR22464 and 90X-SZ183V with Lockheed Martin Energy Research Corp. The author thanks Dr. Jacob Barhen, head of CESAR at ORNL.

REFERENCES

[1] A.A. Amini, R.W. Curwen, and J.C. Gore, "Snakes and Splines for Tracking Non-Rigid Heart Motion," *Computer-Vision: ECCV '96*, B. Buxton and R. Cipolla, eds., pp. 251-261, 1996.

[2] K.J. Astrom and B. Wittenmark, *Adaptive Control*. Addison-Wesley, 1995.

[3] J.L. Barron, D.J. Fleet, and S.S. Beauchemin, "Performance of Optical Flow Techniques," *Int'l J. Computer Vision*, vol. 12, no. 1, pp. 43-77, 1994.

[4] Y. Bar-Shalom and T.E. Fortmann, *Tracking and Data Association*. Orlando, Fla.: Academic Press, 1988.

[5] A. Blake and M. Isard, "3D Position, Altitude and Shape Input Using Video Tracking of Hands and Lips," *Proc. Computer Graphics, SIGGRAPH*, pp. 71-78, 1994.

[6] B. Basle and R. Deriche, "Stereo Matching, Reconstruction and Refinement of 3D Curves Using Deformable Contours," *Proc. Fourth Int'l Conf. Computer Vision*, pp. 421-430, Berlin, 1993.

[7] B. Bascle and R. Deriche, "Energy-Based Methods for 2D Curve Tracking, Reconstruction and Refinement of 3D Curves and Applications," *Proc. SPIE: Geometric Methods in Computer Vision II*, pp. 282-293, San Diego, Calif., 1993.

[8] T.F. Cootes, C.J. Taylor, A. Lanitis, D.H. Cooper, and J. Graham, "Building and Using Flexible Models Incorporating Grey-Level Information," *Proc. Fourth Int'l Conf. Computer Vision*, pp. 242-246, Berlin, 1993.

[9] R. Curwen and A. Blake, "Dynamic Contours: Real-Time Active Splines," *Active Vision*, A. Blake and A. Yuille, eds., pp. 39-58. MIT Press, 1992.

[10] M.P. Dubuisson, S. Lakshmanan, and A.K. Jain, "Vehicle Segmentation and Classification using Deformable Templates," *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol. 18, no. 3, pp. 293-308, 1996.

[11] D.B. Gennery, "Visual Tracking of Known Three-Dimensional Objects," *Int'l J. Computer Vision*, vol. 7, no. 3, pp. 243-270, 1992.

[12] M.S. Grewal and A.P. Andrew, *Kalman Filtering: Theory and Practice*. Upper Saddle River, N.J.: Prentice Hall, 1993.

[13] B.K.P. Horn and B.G. Schunk, "Determining Optical Flow," *Artificial Intelligence*, vol. 17, pp. 185-204, 1981.

[14] M. Isard and A. Blake, "Visual Tracking by Stochastic Propagation of Conditional Density," *Proc. Fourth European Conf. Computer Vision*, pp. 343-356, Cambridge, England, 1996.

[15] M. Kass, A. Witkin, and D. Terzopoulos, "Snakes: Active Contour Models," *Int'l J. Computer Vision*, vol. 1, no. 4, pp. 321-331, 1987.

[16] F. Leymarie and M.D. Levine, "Tracking Deformable Objects in the Plane Using an Active Contour Models," *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol. 15, no. 6, pp. 617-634, 1993.

[17] D.G. Lowe, "Robust Model-Based Motion-Tracking Through the Integration of Search and Estimation," *Int'l J. Computer Vision*, vol. 8, no. 2, pp. 113-122, 1992.

[18] M.R. Luettgen, W.C. Karl, and A.S. Willsky, "Efficient Multiscale Regularization with Applications to the Computation of Optical Flow," *IEEE Trans. Image Processing*, vol. 3, no. 1, pp. 41-64, 1994.

[19] D. Metaxas and D. Terzopoulos, "Shape and Nonrigid Motion Estimation Through Physics-Based Synthesis," *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol. 15, no. 6, pp. 580-591, 1993.

[20] N. Peterfreund, "The Velocity Snake," *Proc. IEEE Nonrigid and Articulated Motion Workshop*, Virgin Islands, 1997.

[21] D. Terzopoulos and R. Szeliski, "Tracking with Kalman Snakes," *Active Vision*, A. Blake and A. Yuille, eds., pp. 3-20. MIT Press, 1992.

[22] J.L. Willems, *Stability Theory of Dynamical Systems*. Nelson, 1970.