2016 23rd International Conference on Pattern Recognition (ICPR)
Cancún Center, Cancún, México, December 4-8, 2016

An Intensity- and Region-Guided Narrow-Band Level Set Model for Contour Tracking

Somenath Das[‡] Suchendra M. Bhandarkar[‡] Ananda S. Chowdhury[†]
[‡]Department of Computer Science, The University of Georgia, Athens, GA 30602-7404, USA
[†]Department of Electronics & Telecommunications Engineering, Jadavpur University, Kolkata 700032, India
Emails: {somenath@uga.edu, suchi@cs.uga.edu, aschowdhury@etce.jdvu.ac.in}

***Abstract*—Level set-based contour tracking methods have generated recent interest in the computer vision community. In this paper, we propose a novel level set-based algorithm for tracking dynamic implicit contours that utilizes minimal prior information. Our solution consists of two main steps. In the first step, a simple first-order Markov chain model is employed for the coarse localization of a target object. In the second step, we evolve level sets within a narrow band to accurately track the target contour. Narrow band curve evolution is guided through color- and region-based terms in the standard Chan-Vese framework. Comprehensive experimentation on a dataset comprising of several publicly available video sequences clearly demonstrate the advantage of the proposed tracking algorithm.**

I. Introduction

Accurate target tracking has been the focus of researchers in the computer vision community for several years with applications drawn from diverse domains such as content-based video indexing, gesture recognition, video-based surveillance, traffic flow monitoring and medical imaging [1]. Objects to be tracked can be represented by points [2], primitive geometric shapes, such as circles, ellipses, and rectangles [3], and articulated shape models and skeletal models [4], [5]. In [6], the tracked object is represented as a Gaussian mixture model.

Statistical methods, such as the Kalman filter [7] and particle filter [8], track the object of interest by computing prior and posterior probabilities for change in inter-frame object location. Kernel-based methods [9] track foreground regions defined for the object(s) using a specific appearance model for the target. Silhouette tracking algorithms [10], on the other hand, track the contour of the target object. Level set-based methods represent a robust subclass of silhouette tracking algorithms. Many contour tracking methods rely heavily on prior information derived from low- or high-level features and/or a motion model. In this paper, we propose a novel two-step algorithm for contour tracking that uses minimal prior information. In the first step, a simple first-order Markov chain model is used for the coarse localization of a target object. In the second step, intensity- and region-guided narrow-band level sets are used to accurately track the target contour.

The remainder of the paper is organized as follows: In Section II, we present the related work and highlight our contributions. In Section III, we describe the proposed tracking algorithm in detail. Experimental results are presented in Section IV. The paper is concluded in Section V with an outline for future research.

II. Related Work

Level set models for tracking differ in many aspects from one another. In [11], level sets evolve using the amount of shift the centroid of the inner region of the object has undergone. In [12], temporal information is incorporated within a level set evolution model. Another class of level set-based tracking methods employ shape priors. For example, [13] tracks multiple regions in an image, based on previously trained static or dynamic shape priors, employing statistically evolving level sets to localize each region in an image. Shape prior-based models have been also used for medical cell cycle analysis [14]. In [15], the authors employ appearance-based models for different image regions and subsequently, based on probabilistic estimation, discriminate between foreground and background pixels and track the object contour. In related work [16], primitive color features are used to discriminate between object and background pixels. Level set-based contour tracking methods work by constantly updating an implicit function for tracking the contour of the object of interest [1], [17], [18]. In [19], prior region- and edge-based cues are considered within a Bayesian framework to enable the level curve to converge to the target contour. In [6], probability density functions defined on texture- and color-based features are used as priors to track target contours using Bayesian inference. An alternative level set-based approach in [20] maintains two interacting level sets to track evolving contours. Coupling between these sets is determined by an image feature-driven probabilistic estimate that is computed for each pixel. Overall, all the aforementioned tracking methods depend heavily on prior information and their applicability is limited by several factors such as the constrained feature space and the velocity profile of target.

The proposed tracking algorithm is based primarily on a level set curve evolution model that minimizes the dependence on prior information while improving the accuracy of the tracked contour. Our model is shown to perform well on widely varying datasets. In the first stage of the proposed algorithm we use a first-order Markov chain which coarsely estimates a rectangular region within each frame that localizes the moving object. This estimate is based on simple primitives such as, the color features and the directional property of the evolving curve. The choice of the Markov chain enables the proposed model to overcome the shortcomings of existing tracking methods such as the aperture problem associated with optical flow computation. Instead, the proposed tracking method provides a very generic setting [21] under which methods such as optical flow computation can be modeled.

In the second stage of the proposed tracking method, we incorporate novel temporal information within the spatial Chan-Vese [17] model for tracking contours of dynamic objects. The first temporal term represents the variation in the histogram within a narrow neighborhood around the initial contour. The second temporal term captures the variation between the areas of the region enclosing the contour over successive frames. The contributions of the proposed tracking method can be summarized as follows: First, we introduce novel intensity- and region-based terms to guide the evolution of narrow-band level sets. Second, we do not use any prior shape information nor any explicit motion model, unlike many level set-based tracking methods that employ a Bayesian inference framework. In sharp contrast, we enable simple coarse localization of the target object using a first-order Markov chain. This limits the search region within which the narrow level set curve is allowed to evolve.

III. Proposed Model

In this section, we provide a detailed description of the proposed model. The two main components of the proposed model are discussed in the following two subsections.

A. *Coarse Localization of the Target*

We first employ a Markov chain based model for a coarse localization of the moving target using a rectangular region (i.e., the bounding box) within each frame as shown in Fig. 1. We describe the procedure for coarse estimation of the bounding box for frame $k+1$ (Fig. 1).

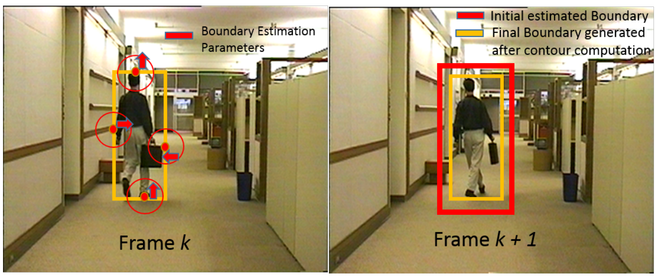

Fig. 1. Parameters for boundary estimation for subsequent frames

In Fig. 1, two successive frames k and $k+1$ are shown. For estimating a rectangular region enclosing the moving object within frame $k+1$, the model depends upon the information from frame k. In Fig. 1, the dependence of the bounding box estimate in frame $k+1$ on the parameters derived from frame k is illustrated. From the tracked contour in frame k, the minimally fitting bounding box is sampled for some reference points (marked in red in Fig. 1). Let us denote by set $S = \{p_1, p_2, ..., p_n\}$, n such points from frame k. It is to be noted that this bounding box (and hence these points) in frame k can be easily determined from the tracked contour. Subsequently, from each point $p_i \in S$ we determine the transition probabilities $T^i_{p_R} = \exp\left(-\frac{(I_{p_R} - I_{p_i})^2}{2}\right)$ for some points p_R within a neighborhood of point p_i. These neighborhoods are illustrated by circular regions in Fig. 1. I_{p_i} denotes the color features at point p_i. Along with $T^i_{p_R}$, we determine the directional inclination function D_{p_i} for each $p_i \in S$. D_{p_i} is a discrete function that indicates the direction of propagation of the level curve at point p_i.

Given $T^i_{p_R}$ and D_{p_i} the model computes the probable updated position of p_i within frame $k+1$ as follows:

$$p_i(k+1) = \underset{p_R}{\operatorname{argmax}} \prod_{m=k-1}^{k} \{T^i_{p_R}(m) + \beta(D_{p_i}(m))\} \quad (1)$$

In eqn. (1), $T^i_{p_R}(m)$ is the transition probability for frame m. The term $D_{p_i}(m)$ is an additional cost function that rewards uniform progress in the level curve movement and penalizes change in direction. Conceptually $T^i_{p_R}(m)$ for frame m should depend on $T^i_{p_R}(m-1)$ that would help localize the corresponding contour pixels between successive frames. However, the curve evolution model is restricted within a narrow band around the previously determined contour, supporting this correspondence automatically. Therefore, the terms $T^i_{p_R}(m)$ and $T^i_{p_R}(m-1)$ can be assumed to be approximately independent of each other, hence they are multiplied in eqn. (1). This assumption does not adversely impact the accuracy of the tracked contour. The proposed model, however, does not eliminate entirely the mutual interdependence of contour points tracked in successive frames. The term $D(m)$ encodes the information about the direction of motion between frames m and $(m-1)$. $D(m)$ assumes a value $\in \{-1, +1\}$ based on backward and forward motion respectively along the direction of level curve propagation. $D_{p_i}(m)$ is positive $(+1)$ for p_i when the direction of propagation of the level curve between subsequent frames is consistent and is negative (-1) otherwise. This enables the level curve to adjust faster to sudden directional changes in contour movement. The coefficient β rewards or penalizes local contour pixel evolution depending on whether there is a detectable change in the direction of evolution when compared to the previous frame.

In the present work, a low value of $\beta(= 0.1)$ was chosen since the dataset in our experiments displayed predominantly unidirectional movement. However, one can certainly assign a higher value to β in the case of datasets wherein the target motion undergoes a higher degree of directional variation. The optimization procedure in eqn. (1) is carried out for all $p_i \in S$. The outcome of this optimization is an initial estimate in the form of a rectangular region that encloses the target object as indicated by the red box in frame $k+1$ in Fig. 1. The model subsequently finds a contour of the target object within this rectangular region. Once the final contour is determined for this frame using the procedure described in the next subsection, one can easily determine a closely fitting bounding box to aid similar optimization for the subsequent frame $k+2$ and so on. The best fitting bounding box for frame $k+1$ is shown in yellow in Fig. 1.

B. Narrow Band Level Set Curve Model

The proposed curve evolution model uses the foreground region estimate or segmentation information to obtain the contour. Segmentation in frame $k+1$ uses the extracted contour information from the previous frame k. For this purpose a term $A(x,y) = f(I_d(x,y), I_h(x,y), I_v(x,y))$ is defined for each point (x,y) on the contour extracted in frame k. The three terms $I_d(x,y)$, $I_h(x,y)$, and $I_v(x,y)$ respectively represent color histograms in the diagonal (and off-diagonal), horizontal and vertical directions over a $n \times n$ neighborhood of a contour point (x,y). Thus, the first term $A(x,y)$ in the proposed curve evolution model has the following form:

$$A(x,y) = I_d(x,y) + I_h(x,y) + I_v(x,y) \tag{2}$$

It is to be noted that I_d in eqn. (2) represents both, the diagonal and off-diagonal histograms. In Fig. 2, we show how a portion of a human silhouette can be tracked over multiple frames based on the three terms in eqn. (2) using a 3×3 neighborhood. Let us assume that in frame k a portion of the shoulder (marked with black points within a box) has been tracked correctly. Then, using eqn. (2), the direction of contour motion for each of these marked points can be computed. These computed directions are used by the level set model to compute the corresponding set of points on the evolved contour in the next frame $k+1$. For instance, let us consider the center cell in Fig. 2(b), which is a contour point in frame k. The complete histogram in the 3×3 neighborhood around this point suggests movement of the contour in the horizontal and top-right directions. This movement is detected by comparing A at the same point within frames k and $k+1$. Note that the first point in frame k belongs to the previously detected contour. From the extracted motion information, an estimate of the corresponding contour point in frame $k+1$ is obtained as shown by the central cell in Fig. 2(c).

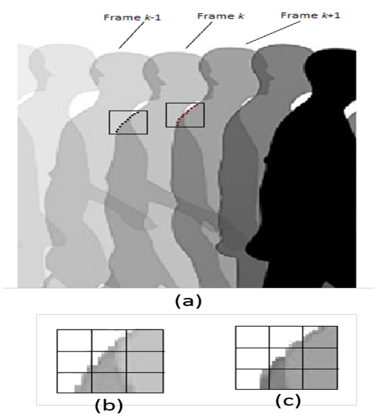

Fig. 2. (a) Human silhouette tracked over frames $k-1$, k and $k+1$. A 3×3 neighborhood around a contour point in (b) frame k and in (c) frame $k+1$

Formally, to extract the motion information at point (x,y) we define the following terms.

$$\begin{aligned} \Delta A_k^+(x) &= A_k(x) - A_k(x+\Delta x) \\ \Delta A_k^-(x) &= A_k(x) - A_k(x-\Delta x) \\ \Delta A_k^+(y) &= A_k(y) - A_k(y+\Delta y) \\ \Delta A_k^-(y) &= A_k(y) - A_k(y-\Delta y) \end{aligned} \tag{3}$$

These four terms represent changes in A at the point (x,y) along positive and negative x and y directions. The maximum possible horizontal and vertical motion can be now defined as:

$$\begin{aligned} \Delta A_k(x) &= \max\{\Delta A_k^+(x), \Delta A_k^-(x)\} \\ \Delta A_k(y) &= \max\{\Delta A_k^+(y), \Delta A_k^-(y)\} \end{aligned} \tag{4}$$

The parameters $\Delta A_{(.)}(.)$ in eqn. (4) are computed for two successive frames k and $k+1$ at the same location (x,y) and are the compared to extract the motion information. This comparison is represented using a vector $\mathbf{T}_{k+1}(x,y) = \left(\frac{\Delta A_{k+1}(x)}{\Delta A_k(x)+\epsilon}, \frac{\Delta A_{k+1}(y)}{\Delta A_k(y)+\epsilon}\right)$. The vector $\mathbf{T}_{k+1}$ is used to guide the level set-based curve evolution within the bounding box and to update the contour in frame $k+1$. The elements of $\mathbf{T}$ represent the ratio of changes in this vector along the x and y directions with respect to the previous frame. For example, non-zero values for both the elements suggest motion in the diagonal and/or off-diagonal direction and so on. A small constant ϵ is added to the denominator of the elements of $\mathbf{T}$ to avoid division by zero. For brevity, we just use the notation $\mathbf{T}$ to denote $\mathbf{T}_{k+1}(x,y)$ and formally express it as:

$$\mathbf{T} = \left(\frac{\Delta A_{k+1}(x)}{\Delta A_k(x)+\epsilon}\right)\hat{\mathbf{x}} + \left(\frac{\Delta A_{k+1}(y)}{\Delta A_k(y)+\epsilon}\right)\hat{\mathbf{y}} \tag{5}$$

The magnitude and phase angle of $\mathbf{T}$ are given by $M = \sqrt{\left(\frac{\Delta \bar{A}_{k+1}(x)}{\Delta A_k(x)+\epsilon}\right)^2 + \left(\frac{\Delta A_{k+1}(y)}{\Delta A_k(y)+\epsilon}\right)^2}$ and $\theta = \tan^{-1}\left(\frac{\left(\frac{\Delta A_{k+1}(y)}{\Delta A_k(y)+\epsilon}\right)}{\left(\frac{\Delta A_{k+1}(x)}{\Delta A_k(x)+\epsilon}\right)}\right)$ respectively. A non-zero value of M denotes a location where the curve dynamics change. A lower value of M denotes a slowly moving region of the contour whereas a higher value denotes faster movement. The phase angle θ denotes the direction along which local changes in the proximity of the contour between subsequent frames are evident.

The second term of the curve evolution model captures a region R_{k+1} in frame $k+1$ that is maximal in size and has undergone minimal changes relative to the region bounded by the object contour in frame k. Essentially, this term serves to ensure that the contour in frame $k+1$ most closely resembles the contour in frame k while simultaneously preserving the maximum contour evolution information. We denote this term as G and express it as:

$$G = \underset{R_{k+1}}{\operatorname{argmin}}\left[\int_{R_{k+1}} (H(\phi(x)) + CH(x))dx - (H(\phi)+CH)_{R_k}\right] \tag{6}$$

In eqn. (6), R_k denotes the region bounded by the object contour in frame k. Numerically, G should not vary considerably with respect to A_k as it signifies minimal inter-frame change in areas. $H(\phi(x))$ is the Heaviside function which is positive if the test point x lies within the object boundary and zero outside; whereas $\phi(x)$ is the implicit level set function computed at point x such that $\phi < 0$, $\phi > 0$, $\phi = 0$ respectively indicate regions outside, inside and on the contour at test point x [18]. $CH(x)$ is the color histogram within the region bounded by the object contour at point x. Integrating the Heaviside function within the region gives us the region size in terms of the number of pixels bounded by the contour. The second term in eqn. (6) is the integrand computed on the target object in frame k. The integration is performed over the evolving contour ϕ at each point $x \in \phi$. In effect, eqn. (6) determines the region in frame $k+1$ that minimizes the variations in the area bounded by the object contour. Note that $(H(\phi) + CH)_{R_k}$ is evaluated for frame k and serves as a constant in eqn. (6).

Let the contour of the optimized region G from eqn. (6) be denoted by C. Combining the terms in eqns. (5) and (6) with the standard Chan-Vese terms [17], we obtain the Euler-Lagrange equation that describes the proposed curve evolution model:

$$\begin{aligned}\frac{\partial \phi}{\partial t} = \delta_\epsilon(\phi)[\mu\, \mathbf{div}\left(\frac{\Delta\phi}{|\Delta\phi|}\right) - \nu - \lambda_1(u_0 - c_1(\phi))^2 + \\ \lambda_2(u_0 - c_2(\phi))^2 + \chi M\cos(\theta) + \sigma|\Delta_t(C)|_{(x,y)}]\end{aligned} \tag{7}$$

The first term $\mathbf{div}\left(\frac{\Delta\phi}{|\Delta\phi|}\right)$ represents the rate of change of the level curve at different points on the curve in the normal direction and acts as a regularizer. The second term ν, which is a constant, results from the integration of the Heaviside function. The third and fourth terms together determine the optimum balance of intensities within and outside the object region. Terms $c_1(\phi)$ and $c_2(\phi)$ respectively represent the average color intensity within and outside the evolving contour ϕ whereas u_0 is the local color intensity at the point of optimization. Thus the aforementioned spatial terms together yield a contour representing an optimal boundary, one that balances color intensity within a narrow band.

The last two terms in eqn. (7) constitute the temporal components necessary for tracking. They force the implicit level set function ϕ to take an updated value in the direction of motion specified by $\mathbf{T}$ (eqn. (5)) and the region-based optimization (eqn. (6)). M and θ are the magnitude and phase angle of $\mathbf{T}$ (eqn. (5)). The last term in eqn. (7) denotes the region-based optimization described in eqn. (6). Following the optimization procedure in eqn. (6) an optimal contour C is extracted for each frame k. The relative changes in C for two successive frames is represented using $\Delta_t(C)$ where the term Δ_t indicates time (t) dependent changes.

As mentioned earlier, an important contribution of this work is the introduction of temporal components within the spatial level set framework. The level set-based model finds an optimum contour within a statistically determined coarse region of the image that encloses the target object. We term the proposed algorithm as the *Intensity- and Region-guided Narrow-Band Level Set* (IRNBLS) tracking algorithm. The IRNBLS algorithm evolves the contour within a narrow band that lies entirely within the coarsely localized region. Let ϕ_k represent the evolved contour in frame k and n_k the number of iterations before the level set model saturates at frame k. The tolerance level for the finally converged contour is denoted by $\tau = 10^{-3}$. The model is not allowed to iterate more than N times. The algorithm returns the cumulative frame-wise tracked contours Y for the entire sequence given three inputs, i.e., the number of frames in video sequence n, an initial contour ϕ_0 in frame 1 and the set of coefficients CF in eqn. (7).

Algorithm 1 IRNBLS

procedure IRNBLS(ϕ_0, CF, N, τ)
  $\phi_1 \leftarrow \phi_0$, $\phi_{-1} \leftarrow \phi_0$ $k \leftarrow 0$, $n_k \leftarrow 0$, $n \leftarrow 3, Y \leftarrow \emptyset$
  for $k = 2$ to $n-1$ **do**
    Construct set S following Fig. 1.
    for all $x_i \in S$ **do**
      $x_i^{'} \leftarrow$ optimize eqn. (1) for x_i.
    end for
    $X \leftarrow \{x_i^{'}\}$ for frame k.
    Compute $c_1(\phi_k)$ and $c_2(\phi_k)$ following [17].
    Compute M and θ from eqn. (5).
    Optimize eqn. (7) in $n \times n$ neighborhood within X.
    $n_k \leftarrow$ steps to converge.
    if $n_k \geq N$ or $\phi \geq \tau$ **then**
      $n \leftarrow n + 2$. Increase neighborhood size.
      Re-optimize eqn. (7).
    end if
    $Y = Y \cup \phi_k$
  end for
  Return $Y = \cup_{k=1}^{n} \phi_k$
end procedure

IV. Results

We show the superiority of the proposed curve evolution model over the standard Chan-Vese framework [17]. The parameter values for the proposed model (eqn. (7)) are experimentally chosen as: $\lambda_1 = \lambda_2 = 0.1$, $\chi = 0.4$, $\sigma = 0.4$,$\mu = 0.001 \times 255^2$, and $\nu = 0.01$.

(a) (b) (c)

Fig. 3. (a) Target with non-uniform intensity distribution. (b) Tracked contour using the Chan-Vese method (shown in red). (c) Tracked contour using the proposed method (shown in blue).

Consider a frame of the *Hall Monitor* video sequence in Fig. 3(a), where the person being tracked is wearing a shirt and pants of different colors, i.e., a target object characterized by varying color and intensity values. The Chan-Vese model [17] fails to give correct results in this situation since the curve evolution procedure based on the Chan-Vese model segments the frames into piecewise uniform regions. As shown in Fig. 3(b), instead of separating the dynamic body (i.e., the man walking in the corridor) as a whole, the Chan-Vese model can only partially identify the dynamic portions (i.e., the relatively darker regions such as the shirt and suitcase) of the frame. The proposed IRNBLS algorithm yields significantly better results, as can be seen in Fig. 3(c), in that it can separate the entire moving target as a whole.

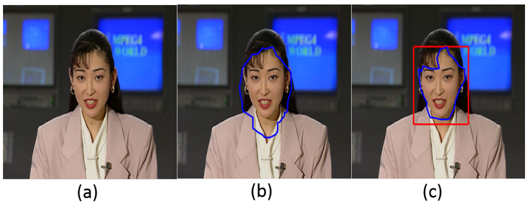

Fig. 4. Effectiveness of coarse target localization: (a) Object (human face) with non-uniform color distribution. (b) Tracked contour without using coarse localization. (c) Tracked contour with coarse localization.

Next, using Fig. 4, we show the usefulness of the coarse target localization procedure where the task is to track a human face. Without the initial bounding box estimate (eqn. (1)), the curve evolution model fails to converge accurately even after several iterations (Fig. 4(b)). In Fig. 4(c), we show that the coarse initial localization results in more accurate tracking.

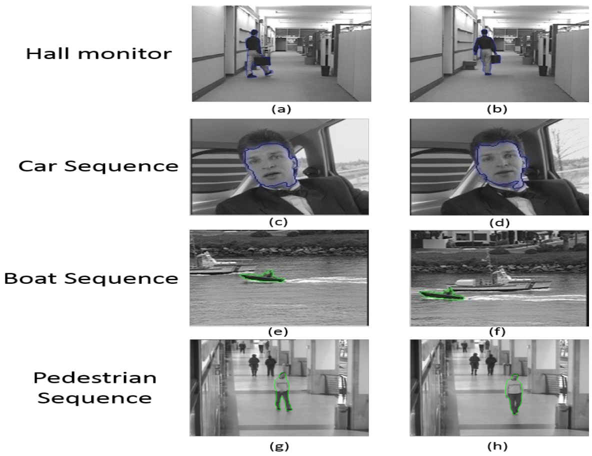

Fig. 5. Tracked contours for different datasets: *Hall Monitor* sequence (a) Frame 61, (b) Frame 82; *Car* sequence (c) Frame 1, (d) Frame 67; *Boat* sequence (e) Frame 35, (f) Frame 93; *Pedestrian* sequence (g) Frame 114, (h) Frame 127.

The proposed IRNBLS algorithm is evaluated using a dataset consisting of 8 different publicly available video sequences [22]. The video sequences exhibit varying degrees of motion, e.g., less motion for the *Car* sequence in Fig. 5(c)-(d) versus considerable motion for the *Boat* sequence in Fig. 5(e)-(f). In some cases the background of the target object is changed by a dynamic entity, such as the larger boat in Fig. 5(e)-(f), or by a panning camera resulting in a

TABLE I

STATISTICAL COMPARISON (MEAN μ AND STANDARD DEVIATION σ) OF CONTOUR HAUSDORFF DISTANCE FOR DIFFERENT TRACKING METHODS

SK [20]		DAM [19]		Yilmaz [6]		IRNBLS	
μ	σ	μ	σ	μ	σ	μ	σ
10.44	4.775	7.34	3.23	6.8188	2.2939	5.428	1.517

dynamic background (Fig. 5(e)-(f)). We compare the results of the proposed IRNBLS algorithm with those of competing level set-based contour tracking methods such as the Yilmaz method [6], discriminative appearance method (DAM) [19], and the Shi-Karl (SK) [20] method. Quantitative comparisons among these methods are performed using the Contour Hausdorff Distance (CHD) measure [23]. The CHD measure for each tracking method is computed by comparing the extracted contours with the manually extracted ground truth contours for the selected frames within each video sequence. Lower CHD values denote higher tracking accuracy. In Fig. 5 we show the tracked contours in selected frames from 4 of the 8 video sequences in our evaluation dataset. In Table I, we display the mean $\pm$ s.d. of the CHD values for each tracking method. The results in Table I clearly show that the proposed tracking method outperforms the competing methods in [6], [19] and [20].

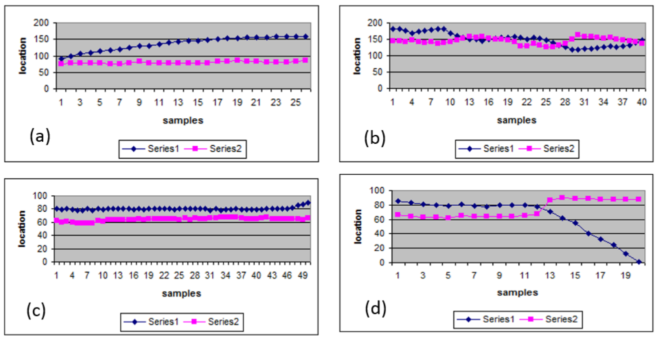

Fig. 6. Variation of x (series 1) and y (series 2) coordinate values of the target objects in the (a) *Hall Monitor* sequence; (b) *Car* sequence; (c) *Boat* sequence; and (d) *Pedestrian* sequence.

The video sequences in our experiments exhibit varying dynamic characteristics. Fig. 6 decpicts the different dynamic behaviors by jointly plotting the frame-wise variations in the x and y coordinates of the centroid of the target objects in different video sequences. The plots in Fig. 6 support our claim that the proposed method can be applied to several widely varying video sequences.

Note that in contour tracking methods, one can potentially obtain a bounding box which minimally encloses the tracked contours (e.g., the yellow box in Fig. 1). Since bounding box-based performance evaluation measures such as the sequence frame detection accuracy (SFDA), multiple-object detection precision (MODP), multiple-object detection accuracy (MODA), and average tracking accuracy (ATA) [26] are used for object tracking methods, we extend our comparison to two such methods [24], [25] those track multiple objects.

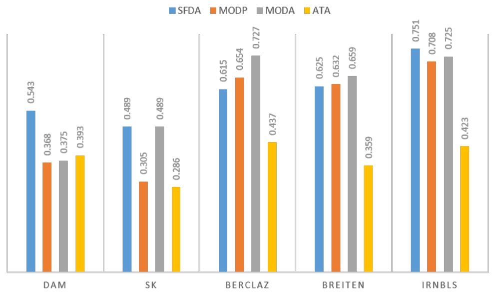

Fig. 7. Graphical comparison using the SFDA, MODP, MODA, and ATA measures between competing tracking methods in [19], [20], [24], [25] and the proposed IRNBLS algorithm

Though used extensively for multiple-object tracking analysis, performance evaluation measures such as MODP and MODA can be easily adapted for single-object tracking analysis. For the proposed method, and competing tracking methods in [19], [20], [24], [25], we compute the bounding boxes for the tracked objects. Fig. 7 depicts the relative performance of the proposed IRNBLS algorithm vis-a-vis the competing tracking methods [19], [20], [24], [25] on the basis of SFDA, MODP, MODA and ATA. Fig. 7 clearly shows that the IRNBLS algorithm outperforms the competing tracking methods in [19], [20], [25] and is comparable to the tracking method in [24]. Specifically, the SFDA and MODP measures of the IRNBLS algorithm are marginally better than those of [24] whereas the MODA and ATA measures are comparable. All the tests were performed on machines with an Intel Core-2 Quad-Core processor running at 2.4 GHz. The average processing time per frame for a video sequence with significant dynamic background was 1.43 seconds. The proposed method with the current execution time can be applied for problems like event analysis in surveillance and video motion capture [27].

V. Conclusion

In this paper, we proposed a color-, intensity- and region-guided narrow-band level set-based method for tracking contours of dynamic objects in a video sequence. The proposed model essentially enhances the standard Chan-Vese framework [17] with the addition of two new terms and is called the *Intensity- and Region-guided Narrow-Band Level Set* (IRNBLS) tracking algorithm. One of the terms captures the variations in histogram information between successive frames whereas the other term captures the region optimally covered by the target contour. The curve evolution model is restricted within a coarsely estimated rectangular region that encloses the target object. Experimental results on standard tracking datasets demonstrate that the proposed method can track contours in a variety of videos very accurately. However, there is scope to further extend the IRNBLS tracking algorithm to videos that vary widely in their target velocity profiles. Another direction of future research will be to extend the proposed framework for detection of single and multiple targets under more difficult imaging conditions like camera motion and uneven illumination.

References

[1] A. Yilmaz et al., “Object tracking: A survey,” *ACM Comp. Surv.*, vol. 38, no. 4, p. 13, 2006.
[2] D. Serby et al., “Probabilistic object tracking using multiple features,” in *Proc. ICPR*, vol. 2, 2004, pp. 184–187.
[3] D. Comaniciu et al., “Kernel-based object tracking,” *IEEE Trans. Patt. Anal. Mach. Intel.*, vol. 25, no. 5, pp. 564–577, 2003.
[4] D. Ballard and C. Brown, *Computer Vision*. Prentice-Hall, 1982.
[5] V. Caselles et al., “A geometric model for active contours in image processing,” *Numerische mathematik*, vol. 66, no. 1, pp. 1–31, 1993.
[6] A. Yilmaz et al., “Object contour tracking using level sets,” in *Proc. ACCV*, vol. 1, 2004.
[7] T. J. Broida and R. Chellappa, “Estimation of object motion parameters from noisy images,” *IEEE Trans. Patt. Anal. Mach. Intel.*, no. 1, pp. 90–99, 1986.
[8] M. S. Arulampalam et al., “A tutorial on particle filters for online nonlinear/non-gaussian bayesian tracking,” *IEEE Trans. Sig. Process.*, vol. 50, no. 2, pp. 174–188, 2002.
[9] S. C. Zhu et al., “Region competition: Unifying snakes, region growing, and energy/bayes/mdl for multiband image segmentation,” in *Proc. ICCV*, 1995, pp. 416–423.
[10] K. Sato and J. K. Aggarwal, “Temporal spatio-velocity transform and its application to tracking and interaction,” *Comp. Vis. Img. Undrstdng.*, vol. 96, no. 2, pp. 100–128, 2004.
[11] S.-H. Lee and M. G. Kang, “Motion tracking based on area and level set weighted centroid shifting,” *IET Comp. Vis.*, vol. 4, no. 2, pp. 73–84, 2010.
[12] W. Fang et al., “Incorporating temporal information into level set functional for robust ventricular boundary detection from echocardiographic image sequence,” *IEEE Trans. Biomed. Engr.*, vol. 55, no. 11, pp. 2548–2556, 2008.
[13] M. Fussenegger et al., “Multiregion level set tracking with transformation invariant shape priors,” in *Proc. ACCV*, 2006, pp. 674–683.
[14] Y. N. Law and H. K. Lee, “Level set based tracking for cell cycle analysis using dynamical shape prior,” in *Proc. Med. Img. Undrstndng. Anal.*, 2012, pp. 137–142.
[15] W. Li et al., “Discriminative level set for contour tracking,” in *Proc. ICPR*, 2010, pp. 1735–1738.
[16] A. S. Chowdhury et al., “Colonic fold detection from computed tomographic colonography images using diffusion-fcm and level sets,” *Patt. Recog. Lett.*, vol. 31, no. 9, pp. 876–883, 2010.
[17] T. F. Chan and L. A. Vese, “Active contours without edges,” *IEEE Trans. Img. Process.*, vol. 10, no. 2, pp. 266–277, 2001.
[18] S. Osher and R. Fedkiw, *Level set methods and dynamic implicit surfaces*. Springer Sc. & Bus. Med., 2006, vol. 153.
[19] X. Sun et al., “Contour tracking via on-line discriminative appearance modeling based level sets,” in *Proc. IEEE ICIP*, 2011, pp. 2317–2320.
[20] Y. Shi and W. C. Karl, “Real-time tracking using level sets,” in *Proc. IEEE Conf. CVPR*, vol. 2, 2005, pp. 34–41.
[21] D. Piao et al., “Computing probabilistic optical flow using markov random fields,” in *Intl. Symp. Comp. Model. Obj. Rep. Img.*, 2014, pp. 241–247.
[22] “http://homepages.inf.ed.ac.uk/rbf/caviardata1/,” http://homepages.inf.ed.ac.uk/rbf/CAVIARDATA1/, accessed: 2016-03-28.
[23] D. P. Huttenlocher et al., “Comparing images using the hausdorff distance,” *IEEE Trans. Patt. Anal. Mach. Intel.*, vol. 15, no. 9, pp. 850–863, 1993.
[24] J. Berclaz et al., “Multiple object tracking using k-shortest paths optimization,” *IEEE Trans. Patt. Anal. Mach. Intel.*, vol. 33, no. 9, pp. 1806–1819, 2011.
[25] M. D. Breitenstein et al., “Online multiperson tracking-by-detection from a single, uncalibrated camera,” *IEEE Trans. Patt. Anal. Mach. Intel.*, vol. 33, no. 9, pp. 1820–1833, 2011.
[26] R. Kasturi et al., “Framework for performance evaluation of face, text, and vehicle detection and tracking in video: Data, metrics, and protocol,” *IEEE Trans. Patt. Anal. Mach. Intel.*, vol. 31, no. 2, pp. 319–336, 2009.
[27] Y. Wei et al., “Interactive offline tracking for color objects,” in *Proc. ICCV*, 2007, pp. 1–8.