Robust Segmentation Based Tracing Using an Adaptive Wrapper for Inducing Priors

Vignesh Jagadeesh, Bangalore S. Manjunath, *Fellow, IEEE*, James Anderson, Bryan W. Jones, Robert Marc, and Steven K. Fisher

***Abstract*— Segmentation based tracing algorithms detect the extent and borders of an object in a given frame I_Z by propagating results from frames $I_{1 \leq z < Z}$. Although application specific tracers have been forthcoming, techniques that automatically adapt across applications have been less explored. We approach this problem by learning a prior model on topological dynamics that encourages segmentation transitions across frames that are most likely for a given application. Further, we augment a generic tracing technique with a locality sensitive prior derived from dense optic flow fields for deformation guidance. The proposed approach comprises two stages where the generic tracer initially yields multiple segmentation transitions when its parameters are perturbed, and the learnt topology prior subsequently propagates high scoring segmentations. Because the learnt topology model wraps around a generic tracer and adapts it by setting its free parameters, the need for careful parameter tuning is completely obviated. Through extensive experimental validation in surveillance, biological and medical image datasets, we verify the applicability of the proposed model while demonstrating good tracing performance under severe clutter.**

***Index Terms*— Tracing, parameter adaptation, electron micrograph, Markov random fields.**

I. Introduction

SEGMENTATION based tracing/tracking is a challenging problem that merits attention due to its wide range of diverse applications. The problem is inherently hard due to wide variability in imaging modalities, background clutter, illumination artifacts, occlusion and various distracters that appear in different problem scenarios. As a result, developing an overarching framework that would reliably work across problem scenarios is an active area of research.

Motivating Example: Let us consider Figure 1 where the segmentation from frame $z-1$ is used as a prior for frame z. Typical segmentation algorithms are formulated as the minimization of an energy function $E(y, P)$, parameterized by the segmentation labels y and to algorithm's parameters P. Perturbation of these parameters P leads to multiple segmentations, as shown in the middle column. Depending on the application of interest what would constitute a good segmentation transition would vary. Let us consider three applications indicated by the red, green and blue arrows drawn next to each other and consider how they would vote for a good segmentation transition. Note that though the segmenter is characterized by an energy function, we utilize the notion of scoring (higher scores map to lower energies), that is more intuitive to explain the voting mechanism. The switch between scores and energies will be obvious from context throughout the paper. The *region term* is an orange color foreground and a dark blue color background. The surveillance application (red solid arrow) would contend that a person cannot split into two and there is a smoothness of motion from one frame to another, thus scoring parameter P_1 highly (low energy). Tracing splitting

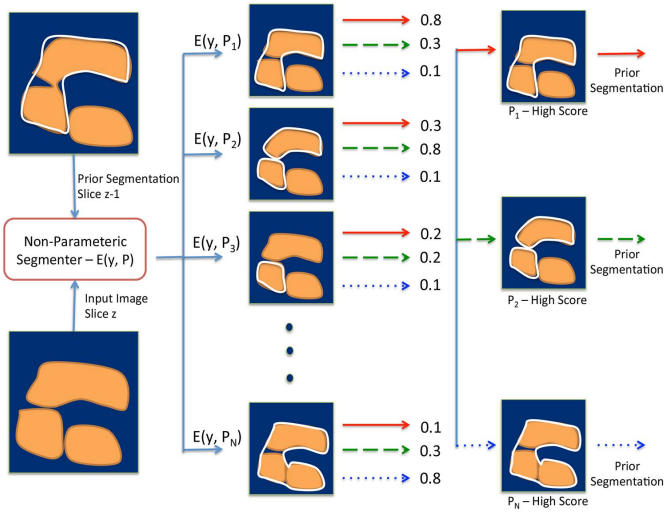


Fig. 1. (Best Viewed in Color) Depending on the application of interest (shown by scores generated by red, green and blue arrows) what constitutes a good segmentation transition from frame $z-1$ to frame z using a segmentation algorithm $E(y, P)$ varies from one application to another. For instance, in surveillance (red solid arrow) there is a prior of connectedness on the object of interest thus preferring the segmentation transition generated by parameter P_1. In cell tracing (green dashed arrow), an object of interest is allowed to split into two since it is a biologically plausible event leading it to prefer the transition generated by P_2. While tracing an expanding organ (blue dotted arrow), there is a prior on increasing surface area, as opposed to splitting or shrinking thereby scoring P_N highly (low energy). The primary aim of this work is to learn the notion of good segmentation transitions in order for the algorithm to seamlessly adapt from one dataset of interest to another.

Manuscript received November 16, 2012; revised April 10, 2013 and July 6, 2013; accepted August 10, 2013. Date of publication August 28, 2013; date of current version October 2, 2013. This work was supported by a grant from the National Science Foundation NSF OIA under Grant 0941717. The associate editor coordinating the review of this manuscript and approving it for publication was Prof. Marios S. Pattichis.

V. Jagadeesh, B. S. Manjunath, and S. K. Fisher are with the Center for Bioimage Informatics, University of California, Santa Barbara, CA 93106 USA (e-mail: vignesh@ece.ucsb.edu; manj@ece.ucsb.edu; fisher@lifesci.ucsb.edu).

J. Anderson, B. Jones, and R. Marc are with the John A Moran Eye Institute, University of Utah, Salt Lake City, UT 84132 USA (e-mail: james.r.anderson@utah.edu; bryan.jones@m.cc.utah.edu; robert.marc@hsc.utah.edu).

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TIP.2013.2280002

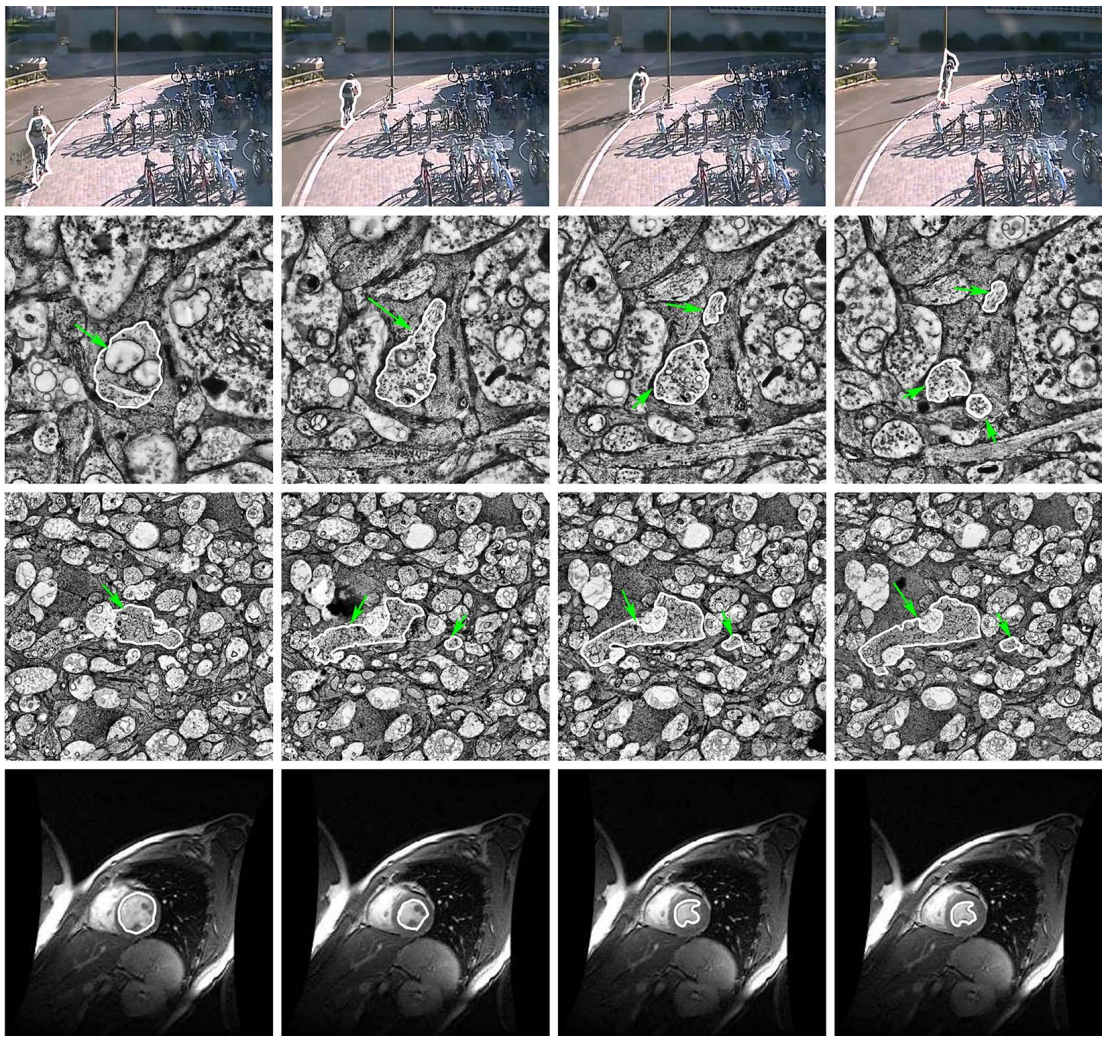


Fig. 2. (Best Viewed in Color) Illustration of the proposed algorithm's applicability in diverse scenarios. The first row is an example of tracking under very poor video quality, illumination artifacts, shadows and occlusions. Rows 2-3 show cross-sectional images sampled from a retinal connectome acquired using scanning section transmission electron microscopy. Some critical challenges that need to be overcome are capturing abrupt deformations across the z-direction, and detection of topological changes such as splitting and merging of regions. Row 2 is an example of an object initially splitting into two regions and further splitting into three. Row 3 illustrates an example of an object splitting on the right and merging on the left. The final row illustrates applicability of the technique to tracking cardiac data in medical imagery.

cells (green dashed arrow) would allow splits as feasible events in case of weak foreground likelihoods between the two regions assigning P_2 a high score (low energy). Finally, while tracking a growing organ (blue dotted arrow) one could be certain that the object cannot split or shrink, and is most likely to expand giving P_N the highest score (low energy). These examples illustrate that good segmentation transitions vary from one scenario to another and that it would be very useful to have a framework that adapts from one scenario to another with minimal human intervention. Figure 2 provides results demonstrating the proposed model's ability to adapt across three different datasets: surveillance in a camera network, organ tracking in medical imagery and neuronal structure tracing in biological image sequences.

The primary contributions of this work include:

- A novel framework capable of inducing high level topology priors for segmentation based tracing, addressed by parameter adaptation (see Definition 1.1) of the segmentation based tracer.
- Introduction of a novel locality sensitive prior (see Definition 1.2) modeling object deformations.

Definition 1.1: Parameter Adaptive Tracer - Consider a segmentation based tracing algorithm operating on an image sequence $\{I_1, I_2, \ldots, I_N\}$, characterized by a sequence of segmentation energies $\{E(y_1, P_1), E(y_2, P_2), \ldots, E(y_N, P_N)\}$, with one segmentation energy corresponding to every image, and $y_i \in \mathcal{Y}$, $P_i \in \mathcal{P}$ denote the label and parameter vectors respectively for the i^{th} image. The segmentation based tracer is parameter adaptive if parameters of the segmentation energy corresponding to the i^{th} image in the sequence are determined by, $P_i \propto f(P_j), 1 \leq j < i$, where f is a functional mapping $f : \mathcal{P} \rightarrow \mathcal{P}$.

Definition 1.2: Locality Sensitive Priors - The segmentation based tracer possesses a locality sensitive prior if the spatial

positioning of the object of interest in frames $i > j \geq 1$, indicated by segmentation labels $y_j, 1 \leq j < i$ is utilized in inferring the segmentation labels in frame i, $y_i \propto g(y_j), i > j \geq 1$, where g is a functional mapping $g : \mathcal{Y} \rightarrow \mathcal{Y}$.

Note on Parameter Adaptation: Most segmentation and tracing algorithms often have free parameters referred to as P_i in Figure 1, that need to be tuned for applicability in a dataset of interest. In the case of static images, these parameters are often learnt from a database of annotated imagery. The hope in such scenarios is that the learnt parameter vector would yield good average performance across the test set. However, a parameter that works well for an image of one class may not work well for an image of another. Hence, adapting parameters of a segmentation algorithm based on image content is crucial. Thus far, we are not aware of any major effort towards parameter adaptation in static image segmentation. The situation in case of videos or image sequences for learning parameters is more difficult due to the size of data involved. Furthermore, a variety of distracters like illumination artifacts, occlusion, scale variations, registration errors may appear as one moves through the third dimension in an image sequence. In such cases, having a fixed parameter for tracing would lead to very poor results, as will become evident in the following sections. For example, in video surveillance, a parameter employed for segmenting an object when it is close to the camera will not be the best parameter to use when it has moved away from the camera and appears much smaller. More importantly, when a tracing algorithm loses a target midway, it is difficult to recover unless an operator intervenes to re-initialize the algorithm. In contrast, a fully automatic tracing algorithm must graciously adapt its parameters according to the image content/immediate history to minimize chances of a target loss. This work primarily focuses on the problem of inducing high level priors using a module wrapping around a generic tracer, and automatically adapting it's parameters. In summary, parameters of the base segmenter are utilized to embed high level priors, and since these parameters are automatically selected by the prior model in each frame there is no need to hand tune them!

Topological Priors Through Parameter Adaptation: The key observation we make in this work is that object dynamics (across time in videos, or depth in 3D stacks) could be made to influence parameter adaptation in image sequence segmentation. Object dynamics refer to behavior of objects (changes they undergo) as one progresses through an image sequence. An example of object dynamics in surveillance or cardiac tracking is the way objects expand or shrink from frame to frame without any splits, while in tracing neuronal structures it is the manner in which structures expand/contract or split/merge. We refer to the above model as the topology prior on structures of interest, which would be capable of scoring the likelihood of a segmentation transition in frame z from $z-1$. Thus far, the two key insights of the paper are utilizing the dynamics of an object to adapt parameters of an image sequence segmenter, and modeling the notion of high level topology priors to capture the behavior of structures across time/depth. Figure 2 illustrates the ability of the proposed model to capture massive object deformations and topological changes. The ability to induce knowledge about the split/merge behavior of a target is made possible by topological priors, and is inherently different from shape/geometric priors as will be discussed. As will become evident, the developed framework can incorporate any generic high level prior (shape/geometric etc.), but we restrict ourselves to topology priors for ease of exposition. Further, as mentioned in the abstract, this work introduces a novel locality sensitive prior derived from dense optical flow fields for accurate tracing through an image sequence, in order to strengthen the base segmenter.

In developing our algorithms (Section II-A) we emphasize generating solutions that conform to priors as opposed to global optimization. Hence, in cases where the feasible parameter set size is large, algorithms yielding local solutions are accepted for tractability. A preliminary version of this work appeared in [1]. While we had proposed the idea of parameters adaptation for Electron Micrograph tracing in [1], the contributions relating to locality sensitive priors, extension to surveillance/medical imaging, studying different inference techniques, comprehensive experimental validation and detailed explanations of the framework are completely new.

II. Background and Related Works

There exists substantial prior work in segmentation based tracing, with frameworks ranging from shortest path based [2], watersheds [3], [4], random walkers [5], active contours based snakes [6], geodesic active contours, vector flows [7], active contours without edges [8], and discrete valued Markov Random Fields (MRFs) inferred using graph cuts [9], [10]. The aim of this paper is to present a method that is capable of wrapping around a non-parametric segmentation technique, thereby achieving two objectives: Firstly, the proposed technique is capable of embedding high level priors into the tracing algorithm using free parameters of the base segmenter. Secondly, free parameters present in the segmentation algorithm are implicitly set by the topology control equations thus annulling the need for hand tuned parameters. In other words, the proposed framework is only applicable to a non-parametric segmenter that implicitly handles topological changes like random fields or level sets. For example, while tracking a cell, biological priors on topology state that cells cannot subdivide into three or more children. In surveillance scenarios, a single human being tracked cannot suddenly split into 5 parts that walk away in different directions. However, we are not aware of an existing segmentation technique that inherently embeds the above topological priors.

In this paper we adopt a segmenter based on Markov Random Fields (MRFs) for illustrating a proof of concept of our ideas. The following discussion provides background on MRFs and introduces the notation used in this work. Markov Random Fields are models formulated to solve the image labeling problem. The aim is to label every pixel p in the set of pixels $\mathcal{P}$ constituting the image with a label y_p from a label set $\mathcal{L} \in \{1, 2, ...|\mathcal{L}|\}$. Each pixel p resides in a grid graph and has data x_p associated with it. Depending on the problem requirements, the number of neighbors with

which a pixel can interact (or has direct edges to) defines the size of its neighborhood ($\mathcal{N}_p$). The goal is to infer the pixel labels conditioned on the data as efficiently and accurately as possible.

The cost function employed for MRFs is given by:

$$E(y) = \sum_{p \in \mathcal{P}} V_p(y_p) + \sum_{p \in \mathcal{P}, q \in \mathcal{N}_p} V_{pq}(y_p, y_q), \qquad (1)$$

where $V_p(y_p) = -\log P(x_p|y_p)$, the negative log likelihood function is commonly known as the unary or terminal cost. $V_{pq}(y_p, y_q) = \lambda_I \exp\left(-\frac{||x_p - x_q||}{\sigma_I^2}\right)\delta(y_p \neq y_q)$ is the interaction or neighborhood cost modeling similarity between adjacent pixels. Inferring y_p from Equation 1 requires its minimization, which is achieved by a mincut [11] on the constructed graph. For purposes in this paper, the latent parameter vector we wish to infer in every frame is given by $\alpha = \{\lambda_I, \sigma_I\}$. A detailed treatment of MRFs can be found in Boykov et al. [9]. The distinction between commonly used priors in segmentation is now clarified. Geometric priors refer to image transformations (eg: liner, affine, projective). Shape refers to an object's appearance modulo rotation, scale, translation and minor deformations. An example is the spherical shape of a football. Topological priors are related to an object's physical appearance in terms of its connectedness. For example, a single cell can split into a maximum of two new connected components, whereas a soccer ball cannot be split into two. Introduction of shape and contextual priors in MRFs has received much attention in recent times [12], [13], [14]. Existing works in the MRF literature introduce topological constraints such as connectedness in a segmentation algorithm [15]–[17], and do not learn a topological model over possible events such as split/merge/expansion/contraction. The proposed technique encodes a topological prior into a non-parametric segmentation framework, and thus differs from deformable shape prior segmenters [18]. Peng and Veksler [19] follow a related method for learning parameters using boosted classifiers for static images.

A. Formulation of Contour Propagation

The main problem being addressed in this work is the construction of a wrapper around a tracing algorithm that embeds high level (topology) priors through parameter adaptation. Given the image I_z from frame z, segmentation Y_{z-1} from frame $z-1$, the goal is to estimate segmentation labels Y_z and parameter α_z in slice z. This is achieved by solving two maximization (equivalently energy minimization) problems, namely Parameter Search: $\underset{\alpha_z}{\text{argmax}} P(\alpha_z|Y_z, I_z, Y_{z-1})$, and Region Search: $\underset{Y_z}{\text{argmax}} P(Y_z|\alpha_z, I_z, Y_{z-1})$. The two stages are now defined and are illustrated in Figure 3:

- *Parameter Search:* Conditioned on the segmentations $P(\alpha_z|Y_z, I_z, Y_{z-1})$, estimate most likely parameters. This is our primary contribution, and Section III focuses on modeling and inferring this quantity.
- *Region Search:* Conditioned on parameters $P(Y_z|\alpha_z, I_z, Y_{z-1})$, estimate most likely segmentations. This portion relates to augmenting a generic MRF segmentation procedure with deformation priors for reliably tracking non-rigid objects. This is a secondary contribution of this paper, and presented in the Section IV. Specifically, the MRF energy is modified to E^{deform}, comprising a modified unary term V_p^{deform} that yields spatial localization and deformation guidance, see Section IV (Equation 14).

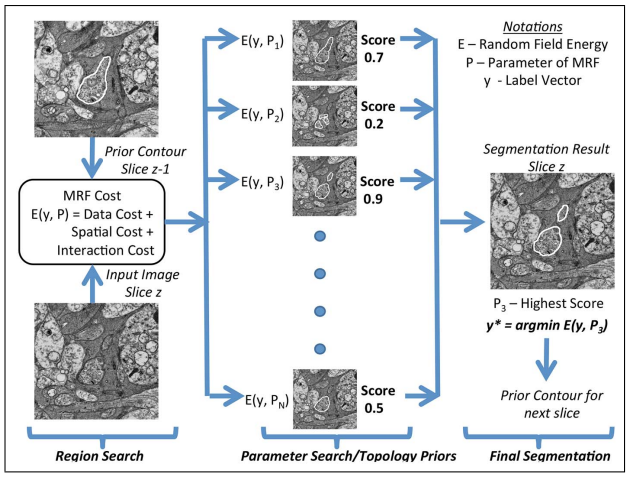


Fig. 3. Illustration of the algorithm flow for a single iteration of the proposed approach. The prior contour is utilized to generate a prediction using contour search. The prediction is used to assign scores (energies) to multiple segmentations obtained by varying MRF parameters P using the regional stability likelihood. It is useful to recall that higher scores correspond to lower energy configurations. The parameter having maximum conformance to topology priors is the segmentation result for the current slice, and acts a prior for the next slice.

This decoupling of the original problem into a segmentation algorithm specific region search, and a generic wrapper in the form of parameter search that can work with any non-parametric segmenter leads to a tractable solution to topology aware tracing. It is this plug and play nature of the construction which makes the model adaptive and more generic than the specific case of MRFs considered.

III. Parameter Search - Formulation

The goodness score for the parameter search module can be written as:

$$\begin{aligned} &P(\alpha_z|Y_z, Y_{z-1}, I_z) \\ &\quad = \frac{P(Y_z, I_z, Y_{z-1}|\alpha_z)P(\alpha_z)}{P(Y_z, I_z, Y_{z-1})} \\ &\quad \propto \underbrace{P(I_z|Y_z, Y_{z-1}, \alpha_z)}_{\text{Data Quality}} \underbrace{P(Y_z|Y_{z-1}, \alpha_z)}_{\text{Plausibility}} P(Y_{z-1}|\alpha_z) \\ &\quad \propto \underbrace{P(I_z|Y_z, Y_{z-1}, \alpha_z)}_{\text{Data Quality}} \underbrace{P(Y_z|Y_{z-1}, \alpha_z)}_{\text{Plausibility}} \end{aligned} \qquad (2)$$

The data quality stated above is measured by the histogram intersection between $f(I_z(Y_z))$ and $f(I_{z-1}(Y_{z-1}))$. Note that the intersection between two histograms $h_{ib}, h_{jb}, b \in \{1, ..B\}$ with B bins is defined by $\sum_{i=1}^{B} \min\{h_{ib}, h_{jb}\}$. The function f can be any function estimating the density of pixel intensity. The simplest form (also used in this work) would be a histogram of pixel values of foreground pixels from frame $z-1$

and z. The topological priors (Plausibility) are modeled as a state transition model as explained below. In the following, note a change in notation from $y_{p,z}$ to Y_z. The higher level topology model works on a more global interpretation (Y_z) of the region as a collection of pixels, while the MRF works at the level of pixels ($y_{p,z}$) by factorizing their probabilities. Without loss of generality, the discussions below refer to a single region undergoing topological changes.

A. Parameter Search - The Topology Prior

In order to model the prior, the essence of the problem reduces to scoring a pair of segmentations, Y_z and Y_{z-1}. In other words, one has to predict how likely Y_z was generated from Y_{z-1}. We construct a simple generative process explaining topological changes in electron micrograph stacks. An object (region) moving down a serial stack can undergo any one of four topological events, namely shrinkage, expansion, split, and merge. Recall that Y_z and Y_{z-1} are binary segmentation masks, with Y_z resulting from Y_{z-1} due to either an (1) expansion event from Y_{z-1} to Y_z, (2) contraction event from Y_{z-1} to Y_z, (3) split event from Y_{z-1} to Y_z, or (4) merge event from Y_{z-1} to Y_z. A shrinkage (expansion), considered a regionally stable event is always assumed to decrease (increase) a region's surface area from one frame to another. A split (merge) is considered a regionally unstable event and is constrained by priors on the nature of split (merge). The events are mutually exclusive, meaning they cannot co-occur for a given region.

The first step is to detect which of the four events occurred while Y_{z-1} transformed itself to Y_z. This can be established using the following two quantities that can be efficiently implemented using morphological image processing. Recall that Y_{z-1} is an estimate of the segmentation in slice $z-1$, and let $\cup_{i=1}^{L} Y_z^i$ be the set of L regions generated in slice z (utilizing parameter α_z), where each of the L regions overlap with Y_{z-1}. Note that $L > 1$ only when a contour splits from slice $z-1$ to slice z.

- *Relative Surface Area (d):* The ratio of region areas from the estimated region in frame $z-1$ and the L overlapping region(s) produced by the segmentation algorithm on frame z. Denoting the number of non-zero elements (foreground pixels) in a binary mask by $|.|$, the relative surface area is defined as,

$$d = \frac{|Y_{z-1}|}{\sum_{i=1}^{L} |Y_z^i|} \tag{3}$$

- *Region Stability (r):* The function is constructed so that if there is expansion or shrinkage (considered stable transitions since the connected component is preserved) from frame $z-1$ to z, r evaluates to a non-negative number, while it is negative for split or merge behavior.

$$r = -\mathcal{I}_S \vee \mathcal{I}_M \tag{4}$$

The variables $\mathcal{I}_S$ and $\mathcal{I}_M$ are indicator variables indicative of a split or merge respectively, and $\vee$ refers to a logical OR operation.

$$\mathcal{I}_S = \begin{cases} 1, & L > 1 \\ 0, & L \leq 1 \end{cases} \qquad \mathcal{I}_M = \begin{cases} 1 - \mathcal{I}_S, & d < 0.5 \\ 0, & \text{otherwise} \end{cases} \tag{5}$$

Regionally stable event is a contour transition that comprises only expansions/contractions and leads to a non-negative regional stability value.

The segmentation (region) transition prior is modeled under the assumption that transitions corresponding to different topological events are normally distributed with respect to d. The decomposition of probabilities with events $T = \{1, 2, 3, 4\}$ corresponding to shrinkage, expansion, split, and merge are given by:

SHRINKAGE $r \geq 0, d \geq 1$,

$$P(T = 1) = \mathcal{N}(d; 1 - \mu_1, \sigma_1) H(d-1) H(r)$$

EXPANSION $r \geq 0, d < 1$,

$$P(T = 2) = \mathcal{N}(d; 1 + \mu_1, \sigma_1)(1 - H(d-1)) H(r)$$

SPLIT $r < 0, \mathcal{I}_S = 1$,

$$P(T = 3) = \mathcal{N}(d; \mu_2, \sigma_2) H(L_{prior} - L) \mathcal{I}_S (1 - H(r))$$

MERGE $r < 0, \mathcal{I}_M = 1$,

$$P(T = 4) = \mathcal{N}(d; \mu_3, \sigma_3) \mathcal{I}_M (1 - H(r))$$

In the above equations, H refers to the Heaviside function that evaluates to one if the argument is non-negative. Once the event has been detected, we characterize the likelihood of contour transition using Gaussian distributions, parameterized by relative surface area d. μ_i, σ_i, where $1 \leq i \leq 3$ are parameters of a normal distribution learnt respectively for shrinkage/expansion, split and merge events. L_{prior} is the maximum number of regions that can result from a split, as observed from the training data. The state transition distribution modeling topological events is given by $P(Y_z|Y_{z-1})$, and is a part of Equation 2.

The probability of a topological change occurring, without any image dependent information is given by:

$$P(Y_z|Y_{z-1}) = \sum_{i=1}^{4} P(T = i) \tag{6}$$

The parameters of the topology prior model $\{\mu_1, \sigma_1, \mu_2, \sigma_2, \mu_3, \sigma_3, L_{prior}\}$ are learnt using standard techniques for fitting Gaussian distributions to training data. In the tracing scenario, training data refers to a set of pre-annotated image sequences where all events split, merge, expansion and contraction are assumed to occur.

The above discussion assumes that splits and merges are plausible events across the third dimension. However, in the case of tracking the endocardium in MR sequences and people in surveillance scenarios, there are two strong region level priors that can be exploited. Firstly, there cannot be a topology change from one frame to another. In other words, there is a requirement that a single connected component must persist through the image sequence. Secondly, there is a smoothness in motion from one frame to another due to the nature of image acquisition. In order to induce priors for this setting, the only modification that is required from the Electron Micrograph application is a change in the scoring function for segmentation transitions. We propose the quality of segmentation

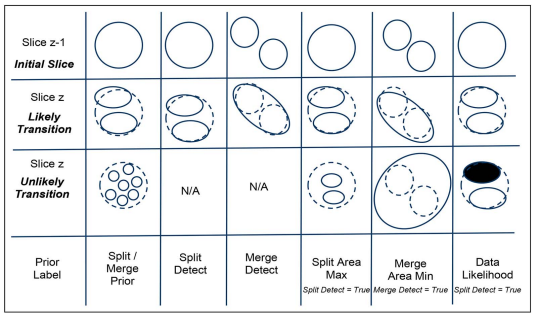


Fig. 4. Illustration of the prior information available pertaining to topological changes. The scoring function for segmentation transitions is constructed based on the above illustration. See text for details.

(region) transitions for endocardium and person tracking to be scored by the following expression,

$$P(Y_z|Y_{z-1}) = \mathcal{N}(1-\mu_1, \sigma_1)H(r) \tag{7}$$

In the above equation, a segmentation transition is rendered infeasible if there is a split since there is a strong prior on connectedness for tracking people and the endocardium. Further, since the frame rate of image capture is high, the mean of the distribution governing surface area stays close to one, which is parameterized by $\{\mu_1, \sigma_1\}$.

The topological model can be thought of as a dynamic regularizer. For instance, splitting a region into numerous child regions could serve to yield maximum data likelihood, but is regularized by the topological prior on splitting. In some applications, the topological prior enforces an infinity cost for splitting (cardiac MR, human tracking) while in others (electron micrograph tracing), it can be learnt from data.

B. Generating Multiple Segmentations

In the case of Markov Random Fields with clique size two, stability of segmentation can be modified by varying regularization λ_I and edge strength parameters σ_I defined in Equation 1. It is well known that the output of segmentation gradually changes from undersegmentation to oversegmentation as the effect of the interaction and edge strength parameters are varied. The greater the value of σ_I, larger is the variance of the contrast sensitive potential V_{pq} thus favoring only very strong edges, while reduction in value begins favoring weaker edges. On a similar note, λ_I is parameter controlling the relative importance of unary and interaction terms. Since the parameter variations control the regional stability of the segmentation outcome, we propose to search over the 2D parameter spanned by $\alpha_z = (\sigma_I, \lambda_I)_z$ for most probable segmentations in slice z.

Some generic priors related to the problem at hand that need to be formalized are (see Figure 4):

- *Split/Merge Prior:* The chance of a region splitting (merging) into (from) more than three sub-regions is extremely low.
- *Split Area Max:* In case the area of a region in frame z decreases in comparison to area in frame $z-1$ (in case of split or shrinkage), area of overlap between regions in frame $z-1$ and z must be maximized.

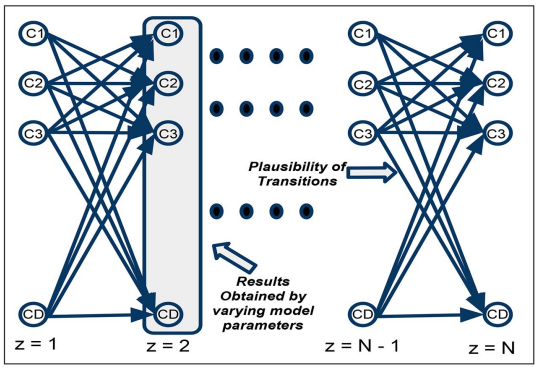


Fig. 5. An illustration of the higher level graphical model, the hidden variables being the parameters that need to be inferred from the data. Each node C, corresponds to a segmentation generated by varying MRF parameters. Note that the cardinality of the discrete parameter space is D, resulting in D segmentations per frame. The nodes (or emission probabilities) denote likelihood of a particular assignment, while the directed edges (transition probabilities) denote how plausible transitioning from one assignment to another is.

- *Merge Area Min:* The chance of a region merging with another region must result in a region with area around the sum of the two parent regions.
- *Split Detect:* Split can be detected if there are multiple overlapping connected components between frame $z-1$ and z.
- *Merge Detect:* Merge can be detected if there is a massive change in region areas between frame $z-1$ and z.
- *Data Likelihood Agreement:* The agreement of data likelihoods between successive overlapping regions must be maximized.

Inferring the proposed model in Eq. 2 yields the desired pixel wise segmentation for every frame in the image stack. Depending on the parameter state space, the inference technique adopted can vary.

C. Parameter Search - Naive Inference

Naive Inference is a greedy approach performing posterior maximization over every frame. Assume M to denote intermediate regions and Θ to be a fixed set containing feasible parameters over which optimization is performed. The aim is to pick a maximizer from Θ, and Algorithm 1 outlines the procedure for naive inference. In order to avoid notational clutter, the MRF energy is expressed in terms of region labels Y_z in the ensuing discussions. We also note here that dynamic programming [20] is a technique that does not take decisions greedily, and waits for the entire data before inference. Figure 5 illustrates the method. Each node in the trellis at a particular frame z refers to a segmentation generated by varying parameters of the MRF using region search. The goal is to infer the most probable state sequence over the entire image stack using Parameter Search. We experimented with Dynamic Programming, but realized that the time taken for computing global solutions make it impractical for our application.

D. Parameter Search - Particle Filtering Inference

Particle Filtering [21] is a very effective state space filtering technique that approximates the true posterior by a sum of

Algorithm 1 Parameter Search: Naive Inference

Require: $I_{1:N}, Y_1$
  for $z = 2 : N$ **do**
    for $\alpha \in \Theta$ **do**
      $Y_z^\alpha = \underset{M_z^\alpha}{\text{argmin}} \; E^{deform}(M_z^\alpha | I_z, Y_{z-1}, \alpha)$, using Equation 14
    end for
    $\{\alpha_z, Y_z^*\} = \underset{Y_z^\alpha, \alpha \in \Theta}{\text{argmax}} \; P(I_z|Y_z^\alpha)P(Y_z^\alpha|Y_{z-1})$, using Equation 6 in Equation 2
  end for

Algorithm 2 Parameter Search: Particle Filter Inference. NOTE: α, w represent particles (MRF parameters) and their weights

Require: $I_{1:N}, Y_1, K, q$ (importance density)
  $\alpha_1^i \sim q(\alpha_z|\alpha_{z-1}) \quad w_1^i = 1/K, \; 1 \le i \le K$
  for $z = 2 : N$ **do**
    Resample: $\{w_{z-1}^i\}_{1 \le i \le K}$, $w_{z-1}^i = 1/K \; \forall i \in 1..K$
    for $i = 1 : K$ **do**
      $\alpha_z^i \sim q(\alpha_z|\alpha_{z-1})$
      $Y_z^{\alpha^i} = \underset{M_z}{\text{argmin}} \; E^{deform}(M_z|I_z, Y_{z-1}, \alpha_z^i)$, using Equation 14
      $w_z^i = w_{z-1}^i \frac{P(I_z|Y_z^{\alpha_z^i})P(Y_z^{\alpha_z^i}|Y_{z-1})}{q(\alpha_z^i|\alpha_{z-1})}$, using Equation 6 in Equation 2
    end for
    Normalize Weights $w_z^i = \frac{w_z^i}{\sum_{j=1}^K w_z^j}, 1 \le i \le K$
    $\alpha_z^* = \sum_{i=1}^K w_z^i \alpha_z^i$
    $Y_z^* = \underset{Y_z}{\text{argmin}} \; E^{deform}(Y_z|I_z, Y_{z-1}, \alpha_z^*)$
  end for

weighted particles, see Algorithm 2. Particles in this case refer to parameter vectors of the random field. In many scenarios, it might not be feasible to enumerate the entire parameter space for search. In order to address this issue, the idea is to efficiently sample the parameter space for candidate particles (parameter vectors) that maximize the posterior. Assuming $q(\alpha_z|\alpha_{z-1})$ to be the system model governing how parameters of the MRF change over frames, the goal is to infer hidden parameters of the MRF α_z. The basic idea behind particle filtering approach is to draw from a posterior distribution represented as a set of particles α_z^i with corresponding weights w^i where $1 \le i \le K$, where K is the number of particles. The particles are sampled from $q(\alpha_z|\alpha_{z-1}) = \mathcal{N}(\alpha_z; \alpha_{z-1}, \sigma_p)$, where σ_p is usually fixed.

The particle filtering approach is favorable compared to greedy methods since they search the state space efficiently and yield high accuracies with significantly lower computational cost. Our experimental findings confirmed this with a particle filter operating using 30 particles outperforming a fixed size greedy technique on the F-Measure for evaluating tracing quality.

IV. Region Search - Formulation

The region search procedure comprises of a tracing algorithm capable of generating good candidate traces for a specific application. For the applications considered in this work (EM tracing, cardiac and human tracking) we are not aware of an off the shelf MRF tracer that produces good candidate solutions. As a result, this section augments a generic MRF segmenter with spatial and deformation priors to generate good candidate traces that can be used by parameter search. Region search can also be thought of as a black-box tracer, a name attributed to the fact that the parameter search module does not have any information on the internal workings of the tracer, except having access to its free parameters.

The basic idea behind the region search is to segment the object in frame z under the assumption that it must be in the vicinity of its estimated location in frame $z-1$. This requires the introduction of a spatial variable (pixel coordinate) exercising control over the labeling configuration. The following discussion derives a probabilistic interpretation of introducing a spatial localization variable into the likelihood construction. In the context of tracing, it can be looked at as a localized shape estimate constraining excessive deformations of the object from one frame to another. Please note that X_z (feature descriptor constructed from I_z) will be used in place of I_z in this section to stay consistent with MRF terminology. The posterior distribution of the random field can be decomposed as follows:

$$\begin{aligned} P(Y_z|X_z, Y_{z-1}, \alpha_z) &= \frac{P(Y_{z-1}, X_z, \alpha_z|Y_z)P(Y_z)}{P(Y_{z-1}, X_z, \alpha_z)} \\ &\propto \prod_p P(y_{p,z-1}, x_{p,z}|y_{p,z}) \\ &\propto \prod_p P(x_{p,z}|y_{p,z-1}, y_{p,z})P(y_{p,z-1}|y_{p,z}) \end{aligned} \tag{8}$$

Ignoring constants, the maximization problem of the random field now reduces to $Y_z^* = \underset{y_z}{\text{argmax}} P(Y_z|X_z, Y_{z-1})$ thus yielding a maximum likelihood estimate given by:

$$Y_z^* = \underset{Y_z}{\text{argmax}} \prod_p P(x_{p,z}|y_{p,z-1}, y_{p,z})P(y_{p,z-1}|y_{p,z}) \tag{9}$$

We now have to model $P(y_{p,z-1}|y_{p,z})$ for spatial localization to take effect, and we accomplish the same using dense optical flow fields.

The proposed formulation achieves accurate spatial localization by introducing dense flow fields that accurately capture object deformations from one frame to another. Denoting the dense optical flow field between the two frames (I_{z-1} and I_z) by $\mathbf{u}(I_z, I_{z-1})$, it can be shown [22] that the motion residual or shape deformation reduces to a dot product between the gradient of a signed distance function ϕ_{z-1} of Y_{z-1} with the flow field $\mathbf{u}$, and is given by $e_z = \mathbf{u}^T(I_z, I_{z-1})\nabla\phi_{z-1}$. ϕ_{z-1} is a smooth function evaluating the shortest path magnitude from any pixel p to the region boundary.

Motion Residual: Formally, motion residual is derived from optical flow fields, and can be defined as the vector field that displaces a target contour in frame $z-1$ to its current form in frame z. In other words, it is a function that transforms a contour in frame $z-1$ to frame z by utilizing information from images I_{z-1} and I_z. Considering the estimate a region in the current frame z to be a signed distance function $\hat{\phi}_z$, it can be proved using a brightness constancy constraint on the signed distance functions (omitting higher order terms in the resulting Taylor expansion) that $\hat{\phi}_z = \phi_{z-1} - e_z$.

We now propose the following form for $P(y_{p,z-1}|y_{p,z})$, utilizing estimates obtained from dense flow fields. The following model biases likelihood potentials to assume shapes that resemble previous segmentations. Intuitively, Equation 9 aims to localize the search for segmentation in slice z in the vicinity of segmentation in $z-1$. For this purpose, a signed distance function ϕ_z assigns a very low probability to foreground if a pixels distance from contour in $z-1$ is large, while assigning a high probability score to foreground to pixels close to contour in $z-1$, and vice-versa for the background.

It is useful to note here that $\hat{\phi}_z = \phi_{z-1} - e_z$ is derived from the motion residual e_z, thus making the spatial localization deformation aware:

$$P(y_{p,z-1}|y_{p,z}) = \exp\left(-\frac{\hat{\phi}_{p,z}H(\hat{\phi}_{p,z})}{\sigma_s}\right)y_{p,z} + \left(1-\exp\left(-\frac{\hat{\phi}_{p,z}H(\hat{\phi}_{p,z})}{\sigma_s}\right)\right)(1-y_{p,z}) \quad (10)$$

where σ_s refers to a smoothing parameter for the region prediction in frame z and H is a Heaviside function defined by,

$$H(x) = \begin{cases} 0 & \text{if } x < 0 \\ 1 & \text{if } x \geq 0 \end{cases} \quad (11)$$

The unary potentials in terms of the MRF cost function are now changed to,

$$\begin{aligned} \sum_p V_p^{\text{deform}}(y_p) &= -\log\prod_p P(x_{p,z}, y_{p,z-1}|y_{p,z}) \\ &= -\log\prod_p P(x_{p,z}|y_{p,z}, y_{p,z-1}) \quad (12) \\ &\quad -\log\prod_p P(y_{p,z-1}|y_{p,z}) \\ &= \sum_p \underbrace{V_p^{\text{data}}(y_{p,z})}_{\text{Data Likelihood}} + \underbrace{V_p^{\text{spatial}}(y_{p,z})}_{\text{Displacement Cue}} \quad (13) \end{aligned}$$

The likelihood probability can be written as $P(x_{p,z}, y_{p,z-1}|y_{p,z})$. This contains a term of data conditioned on labels in current and previous frames. By chain rule of probability the likelihood can be decoupled into a purely appearance based term $P(x_{p,z}|y_{p,z}, y_{p,z-1})$, and one term based on spatial localization $P(y_{p,z-1}|y_{p,z})$. In other words, the appearance term models how an object looks like while the spatial localization term focuses on probable pixel locations of the object of interest. The augmented tracer with deformation guidance is characterized by the energy function defined by:

$$E^{\text{deform}}(y_z) = \sum_{p\in\mathcal{P}} V_p^{\text{deform}}(y_{p,z}) + \sum_{p\in\mathcal{P},q\in\mathcal{N}_p} V_{pq}(y_{p,z}, y_{q,z}), \quad (14)$$

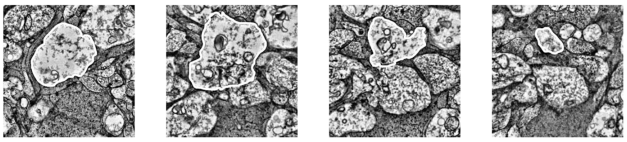


Fig. 6. Results employing the proposed region search procedure. Parameters are not inferred online to demonstrate the stand alone applicability of the locality sensitive displacement prior. Notice the ability of the algorithm to latch on to rapid object deformations and appearance variability.

Figure 6 illustrates the working of the region search procedure outlined in Algorithm 3. As can be seen, the technique does a fair job in keeping track of the object as once moves through the z-direction. Now that the base tracing scheme has been augmented with a strong spatial localizer providing deformation guidance, we turn our attention to introducing topological priors using parameter search.

V. Experimental Validation

This section initially presents experimental results on (V-A) Electron Micrograph data. Subsequently, experiments on endocardium tracking and surveillance are presented (V-B).

A. Experimental Results on Electron Micrograph Tracing

There has been a lot of interest in Electron Micrograph segmentation and tracing [5], [10], [23]–[26]. The primary motivation behind the proposed technique is the 3D tracing problem in retinal connectome data [27], a problem (data) not solved by any of the above referred techniques. The data in consideration comprises of rich textures and edge profiles that need to be jointly modeled, along with the need for handling abrupt deformations and topological changes. This section clearly demonstrates advantages of using the proposed model using extensive qualitative and quantitative validation.

The electron micrograph data contains distractions that could be reduced using established algorithms. We employ adaptive histogram equalization as a preprocessing step. Considering x_p^z to be a feature response at pixel p comprising smoothed intensity values after Gaussian filtering with multiple scales, the unary potential is formed as, $V_p(y_p^z = 1) = -log(p(x_p^z|\mathbf{FG^{z-1}}))$, $V_p(y_p^z = 0) = -log(p(x_p^z|\mathbf{BG^{z-1}}))$. Here **FG** and **BG** refer to the set of foreground and background pixels respectively segmented from the previous frame. This captures the notion of multi scale neighborhood averages with a total of three scales ($S = 3$) employed for the Gaussian kernel, while the costs were evaluated using standard histogram techniques. In the case of surveillance feeds, color histograms in RGB space were utilized to construct foreground and background potentials.

A total of 9 image stacks were used in quantitative validation. Seven of these image stacks (6 Electron Micrograph

Algorithm 3 Tracking using Region Search

Require: I_z, I_{z-1}, y_{z-1}

1: **for** z = 2:N **do**
2: $e_m = \mathbf{u}^T(I_z, I_{z-1})\nabla\phi_{z-1}$
3: $\hat{\phi}_z = \phi_{z-1} - e_z$
4: Compute V_p as discussed
5: $y_z^* = \underset{y_z}{\text{argmin}} \; E^{\text{deform}}(y_z)$ from Equation 14
6: **end for**

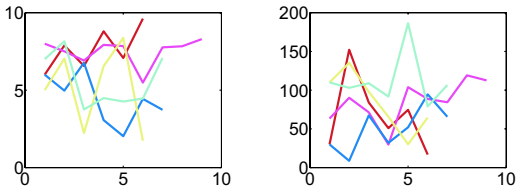


Fig. 7. (X-axis: Frame Number, Y-axis: Parameter Value) Observe parameter variations (each curve corresponds to a different stack) resulting from likelihood maximization. (Left) Contrast parameter variations over z for electron micrograph stacks, (Right) Interaction weight variations across the stack. As is evident a single fixed parameter, even if learnt offline would not consistently have high scoring segmentation transitions. The wide variations in the parameter vector justifies need for the proposed method.

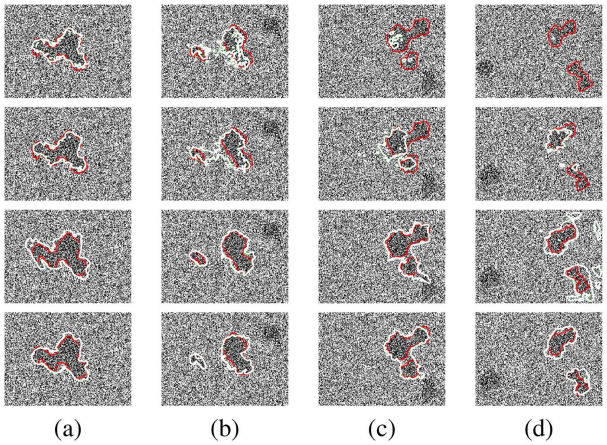


Fig. 8. (Best Viewed in Color) Row1 (Level Set Tracker with low regularization), Row2 (Level Set Tracker with high regularization), Row3 (Region Search), Row4 (Topology Aware MRF) show results sampled from different frames in a synthetic data. Since the top three rows do not have topological priors, they are allowed to undergo arbitrary topological changes. The last row shows the applicability of the topology aware MRF having the best quantitative performance among all methods. In the above figures, red contours are ground truth, while white contours are generated by the respective algorithms.

+ 1 Synthetic) had a splits or merges, which are much more challenging to detect than simple expansion/contraction. The other two image sequences are from cardiac data. On an average the F-measures of proposed technique is greater by more than 10 percent by nearest competing method. Further, in spite of existing metrics not explicitly accounting for split/merge error the proposed method still has a significant 10 percent lead. Penalizing split/merge errors would still lead to much lower performance on prior techniques.

Figure 7 demonstrates the variations in parameter values while segmenting different stacks in the connectome. The wide parameter variations in maximizing the segmentation transition score provides intuitive justification for flexibly tuning parameters across the stack. In order to ensure a fair comparison, the proposed technique is employed to generate a list of optimal parameter values, and the average of optimal parameters are used for the corresponding region search tracing procedure. This would be the best case scenario in comparison to hand tuned or random parameter initialization in the first frame.

A proof of concept for the proposed approach is illustrated using synthetic image sequences generated by hand to mimic topological changes typical in real data. As shown in Figure 8, a random shape having a black color is corrupted throughout the domain using random Gaussian noise of high variance. This makes sure that the data term (spatial homogeneity) of the foreground object is non uniform justifying the use of MRFs. Figure 8 illustrates a comparative performance between a traditional level set tracker and the proposed locality constrained and topology aware tracing. As can be observed from the figure, the topology aware model consistently outperforms the other models. The level set tracker implicitly handles topological changes, but begins splitting into many smaller regions to explain the data better. The first row shows results of level sets with a small regularization parameter, while the second row demonstrates results with higher regularization. Results from applying the standalone region search tracer are shown in the third row, where the region again begins splitting towards the end of the stack (frame c-d). In comparison, the proposed model in the fourth row performs the best in adhering to required topological dynamics.

For performance analysis, we consider the F-measure of the segmentations provided by a level set trackers and proposed MRF models in Figure 10. Note that the F-measure is defined by $F = \frac{2PR}{P+R}$, where precision (P) is the ratio of true positives to the sum of true positives and false positives, while recall (R) is the ratio between true positives and sum of true positives and false negatives.

Sample Electron Micrograph images to be segmented are is shown in Figure 9. The main difficulties in segmentation of Electron Micrograph stacks are massive deformations, abrupt topological changes and extremely noisy textures. It is assumed that the object location is provided to the algorithm in the first frame of a stack to initialize tracing. An important clarification pertains to the usage of frame-frame versus true 3D segmentation. Usually 3D segmentation models have high memory complexity and require a homogenous data term through the volume. This is not guaranteed in the application of interest due to severe noise in the foreground data term.

Figure 9 illustrates the working of the entire model on complex topological changes. Observe massive object elongation that the algorithm is able to trace with the region search tracer and subsequently utilize the segmentation transition scores to detect topological changes and trace each region over depth. Figure 10 reports evaluation of proposed scheme in comparison to traditional methods (including traditional MRF cost and level set based trackers) on synthetic data and sampled stacks from the connectome. Significant improvements in the

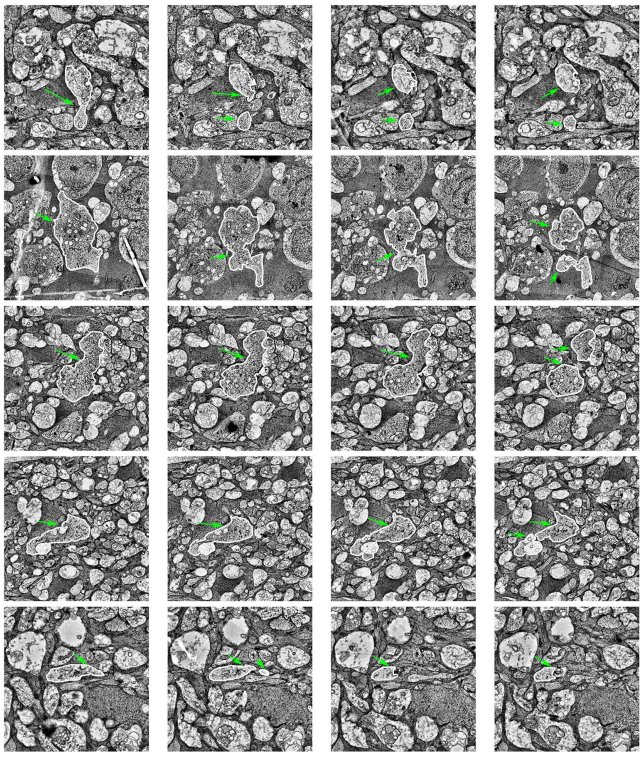


Fig. 9. (Best Viewed in Color) Results indicating the applicability of the full model (region/parameter search). Each row in the above figure illustrates performance of the algorithm on regions sampled from different parts of the connectome volume. Observe the method's ability to latch on to topological events including significant object deformations and splitting.

F-measures in comparison to state of the art justify the need for the proposed technique.

B. Experimental Results on Endocardium Tracking and Surveillance

Endocardium Tracking: The proposed method reliably (Figure 11) tracks the deforming surface of the heart throughout the image sequence. Non-adaptive MRFs perform well when the object sizes are fairly constant. For example, a particular parameter setting might work when the endocardium relaxes (has large area) and might fail during contraction (has small area). Automatic parameter tuning helps overcome this drawback by automatically adapting parameters across the image sequence. Ground truth data that comes with the dataset is used to validate performance of the algorithm. Table I and Table II illustrate comparative performance of the non-adaptive MRF model with the proposed approach. The proposed approach yields very promising results in comparison to the non-adaptive MRF which at times completely loses the target, see Table II where the non-adaptive model loses the difficult target after frame 6.

We would like to point out that the proposed method performs better on the average with the easy target too. The proposed technique searches over a very large parameter space, as opposed to conventional schemes that have fixed parameters. In some very rare scenarios, it is possible that the proposed approach might get stuck in a locally optimal solution, while a global solution corresponds to the fixed parameter

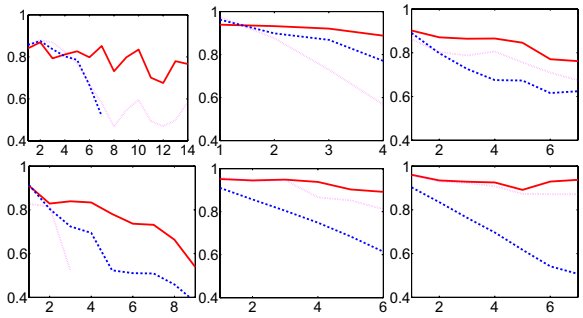


Fig. 10. (X-axis: Frame Number, Y-axis: F Measure), (Top Left) Performance of proposed technique (red-solid) on the noisy synthetic dataset (Figure 8) is compared to a level set tracker with high (magenta-dotted) and low (blue-dashed) regularization parameters. (Top Right - Bottom) F-measure plots comparing proposed topology aware model (red-solid) with the standalone region search tracer (magenta-dotted) and level set tracker (blue-dashed). As can be observed, the topology aware model consistently outperforms the stand alone region search tracer and level set based trackers.

TABLE I

EASY TARGET ON CARDIAC SEQUENCE: THE PROPOSED TECHNIQUE PERFORMS BETTER THAN TRADITIONAL TECHNIQUES THUS YIELDING MORE ACCURATE SEGMENTATIONS RESEMBLING THE GROUND TRUTH

Frame Number	3	6	9	12	15	18
Traditional	.82	.75	.74	**.72**	**.84**	.92
Proposed	**.92**	**.82**	**.76**	.68	.83	**.93**

TABLE II

DIFFICULT TARGET ON CARDIAC SEQUENCE: THE PROPOSED TECHNIQUE COMPREHENSIVELY BEATS TRADITIONAL APPROACHES SINCE TRADITIONAL APPROACHES LOSE DIFFICULT TARGETS VERY EARLY ON AND CANNOT RECOVER

Frame Number	3	6	9	12	15	18
Traditional	.90	.65	lost	lost	lost	lost
Proposed	**.90**	**.91**	**.82**	**.75**	**.87**	**.90**

setting held by conventional methods. Frames 3 and 4 for the easy target in Table II are indicative of the above phenomenon.

Surveillance: Tracking people in surveillance feeds is a challenging problem due to illumination artifacts, occlusion, shadows and clutter. Smoothness priors across time can be efficiently utilized by the proposed framework to track targets reliably across video feeds. Figure 12, 13 illustrates the applicability of the proposed model to accurately track and resist occlusions in challenging outdoor scenarios, along with the drawbacks of traditional methods. It is useful to note here that traditional background subtraction or optic flow based trackers would fail due to contour leakage caused by shadows/illumination artifacts and packet drops in video transmission.

Failure Scenarios: The algorithm is designed to tolerate minor registration errors due to the computation of dense flow fields. However, this relies on the manner in which optical flow is computed, and it may not always be possible for the algorithm to directly control. In our implementation, dynamic programming does not seem to provide a significant performance gain, though it is theoretically more sound than the greedy method. This issue warrants further investigation.

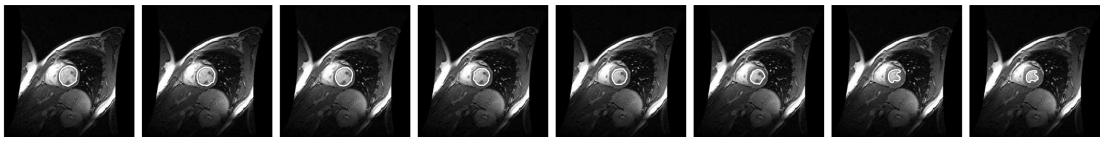


Fig. 11. (Best Viewed in Color) Performance of proposed technique on time varying cardiac data. The primary challenge is in tracking a deforming surface where the topology constraint is that an endocardium cannot split into smaller structures at any point in time. During the process of maximizing the segmentation transition scores, segmentation labels that do not conform to a single connected component are eliminated as infeasible.

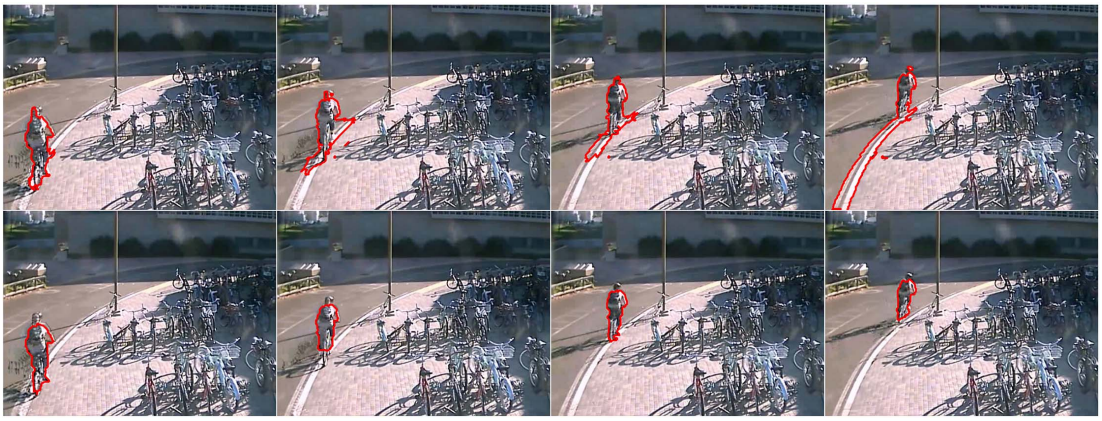


Fig. 12. (Best Viewed in Color) Applicability of the proposed method in outdoor surveillance. The top row illustrates application of the traditional approach without parameter adaptation. As can be observed, a contour leak in one frame keeps propagating as tracking progresses. On the other hand, the adaptive model (bottom row) ensures a smooth segmentation transitions in the time direction thus exploiting obvious priors available in tracking scenarios.

Fig. 13. The above figure illustrates ability of the proposed model to resist occlusions that commonly occur in cluttered scenes.

VI. Conclusion

In summary, this paper has presented an algorithm to make tracing algorithms topology aware. The region search tracer constrains region displacement across the z direction, while the parameter search procedure embeds learnt topology priors. The algorithm does not resort to complex graph constructions and exploits the inherent free parameters of traditional MRFs for stack segmentation using the region search tracer. The highlight of the proposed technique lies in using free parameters (which are otherwise hand tuned) for detecting topological changes using an efficient search procedure. Promising experimental results were presented on ssTEM stacks, cardiac image sequences and surveillance video. Utilizing higher order potentials [28] is part of future work. Higher order models aid in modeling higher order spatial interactions and can be used for embedding topological priors, an aspect that has been unexplored in the literature. Finally, extending the proposed approach to other microscopic imaging modalities like confocal/light microscopy is another promising future direction.

Acknowledgment

The authors would like to thank the anonymous reviewers for their valuable feedback.

References

[1] V. Jagadeesh, B. Manjunath, J. Anderson, B. Jones, R. Marc, and S. K. Fisher, "Online parameter estimation in dynamic Markov random fields for image sequence analysis," in *Proc. 19th IEEE ICIP*, Mar. 2012, pp. 301–304.

[2] E. Jurrus, R. Whitaker, B. Jones, R. Marc, and T. Tasdizen, "An optimal-path approach for neural circuit reconstruction," in *Proc. 5th IEEE ISBI*, May 2008, pp. 1609–1612.

[3] L. Vincent and P. Soille, "Watersheds in digital spaces: An efficient algorithm based on immersion simulations," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 13, no. 6, pp. 583–598, Jun. 1991.

[4] C. Couprie, L. Grady, L. Najman, and H. Talbot, "Power watersheds: A new image segmentation framework extending graph cuts, random walker and optimal spanning forest," in *Proc. IEEE 12th Int. Conf. Comput. Vis.*, Sep. 2009, pp. 731–738.

[5] V. Jagadeesh, M.-C. Shih, B. Manjunath, and K. Rose, "Scalable tracing of electron micrographs by fusing top down and bottom up cues using hypergraph diffusion," in *Proc. MICCAI*, Oct. 2012, pp. 321–328.

[6] M. Kass, A. Witkin, and D. Terzopoulos, "Snakes: Active contour models," *Int. J. Comput. Vis.*, vol. 1, no. 4, pp. 321–331, 1988.

[7] C. Xu and J. L. Prince, "Snakes, shapes, and gradient vector flow," *IEEE Trans. Image Process.*, vol. 7, no. 3, pp. 359–369, Mar. 1998.

[8] T. F. Chan and L. A. Vese, "Active contours without edges," *IEEE Trans. Image Process.*, vol. 10, no. 2, pp. 266–277, Feb. 2001.

[9] Y. Y. Boykov and M.-P. Jolly, "Interactive graph cuts for optimal boundary & region segmentation of objects in N–D images," in *Proc. 8th IEEE ICCV*, vol. 1. Jul. 2001, pp. 105–112.

[10] V. Jagadeesh, N. Vu, and B. Manjunath, "Multiple structure tracing in 3D electron micrographs," in *Proc. MICCAI*, Sep. 2011, pp. 613–620.

[11] Y. Boykov and V. Kolmogorov, "An experimental comparison of min-cut/max-flow algorithms for energy minimization in vision," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 26, no. 9, pp. 1124–1137, Sep. 2004.
[12] N. Vu and B. Manjunath, "Shape prior segmentation of multiple objects with graph cuts," in *Proc. IEEE Comput. Soc. Conf. CVPR*, May 2008, pp. 1–8.
[13] J. Winn and J. Shotton, "The layout consistent random field for recognizing and segmenting partially occluded objects," in *Proc. IEEE Comput. Soc. Conf. CVPR*, vol. 1. Apr. 2006, pp. 37–44.
[14] J. Malcolm, Y. Rathi, and A. Tannenbaum, "Multi-object tracking through clutter using graph cuts," in *Proc. IEEE 11th ICCV*, Oct. 2007, pp. 1–5.
[15] Y. Zeng, D. Samaras, W. Chen, and Q. Peng, "Topology cuts: A novel min-cut/max-flow algorithm for topology preserving segmentation in N–D images," *Comput. Vis. Image Understand.*, vol. 112, no. 1, pp. 81–90, Oct. 2008.
[16] S. Nowozin and C. H. Lampert, "Global interactions in random field models: A potential function ensuring connectedness," *SIAM J. Imaging Sci.*, vol. 3, no. 4, pp. 1048–1074, Dec. 2010.
[17] S. Vicente, V. Kolmogorov, and C. Rother, "Graph cut based image segmentation with connectivity priors," in *Proc. IEEE Conf. CVPR*, Jun. 2008, pp. 1–8.
[18] D. Cremers, "Dynamical statistical shape priors for level set-based tracking," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 28, no. 8, pp. 1262–1273, Aug. 2006.
[19] B. Peng and O. Veksler, "Parameter selection for graph cut based image segmentation." in *Proc. Brit. Mach. Vis. Conf.*, 2008, pp. 1–10.
[20] L. Rabiner and B. Juang, "A tutorial on hidden Markov models," *Proc. IEEE*, vol. 77, no. 2, pp. 257–286, Feb. 1989.
[21] M. S. Arulampalam, S. Maskell, N. Gordon, and T. Clapp, "A tutorial on particle filters for online nonlinear/non-Gaussian Bayesian tracking," *IEEE Trans. Signal Process.*, vol. 50, no. 2, pp. 174–188, Feb. 2002.
[22] P. Ghosh, M. E. Sargin, and B. S. Manjunath, "Robust dynamical model for simultaneous registration and segmentation in a variational framework: A Bayesian approach," in *Proc. IEEE 12th ICCV*, Sep. 2009, pp. 709–716.
[23] D. B. Chklovskii, S. Vitaladevuni, and L. K. Scheffer, "Semi-automated reconstruction of neural circuits using electron microscopy," *Current Opinion Neurobiol.*, vol. 20, no. 5, pp. 667–675, Oct. 2010.
[24] E. Jurrus, M. Hardy, T. Tasdizen, P. T. Fletcher, P. Koshevoy, C.-B. Chien, W. Denk, and R. Whitaker, "Axon tracking in serial block-face scanning electron microscopy," *Med. Image Anal.*, vol. 13, no. 1, pp. 180–188, 2009.
[25] B. Andres, U. Köthe, M. Helmstaedter, W. Denk, and F. A. Hamprecht, "Segmentation of SBFSEM volume data of neural tissue by hierarchical classification," in *Pattern Recognition*. New York, NY, USA: Springer-Verlag, Jun. 2008, pp. 142–152.
[26] A. Lucchi, K. Smith, R. Achanta, G. Knott, and P. Fua, "Supervoxel-based segmentation of mitochondria in EM image stacks with learned shape features," *IEEE Trans. Med. Imaging*, vol. 31, no. 2, pp. 474–486, Feb. 2012.
[27] J. Anderson, S. Mohammed, B. Grimm, B. Jones, P. Koshevoy, T. Tasdizen, R. Whitaker, and R. Marc, "The Viking viewer for connectomics: Scalable multi-user annotation and summarization of large volume data sets," *J. Microscopy*, vol. 241, no. 1, pp. 13–28, Jan. 2011.
[28] P. Kohli, L. Ladickỳ, and P. Torr, "Robust higher order potentials for enforcing label consistency," *Int. J. Comput. Vis.*, vol. 82, no. 3, pp. 302–324, Mar. 2009.

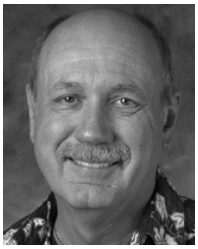
Vignesh Jagadeesh received the Ph.D. and M.S. degrees in electrical and computer engineering from the University of California, Santa Barbara, CA, USA in 2013 and 2009, respectively. His research goals are centered around developing reliable computer vision systems capable of wide deployment. He has worked with the vision group at True Vision 3-D surgical on surgical tracking/augmentation, with the team at Mayachitra, Inc., on image/video segmentation, and with eBay Research on visual recommender systems. His research interests include image segmentation, tracking, object detection, recognition, and large scale image analysis. He is currently a Research Scientist with eBay Research.

Bangalore S. Manjunath (F'05) received the M.S. degree in systems science and automation from the Indian Institute of Science, India, and the Ph.D. degree in electrical engineering from the University of Southern California, Los Angeles, CA, USA. He is currently a Professor of electrical and computer engineering with the University of California, Santa Barbara, CA, USA. His research interests include image processing and computer vision. He has published more than 250 peer reviewed papers on image segmentation, registration, texture analysis, content-based image and video retrieval, image steganography and forensics, camera networks, and bio-image analysis. He is a coinventor on 24 U.S./international patents. He was an Associate Editor of the IEEE TRANSACTIONS ON PATTERN ANALYSIS AND MACHINE INTELLIGENCE.

James Anderson received the Ph.D. degree from the University of Utah, Salt Lake City, UT, USA, in 2010, on the Viking viewer for scalable annotation and browsing of connectomes. He is currently a Researcher with the Moran Eye Center, University of Utah. He was a Software Engineer with Microsoft.

Bryan W. Jones received the Ph.D. degree in neurophysiology from the University of Utah, Salt Lake City, UT, USA, and was a Post-Doctoral Fellow with the Moran Eye Center. He is currently a Research Assistant Professor in ophthalmology with the University of Utah. His research interests include retinal connectomics and computational molecular phenotyping. He is also the Webmaster of Webvision.

Robert Marc received the Ph.D. degree with the University of Texas at Houston, Houston, TX, USA. He was an NIH Post-Doctoral Fellow with the University of California, Los Angeles, CA, USA. He is currently a Distinguished Professor of ophthalmology with the University of Utah, Salt Lake City, UT, USA, School of Medicine, John A. Moran Eye Center. His global research objectives are building retinal and brain connectomes - 10-100 terabyte-scale maps of neural networks, creating new imaging strategies and molecular probes to reveal cellular physiological states, and proofing optogenetics therapies for blinding diseases.



Steven K. Fisher received the Ph.D. degree from Purdue University, West Lafayette, IN, USA. He was a Post-Doctoral Fellow with Johns Hopkins University, California, MD, USA. He is currently a Professor Emeritus of Molecular, Cellular, and Developmental Biology, and Research Professor of neuroscience with the University of California, Santa Barbara (UCSB), CA, USA. He served as the Director of the Neuroscience Research Institute, UCSB, from 1984 to 2001. His research group is dedicated to studying the cell biology of vertebrate retina using a variety of cell and molecular techniques but with an emphasis on biological imaging.