Swansea University
Prifysgol Abertawe

Cronfa
Setting Research Free

Cronfa - Swansea University Open Access Repository

This is an author produced version of a paper published in:
International Journal for Numerical Methods in Biomedical Engineering

Cronfa URL for this paper:
http://cronfa.swan.ac.uk/Record/cronfa49788

Paper:
Essa, E., Jones, J. & Xie, X. (in press). Coupled s-Excess HMM For Vessel Border Tracking and Segmentation.
International Journal for Numerical Methods in Biomedical Engineering

---

This item is brought to you by Swansea University. Any person downloading material is agreeing to abide by the terms of the repository licence. Copies of full text items may be used or reproduced in any format or medium, without prior permission for personal research or study, educational or non-commercial purposes only. The copyright for any work remains with the original author unless otherwise specified. The full-text must not be sold in any format or medium without the formal permission of the copyright holder.

Permission for multiple reproductions should be obtained from the original author.

Authors are personally responsible for adhering to copyright and publisher restrictions when uploading content to the repository.

http://www.swansea.ac.uk/library/researchsupport/ris-support/

Coupled s-Excess HMM For Vessel Border Tracking and Segmentation

Ehab Essa[1,2*], Jonathan-Lee Jones[2], Xianghua Xie[2†]

[1]Department of Computer Science, Faculty of Computers and Information, Mansoura University, Egypt
[2]Department of Computer Science, Swansea University, UK
April 16, 2019

Abstract

In this paper, we present a novel image segmentation technique, based on hidden Markov model (HMM), which we then apply to simultaneously segment interior and exterior walls of fluorescent confocal images of lymphatic vessels. Our proposed method achieves this by tracking hidden states, which are used to indicate the locations of both the inner and outer wall borders throughout the sequence of images. We parameterize these vessel borders using radial basis functions (RBFs), thus enabling us to minimize the number of points we need to track as we progress through multiple layers, and therefore reduce computational complexity. Information about each border is detected using patch-wise convolutional neural networks (CNN). We use the softmax function to infer the emission probability and use a proposed new training algorithm based on s-excess optimization to learn the transition probability. We also introduce a new optimization method to determine the optimum sequence of the hidden states. Thus we transform the segmentation problem into one that minimizes an s-excess graph cut, where each hidden state is represented as a graph node and the weight of these nodes are defined by their emission probabilities. The transition probabilities are used to define relationships between neighboring nodes in the constructed graph. We compare our proposed method to the Viterbi and Baum-Welch algorithms. Both qualitative and quantitative analysis show superior performance of the proposed methods.

1 Introduction

Object detection and tracking of non-rigid shapes (especially those in low quality, noisy images) is a difficult task for traditional "bounding box" style methods [1]. Using a contour based model can improve this accuracy [2] and in sequential images tracking the contour improves overall quality of results compared to bounding box based methods [3]. Object based tracking often invovles image segmentation in order to delineate the object of interest. There are various techniques of segmentation methods that have been developed, such as region-based approaches, e.g., thresholding, region growing, and clustering-based segmentation [4, 5, 6], energy minimization approaches, including combinatorial optimization [7] and deformable models [8, 9];,and machine learning based approaches [10, 11].

In this paper, we propose a novel solution, based on hidden Markov model (HMM), to the problem of tracking and segmenting anatomical borders in lymph vessels. This is applied over a series of cross-sectional images obtained from confocal microscope sections of lymphatic vessels in order to obtain a coherent 3D segmentation of the vessel.

HMM is a stochastic model that satisfies Markov property in a finite set of hidden states. HMM is used in a variety of applications, especially those dealing with time-varying or sequential series [12], such as speech recognition [13], object classification [14, 15, 16], and object tracking [17, 18, 19, 20, 21, 22]. In [14], the authors used two HMMs (one for each class) to detect anomalies in local wall motion during stress echocardiography, utilizing the suitability of HMMs in dealing with time-varying properties to study the continuous movements in the cardiac cycle. In [16], the authors used an HMM with convolutional neural networks (CNN) to identify faults in mechanical systems. CNN is used to extract features from raw signal data then an HMM model is used as a classifier to recognize faults. The ability to track and recognize these sequential events is part of the reason HMMs have such a wide variety of applications. In [22], for example, an HMM is proposed to track 3D catheter tip inside the 3D vessel tree of the patient. In [17], the authors perform a person tracking using pseudo 2-dimensional (P2D) HMM with a Kalman filter. P2DHMM consists

*ehab_essa@mans.edu.eg
†x.xie@swansea.ac.uk

of a number of 1D HMMs where image columns are modeled by super hidden states, each of these super states consist of another number of HMM hidden states. In [19], the authors use an HMM to track hand gesture accompanied with a particle filter. In this case, the particle filter is used to provide an estimate as to which hand region is likely to be detected. After the particle filter performs this estimation, the HMM can then estimate the hand shape, using the Viterbi algorithm, where the hidden state are given by a set of quantized pre-learned examples. However, one problem that arises from these methods is that it could be very computationally expensive depending on the complexity of the object being studied.

One solution to this was proposed in [18, 20] where the authors sampled the contours into a set of discrete points. They incorporated region and edge features into the HMM, which were, in turn, extracted along the normal direction at each of the sampled contour points. In doing so, the complexity of the problem was greatly reduced, leading to a reduction of the overall computational expense involved in calculating the solution. An unscented Kalman filter is then utilized to track an elliptical shape which approximates the contour. This work was extended by Sargin et al. [21] in order to work with open contours of differing lengths. The observational probabilities for the HMMs were defined by utilizing an arc emission, rather than state emission as in traditional approaches. Again, the Viterbi algorithm was used in order to obtain the sequence of optimal hidden states. One drawback of this work is that occlusions in the image could be detrimental to the search process. In these cases, the reliability of the emission probability is compromised and transition probability fails to guarantee a smooth transition of the border. Furthermore, a problem occurs when the Viterbi algorithm is used in cases where the starting and ending points of the segmentation are the same or multiple borders are being tracked. Both such instances occur when tracking vessel walls.

In this paper, we develop an HMM-based, coupled border tracking and segmentation method that can deal with both noise and occlusions and provide a satisfactory segmentation result. We demonstrate our technique on lymphatic vessel confocal images, through simultaneously delineating the inner and outer borders. Lymph vessels are of clinical significance, as fluid build-up in this system decreases post operative recovery rates. Segmenting the lymphatic images is a challenge due to high levels of noise, inconsistent intensity throughout, and occlusions of various sizes.

We propose a novel optimization method using the s-excess graph to find the optimal sequence of the hidden states instead of the more commonplace Viterbi algorithm, which has weaknesses in this area. By transforming the problem into solving the minimum-cost s-excess graph, with nodes of this graph corresponding to the desired border, we can solve it in polynomial time. The transition probability is inferred based on the proposed s-excess optimization instead of the Baum-Welch algorithm. Efficiency is maintained despite the complexity of the object by taking discrete samples. These are taken by transforming the image into polar coordinates to allow the use of radial basis functions (RBFs) to parameterize the borders, creating a reduced sample set to work with and thus drastically reducing the complexity from a large number of points to a much smaller number of RBF centers. The emission probability is defined via the softmax function obtained from a trained CNN on the vessel border and background information.

This paper extends our previous work [23, 24] that developed the initial approach. The main contributions of this paper can be summarized as follows: (1) we propose a novel optimization technique to find the optimal sequence of HMM hidden states based on s-excess optimization to segment and track object contours (Sections 3.2 and 3.3); (2) the proposed optimization allows us to segment multiple borders simultaneously and enforce some geometric constraints (Sections 3.3.1 and 3.3.2); (3) the transition probability (Section 3.4) is learned based on the s-excess optimization in a similar manner to Viterbi training; (4) the emission probability is computed based on a patch-wise CNN that helps to differentiate between vessel borders and background (Section 3.5); and (5) the experimental result demonstrate this to be an effective approach to segment lymphatic vessel walls in noisy confocal images.

2 Related Works

Minimization of closure graph and s-excess problems have been previously applied in image segmentation, e.g. [25] and more recently in [26]. The problem is formulated as a Markov random field (MRF) and is solved using the minimum s-t cut algorithms.

In [25], the authors investigated segmentation of volumetric images. This was achieved by formulating the problem as a minimization of a closed graph. This was then used on synthetic models to find the optimal surfaces (they investigated both single surface and multiple surfaces in their work) through d-dimensional multi-column graphs ($d \geq 3$). One limitation of this work however is making the assumption that the surface is changing within

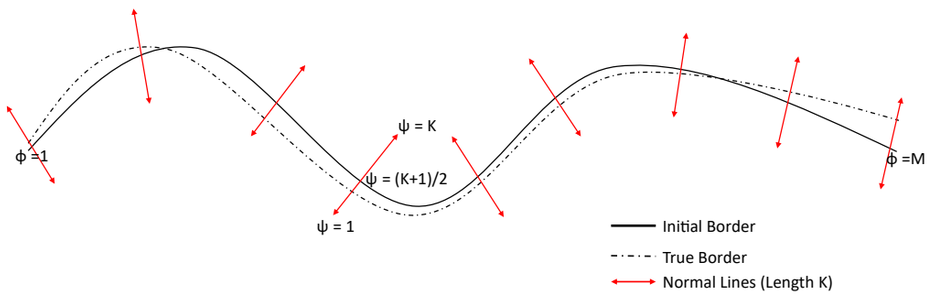


Figure 1: An illustration showing an initial segmenting border and the true border that can be obtained by using the proposed HMM. We divide the border into a series of M points that are evenly distributed and draw a normal line of length K at each point. The final position of border can then be interpolated using RBFs.

the constraint of a global constant. Although this allowed their method to use weighted graph nodes, it is difficult to define this constant and errors in this may lead to sudden and undesirable changes in the resultant surface as there is a great deal of sensitivity to changes in nodal weights.

Recently [25] has been extended, utilizing a different graph construction which allows it to be used in situations such as volumetric data without image unraveling. Petersen et al. [27] built on this method to use graph columns derived from flow lines. These flow lines are non-intersecting, non-parallel lines which they obtain from an initial segmentation. By using this method, they were able to guarantee their segmentations do not self-intersect and manage to cope with high curvature in their target surfaces.

In [26], the use of prior information is incorporated into the model. Through the use of a set of convex functions, the prior shapes are encoded as a series of pairs of adjacent columns. Our method in this paper uses a similar approach to Song et al. [26] to include priors, however, in our work the problem is formulated through the tracking of borders across consecutive frames. We therefore use the transition probability to define the pair-wise cost in our method. This transition probability is acquired from the training set which we treat as a stochastic process. This differs from the optimal surface method, which tries to identify the border as a continuous surface in a volumetric image. Optimal surface segmentation is, therefore, much more computationally expensive. Another advantage of our proposed method is that it is more versatile with regards to the shapes it can segment. This allows us to investigate more complex shapes, for example, both open and closed contours, as well as those of variable length. These advantages are highly suitable for our application, due to the nature of the objects being segmented.

3 Proposed Method

The proposed method integrates contour segmentation and tracking in an HMM-based framework. The HMM problems can be solved in a novel way by converting the optimization problems into minimizing an s-excess graph. The border is parameterized using RBF functions to be reduced to a set of sample point that needs to be tracked and segmented. At each RBF center, a normal line is defined where its points will be represented as hidden states that quantify the border location. Finding the best sequence of hidden states is transformed into an s-excess graph minimization. The transition probability is learned through solving s-excess minimization problem on the training set. CNN is employed to learn the emission probability. The block diagram of the proposed model is shown in Fig. 2. The proposed method is compared to traditional HMM in order to segment and track lymphatic vessel on confocal images.

3.1 Contour parameterization

In our proposed method, we are able to handle both open and closed curves/surfaces in order to segment our border of interest. In order to ensure our graph optimization is efficiently carried out, we "unravel" the images, in effect converting them to polar coordinates. The border (or surface) of interest thus appears to be an open contour in 2D (or an open surface in the case of a volume). The first column(s) are then connected to the last column(s) of the image data. This ensures that continuity in the original object is preserved in the process (as these columns would in fact be neighbors prior to the transfomration). This method is particularly suitable for segmenting vessel-like structures; as these, once transformed in this

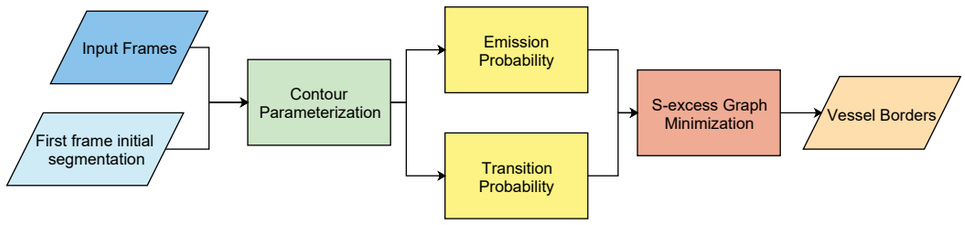


Figure 2: An overview of the proposed method.

way, provide an open surface that represents our desired border(s). We can then treat the problem of segmenting both the inner and outer vessel walls in 3D through tracking two open contours in which there is no overlapping or crossing. These are aligned longitudinally along our unwrapped image. This method has an advantage over optimal surface methods such as [25] in that by using this approach in our method the computational overhead when dealing with 3D data is greatly reduced, even though the optimal surface method in this case also uses a similar image unraveling system. This improved efficiency is due to only requiring adjacent cross-sectional slices to be used to predicate the next position, rather than all cross-sections in order to build the graph for segmentation. Computational complexity is further simplified through RBF based parameterization when investigating the border of interest in 2D cross sections, This is possible as it is unnecessary to track each point along the border and the sample points can be used to interpolate the resultant border in the final stages.

As shown in Fig. 1, the border is sampled into a series of M points, which are evenly distributed along the border of interest. At each of these points, we then extend a line segment perpendicular to the border (a normal line). Each of these normal lines has K points. The final border is approximated by RBFs with the potential RBF centers being given by the hidden states of the HMM. We can therefore index the points on normal lines segments and RBF centers by $\psi \in \{1, \ldots, K\}$ and $\phi \in \{1, \ldots, M\}$, respectively. The initial border for the current cross-section or frame is defined based upon the segmentation result of the previous frame, as shown in Fig. 1.

3.2 HMM formulation

The contour segmentation is performed by tracking the RBF centers located across the longitudinal direction. In the HMM, one possible sequence of hidden states $Q = \{q_1, \ldots, q_\phi, \ldots, q_M\}$ corresponds to one possible RBF center locations along a set of M normal lines. A sequence of HMM observations drawn from the normal lines is denoted as $O = \{o_1, \ldots, o_\phi, \ldots, o_M\}$. HMM is characterized by three probability measures for the transition, emission and initial states. The joint distribution of a sequence of states Q and observations O can be defined according to Markov properties as follows:

$$P(Q, O) = \prod_{\phi=1}^{M} P(q_\phi | q_{\phi-1}) P(o_\phi | q_\phi) \tag{1}$$

where $P(q_1|q_0)$ is the initial probability, $P(q_\phi|q_{\phi-1})$ is the transition probability between states q at two normals ϕ and $\phi-1$ and $P(o_\phi|q_\phi)$ is the emission probability of an observation o_ϕ being generated from a state q_ϕ.

Finding the optimal state sequence with respect to the maximum a posteriori (MAP) criterion is usually attained by the Viterbi algorithm. However, [28, 29] showed the HMM can be represented by Gibbs distribution and its energy function. Thus, The MAP can be expressed as a minimization of an energy function. This energy function is analogous to hidden MRF energy function. In the next section, we show that finding the optimal sequence of states can be established using s-excess optimization.

3.3 Minimum s-excess graph optimization

The inference problems of the HMM can be categorized into realizing the hidden sequence of states given a set of observations and estimating the model parameters. Viterbi algorithm offers feasible solutions for both decoding and parameter estimations. However, the Viterbi algorithm does not include prior information e.g. the spatial relationship between states which is important when performing contour tracking and segmentation. Alternatively,

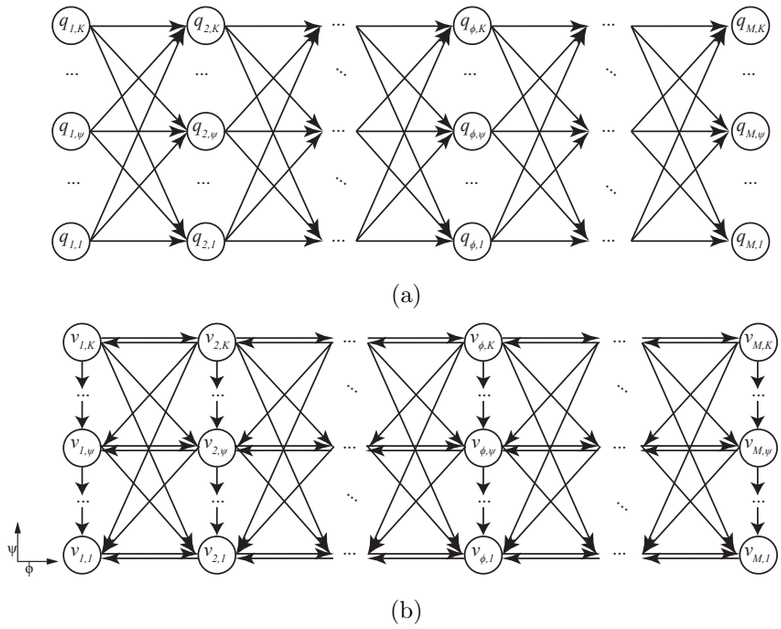


Figure 3: An example of the traditional HMM construction (a), and the proposed HMM shown in (b).

HMM can be generalized to hidden MRF (HMRF) [30] by converting conditional dependencies into potential cliques, thus instead of applying the MAP, it can be solved in polynomial time through an energy minimization method, e.g. graph cut. Here, we introduce an optimization method using the s-excess graph to find the underlying hidden sequence of states and the HMM parameters i.e. transition probability. We express HMM inference problem as a minimum s-excess problem and it is effectively solved using s-t graph cut.

Given a set of hidden random variables X under a set of observations O, where each hidden state is associated to a variable referring to its label $x_i \in L$ and $L = \{0, 1\}$. According to MAP, estimating the labeling of each state is achieved by maximizing the posterior probability defined as follows:

$$P(X|O) = P(O|X)P(X) \tag{2}$$

where $P(X)$ is the prior probability defined according to Gibbs distribution. $P(O|X) = \prod_{\phi=1}^{M} P(o_\phi|x_\phi)$ is the emission probability of the observation under the states labels. Therefore, eq. (2) can be written as:

$$P(X|O) = \frac{1}{Z} \prod_{\phi=1}^{M} P(o_\phi|x_\phi) \prod_{c \in C} \xi_c(X) \tag{3}$$

where Z is a normalization term, $\xi_c(X) = \exp(-U(X))$ is the potential function defined on the clique c, and C is the set of all possible cliques. In HMRF, each state is conditionally depended on the previous state. Then the clique potential is defined on every pair of states. We construct a graph of all the hidden states $\mathcal{Q}$, where $\mathcal{N}$ is a set of pair of states between normal lines. The problem is then converted to an energy minimization by taking the negative log transform:

$$E(X) = \sum_{i \in \mathcal{Q}} D_i(x_i, o_i) + \sum_{(i,j) \in \mathcal{N}} T_{ij}(x_i, x_j) \tag{4}$$

where $D_i(x_i, o_i)$ is the data term, defined as a function derived from the observed data and measures the cost of assigning the label x_i to the state i based on the emission probability. The second term is the smoothness term $T_{ij}(x_i, x_j)$ that measures interaction of each pair of states i, j by penalizing label discontinuities between them based on the transition probability. The two terms are defined as:

$$D_i(x_i) = -\ln P(o_i|x_i) \tag{5}$$

$$T_{ij}(x_i, x_j) = -\ln \exp(-\hat{A}(i,j))|x_i - x_j| \tag{6}$$

where $\hat{A}$ is a symmetric transition matrix. The computation of terms $P(o_i|x_i)$ and $\hat{A}$ are explained later in Sections 3.4 and 3.5.

Here, The MRF problem is solved through the optimization of the s-excess graph. The minimum s-excess problem [31] is a relaxation of minimum closure set problem. An s-excess set is different from a closed set as it lets some successor nodes be absent from the closed set with some assigned cost. Let $G(V, E)$ be a directed graph representing the hidden states trellis, each node $i \in V$ has a cost $w(i)$ equivalent to data term in HMRF and on each edge $e \in E$ there is a positive cost defined by the smoothness term. The goal is to minimize the energy function $\mathfrak{E}(S)$ by finding a subset of graph nodes $S \subset V$ with the minimum total nodal cost added to the separating cost of the set S from the rest of the graph $S' = V - S$.

$$\mathfrak{E}(S) = \sum_{i \in S} w(i) + \sum_{\substack{(i,j) \in E \\ i \in S, j \in S'}} t(i,j). \tag{7}$$

where $w(i) = D_i(1) - D_i(0)$, and $t(i,j) = T_{ij}(1,0)$. As shown in [32] the eq. (4) and eq. (7) are equivalent.

3.3.1 Graph Construction

A special graph construction of the trellis diagram is needed to transform the decoding problem into computing a minimum s-excess graph. The normal lines M are collected together to form a trellis graph G with dimension $M \times K$, where K is the length of the normal line. Each vertex in the graph associates with one hidden state. There are two types of arcs in the graph: intra-chain and inter-chain.

The intra-chain arcs are defined along every chain to prevent the desire border from intersecting the chains more than once. This is similar to Viterbi algorithm, which returns single state from each chain that has the maximum probability. Every node $v(\phi, \psi)$ $(\psi > 1)$ is connected to the below node $v(\phi, \psi - 1)$ with a directed arc of cost $+\infty$ weight.

For inter-chain arcs, a set of directed arcs are constructed from each node $v(\phi_1, \psi_1)$ in chain $C(\phi_1, \psi_1)$ to nodes $v(\phi_2, \psi_2')$ in adjacent chain $C(\phi_2, \psi_2)$, where $1 \leq \psi_2' \leq \psi_1$. Similarly, every node $v(\phi_2, \psi_2) \in C(\phi_2, \psi_2)$ has directed arcs to nodes $v(\phi_1, \psi_1') \in C(\phi_1, \psi_1)$ where $1 \leq \psi_1' \leq \psi_2$. The inter-chain arcs control the smoothness of the border, and the cost of these arcs is defined based on eq. (6).

The cost of outgoing edges to nodes that lie outside the closed set being counted for through the minimization of the energy function. The inter-chain arcs directions go from upper node on one side to lower or the same level nodes on the other side. Thus, the transition matrix $\hat{A}$ is redefined by adding the upper to the transpose of the lower triangular parts of the matrix. The last sequence of states in the trellis graph is tied up to each other to create the foundation of the closed graph.

In Fig. 3 the comparison between our proposed HMM construction method and the traditional HMM method is shown. It is essential to have both the intra-chain arcs and the last row arcs in order for the closed set graph to be computed. We can thus define the transition between states by using a set of inter-chain arcs from left-to-right and vice versa, which then ensures to maintain proportionality between the cost of our cut and that of the distance of border location between neighboring chains.

3.3.2 Coupled Borders Optimization

S-excess optimization has one more advantage, that is allowing to find multiple borders at the same time in accordance with some geometric constraints. Assume we have n borders, for each border S_h, a graph is constructed as described in the previous section. A new inter-borders penalization term $f(S_i(p) - S_j(p))$ is added to control the inter-border interactions. The energy function eq. (4) can be rewritten as:

$$\begin{aligned} E(X) = \sum_{h=1}^{n} \sum_{i \in \mathcal{Q}_h} D_i(x_i, o_i) + \sum_{h=1}^{n} \sum_{(i,j) \in \mathcal{N}_h} T_{ij}(x_i, x_j) + \\ \sum_{p \in \mathcal{N}_s} f(S_i(p) - S_j(p)) \end{aligned} \tag{8}$$

where $\mathcal{N}_h$ is a set of all chains representing normal lines for a single surface and $\mathcal{N}_s$ contains all pairing chains in every adjacent surfaces S_i, and S_j. The inter-border function f is imposed with the distance constraint between two vessel borders, i.e. $\underline{\Delta} \leq |S_i(p) - S_j(p)| \leq \bar{\Delta}$:

$$f(S_i(p) - S_j(p)) = \begin{cases} 0 & \text{if } \underline{\Delta} \leq |S_i(p) - S_j(p)| \leq \bar{\Delta}, \\ \infty & \text{otherwise.} \end{cases} \tag{9}$$

A set of inter-border arcs are defined in such a way in order to enforce the inter-border function, maintaining separation between two vessel borders. Let $C_1(\phi, \psi)$ and $C_2(\phi, \psi)$ denote two corresponding chains obtained from G_1 and G_2 which have borders denoted by S_1 and S_2 respectively. Likewise, let $\hat{\psi}$ represent the spatial coordinate equivalent to ψ in the graph domain, and $X \times Y$ denote the size of the spatial domain. Therefore, it follows that for any node $v_1(\phi, \hat{\psi})$ in $C_1(\phi, \psi)$, we can establish a direct arc between $v_1(\phi, \hat{\psi})$ and $v_2(\phi, \hat{\psi} - \bar{\Delta})$ where $\hat{\psi} \geq \bar{\Delta}$ and $v_2 \in C_2(\phi, \psi)$. Also, nodes $v_2(\phi, \hat{\psi})$ in $C_2(\phi, \psi)$ are linked to $v_1(\phi, \hat{\psi} + \underline{\Delta})$ so that $\hat{\psi} < Y - \underline{\Delta}$ and $v_1 \in C_1(\phi, \psi)$. These arcs have a cost of $+\infty$ assigned to them, with hard constraints defining both the maximum and minimum distance between them.

3.3.3 Computing min-cut/max-flow

For each node, the unary cost w that is defined based on the emission probability as shown in Section 3.3 is transformed by $u(\phi, \psi) = w(\phi, \psi) - w(\phi, \psi - 1)$, except for the bottom chain, i.e. $u(\phi, 1) = w(\phi, 1)$ to facilitate the energy minimization using s-t cut algorithm.

Hochbaum [33] showed the equivalence of the minimum s-excess problem to that of solving for the minimum s-t cut [34] as defined on a proper graph. It therefore follows that an s-t cut algorithm can be used to find the minimum closed set in polynomial time, with finding the minimum for this binding the computational time $T(n, m)$, where we can define n number of graph nodes and m number of edges. Two special nodes are defined for source and sink. The positive nodes u^+ are connected to the source and the negative nodes u^- to the sink with weight $|u|$. We can therefore define the cost of the cut (i.e the minimum s-excess graph) as the total edge cost in separating the graph into source and sink sets.

3.4 Transition Probability

There are two well-known parameter estimation methods for HMM: Baum-Welch and Viterbi training. The Viterbi training is an approximation of the Baum-Welch method and is computationally much faster. Several modifiled parameter estimation methods have also been introduced. For example, in [35], smoothness constraints were used in order to define a transition probability to limit the abrupt changes for the contour by penalizing larger displacements from the initial contour. In [21], transition probabilities are computed through modeling the measurement of bending between every pair of observations and the contour position displacement using two Gaussian distributions with weighted coefficients defined for each transition. Then, the optimal value of these coefficients is calculated using a fixed-point iterative method.

In this paper, we propose an s-excess based optimization for inferring both the transition and prior probabilities. Drawing inspiration from Viterbi training algorithm, we use the s-excess optimization to find the optimal sequence of hidden states given certain observations. The s-excess training algorithm starts with random initializations of both transition and initial probabilities and for each training example. It applies s-excess optimization to find the most likely sequence of states as described in Section 3.3, then updates the HMM parameters. The transition matrix $A(i, j)$ between states i and j is defined by counting the number of times a_{ij} the state j follows the state i between every pair of neighboring normal lines in the most likely sequence, normalized by the total number of transitions from the state i:

$$A(i, j) = \frac{a_{ij}}{\sum_{j'} a_{ij'}} \tag{10}$$

Initial probability π_i is defined similarly by counting the number of transitions from each state i in the first normal line. Since the s-excess optimization by the graph cut provides a global solution, the training algorithm is only required for one iteration. The symmetric transition matrix $\hat{A}$ can be computed by summing upper and lower triangular parts of the transition matrix A and then apply normalization.

3.5 Emission Probability

Traditionally, emission probability is defined by a Gaussian distribution. However, the underlying observations may not fit this parametric model making the output probability inconsistent. A hybrid system using an HMM and neural networks has been used, such as by [36, 37], to allow the discriminative power of the neural networks to infer the emission probability. In this paper, the emission probability is defined by using convolutional neural networks and then encoded in the HMRF data term.

CNN approximates the emission probability by classifying a small patch, centered around each state, as a border of interest or background. The proposed CNN architecture consists of one convolutional layer, followed by a rectified linear unit, max pooling layer, fully connected

Table 1: The architecture of the proposed CNN for inferring the emission probability.

Layer	Type	Input	#filters	Kernel	Stripe	Pad
Data	input	$60 \times 20 \times 1$	-	-	-	-
conv1	convolution	$60 \times 20 \times 1$	64	3×3	1	1
Relu	Rectifier Unit	$60 \times 20 \times 64$	-	-	-	-
pool1	max-pooling	$60 \times 20 \times 64$	-	2×2	2	0
fc1	Fully connected	$30 \times 10 \times 64$	-	-	-	-
Relu	Rectifier Unit	500×1	-	-	-	-
fc2	Fully connected	500×1	-	-	-	-
smax	softmax	2×1	-	-	-	-

Table 2: Classification accuracy (mean value and standard deviation) for inner border.

Accuracy (%)	20×20	40×20	60×20	80×20
Inner vessel wall	84.34 ± 2.45	88.81 ± 1.33	90.06 ± 0.99	88.06 ± 1.19

layer, and finally ended with 2-way softmax layer. The filter size of the convolution layer is 3×3. The max pooling size is 2×2. The fully connected layer has 500 hidden units. Table 1 shows the full architecture of the proposed CNN.

The output layer of the CNN is using the softmax function to compute the posterior probability $P(x_i|o_i)$ of class x_i given the input observations. It has been shown [38, 39] that by using Baye's rule and emission probability is the computed by the scaled likelihood:

$$P(o_i|x_i) = \frac{P(x_i|o_i)}{P(x_i)} \tag{11}$$

where $P(x_i)$ is the prior probability of class x at state i and it is computed by counting the number of state class in the training examples.

$$P(x_i|o_i) = y_x = \frac{\exp(z_x)}{\sum_{x'} \exp(z_{x'})} \tag{12}$$

where z_x is the output neurons from the previous layer multiplied by the weights.

4 Application and Results

We applied the proposed method on a set of lymphatic images to segment the vessel boundaries. These images were acquired using confocal microscopy, which can be used to visualize 3D volume objects that are thicker than the focal plane, a problem for a conventional light microscope. We obtained 8 *in vitro* volumes using confocal microscopy after charged with a fluorescent dye and pressure applied to the sample to ensure valve opening. Each volume has dimensions $512 \times 512 \times 512$. For the purpose of our evaluation, we carry out leave-one-subject-out cross-validation.

As a pre-processing step, we apply anisotropic vessel diffusion [40] in order to enhance the image. By using this method, image features are enhanced, alongside improved vessel connectivity and a reduction in the level of noise typically present in this type of images. Then the images are unravelled, converting them into polar coordinates of size 200×1257.

Every frame where manual labeling has been performed is used in the evaluation process, with five different evaluation metrics. These metrics include the absolute mean difference (AMD), Hausdorff distance (HD) in pixel, area overlap (AO %), sensitivity (Sens. %), and

Table 3: Mean (standard deviation) of five different evaluation metrics for inner border segmentation.

	AMD	HD	AO	Sens.	Spec.
Optimal Surface [25]	7.10(3.50)	50.60(22.10)	90.10(5.60)	91.20(5.60)	**99.30**(**0.60**)
BW Training & VT decoding	4.14(2.45)	10.06(4.54)	94.28(3.91)	95.61(4.37)	99.10(1.58)
VT Training & VT decoding	5.46(10.71)	12.02(12.85)	93.16(11.90)	95.20(12.35)	98.75(1.06)
S-E Training & VT decoding	3.11(1.60)	8.48(3.32)	95.82(2.41)	97.40(2.65)	98.95(1.27)
BW Training & S-E decoding	3.60(4.03)	14.13(16.46)	95.12(5.13)	96.68(5.19)	99.05(0.86)
VT Training & S-E decoding	2.58(1.41)	8.04(3.43)	96.44(2.25)	98.26(2.27)	98.91(0.96)
S-E Training & S-E decoding	2.39(1.32)	7.27(2.93)	96.69(2.12)	98.45(2.12)	98.95(0.92)
S-E Training & CS-E decoding	**2.23**(**1.08**)	**6.95**(**2.57**)	**96.94**(**1.73**)	**98.58**(**1.60**)	99.04(0.89)

$p < 0.0001$

Table 4: Mean (standard deviation) results of five different evaluation metrics for outer border segmentation.

	AMD	HD	AO	Sens.	Spec.
Optimal Surface [25]	7.00(5.90)	54.90(26.10)	91.70(7.30)	92.70(7.50)	**99.20**(**0.50**)
BW Training & VT decoding	3.06(1.59)	8.74(4.76)	96.54(1.76)	98.85(1.66)	98.12(1.69)
VT Training & VT decoding	3.79(1.25)	11.27(4.08)	95.71(1.46)	98.24(1.82)	97.96(1.40)
S-E Training & VT decoding	2.78(1.27)	8.20(3.59)	96.78(1.55)	97.59(1.80)	99.36(0.56)
BW Training & S-E decoding	2.73(5.37)	8.80(13.21)	96.96(5.22)	98.03(5.38)	99.17(0.70)
VT Training & S-E decoding	2.11(0.95	7.11(3.32)	97.53(1.25)	98.71(1.39)	99.09(0.78)
S-E Training & S-E decoding	1.92(0.80)†	6.84(3.03)	97.75(1.06)‡	98.86(1.18)	99.14(0.69)
S-E Training & CS-E decoding	**1.90**(**0.81**)	**6.53**(**3.11**)	**97.77**(**1.05**)	**98.85**(**1.23**)	99.16(0.65)

$p < 0.0001, \dagger p = 0.0031, \ddagger p = 0.0001$

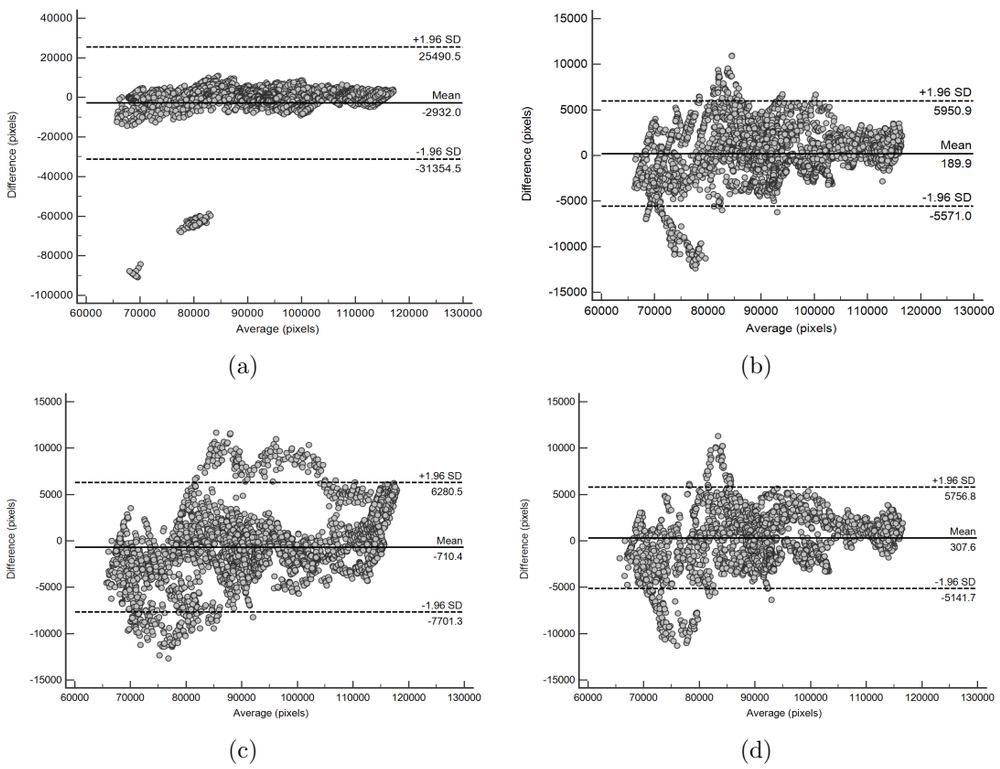


Figure 4: Bland-Altman plots for segmentation results of the inner border using (a) VT training and VT decoding. (b) VT training and SE decoding. (c) SE training and VT decoding. (d) SE training and SE decoding.

specificity (Spec. %), which are defined as:

$$AMD = \frac{1}{N}\sum_{i=1}^{N}|y_{AT}(i) - y_{GT}(i)|, \tag{13}$$

$$HD = \max_{a \in y_{AT}}\{\max_{b \in y_{GT}}[dis(a,b)]\}, \tag{14}$$

$$AO = \frac{TP}{TP + FN + FP}, \tag{15}$$

$$Sens. = \frac{TP}{TP + FN}, \tag{16}$$

$$Spec. = \frac{TN}{TN + FP}, \tag{17}$$

where N is the number of border points, y_{AT} is the automatically segmented border, y_{GT} is the groundtruth border, and $dis(a, b)$ is the Euclidean distance between a and b sets of points of the borders y_{AT} and y_{GT}. For the purpose of sensitivity and specificity measurements, we considere the inner region as true positive area TP of the vessel, FN denotes false negative, FP denotes false positive, and TN is true negative. Using the polar images, we measured normal line segments at a length of 101 pixels, and we use 42 regularly spaced RBF centers, i.e. one RBF center every 30 pixels.

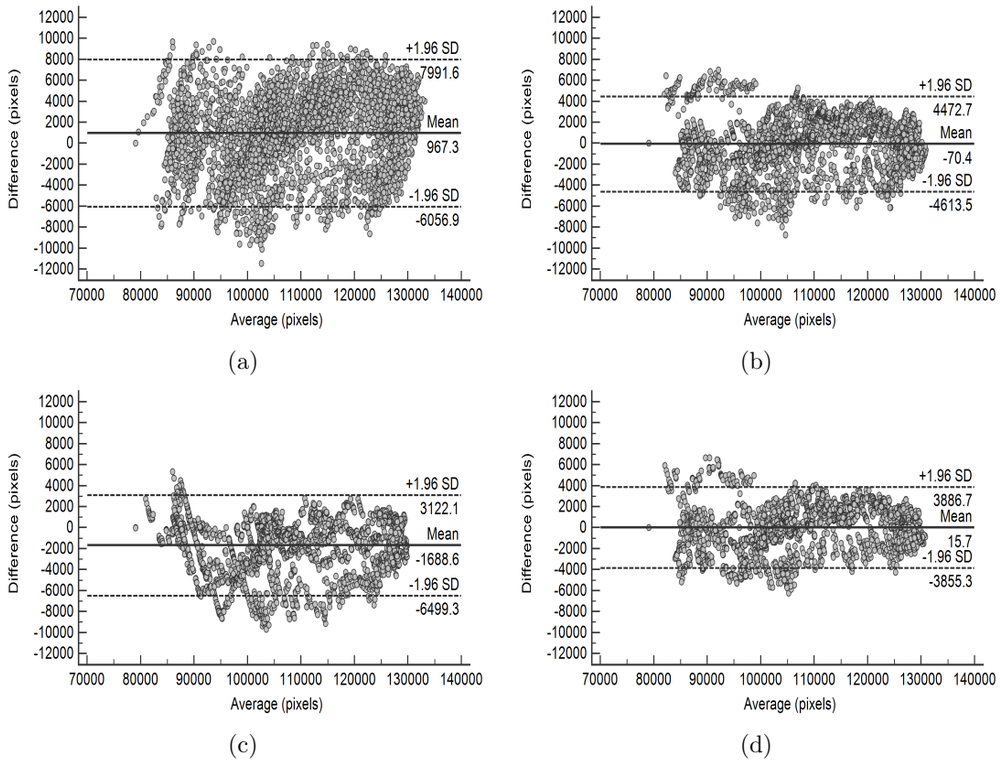


Figure 5: Bland-Altman plots for segmentation results of the outer border using (a) VT training and VT decoding. (b) VT training and SE decoding. (c) SE training and VT decoding. (d) SE training and SE decoding.

We evaluated different input patch sizes for the CNN to classify the inner vessel border at 20×20, 40×20, 60×20, and 80×20 as illustrated in Table 2. The mean and standard deviation of CNN accuracy at patch size 60×20 gives the best result 90.06 ± 0.99. This patch size is adapted in all following experiments.

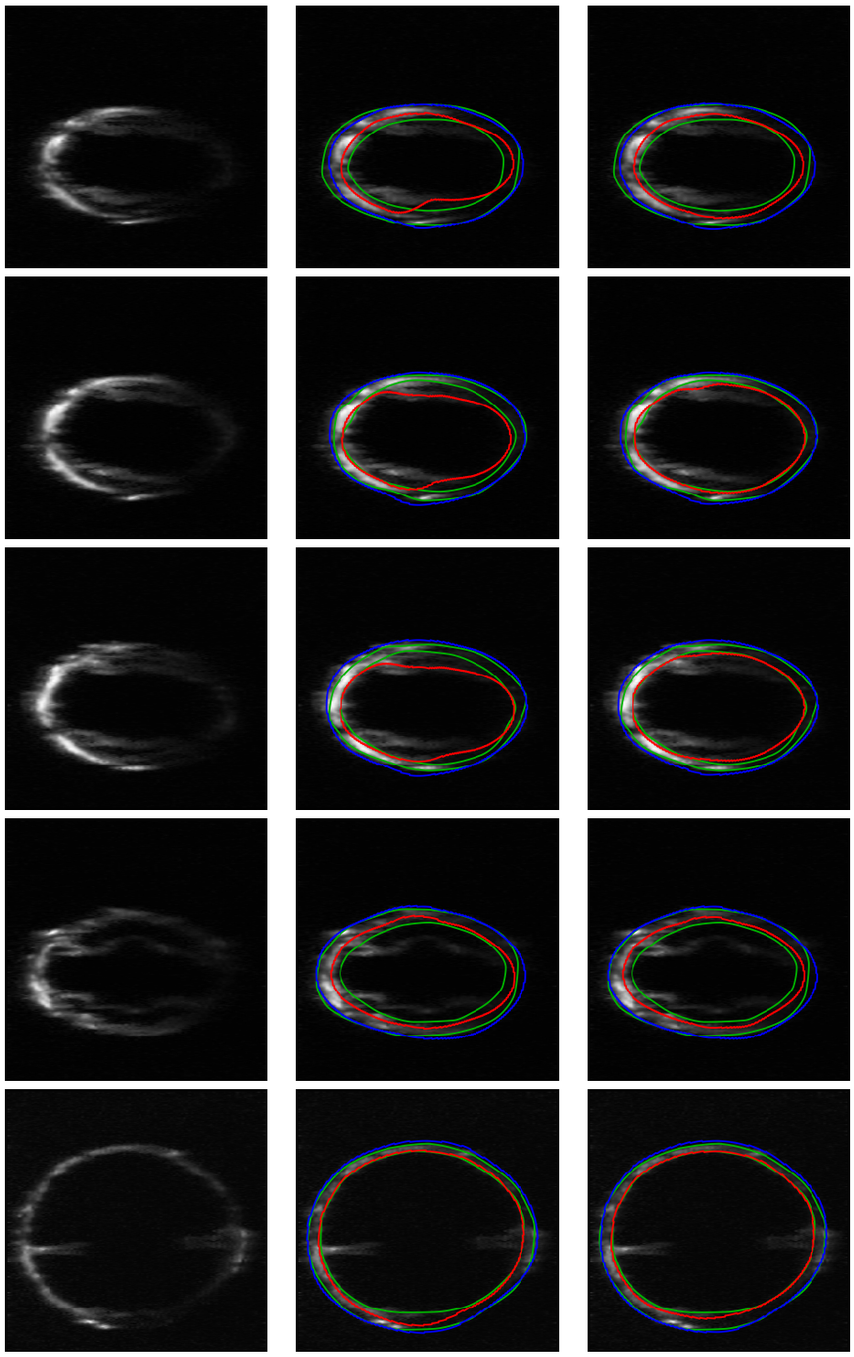
The experiments are designed to evaluate both the proposed HMM training and decoding algorithms based on the s-excess optimization. Tables 3 and 4 show the quantitative results of segmenting the inner and outer lymphatic vessel walls, respectively. A paired t-test was used to test for significant differences in the computed AMD, HD, and AO between the borders obtained by the proposed training and decoding methods and all other methods (p-values < 0.05 are considered significant). p-values is presented in Tables 3 and 4.

In order to evaluate the effect of the HMM training algorithm to infer the transition matrix, we compared the proposed s-excess (S-E) based training method to two popular HMM training methods: Baum-Welch (BW) and Viterbi (VT). Viterbi decoding is used after the three training methods to find the optimal sequence of states as shown in Tables 3 and 4 rows (2-4). It is evidently clear that the proposed S-E training method achieved better performance in segmenting both inner and outer vessel walls, e.g. the absolute mean difference of the inner border for the proposed method is 3.11 compared to 5.46 and 4.14 for VT training and BW training respectively and for the outer border the absolute mean difference of the proposed method is 2.78 compared to 3.79 and 3.06 for VT training and BW training respectively.

For evaluating the decoding method, we test the proposed S-E decoding based on a transition matrix acquired by using three different training methods (BW, VT and S-E) as shown in Tables 3 and 4 rows (5-7). The S-E decoding in conjunction with the S-E training gives the best result. It also worth noting that the S-E decoding outperforms the VT decoding even if a transition matrix is based on the VT training, e.g. the absolute mean difference of the inner border for the S-E decoding method is 2.58 compared to 5.46 and 4.14 for VT decoding and for the outer border the absolute mean difference of the S-E decoding method is 2.11 compared to 3.79 for VT decoding.

The proposed S-E decoding can segment coupled borders simultaneously (CS-E). The minimum $\underline{\Delta}$ and maximum $\bar{\Delta}$ distance between two borders are limited to 5 and 15 respectively. Coupled S-E decoding shows some improvement over single S-E decoding (i.e. segmenting the borders separately) as shown in Tables 3 and 4 rows (7-8).

BlandAltman plots of inner and outer vessel borders are shown in Fig. 4 and 5 to assess agreement between the groundtruth and the automatically segmented vessel by measuring



(a) (b) (c)

Figure 6: Comparison of segmentation results (red: inner, blue: outer) and ground-truth (green) for both proposed methods using S-E Training. (a) original image (b) S-E decoding method. (c) CS-E decoding method.

the difference versus average of the vessel area. The segmented vessel border is obtained by using (a) VT training and VT decoding, (b) VT training and S-E decoding, (c) S-E training and VT decoding, and (d) S-E training and SE decoding. By using S-E training and decoding, all datasets inside the acceptable range of 1.96 times the standard deviation except some outliers with a minimal bias on the results. Our proposed s-excess optimization can be clearly seen to outperform the Viterbi algorithm when used to tracking both borders. Viterbi based method appears to be easily distracted by lymphatic valves attached to the inner wall which lead to a high distance error and low sensitivity. A much larger amount of data had to be used to train the transition and prior probabilities in the Viterbi method when compared to our proposed method.

Some typical segmentation results of single and coupled S-E decoding are shown in Fig. 6. Simultaneous segmentation of both borders was also found to produce results better than segmenting the borders separately. This was especially true in the case of segmenting the

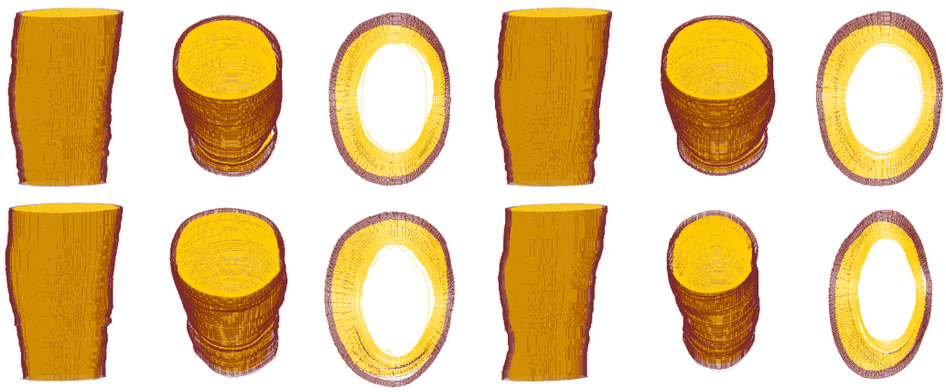


Figure 7: 3D visualization of results obtained using the proposed coupled s-excess method (inner wall: yellow, outer wall: red).

inner border (shown in Table 3). The performance was greatly improved with the addition of the inter-border constraint, by decreasing its tendency of being distracted by lymphatic valves in the vicinity.

We also carried out comparisons of the proposed method against the optimal surface segmentation [25] in Tables 3 and 4. We found that the results of using optimal surface method showed a particular sensitivity to the distance parameters set between columns. The proposed method has fully connected arcs between chains, which improved the consistency and accuracy of the results.

In Fig. 6, a comparison between the proposed method using S-E decoding and CS-E decoding is shown. CS-E decoding method provides a smoother border transition particularly in cases of occlusion or missing features. Fig. 7 shows 3D visualizations of some typical results of the proposed method from three different angles. The results illusrtrate the robustness and consistency of the proposed method for segmenting the lymphatic vessel.

5 Conclusions

In this paper, an HMM-based method was presented in order to simultaneously segment the inner and outer vessel borders in Lymphatic vessels, obtained from confocal images. Our proposed method uses a set of normal lines, which are calculated from the previous frame's segmentation, to search for the position of the border in the current frame. In our method, the emission probability is inferred from the softmax function of a CNN. We show that the optimal sequence of the hidden states can be calculated via solving a minimum s-excess problem using *s-t* cut algorithm, and this can be later used to train the model similar to Viterbi training. This sequence of the hidden states corresponds to RBF centers of the border and can therefore be used to interpolate the position of the true border. We provide both qualitative and quantitative results obtained from a set of lymphatic image sequences showing our proposed method achieving better performance compared to other traditional HMM approaches and the optimal surface technique.

References

[1] Ying G, Liu S, Liu Z, Jin Y. Accurate natural contour tracking for non-rigid object? *2015 IEEE International Conference on Information and Automation*, 2015; 1382–1387, doi:10.1109/ICInfA.2015.7279502.

[2] Sridevi S, Sundaresan M. A new approach for coronary plaque layer analysis based on ivus. *2014 International Conference on Computing for Sustainable Global Development (INDIACom)*, 2014; 490–494, doi:10.1109/IndiaCom.2014.6828186.

[3] Heo S, Koo HI, Cho NI. Open contour tracking using a new state space model and nonrigid motion training. *IEEE Transactions on Circuits and Systems for Video Technology* 2016; **PP**(99):1–1, doi:10.1109/TCSVT.2016.2592321.

[4] Wang CD, Lai JH. Energy based competitive learning. *Neurocomputing* 2011; **74**(12):2265 – 2275, doi:https://doi.org/10.1016/j.neucom.2011.03.013.

[5] Wang C, Lai J, Zhu J. Graph-based multiprototype competitive learning and its applications. *IEEE Transactions on Systems, Man, and Cybernetics, Part C (Applications and Reviews)* 2012; **42**(6):934–946.

[6] Cai H, Yang Z, Cao X, Xia W, Xu X. A new iterative triclass thresholding technique in image segmentation. *IEEE Transactions on Image Processing* 2014; **23**(3):1038–1046, doi:10.1109/TIP.2014.2298981.

[7] Essa E, Xie X. Automatic segmentation of cross-sectional coronary arterial images. *Computer Vision and Image Understanding* 2017; **165**:97 – 110.

[8] Yeo SY, Xie X, Sazonov I, Nithiarasu P. Geometrically induced force interaction for three-dimensional deformable models. *IEEE Transactions on Image Processing* 2011; **20**(5):1373–1387.

[9] Zhang H, Xie X. Divergence of gradient convolution: Deformable segmentation with arbitrary initializations. *IEEE Transactions on Image Processing* 2015; **24**(11):3902–3914, doi:10.1109/TIP.2015.2456503.

[10] Ronneberger O, Fischer P, Brox T. U-net: Convolutional networks for biomedical image segmentation. *Medical Image Computing and Computer-Assisted Intervention – MICCAI 2015*, Springer International Publishing, 2015; 234–241.

[11] Essa E, Xie X. Phase contrast cell detection using multilevel classification. *International Journal for Numerical Methods in Biomedical Engineering* 2018; **34**(2):e2916, doi:10.1002/cnm.2916.

[12] Dileep AD, Sekhar CC. Hmm based intermediate matching kernel for classification of sequential patterns of speech using support vector machines. *IEEE Transactions on Audio, Speech, and Language Processing* Dec 2013; **21**(12):2570–2582, doi:10.1109/TASL.2013.2279338.

[13] Rabiner L. A tutorial on hidden markov models and selected applications in speech recognition. *Proceedings of the IEEE* 1989; **77**(2):257–286.

[14] Mansor S, Noble J. Localwall motion classification of stress echocardiography using a hidden markov model approach. *ISBI*, 2008; 1295–1298.

[15] Saini R, Roy PP, Dogra DP. A segmental hmm based trajectory classification using genetic algorithm. *Expert Systems with Applications* 2018; **93**:169 – 181, doi:https://doi.org/10.1016/j.eswa.2017.10.021.

[16] Wang S, Xiang J, Zhong Y, Zhou Y. Convolutional neural network-based hidden markov models for rolling element bearing fault identification. *Knowledge-Based Systems* 2018; **144**:65 – 76, doi:https://doi.org/10.1016/j.knosys.2017.12.027.

[17] Breit H, Rigoll G. Improved person tracking using a combined pseudo-2D-HMM and kalman filter approach with automatic background state adaptation. *ICIP*, vol. 2, 2001; 53–56.

[18] Chen Y, Rui Y, Huang T. JPDAF based HMM for real-time contour tracking. *CVPR*, vol. 1, 2001; I–543–I–550.

[19] Fei H, Reid I. Joint bayes filter: A hybrid tracker for non-rigid hand motion recognition. *ECCV*, 2004; 497–508.

[20] Chen Y, Rui Y, Huang T. Multicue HMM-UKF for real-time contour tracking. *IEEE PAMI* 2006; **28**(9):1525–1529.

[21] Sargin M, Altinok A, Manjunath B, Rose K. Variable length open contour tracking using a deformable trellis. *IEEE TIP* 2011; **20**(4):1023–1035.

[22] Ambrosini P, Smal I, Ruijters D, Niessen WJ, Moelker A, Walsum TV. A hidden markov model for 3d catheter tip tracking with 2d x-ray catheterization sequence and 3d rotational angiography. *IEEE Transactions on Medical Imaging* 2017; **36**(3):757–768, doi:10.1109/TMI.2016.2625811.

[23] Essa E, Xie X, Jones JL. Minimum s-excess graph for segmenting and tracking multiple borders with hmm. *Medical Image Computing and Computer-Assisted Intervention (MICCAI)*, Springer International Publishing, 2015; 28–35.

[24] Essa E, Xie X, Jones JL. Graph based lymphatic vessel wall localisation and tracking. *Graph-Based Representations in Pattern Recognition: 10th IAPR-TC-15 International Workshop*, 2015; 345–354.

[25] Li K, Wu K, Chen D, Sonka M. Optimal surface segmentation in volumetric images-a graph-theoretic approach. *IEEE PAMI* 2006; **28**(1):119–134.

[26] Song Q, *et al.*. Optimal multiple surface segmentation with shape and context priors. *IEEE TMI* 2013; **32**(2):376–386.

[27] Petersen J, *et al.*. Optimal surface segmentation using flow lines to quantify airway abnormalities in chronic obstructive pulmonary disease. *MIA* 2014; **18**:531–541.

[28] Zhao Y, Atlas LE, Zhuang X. Application of the gibbs distribution to hidden markov modeling in speaker independent isolated word recognition. *IEEE Transactions on Signal Processing* 1991; **39**(6):1291–1299.

[29] Gravier G, Sigelle M, Chollet G. A markov random field model for automatic speech recognition. *Proceedings 15th International Conference on Pattern Recognition. ICPR-2000*, vol. 3, 2000; 254–257 vol.3.

[30] Blake A, Kohli P, Rother C. *Markov Random Fields for Vision and Image Processing*. The MIT Press, 2011.

[31] Hochbaum D. Anniversary article: Selection, provisioning, shared fixed costs, maximum closure, and implications on algorithmic methods today. *Manage. Sci.* 2004; **50**(6):709–723.

[32] Bai J, Shah A, Wu X. Optimal multi-object segmentation with novel gradient vector flow based shape priors. *CoRR* 2017; **abs/1705.10311**. URL http://arxiv.org/abs/1705.10311.

[33] Hochbaum D. An efficient algorithm for image segmentation, markov random fields and related problems. *J. ACM* 2001; **48**(4):686–701.

[34] Boykov Y, Kolmogorov V. An experimental comparison of min-cut/max-flow algorithms for energy minimization in vision. *IEEE PAMI* 2004; **26**(9):1124–1137.

[35] Chen Y, Huang T, Rui Y. Parametric contour tracking using unscented kalman filter. *ICIP*, vol. 3, 2002; 613–616 vol.3.

[36] Bourlard HA, Morgan N. *Connectionist Speech Recognition: A Hybrid Approach*. Kluwer Academic Publishers: Norwell, MA, USA, 1993.

[37] Wilinski P, Solaiman B, Hillion A, Czarnecki W. Toward the border between neural and markovian paradigms. *IEEE Transactions on Systems, Man, and Cybernetics, Part B (Cybernetics)* 1998; **28**(2):146–159, doi:10.1109/3477.662756.

[38] Guo Q, Wang F, Lei J, Tu D, Li G. Convolutional feature learning and hybrid cnn-hmm for scene number recognition. *Neurocomputing* 2016; **184**:78 – 90, doi:http://dx.doi.org/10.1016/j.neucom.2015.07.135.

[39] Koller O, Zargaran O, Ney H, Bowden R. Deep sign: Hybrid cnn-hmm for continuous sign language recognition. *The British Machine Vision Conference (BMVC)*, 2016.

[40] Manniesing R, Viergever M, Niessen W. Vessel enhancing diffusion: A scale space representation of vessel structures. *MIA* 2006; **10**(6):815 – 825.