ONLINE PARAMETER ESTIMATION IN DYNAMIC MARKOV RANDOM FIELDS FOR IMAGE SEQUENCE ANALYSIS

Vignesh Jagadeesh[1], *B.S Manjunath*[1], *James Anderson*[2], *Bryan Jones*[2], *Robert Marc*[2], *Steven K. Fisher*[1]

[1]Center for Bioimage Informatics, UCSB [2] John A. Moran Eye Institute, Univ of Utah

ABSTRACT

Markov Random Fields (MRF) have proven to be extremely useful models for efficient and accurate image segmentation.Recent literature points to an increased effort towards incorporating useful priors (shape, geometry, context) in a MRF framework. However, topological priors, considered extremely crucial in biological and natural image sequences have been less explored. This work proposes a strategy wherein free parameters of the MRF are used to make it topology aware using a semantic graphical model working in conjunction with the MRF. Estimation of free parameters is constrained by prior knowledge of an object's topological dynamics encoded by the graphical model. Maximizing a regional conformance measure yields parameters for the frame under consideration. The application motivating this work is the tracing of neuronal structures across 3D serial section Transmission Electron Micrograph (ssTEM) stacks. Applicability of the proposed method is demonstrated by tracing 3D structures in ssTEM stacks.

Index Terms— MRF, Parameter Estimation, Image Sequence Analysis

1. INTRODUCTION

Markov Random Field (MRF) models have found wide applicability in image analysis due to their ability to fuse prior knowledge with the observed data efficiently and accurately. Classical formulations predominantly restrict themselves to smoothness priors driven by intensity (color, texture) likelihoods. Recently, there has been effort focussed on embedding contextual and geometric priors into random fields with impressive results. This paper explores a novel topology prior, an area that is gaining renewed interest in the MRF literature. *In particular this work concerns itself with a setting where image characteristics change with respect to some parameter.* In case of tracking applications, the aforementioned parameter is time, while it is the z-direction for image stacks. In other words, given N images, the problem is to localize the object in each image by handling prior knowledge of topological changes an object may undergo. The application motivating this work is the tracing of 3D structures from serial section Transmission Electron Micrograph (ssTEM) stacks obtained from a retinal connectome [1]. The requirement of automated or even human assisted semi-automated tracing methods for connectomes cannot be overemphasized. The single connectome under consideration could take biologists a year or so to fully annotate. The challenges involved in segmenting ssTEM stacks are coping with excessive deformations across depth, capturing topological changes (shrinkage, expansion, merging and splitting) that may occur at any depth, and building appearance models from noisy textures. Figure 1 illustrates the ability of the proposed topology aware MRF to capture massive object deformations and topological changes (contour splitting and merging). 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.

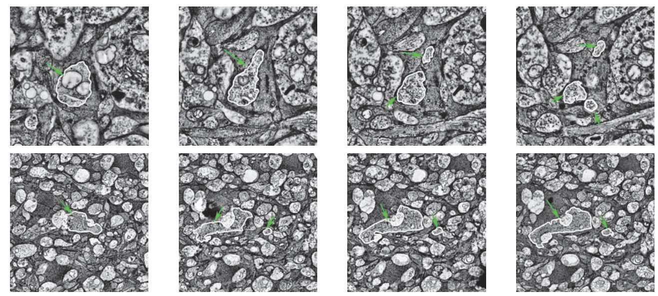

Fig. 1. (Best Viewed in Color) The above figure shows cross-sectional images sampled from a retinal connectome acquired using scanning section transmission electron microscopy. The images are 2D crosssections of neuronal structures present in the three dimensional volume. As can be seen, manual markup of these structures is challenging considering the number of structures to be traced (volume of data). The image analysis problem is to reconstruct these neuronal structures in three dimensions, given an initial contour in the first slice. Some critical challenges that need to be overcome are capturing abrupt contour deformations across the z-direction, and detection of topological changes such as splitting and merging of contours. The first row in the above figure shows an example of a contour initially splitting into two contours (frame 2-3) and further splitting into three contours (frame 3-4). The second row illustrates an example of a contour splitting on the right (frame 1-2) and merging on the left (frame 2-3). Developing a technique that overcomes the challenges stated above and achieves results as illustrated is central to interpreting electron micrograph data. The contours shown are not annotated by a human, but automatically generated by the technique proposed in this work.

The basic idea behind the proposed technique is to learn a prior model (topological dynamics) that auto-tunes MRF parameters as one moves through an image sequence. A little thought must convince the reader of advantages offered by auto-tuning parameters across the image stack. One could attempt to learn parameters using (state of the art) pseudolikelihood or max margin techniques. However, as the object's appearance and topology change from one frame to another, it is not obvious how one adopts a single learnt parameter vector (or its distribution) for the frame under consideration. The proposed method utilizes topological dynamics of the object to constrain parameter variations. This is implicitly achieved by using free parameters of the MRF to control topological dynamics of the object. This is an important difference in the context of image sequence analysis, since there are changes from one frame to another (illumination, occlusion, object deformations to name a few) that cannot be accounted for by a single parameter vector for the entire image sequence. Hence, it makes sense to learn a prior model lending flexibility to parameter variations rather than pick a parameter vector learnt offline. The aim of this paper is to present a method that is capable of wrapping around a non-parametric segmentation technique (eg: mrfs, level sets), thereby achieving two objectives: Firstly, embedding topological priors using free parameters of the algorithm. Secondly, auto-tuning the free parameters by topology control equations thus annulling the need for hand tuned parameters. The primary contributions of this work include:

- *A framework for online parameter estimation in Dynamic Markov Random Fields*
- *Enforcing prior knowledge of learnt topology to improve image sequence analysis*

In developing our algorithms (Section II) we emphasize generating solutions conforming to prior topology as opposed to global optimization. Hence, in certain cases (when feasible parameter set size is large), local solutions are accepted for tractability.

Related Works: A detailed treatment of MRFs can be found in Boykov et al. [2] and is not discussed here. Vu et al. [3] introduced shape, Winn et al. [4] introduced contextual priors in MRFs. Work dealing with topology models deals with constraining object topology [5], and does not learn a topological model over possible events. Further, the idea of modeling topological priors using free parameters of the dynamic MRF is inherently different from the previous formulations. The proposed technique encodes a topological prior (see Figure 3) into a non-parametric segmentation framework, in contrast to deformable shape prior segmenter such as the work of Cremers et al [6]. The primary motivation of both works are different, while a unification would intuitively lead to a stronger segmenter. Papers by [7, 8] illustrate different approaches for Electron Micrograph segmentation/tracing.

2. FORMULATION

Notations: The following discussions conform to, $I_{1:N}$: Set of N images comprising the sequence, z : Iterator that moves over depth/time, C_z^α : Segmentation partitioning slice z with parameter α, $y_{p,z}, x_{p,z}$: Label and Data at pixel p in slice z.

MRFs are models formulated to solve the image labeling problem. The aim is to label every pixel $p \in \mathcal{P}_z$ in an image with a label y_p from a label set $\mathcal{L} \in \{1, 2, ...|\mathcal{L}|\}$. Each pixel p resides in a planar graph and has data $x_{p,z}$ associated with it at slice z. 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:

$$E(y_z) = \sum_{p \in \mathcal{P}_z} V_p(y_{p,z}) + \sum_{p \in \mathcal{P}_z, q \in \mathcal{N}_p} V_{pq}(y_{p,z}, y_{q,z}) \quad (1)$$

Unary Potentials: $V_p(y_{p,z}) = -\log P(x_{p,z}, y_{p,z-1}|y_{p,z})$, the negative log likelihood function is commonly known as the unary or terminal cost. In order to spatially localize an object of interest in slice z, we propose the following form for the spatial localization prior, which acts as a rough shape prior. The following equation biases likelihood potentials (in slice z) to assume shapes that resemble previous segmentations(in slice $z-1$). In words, the farther a pixel is from a previous segmentation in $z-1$, the less likelihood it has of being foreground. $P(y_{p,z-1}|y_{p,z}) = \exp(-\hat{\phi}_{p,z} H(\hat{\phi}_{p,z})/\sigma_s) y_{p,z} + (1 - \exp(-\hat{\phi}_{p,z} H(\hat{\phi}_{p,z})/\sigma_s))(1 - y_{p,z})$, where σ_s is a smoothing parameter. $\hat{\phi}_{p,z}$ is the signed distance function of the segmentation result in $z-1$, and acts as a prior for segmenting slice z. Unary potentials are constructed from multiplying the spatial localization prior with pixel wise likelihoods. ***Interaction Potentials:*** are neighborhood potentials modeling pixel similarity $V_{pq}(y_p, y_q) = \lambda_I \exp(-||x_p - x_q||/\sigma_I^2)\delta(y_p \neq y_q)$. Inferring y_p from equation 1 requires its minimization, which is achieved by a mincut [9] on the constructed graph.

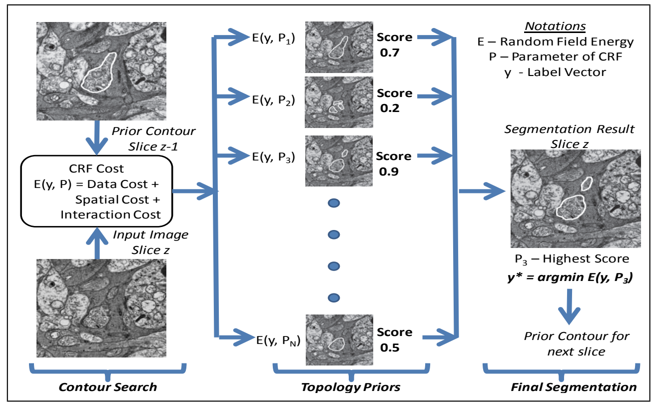

Fig. 2. Illustration of the algorithm flow for a single iteration of Topology aware MRFs. The prior contour is utilized to generate a prediction using spatially constrained graph cuts. The prediction is used to score (regional stability likelihood) multiple segmentations obtained by varying MRF parameters P. The parameter having maximum conformance to topology priors is the segmentation result for the current slice, and acts a prior for the next slice.

As mentioned earlier, the main problem being addressed in this work is to make the segmentation algorithm topologically aware of expansion/shrinkage and split/merge events. Figure 2 illustrates the proposed workflow. The novelty of the proposed formulation lies in the introduction of scoring functions for topological transitions using a regional stability likelihood. Initially, a spatially constrained graph cut (constraining the search space of a contour from one frame to another) aids in enforcing 3D smoothness. Subsequently, topological priors are incorporated by maximizing a likelihood function learnt from training data.

2.1. Topological Prior as a State Transition Model

We now introduce the notion of regional stability. Regional stability for purposes in this paper is the stability of a segmentation result to variations in free parameters of the algorithm. In the case of Markov random fields with clique size 2, regional stability can be determined by finding out the stability of segmentation by varying regularization (λ_I) and edge strength parameters (σ_I). 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 can be seen as a parameter controlling the relative importance of unary and interaction terms. In order to detect a split or merge event, a search is carried out in the parameter space θ_z

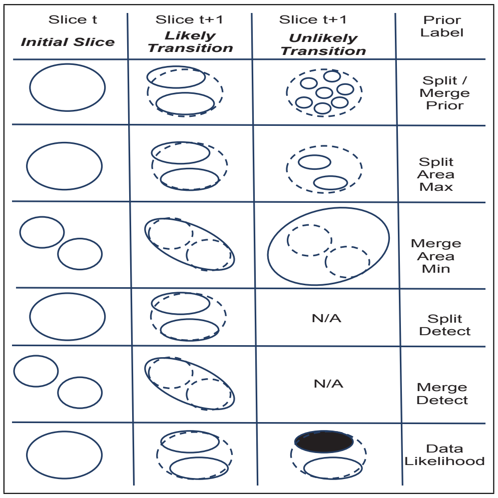

Fig. 3. Illustration of the prior information available pertaining to topological changes. The regional conformance likelihood is constructed based on the above illustration. See text for an explanation of the different rows in the above figure.

= $(\sigma_I, \lambda_I)_z$ to ascertain regional stability. Since the parameter variations control the regional stability of the segmentation outcome, the method to search over this 2D parameter space for stable regions in the vicinity of the previously segmented contour is referred to as region search. In the following, note a change in notation from $y_{p,z}$ to C_z. The higher level graphical model works on a more global interpretation (C_z) of the contour as a collection of pixels, while the MRF works at the level of pixels ($y_{p,z}$) by factorizing their probabilities as is evident from the previous sections. The regional stability likelihood R, used for constructing the higher level graphical model is defined as:

$$R = P(C_z, I_z|C_{z-1}) = \underbrace{P(I_z|C_z, C_{z-1})}_{Data\ Quality}\underbrace{P(C_z|C_{z-1})}_{Plausibility} \quad (2)$$

The data quality is measured by the histogram intersection between $f(I_z(C_z))$ and $f(I_{z-1}(C_{z-1}))$. Note that the histogram 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 available for the problem at hand are now formalized, (see Figure 3): *Split/Merge Prior*: Chance of contour splitting (merging) into (from) more than three sub-contours is extremely low. *Split Area Max*: If area of contour in z decreases in comparison to area in $z-1$ (split or shrinkage), area of overlap between contours in $z-1$ and z must be maximal. *Merge Area Min*: Chance of contour merging with another contour results in a contour with area around the sum of two parent contours. *Split Detect*: Splits occur when there are multiple overlapping connected components between frame $z-1$ and z. *Merge Detect*: Merges occur when there is a massive change in contour areas between frame $z-1$ and z. *Data Likelihood Agreement*: The agreement of data likelihoods between successive overlapping contours must be maximized. In short, a contour moves down a serial stack and can undergo any one of four topological events, namely shrinkage, expansion, split, and merge. A shrinkage(expansion), considered a regionally stable event is always assumed to decrease(increase) a contour's surface area from one frame to another. Further, a split(merge) is considered a regionally unstable event with prior constraints on the nature of split (merge). The events are mutually exclusive, meaning they cannot co-occur for a given contour. The state transition distribution modeling topological events is $P(C_z|C_{z-1})$. Consider C_{z-1} to be an estimate of the contour using an estimation procedure, and $\cup_{i=1}^{L} C_z(i)$ to be the set of L contours generated by a parameter setting of the segmentation algorithm. We define two important quantities,

- ***Relative Surface Area***(d): The ratio of contour areas from the estimated contour from slice $z-1$ and the L overlapping (across slices) contour(s) produced by the segmentation algorithm on slice z, $d = \frac{|C_{z-1}|}{\cup_{i=1}^{L}|C_z(i)|}$
- ***Region Stability***(r): Regional stability as one transitions from frame $z-1$ to the current frame z. The function is constructed so that if there is expansion or shrinkage (considered stable transitions since the connected component is preserved), r evaluates to a non-negative number, while it is negative for split or merge behavior. $r = -(\mathcal{I}_S \vee \mathcal{I}_M)$. 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.

$$\mathcal{I}_S = \begin{cases} 1, & 1 < L < L_{prior} \\ 0, & \text{otherwise} \end{cases} \quad \mathcal{I}_M = \begin{cases} 1, & d < 0.5 \\ 0, & \text{otherwise} \end{cases} \quad (3)$$

The probability of a topological change occurring, without any image dependent information is given by: $P(C_z|C_{z-1}) = \sum_{i=1}^{4} P(T = i)$. The contour 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 is given by: $P(T = 1) = \mathcal{N}(1 + \mu_1, \sigma_1)H(d-1)H(r)$, $P(T = 2) = \mathcal{N}(1-\mu_1, \sigma_1)H(1-d)H(r)$, $P(T = 3) = \mathcal{N}(\mu_2, \sigma_2)\mathcal{I}_S$, $P(T = 4) = \mathcal{N}(\mu_3, \sigma_3)\mathcal{I}_M$. In the above equation, H refers to the Heaviside function that evaluates to one if the argument is non-negative. μ_i, σ_i, where $1 \leq i \leq 3$ are parameters of a normal distribution learnt respectively for shrinkage/expansion, split and merge. L_{prior} is the maximum number of contours that can result from a split, as observed from the training data. Given the result from the MRF segmentation and a learnt

Algorithm 1 Topology Aware MRF: Particle Filter Inference. NOTE: α, w represent particles (MRF parameters) and their weights

Require: $I_{1:N}, C_1, K, q$ (importance density)
  $\alpha_1^i \sim q(\alpha_z|\alpha_{z-1}) \quad w_1^i = 1/K, 1 \leq i \leq K$
  for $z = 2 : N$ **do**
    Resample: $\{w_{z-1}^i\}_{1 \leq i \leq K}, w_{z-1}^i = 1/K \ \forall i \in 1..K$
    for $i = 1 : K$ **do**
      $\alpha_z^i \sim q(\alpha_z|\alpha_{z-1})$
      $C_z^{\alpha^i} = \underset{M_z}{\operatorname{argmin}} \quad E(M_z|I_z, C_{z-1}, \alpha_z^i)$
      $w_z^i = w_{z-1}^i \frac{P(I_z|C_z^{\alpha_z^i})P(C_z^{\alpha_z^i}|C_{z-1})}{q(\alpha_z^i|\alpha_{z-1})}$
    end for
    Normalize Weights $w_z^i = \frac{w_z^i}{\sum_{j=1}^{K} w_z^j}, 1 \leq i \leq K$
    $\alpha_z^* = \sum_{i=1}^{K} w_z^i \alpha_z^i$, $C_z^* = \underset{C_z}{\operatorname{argmin}} \quad E(C_z|I_z, C_{z-1}, \alpha_z^*)$
  end for

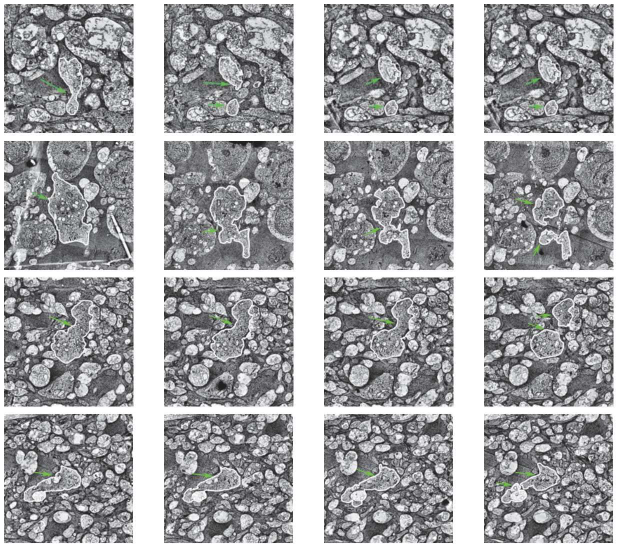

Fig. 4. (Best Viewed in Color) Results indicating the applicability of the full model latching on to topological events, including drastic deformations and splitting. Each row illustrates performance on different stacks.

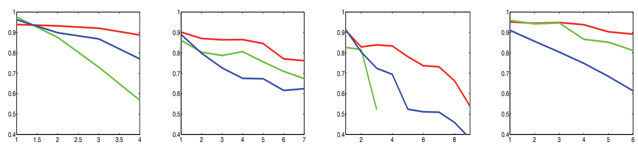

Fig. 5. (X-axis: Frame Number, Y-axis: F Measure), F-measure plots comparing proposed topology aware model (red) with traditional graph cuts alone (green) and level set tracker (blue). The topology aware model consistently outperforms traditional graph cuts and level set based trackers.

topology model, the challenge is to infer the optimal parameter α_z for every frame. The algorithm for inferring the joint model is presented in Algorithm 1, where a particle filter [10] estimates optimal parameter values using a set of K particles.

3. EXPERIMENTS

Experimental results are reported on 3D electron micrographs, where a subset of 3D stacks with annotations were used for learning parameters of the topology model. Subsequently, the learnt parameters were utilized for tracing, as shown in Figure 4. ***Preprocessing:*** The electron micrograph data is inherently noisy and may contain distractions. We employ the Contrast Limited Adaptive Histogram Equalization procedure for preprocessing. ***Likelihood Potentials:*** Electron Micrograph data is rich in texture, but not of the sort one would find in traditional texture analysis literature. It is used as a valuable cue by biologists, but would appear to be visually very *noisy* for an untrained person. We propose computing multi-scale local histograms from $I_z^s = I_z * g_{\sigma_i}, 1 \leq i \leq S$. The above equation refers to smoothing of image I_z at position z on the stack by a Gaussian kernel g_σ of standard deviation σ_i, where S is the total number of scales used for smoothing. Consider q_p^s to be a pixel response to a smoothing at scale s and let Q_p^z be the feature vector for pixel p. The unary potential functions are denoted denoted by,

$$\begin{aligned} V_p(y_p = 1) &= -log(P(Q_p^z|\mathbf{FG^{z-1}})) \\ V_p(y_p = 0) &= -log(P(Q_p^z|\mathbf{BG^{z-1}})) \end{aligned} \quad (4)$$

Here $\mathbf{FG}$ and $\mathbf{BG}$ refer to the set of foreground and background pixels respectively segmented from the previous frame. This feature vector captures the notion of multi scale neighborhood averages and concatenates the same to form a feature vector. A total of three scales ($S = 3$) were employed for the gaussian kernel, while the costs were evaluated using standard histogram techniques. Figure 4 illustrates the working of the entire model with the regional likelihood maximization, and more complex topological changes. Observe massive contour elongation along the first few slices that the algorithm is able to trace with the displacement prior and subsequently utilize the regional stability measure to detect topological changes and trace each contour over depth. Figure 5 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 variations in the magnitude of F-measures justifies the need for the proposed technique. In conclusion, this work presented an algorithm for utilizing learnt topology priors for enhancing performance on image sequence analysis task by fusing a top down graphical model, with a low level MRF with promising results. Future work includes scaling the proposed technique to simultaneously trace multiple structures.

Acknowledgments: This work was supported by a grant from the National Science Foundation, NSF OIA 0941717

4. REFERENCES

[1] J.R. Anderson, B.W. Jones, and et al, “A computational framework for ultrastructural mapping of neural circuitry,” *PLOS09*.

[2] Y.Y. Boykov and M.P. Jolly, “Interactive graph cuts for optimal boundary and region segmentation of objects in ND images,” in *ICCV 2001*.

[3] N. Vu and BS Manjunath, “Shape prior segmentation of multiple objects with graph cuts,” in *CVPR08*.

[4] J. Winn and J. Shotton, “The layout consistent random field for recognizing and segmenting partially occluded objects,” in *CVPR06*.

[5] Y. Zeng, D. Samaras, and et al, “Topology cuts: A novel min-cut/max-flow algorithm for topology preserving segmentation in ND images,” *CVIU08*.

[6] D. Cremers, “Dynamical statistical shape priors for level set-based tracking,” *PAMI06*.

[7] V. Jain, B. Bollmann, and et al, “Boundary learning by optimization with topological constraints,” in *CVPR10*.

[8] V Jagadeesh, N Vu, and B.S Manjunath, “Multiple structure tracing in 3d electron micrographs,” *MICCAI 2011*.

[9] Y. Boykov and V. Kolmogorov, “An experimental comparison of min-cut/max-flow algorithms for energy minimization in vision,” *PAMI04*.

[10] O. Cappé, S.J. Godsill, and E. Moulines, “An overview of existing methods and recent advances in sequential monte carlo,” *Proceedings of the IEEE07*.