Title

A Marked Poisson Process Driven Latent Shape Model for 3D Segmentation of Reflectance Confocal Microscopy Image Stacks of Human Skin

Journal

IEEE Transactions on Image Processing, 26(1)

ISSN

1057-7149

Authors

Ghanta, Sindhu
Jordan, Michael I
Kose, Kivanc
et al.

Publication Date

2017

DOI

10.1109/tip.2016.2615291

Peer reviewed

Published in final edited form as:

IEEE Trans Image Process. 2017 January ; 26(1): 172–184. doi:10.1109/TIP.2016.2615291.

A Marked Poisson Process Driven Latent Shape Model for 3D Segmentation of Reflectance Confocal Microscopy Image Stacks of Human Skin

Sindhu Ghanta,
Parallel Machines, Sunnyvale, CA, USA

Michael I. Jordan,
University of California, Berkeley, USA

Kivanc Kose,
Memorial Sloan Kettering Cancer Center, NY, USA

Dana H. Brooks,
Department of Electrical and Computer Engineering, Northeastern University, Boston, MA, USA

Milind Rajadhyaksha, and
Memorial Sloan Kettering Cancer Center, NY, USA

Jennifer G. Dy
Department of Electrical and Computer Engineering, Northeastern University, Boston, MA, USA

Abstract

Segmenting objects of interest from 3D datasets is a common problem encountered in biological data. Small field of view and intrinsic biological variability combined with optically subtle changes of intensity, resolution and low contrast in images make the task of segmentation difficult, especially for microscopy of unstained living or freshly excised thick tissues. Incorporating shape information in addition to the appearance of the object of interest can often help improve segmentation performance. However, shapes of objects in tissue can be highly variable and design of a flexible shape model that encompasses these variations is challenging. To address such complex segmentation problems, we propose a unified probabilistic framework that can incorporate the uncertainty associated with complex shapes, variable appearance and unknown locations. The driving application which inspired the development of this framework is a biologically important segmentation problem: the task of automatically detecting and segmenting the dermal-epidermal junction (DEJ) in 3D reflectance confocal microscopy (RCM) images of human skin. RCM imaging allows noninvasive observation of cellular, nuclear and morphological detail. The DEJ is an important morphological feature as it is where disorder, disease and cancer usually start. Detecting the DEJ is challenging because it is a 2D surface in a 3D volume which has strong but highly variable number of irregularly spaced and variably shaped “peaks and valleys”. In addition, RCM imaging resolution, contrast and intensity vary with depth. Thus a prior model needs to incorporate the intrinsic structure while allowing variability in essentially all its parameters. We propose a model which can incorporate objects of interest with complex shapes and variable appearance in an unsupervised setting by utilizing domain knowledge to build appropriate priors of the model. Our novel strategy to model this structure combines a spatial

Poisson process with shape priors and performs inference using Gibbs sampling. Experimental results show that the proposed unsupervised model is able to automatically detect the DEJ with physiologically relevant accuracy in the range 10 – 20μ*m*.

Keywords

Segmentation; shape model; 3D Bayesian model; Poisson process; dermis-epidermis

I. Introduction

Detecting and segmenting objects of interest from 3D datasets is an active area of research [1]. In medical images, structures of interest often have a complex and highly variable shape, making segmentation challenging [2]. Especially in microscopy images, this problem is compounded by subtle variation in contrast, resolution and intensity often making appearance based classification difficult [3], [4].

In this paper, we present a novel unsupervised generative framework that can be used for delineating the complex 3D boundary of an object. This probabilistic model can segment/detect multiple complex-shaped objects simultaneously. The framework of our model can be adapted to different applications by choosing the appropriate shape prior based on domain knowledge. In developing this model, we were inspired by the problem of finding the boundary between superficial cellular epithelium and underlying connective tissue, a commonly encountered problem in tissue microscopy [5], [1], [6]. Specifically we address a special case of this general application; delineation of the boundary between epidermis (i.e., epithelium) and dermis (i.e., connective tissue) in 3D image stacks of reflectance confocal microscopy (RCM) images of human skin *in vivo*. We chose this application problem not only because of its practical relevance to detection and treatment of skin cancer, but also because difficulties it presents in terms of the high degree of variability of all relevant parameters makes it a good test case for a much wider class of 3D microscopy segmentation problems of interest.

In our proposed model, the DEJ is modeled as a union of an unknown number of "hill" shapes emerging at random locations using a Marked Poisson Process. These hill tops in human skin are irregular in both depth and lateral location with variability across individuals and even across different sites from the same patient. In addition, these hill tops that we model are very sparse compared to the 3D volume of space which motivates the use of Poisson as an appropriate prior. This model has three flexible components: a random number of objects at random locations, a parametrized shape model and a parameterized appearance model. Details of each component of the model are briefly described later in this Section and then formalized in Section II.

RCM is a relatively new non-invasive technique that is capable of capturing images of human skin down to the depth of 100 – 150μ*m* [7]. *En-face* slices (parallel to the surface) are captured sequentially for increasing depths in skin and recorded as a z-stack of images (as illustrated in Figure 1). Skin biologists and clinicians are interested in the location of the

dermal-epidermal junction (DEJ) in these z-stacks because disorder and disease, including cancer usually start at this boundary [8].

The DEJ in skin typically has a 3D structure, marked by multiple hills and valleys undulations, called dermal papillae (hills) and epidermal rete ridges (valleys) (Figure 2). This structure is flexible, irregular, and subject to a high degree of inter and intra-subject (or specimen) variability [9]. In skin biology and clinical pathology, the presence, location, and morphology of the DEJ is a key attribute and changes in the architecture of DEJ exhibit key diagnostic information [10], [11]. Morphologic changes at the DEJ are the first and most important features that allow diagnosis of benign melanocytic lesions and early detection of their progression toward melanoma. For example, broadening of the dermal papillae and elongation of rete ridges in conjunction with proliferation of melanocytes (a brown pigment typically found in the epidermis) along the boundary is one of the earliest and most important indicator of melanoma, the most deadly form of skin cancer [12], [13]. As cancer spreads laterally along this boundary and/or in depth, the form of DEJ, as well as morphological structures around it, becomes a key determinant in cancer staging. The location of abnormal melanocytes relative to the DEJ (above in epidermis or below in dermis), determines the stage (depth of invasion) of the disease, which then guides diagnosis and treatment. Thus, the architecture of the DEJ is an important focus of biological and pathological work.

However, RCM images can be challenging to read and interpret. The images are in *en-face* orientation and record reflectance contrast, typically visualized in grayscale, whereas pathology is examined in orthogonal sections with the benefit of color stains. The signal-to-noise and contrast in RCM images, while sufficient to reveal critical features, varies with skin pigmentation. Thus, the ability to delineate the DEJ and analyze features and morphologic patterns remains limited, at present, to well-trained experts [5]. Currently, the DEJ is identified in RCM by expert readers, who manually and laboriously read the images in sequence through the z-stacks. These experts form an international cohort of ‘early adopter’ clinicians, who have worked with RCM technology during the past decade and have become highly skilled in reading images. For novice (non-expert) clinicians in the wider cohort, however, who are keen to adopt RCM, learning to read images is challenging and requires substantial effort and time. This purely visual approach results in variable accuracy and repeatability [14], [15]. This is a major technical barrier limiting both wider adoption of RCM imaging and training for current and future dermatologists. Both the potential reward of RCM imaging and the importance of overcoming this barrier are being increasingly recognized by leaders in dermatology community. Recently, our group has introduced an automated image processing technique to assist clinicians in identifying the DEJ [4], [16]. However, this approach works on dividing the RCM slices into tiles, which are then classified, followed by post-detection interpolation, rather than model the structure of DEJ itself. Up to now we have not succeeded in developing a global generative model of the normal DEJ that utilizes its known physical/architectural structure. In this paper, a novel unsupervised algorithm that captures the structure of DEJ while allowing inter and intra-subject variability is presented and validated against RCM z-stacks, to automatically detect the DEJ and assist biologists and clinicians in reading images with improved accuracy and repeatability.

Figure 3 shows slices from two example stacks, one from dark and one from fair skin. As can be appreciated from these images, identifying the DEJ in RCM images is a difficult task due to subtle optical variations in resolution, intensity and low contrast that are hard to recognize. This results in subjectivity and high variability among readers. The complexity of the problem varies with skin type – with dark skin type being easier than fair due to the high reflectance contrast at the DEJ due to the presence of melanin, while in fair skin, the contrast is relatively low due to the lack of melanin, as illustrated in Figure 3. In our group's previous work [4], [16], we adopted a fine-grained local strategy to first recognize DEJ characteristics at the scale of $16\mu m \times 16\mu m$ tiles for dark skin and $25\mu m \times 25\mu m$ tiles for fair skin, followed by post-detection interpolation. Larger tiles were used for fair skin in order to be able to extract statistically reliable textural features in the absence of reflectance contrast. Because of the intrinsic difficulty of the problem, we did not employ a generative 3D model of the DEJ itself, but instead classified tiles as being in the epidermis (above DEJ), in the dermis (below it), or in an "uncertainty region" around it, all based on discriminative features. Although we were successful at finding normal DEJ, any new type of abnormality would require re-thinking the approach. Thus we have been searching for a method which would allow us to model the DEJ with all its complexity and variability. In this work, we solve this problem by developing a global model of the DEJ which leverages what is known about its structure and has sufficient flexibility to incorporate inter and intra-subject (or specimen) variability. Our approach is based on Bayesian modeling, where domain knowledge is carefully incorporated into the prior probability distributions of the model parameters.

Marked *Point* Processes have been used in the past for detection and segmentation of multiple arbitrarily shaped objects (for example, cell nuclei) in microscopy images [6], [17]. A Poisson process is a special case of the point process and has several advantages, such as its independence property, that make a generative Bayesian setting possible. Our model will allow us to overcome limitations of more traditional segmentation methods which tend to segment small collections of single objects rather than a globally-defined surface. In summary, we designed a novel probabilistic model based on a marked spatial Poisson process that can take uncertainty in number, location, shape and appearance into account. Our model allows incorporation of domain knowledge information regarding the general shape of objects, while allowing shape parameter to vary across objects. This modeling strategy contributes to the field of probabilistic models for segmentation in 3D. Inference on such a model, involving trans-dimensional jumps, is typically performed using reversible jump Markov chain Monte Carlo (RJMCMC). Instead, we developed a Gibbs sampling framework which is another major contribution of this work. We demonstrate the performance of our model on real RCM image stacks from human skin, segmenting the DEJ for both dark and fair skin.

II. Model Formulation

In this section, we describe our notation and each component of the model in detail. Let the 3D image stack of skin lie in a continuous space R^3. We represent this continuous space by a piecewise constant basis. We only need samples at these discrete locations to represent the coefficients of the basis. RCM captures samples at these locations resulting in a data tensor $X \in \mathbb{R}^{N_r \times N_c \times N_z}$, where N_r, N_c and N_z are the number of rows, columns and slices of the

image respectively. Each element in this tensor is a voxel, represented by a scalar (e.g., grayscale intensity value) or a vector (e.g., texture features). Let the total number of voxels in the image be denoted by $N = [N_r \times N_c \times N_z]$.

One sample from a spatial Poisson process consists of a random number of points at random locations based on the underlying spatial Poisson intensity prior parameter. Let the number of these points generated from the Poisson process be denoted by M and the location of these M points be denoted by $\mathbf{L} = [\mathbf{l}_1, \ldots, \mathbf{l}_M]$, where $\mathbf{l}_m = [l_{x,m}, l_{y,m}, l_{z,m}]$ which represent the location of m^{th} object in 3D. Let the Poisson intensity be denoted by $\beta(\boldsymbol{\tau})$, where $\boldsymbol{\tau}$ represents the location in $3D$ space of the RCM stack. Let the shape parameters at each location ($\mathbf{L} = [\mathbf{l}_1, \ldots, \mathbf{l}_M]$) be represented by $\mathbf{S} = [\mathbf{s}_1, \ldots, \mathbf{s}_M]$, where each $\mathbf{s}_m$ is a vector consisting of shape parameters (a probabilistic model of hills in this case). The hills (shape) are *marks* associated with the outcome of a spatial Poisson process and hence the model is called a marked spatial Poisson process. The exact shape of each hill determined by the value of shape parameters is unknown and inferred by the model. Therefore, the marks are latent and we have a latent marked Poisson process. Given the location and shape of each hill defining a boundary (DEJ) here, the appearance model of voxels above and below the boundary are generated from an appearance model of epidermis and dermis respectively. Note that the number of hills (M) is random and open ended making it possible to generate a flexible model of the DEJ. We summarize the generative process as follows:

1. Generate M points along with their location from a spatial Poisson Process (Figure 4(a)).
2. For each object, given its location, generate object boundaries from the shape model (Figure 4(b)).
3. Given the location of objects and their shape boundaries, data below the boundary is generated from a dermis appearance model and data above the boundary is generated from a epidermis appearance model (Figure 4(c)).

In the following subsections, we describe each of the three components of the model in detail.

A. Spatial Poisson Process

A Poisson process in general is defined on a locally compact metric space $\mathcal{S}$ with intensity measure Λ (which is finite on every compact set and has no atoms), as a point process on $\mathcal{S}$ such that, for every compact set $\mathcal{B} \subset \mathcal{S}$, the count $N(\mathcal{B})$ has a Poisson distribution with mean $\Lambda(\mathcal{B})$. In a spatial setting, one realization of the Poisson process consists of a random number of events (based on the underlying intensity value in the space) at random locations in the space R^3. When $\beta(\boldsymbol{\tau}) = \beta$, it is called a homogeneous Poisson process. In a non-homogeneous Poisson process, the intensity parameter β is not uniform across the region $\boldsymbol{\tau}$.

A Poisson process in general is used to model point data. It provides a likelihood value to every possible spatial pattern of points in a given space. One can incorporate prior knowledge about what kind of patterns are more likely to occur for a given application using this process. The Poisson intensity β is a function of the space on which the Poisson process is defined. For example, in our case, the space is 3D. The hill peaks that we want to model

using a Poisson process occur at varying depths in different parts of the stack. We therefore need a Poisson intensity defined across the 3D space. There are many ways in which this Poisson intensity parameter $\beta(\tau)$ can be specified. The simplest way is to assume that $\beta(\tau)$ is a constant across the space τ. This means that the probability of occurrence of a point is equal in all parts of the space. If one knows that specific parts of the space τ have a higher probability of occurrence of a point, a higher β value can be provided for that part. An example where was chosen as the prior on Poisson intensity along the depth of the z-stack is shown in Figure 5. The expected number of points in the space τ is given by $\boldsymbol{\beta} = \int_{\tau} \beta d\tau$. We take advantage of these properties of a Poisson process resulting in a novel natural formulation for shape.

B. Shape Model

Each location (determined from the outcome of a Poisson process) is associated with a shape. In our case, the dermal-epidermal junction (DEJ) is composed of hills and valleys (see Figure 2). One can model the shape prior to be a hill peak. One could alternatively model the DEJ as a union of valleys (inverted hills). However, the basin of these inverted hills (valleys) is not always present in the RCM stack, as the stack is sometimes not deep enough to include the bottom of the valley. This makes modeling and inference harder, so we chose to model the hill peak. The height, orientation and size of the hill associated with each peak location is variable and this has to be captured in the shape prior of the model.

On the *en-face* ($x-y$) plane, the hill is an ellipse. The major and minor axis are variable and also change as a function of depth (they increase with depth in the z-direction). We propose the following 3D shape model for each hill to capture these properties, where each hill peak location is denoted by $[l_x, l_y, l_z]$ (outcome of the Poisson process). The ellipse has a center ($[l_x, l_y, l_z]$), a major axis (ρ_{ma}), a minor axis (ρ_{mi}), and an orientation (ρ_o). Let $\rho_{ma,o}$, $\rho_{mi,o}$ both greater than zero represent the major and minor axis value at the peak of the hill. Therefore, the hill is assumed to have a flat top with strictly positive initial major and minor axis values. As we go down the hill along the z-stack, the major and minor axes parameters increase as a function of depth. Note that this increase is nonlinear (Figure 1) where the highest rate of increase occurs closer to the tip of the peak. This structure is captured by utilizing a log rate of increase in the major and minor axis, within which variation is controlled by parameters $\rho_{ma,z}$ and $\rho_{mi,z}$. The value of major ($\rho_{ma,z}$) and minor ($\rho_{mi,z}$) axis at each depth z of the RCM stack is given by Equation 1.

$$\rho_{ma,z}=\rho_{ma,o}+(w_{ma} \times \log(l_z - z)), 1 \leq z \leq l_z \quad (1)$$

$$\rho_{mi,z}=\rho_{mi,o}+(w_{mi} \times \log(l_z - z)), 1 \leq z \leq l_z$$

The shape parameters of each hill are represented by $\mathbf{s} = [\rho_{ma,o}, \rho_{mi,o}, \rho_o, w_{ma}, w_{mi}]$. A prior distribution on these parameters is chosen to capture the domain knowledge of possible shapes. An example of a random sample drawn from one hill shape model is shown in Figure 6. Note that we model the shape of the actual DEJ as a union of several hills. This is

consistent with the natural structure of DEJ. An example random sample from this union model presenting a possible DEJ boundary is shown in Figure 7.

C. Appearance Model

An appearance model captures the likelihood of a given voxel/pixel to belong to either epidermis or dermis. One can either employ the intensity values directly or utilize higher level features such as texture to calculate this likelihood value.

Because of the high degree of variability in these data, parameters of the appearance model are determined adaptively for every image stack and inference is performed in an unsupervised setting. A given RCM stack is determined to be from either “dark” or “fair” skin type by the clinician when it is captured. For dark skin, intensity values themselves are used as appearance features. Intensity in the range [0, 255] is utilized. For fair skin, higher level features are extracted because fair skin images have lower intensity contrast than dark skin, making segmentation difficult based on intensity alone. Clinicians rely on the variation in texture of the voxels corresponding to dermis and epidermis to locate DEJ visually in fair skin stacks, and that motivated us to extract texture parameters for our appearance model. We extracted these standard features over a sliding window of size [50 × 50] since that would cover at least two basal cell thicknesses (which is $10 – 15 \mu m$ in diameter each corresponding to 20 – 30 pixels). Standard deviation, contrast, energy and two Gabor filter features are used (total of 5 features) for fair skin.

We assumed a Gaussian distribution for the appearance model (a multinomial Gaussian distribution for the fair skin type model). In order to calculate the parameters of this model (mean and variance/covariance), K-means clustering was performed on the features from bottom 50% of slices of the z-stack since the slices at the bottom of the z-stack are generally both dermis and epidermis voxels. Two clusters are obtained using K-means and the cluster that shows higher probability value for pixels in the bottom layer of the z-stack is chosen as the one belonging to dermis and is used to calculate the dermis appearance model parameters. Voxels in the other cluster are used to estimate the parameters of epidermis appearance model. The clustering result from K-means might be approximate and not fully consistent with respect to the structure of DEJ. However, this sensitivity is mitigated by the integration of appearance with the shape and spatial Poisson models.

To put this all together with more precision, let $\mathbf{x}_n$ represent the feature vector of each voxel. Let the parameters of the Gaussian distribution representing the epidermis appearance model (voxels above the DEJ boundary) be given by mean $\boldsymbol{\mu}_e$ and variance $\boldsymbol{\Sigma}_e$ and the dermis appearance model be given by mean $\boldsymbol{\mu}_d$ and variance $\boldsymbol{\Sigma}_d$. Let the appearance parameters of the model be denoted by $\alpha_e = [\boldsymbol{\mu}_e, \boldsymbol{\Sigma}_e]$ and $\alpha_d = [\boldsymbol{\mu}_d, \boldsymbol{\Sigma}_d]$, and all the appearance parameters be denoted by $\alpha = [\alpha_e, \alpha_d]$ for notational convenience. Once the location and shape of each hill is known, a boundary/partitioning of the $3D$ space is defined. The likelihood of the image given all the parameters is given by Equation 2.

$$p(\mathbf{X}|M,\mathbf{L},\mathbf{S},\alpha)=\prod_{n=1}^{N}[p(\mathbf{x}_n|\alpha_d)]^{I_n}[p(\mathbf{x}_n|\alpha_e)]^{1-I_n} \quad (2)$$

where, I_n is a random indicator variable that takes a value 1 when the pixel n is below the boundary (dermis) and 0 otherwise. Note that this random variable is a deterministic function of **L** and **S**. Therefore, the likelihood here is a product of the probability of each pixel in the RCM stack to belong to dermis/epidermis, based on the boundary defined by **L** and **S**.

D. Overall Model

For notational convenience, let $\theta = [\beta(\tau), \zeta, \alpha]$. The joint likelihood of the model is given by Equation 3.

$$p(M,\mathbf{L},\mathbf{S},\mathbf{X}|\theta)=p(M,\mathbf{L}|\beta(\tau))p(\mathbf{S}|M,\mathbf{L},\zeta)\ p(\mathbf{X}|M,\mathbf{L},\mathbf{S},\alpha) \quad (3)$$

where, ζ represents the hyper-parameters of the prior distribution on shape parameters. We call our new model **MPP-skin**.

III. Inference

Inference of the DEJ boundary, given the observation **X** involves inferring the posterior distribution of the parameters *M,* **L** and **S**, which is described in this section. Note that we do not know the dimension of the parameters **L, S** as the value of *M* is unknown, and a part of the inference. Reversible jump Markov chain Monte Carlo (RJMCMC) sampling is typically utilized in such a setting. However, it involves designing appropriate proposal distributions and selection of parameters for good acceptance ratio. Gibbs sampling on the other hand involves sampling directly from the conditional posterior distribution assuming that all the other random variables are known. This is a special case of Metropolis-Hastings algorithm where all proposals are always accepted [18]. We therefore make a design choice of using Gibbs sampling for inference on our model. Unfortunately, Gibbs sampling does not naturally lend itself to trans-dimensional jumps. In order to utilize Gibbs sampling for inference in our model, we leverage the independence property of the Poisson process.

A. Independence property of Poisson process

One of the useful properties of the Poisson process is that it follows the independence property [19]. This property states that, *if the Poisson process is defined on a locally compact metric space* $\mathcal{S}$, *where* $\mathcal{B} \in \mathcal{S}$ *and* $\mathcal{B}_1, \ldots, \mathcal{B}_m$ *are disjoint compact sets, then* $N(\mathcal{B}_1), \ldots, N(\mathcal{B}_m)$ *are independent* [20]. This property is presented in Definition 1.3 of [20]. In the current setting, this implies that the number of points (hill s) in any small volume ($\Delta V \subseteq R^3$) of the 3D space R^3, follow a Poisson distribution with intensity parameter $\beta = \int_{\Delta V} \beta(\tau) d\tau(\Delta\tau)$. Let the disjoint compact volume in the current setting belong to each voxel in the 3D space R^3.

This implies that the probability of occurrence of a peak at any voxel n can be modeled as an independent random variable (that follows a Poisson distribution) with Poisson intensity parameter $\beta(n)$, where $\beta(n) = \int_{\tau(n)}\beta(\tau)d\tau(n)$. Let $\mathbf{G} = [g_1, \ldots, g_N]$ be a random variable representing the Poisson outcome at each voxel n. Each g_n is therefore a discrete random variable representing the number of hill peaks at a particular index n. Note that the dimension of G is the same as the number of voxels in the image (N) and each g_n indicates the number of hill peaks at voxel n. Note also that objects can overlap at a given voxel location.

The resulting model, after introduction of **G** is given by Equation 4.

$$p(\mathbf{G},\mathbf{S},\mathbf{X}|\cdot)=p(\mathbf{G}|\beta(\tau))p(\mathbf{S}|\mathbf{G},\boldsymbol{\zeta})p(\mathbf{X}|\mathbf{G},\mathbf{S},\alpha). \quad (4)$$

The number of hills in the z-stack M is now given by $\sum_{n=1}^{N} g_n$. Note that **L** is deterministic given **G**. Let $\boldsymbol{\beta} = \int_{\tau}\beta(\tau)d(\tau)$. A graphical representation of this model is shown in Figure 8.

$$p(\mathbf{G}|M)=\frac{p(\mathbf{G})}{p(M|\beta(\tau))}=\frac{\prod_{n=1}^{N}\frac{\exp(-\beta_n)\beta_n^{g_n}}{g_n!}}{\frac{\exp(-\boldsymbol{\beta})(\boldsymbol{\beta})^{M}}{M!}}=\frac{M!}{g_1!\ldots,g_N!}\left(\frac{\beta_1}{\boldsymbol{\beta}}\right)^{g_1}\ldots\left(\frac{\beta_N}{\boldsymbol{\beta}}\right)^{g_N}$$

$$p(g_n|\mathbf{G}_{-n},\boldsymbol{\beta})=\frac{\exp(-\beta_n)\beta_n^{g_n}}{g_n!} \quad (5)$$

B. Gibbs Sampling

Gibbs sampling is a special case of the Metropolis Hastings algorithm where inference involves drawing samples from the posterior distribution of each random variable assuming all the other random variables of the model are known. This is performed sequentially until convergence is achieved. We describe the sampling process for each random variable (**G, S**) of the model in this section.

$p(g_n|\mathbf{G}_{-n}, \beta_{1:N})$ is a Poisson random variable with intensity β_n and is given by Equation 5. The sampling step for the variable g_n given all the other parameters and observations of the model ({**X, S,** $\boldsymbol{\zeta}$, α}) is given by Equation 6.

$$p(g_n=k|\mathbf{G}_{-n},\cdot)\propto\begin{cases} p(k|\mathbf{G}_{-n},\beta_n)p(\mathbf{X}|\cdot), \text{if } k\leq k_p \\ p(k|\mathbf{G}_{-n},\beta_n)\cdot \\ \int p(\mathbf{X}|\cdot)p(\mathbf{s}_{M+1}|\boldsymbol{\zeta})ds, \text{if } k=k_p+1 \end{cases} \quad (6)$$

where k_p depicts the number of objects in the current state at voxel g_n. Sampling the random variable g_n involves three cases: remaining in the same state ($g_n = k_p$), jumping to a state ($g_n = k_p - 1$) with lower number of points (assuming $k > 0$) and adding a point ($g_n = k_p + 1$). Note that k here represents the number of hill peaks in the current state at voxel g_n and k_p

represents the number of hill peaks observed in the previous state. This involves calculating the likelihood for the three cases followed by normalization. After normalization, we have a discrete probability distribution. g_n is sampled from this distribution. Equation 6 shows the distributions for deleting and adding a point. For a case where there are multiple points at g_n, a random point is chosen to be deleted. Note that this type of sampling is very similar to the Gibbs sampling framework for Dirichlet process mixtures [21] also popularly known as the Chinese restaurant process (CRP). In the CRP, the number of clusters are unknown (modeled by a Dirichlet process prior) and each cluster is associated with a fixed number of parameters (mean and variance). Our method on the other hand consists of an unknown number of peaks at random locations modeled as a spatial Poisson process prior, and each location/peak is associated with shape parameters.

The sampling step for the variable $\mathbf{s}_m$ is given by Equation 7.

$$p(\mathbf{s}_m|\mathbf{S}_{-m},\cdot) \propto p(\mathbf{X}|\mathbf{G},\mathbf{S},\alpha)p(\mathbf{s}_m|\zeta) \quad (7)$$

The shape variables are assumed to be independent of each other. The shape prior $p(\mathbf{s}_m|\zeta)$, is not conjugate to the likelihood of the model, $p(\mathbf{X}|\mathbf{G}, \mathbf{S}, \alpha)$. This implies that there is no closed form solution to Equation 7 and thus a sampling-resampling technique is utilized to approximate the posterior distribution. Convergence is defined experimentally as when the first order statistics of the random variables being sampled do not change more than a preset small threshold ε. Inference of the parameters G and S determines the segmentation of the dermal-epidermal junction in RCM images.

Implementation Details—For approximating the posterior distribution of shape variables, a fixed number (300) of shapes from the prior distribution were generated and their corresponding likelihood value calculated. Shape parameters were then sampled from these candidate shapes, by sampling from a discrete distribution (representing each shape model) whose weights were the normalized scores of Equation 7.

When sampling for **G**, we used a hybrid Gibbs sampler, starting with a blocked Gibbs sampler and then transitioning to single site sampling. Blocks of pixels in each slice of the 3D stack were sampled for the value of **G** at first. Blocks that show very low likelihood for the occurrence of a hill peak were removed from further sampling. In other words, these were the pixels whose appearance is closer to background than foreground model. This is a way to get rid of obvious non-object center locations and helps save time on unnecessary sampling of obvious non-object pixels for the value of **G**. After one sweep of a blocked sampling approach on the entire stack, we were left with a set of blocks that have higher potential for occurrence of a hill peak in them. As a next step, we sampled every pixel location in these blocks for the value of **G**. The algorithm is implemented in MATLAB on a 12GB RAM, 2.40GHz Intel quadcore computer with Windows OS and processing time varies from 3–12 hours depending on the data where each stack consisted of 45 – 70 slices with a resolution of 1000 × 1000 pixels in each slice.

IV. Experiments and Results

We validated our proposed model on nine stacks of dark skin and fifteen stacks of fair skin, for all of which we have ground truth (expert labeling of the DEJ). These stacks of images were collected *in vivo* with a reflectance confocal microscope. Imaging was done on healthy skin of volunteers. No special preparation was needed before imaging. We only apply baby oil on the imaging area for optical index matching purposes. We used a commercial version of the microscope developed for skin imaging (Vivascope 1500, Caliber Imaging and Diagnostics, Rochester, NY), which is routinely used for clinical imaging and reported clinical studies [22], [23], [11], [24]. Images used in the study are collected from arm and trunk of relatively young and healthy skin volunteers between the age of 20–50. This version of the microscope has an imaging window that sticks to the area of interest. The field of view of a single RCM image is $0.5\mu m \times 0.5\mu m$.

Each stack is from a single skin site and consists of 45–70 image slices, where each slice is 1000×1000 pixels, with a pixel (lateral) resolution of $0.5\mu m$ and optical sectioning (axial resolution) of $3\mu m$. The axial distance between consecutive slices (z-step) is $1\mu m$, which results in higher sampling than the optical sectioning thickness of the imaging system. The values of the hyperparameters (parameters of Poisson and shape priors) were chosen based on these experimental parameters, as we describe next.

A Gamma distribution was chosen to model the variation of Poisson intensity $\beta(\tau)$ along the depth of z-stack as its shape captured the prior information we have about the depth range where the DEJ usually starts. The DEJ starts in the range $50–100\mu m$ below the skin surface, which is where we expect to see the hill peaks. The Poisson intensity $\beta(\tau)$ should be such that it gives a higher probability to the RCM slices in this range (which corresponds to hill peaks approximately in the range 25–50th slice in depth here, as the acquisition of images starts well into the epidermis). Therefore, we chose to use the hyper-parameter values 10, 3 (shape and scale parameter of a Gamma distribution) for the Gamma distribution ($\beta(\tau_z) \sim \Gamma(10, 3)$) prior. This corresponds to a shape which gives maximum probability to slices in the range 25 – 50, as illustrated in Figure 5. Since in general we did not have prior knowledge about the regions in each slice where the DEJ has high/low probability of occurrence of hill peaks, we chose a uniform Poisson intensity prior in the $x - y$ plane. This implies that two voxels belonging to different layers have different Poisson intensity values while the voxels belonging to the same layer have the same value.

In our experiments, shape parameters $\mathbf{s} = [\rho_{ma,o}, \rho_{mi,o}, \rho_o, w_{ma}, w_{mi}]$ were modeled by prior distributions that can capture the prior knowledge of all possible shapes. Parameters of the shape model, minor (ρ_{mi}) and major axis (ρ_{mi}) were assumed to be generated from a Gamma distribution. In our case, the average thickness of a basal cell is 10 – $15\mu m$. Basal cells form the basal layer which is the bottom layer of epidermis, lying just above the DEJ. We assumed a Gamma distribution with the hyper-prior values 5 and 5 as this ensures a positive value with highest probability at 10.

$$\rho_{ma} \sim \Gamma(\lambda_{ma}, \eta_{ma}) \quad (8)$$

$$\rho_{mi} \sim \Gamma(\lambda_{mi}, \eta_{mi})$$

There is no preferential distribution of orientations and rate of increase of hills. Orientation (ρ_o) and rate of increase of major and minor axes (w_{ma}, w_{mi}) are therefore assumed to be generated from a uniform distribution. Uniform distributions for w_{ma}, w_{mi} were chosen to be between (2, 20) since a range beyond 20 as the upper limit did not bring significant change in the shape of the ellipsoid.

$$\rho_o \sim U[0, \pi] \quad (9)$$

$$w_{ma} \sim U[w_{\min}, w_{\max}]$$

$$w_{mi} \sim U[w_{\min}, w_{\max}]$$

Values of all the hyper-parameters used in the experiments are reported in Table I

Performance of the algorithm was measured in terms of the absolute value of the error in localization of the DEJ along the z-axis for each voxel compared to ground-truth and its mean and variance were reported. Ground-truth annotations were given by consensus labeling by three experts.

Table II reports the results of our algorithm (*MPP-skin*) for dark skin RCM image stacks and Table III presents the results on the fair skin stacks. The tables report the mean and variance of the absolute value of distance between DEJ as detected by our algorithm and the expert labels. Down arrows in the table indicate that lower values are desirable. For comparison, we include the results from our previous approach *Kurugol et al.* [4], [16]. For reference, we also ran a standard graphcuts algorithm [25] for both skin types and report those results as well. Results from Table II and III showed that the graphcuts algorithm performs poorly compared to MPP-skin and [4], [16] for both dark and fair skin types. Results from Table II show that the proposed *MPP-skin has equivalent performance compared to Kurugol et al.* [4], [16], *with a small advantage in average accuracy. For the fair skin case, the epidermis-transition and transition-dermis error values were reported in* [4], [16] *rather than a single DEJ.* However, we now have groundtruth labels available for a single DEJ junction rather than two junctions with uncertainty region, which was not the case at the time of the previous publication. We calculated the error value of the junctions reported in [4], [16] to this new single DEJ junction and reported the results in Table III. The average error values for fair skin are 9.71μm for epidermis-transition boundary (to the true DEJ boundary) and 9.67μm for the transition-dermis boundary (to the true DEJ boundary). On the other hand, our algorithm reports an average error of 12.1μm. All of these errors are less than the size of single basal cell away from the boundary (basal cells are generally approximately 10 to 15 microns in height). Thus the accuracies can be considered comparable.

Figures 9 and 10 show an example DEJ boundary marked by an expert and automatically detected by our algorithm, *MPP-skin* on an *en-face* slice of image stack 1 and stack 12 in dark and fair skin respectively. Figure 11 shows a surface $3D$ plot of the DEJ detected by our algorithm on a sample dark and fair skin stack. Color indicates how far the DEJ is from the expert labels. Blue indicates smaller error and red indicates larger error.

V. Conclusion

Segmenting boundaries of interest in 3D microscopy images is often challenging due to high intra and inter-subject (or specimen) variability, and the complexity of the boundary structures. Incorporation of shape information about the object of interest can often help improve segmentation performance. In addition, the detection problem usually involves locating multiple occurrences of similar objects, where the locations are sufficiently variable that stochastic models are required. We present a novel strategy utilizing a marked Poisson process to address this problem and report on our application for automatically detecting the DEJ in 3D reflectance confocal microscopy images of human skin. The multiple hills and valleys topography of the DEJ was addressed by developing an unsupervised probabilistic model that incorporates not only appearance of the skin in RCM images, but also the shape, using a marked spatial Poisson process. Inference on the algorithm was accomplished using Gibbs sampling. Experimental results on *in vivo* measured data from human skin are presented for both dark (less challenging) and fair (more challenging) skin types. We show that this model is capable of detecting the DEJ for both skin types with an average error of $5.41\mu m$ and $12.1\mu m$ for dark and fair respectively, compared to expert labeling, well within the range of clinical utility.

MPP-skin offers several advantages over our previous [4], [16] approach. First, Kurugol et al.'s approach does not detect the DEJ in fair skin and instead detects an uncertainty region between the epidermis and dermis. In contrast, our MPP-skin approach detects the DEJ in both skin types. Second, Kurugol et al.'s approach involves classification of tiles based on the extracted features followed by smoothing rather than a model for the DEJ, whereas, the MPP-skin provides an explicit model of the boundary. Third, [4], [16] is a supervised method (involving several SVM components for the fair skin model), whereas MPP-skin requires selection of distributions and parameters based on prior knowledge, but does require detailed labeling for parameter estimation. This modeling can be based on physically and physiologically meaningful structure rather than "blindly" learning parameters from data. Thus, if the structures are different in other data (stacks with different tissue structure), these differences can be incorporated meaningfully. Indeed, although the results reported here for the older method are based on a strict separation of training and test data, our experience has been that the method, because it relies on training a classifier, has been at times unreliable to train appropriately. On the other hand, the MPP-skin model relies only on the basic structural architecture of the DEJ with only a small number of physiologically meaningful fixed parameters, and thus should generalize more reliably. If there is a reason to believe that DEJ structure varies (e.g. with age or location), adapting MPP-skin to such settings would be very straightforward.

The differences may be even more significant when the approaches are applied to diseased skin, where the DEJ may be disrupted or otherwise abnormal. In case of the supervised method, if this causes unexpected results, we anticipate that it will be difficult to distinguish between the presence of disease due to difficulty in appropriately training the model to capture normal variability. On the other hand, with the current generative model, non-standard results or difficulty in fitting the model to data may well be a useful biomarker of abnormality in the skin structure.

Automation in DEJ delineation is a major step in the standardization of the image acquisition process. In the short term, standardization of mosaicing will lead to more rapid and repeatable imaging, both across locations on the body and longitudinally, within a single subject and across subjects. Shorter imaging sessions will increase patient comfort and save expense. In the longer term, standardization in data acquisition will make the adaptation of RCM imaging easier and enable the widespread use of this technology in clinics. Moreover, data collected in a standardized manner will facilitate larger cross-site studies, which will eventually lead to increase in sensitivity and specificity of RCM imaging in diagnosis. In this context, computational time could be a limitation which can be improved by using GPU based parallelization. In addition, the algorithm depends to an unknown degree, on the spacing of the slices in depth, which incurs a tradeoff between that resolution and imaging time. Finally, there are limitations to the shape model employed that we will try to generalize in the future.

Although we designed this model to solve the DEJ segmentation task, the underlying model can be applied to other multiple object segmentation applications utilizing the appropriate shape and appearance models for those domains. For instance, it could be used for detection and segmentation of multiple cells in both 2D and 3D microscopy images [3], [2], [26]. The shape and appearance models for these applications could either be determined from training data or domain knowledge of the application. For example, in 2D/3D cell microscopy or histopathology image data [3], [17], a parametric shape model of ellipses (ellipsoids in case of 3D [2], [26]), and a Gaussian model of intensity for appearance could be utilized. One could alternatively use a point distribution model parameterized by landmark points around the boundary as well [27]. In these examples, the parametrization of shape and appearance would change, while the underlying model of the marked Poisson process remains the same.

Acknowledgments

We thank Miguel Cordova for his help in labeling the RCM stacks that we used in this study, and Dr. Giovanni Pellacani for extensive discussions and guidance. This project was supported in part by several NIH institutes: National Cancer Institute (NCI) (R01CA156773), National Heart, Lung, and Blood Institute (NHLBI) (R01HL089856), National Institute of General Medical Sciences (P41GM103545), NIBIB's Image Guided Interventions program (R01EB012466) and the Provost office of Northeastern University (Graduate Dissertation Completion Fellowship in 2014).

References

1. Decencière E, Tancrède-Bohin E, Doklàdal P, Koudoro S, Pena A-M, Baldeweck T. Automatic 3D segmentation of multiphoton images: a key step for the quantification of human skin. Skin Research and Technology. 2013; 19(2):115–124. [Online]. Available: http://dx.doi.org/10.1111/srt.12019. [PubMed: 23441573]

2. Sarti A, Solorzano COD, Lockett S, Malladi R. A geometric model for 3D confocal microscope image analysis. IEEE Transactions on Biomedical Engineering. 2000; 47:1600–1609. [PubMed: 11125595]
3. de Solorzano CO, Rodriguez EG, Jones A, Pinkel D, Gray JW, Sudar D, Lockett SJ. Segmentation of confocal microscope images of cell nuclei in thick tissue sections. Journal of Microscopy. 1999; (Pt 3, 193):212–226.
4. Kurugol S, Kose K, Park B, Dy JG, Brooks DH, Rajadhyaksha M. Automated delineation of dermal-epidermal junction in reflectance confocal microscopy image stacks of human skin. Journal of Investigative Dermatology. 2014; 135(3):710–717. [PubMed: 25184959]
5. Lagarrigue SG, George J, Questel E, Lauze C, Meyer N, Lagarde J-M, Simon M, Schmitt A-M, Serre G, Paul C. In vivo quantification of epidermis pigmentation and dermis papilla density with reflectance confocal microscopy: variations with age and skin phototype. Experimental Dermatology. 2012; 21(4):281–286. [Online]. Available: http://dx.doi.org/10.1111/j.1600-0625.2012.01451.x. [PubMed: 22417304]
6. Kulikova, M., Veillard, A., Roux, L., Racoceanu, D. Proceedings in SPIE Medical Imaging. San Diego, California, USA: 2012 Feb. Nuclei extraction from histopathological images using a marked point process approach.
7. Gonzalez, S., Gill, M., Halpern, AC. Reflectance Confocal Microscopy of Cutaneous Tumors: An Atlas with Clinical, Dermoscopic and Histological Correlations. Taylor & Francis: 2008. [Online]. Available: http://books.google.com/books?id=tZMTomd9TagC
8. Weedon, D. Chapter 32 - Lentigines, nevi and melanomas. New York: Churchill Livingstone; 1997.
9. Mizukoshi K, Yonekura K, Futagawa M, Nakamura T, Hirayama K, Takahashi K. Changes in dermal papilla structures due to aging in the facial cheek region. Skin Research Technology. Epub 2014 Aug 23.
10. Burgeson RE, Christiano AM. The dermal-epidermal junction. Current Opinion in Cell Biology. 1997; 9(5):0955–0674.
11. Gill M, Longo C, Farnetani F, Cesinaro AM, Gonzalez S, Pellacani G. Non-invasive in vivo dermatopathology: identification of reflectance confocal microscopic correlates to specific histological features seen in melanocytic neoplasms. Journal of the European Academy of Dermatology and Venereology. 2014; 8:1069–1078.
12. Wolff, K., Kibbi, A-G., Mihm, J Martin C. Chapter 4 Basic Pathological Reactions of the Skin. In: Freedberg, IM.Eisen, AZ.Wolff, K.Austen, KF.Goldsmith, LA., Katz, SI., editors. Fitzpatricks Dermatology in General Medicine (6th Edition). New York: McGraw-Hill; 2003.
13. Langley, RGB., Barnhill, RL., Mihm, MC., Fitzpatrick, TB., Sober, AJ. Fitzpatricks Dermatology in General Medicine (6th Edition), Chapter 93 - Neoplasms: Cutaneous Melanoma. Freedberg, IM.Eisen, AZ.Wolff, K.Austen, KF.Goldsmith, LA., Katz, SI., editors. New York: McGraw-Hill; 2003.
14. Rao BK, Mateus R, Wassef C, Pellacani G. In vivo confocal microscopy in clinical practice: comparison of bedside diagnostic accuracy of a trained physician and distant diagnosis of an expert reader. Journal of the American Academy of Dermatology. 2013; 6(69):295–300.
15. Stevenson AD, Mickan S, Mallett S, Ayya M. Systematic review of diagnostic accuracy of reflectance confocal microscopy for melanoma diagnosis in patients with clinically equivocal skin lesions. Dermatology Practical and Conceptual. 2013; 3(4):19–27. [PubMed: 24282659]
16. Kurugol S, Dy JG, Brooks DH, Rajadhyaksha M. Pilot study of semiautomated localization of the dermal/epidermal junction in reflectance confocal microscopy images of skin. Journal of Biomedical Optics - SPIE. 2011; 3(16)
17. Avenel C, Kulikova MS. Marked point processes with simple and complex shape objects for cell nuclei extraction from breast cancer H&E images. Proceedings in SPIE Medical Imaging. 2013; 8676:86 760Z–86 760Z–7. [Online]. Available: http://dx.doi.org/10.1117/12.2006634.
18. Murphy, KP. Machine Learning: A Probabilistic Perspective. Cambridge, MA: The MIT Press; 2012.
19. Cox, D., Isham, V. ser. Monographs on applied probability and statistics. London, New York: Chapman and Hall; 1980. Point Processes.

20. Baddeley, A. Spatial point processes and their applications. In: Weil, W., editor. ser. Stochastic Geometry: Lectures given at the C.I.M.E. Summer School held in Martina Franca, Italy, September 13–18, 2004. Berlin, Heidelberg: Springer Berlin Heidelberg; 2007.
21. Neal RM. Markov chain sampling methods for dirichlet process mixture models. Journal of Computational and Graphical Statistics. 2000; 9(2):249–265.
22. Nori S, Rius-Diaz F, Cuevas J, Goldgeier M, Jaen P, Torres A, Gonzalez S. Sensitivity and specificity of reflectance-mode confocal microscopy for in vivo diagnosis of basal cell carcinoma: A multicenter study. Journal of American Academy of Dermatology. 2004; 51(6):923–930. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S0190962204017281.
23. Guitera P, Menzies SW, Longo C, Cesinaro AM, Scolyer RA, Pellacani G. In vivo confocal microscopy for diagnosis of melanoma and basal cell carcinoma using a two-step method: analysis of 710 consecutive clinically equivocal cases. Journal of Investigative Dermatology. 2012; 132(10):2386–2394. [PubMed: 22718115]
24. Alarcon I, Carrera C, Palou J, Alos L, Malvehy J, Puig S. Impact of in vivo reflectance confocal microscopy on the number needed to treat melanoma in doubtful lesions. British Journal of Dermatology. 2014; 4(170):802–808.
25. Yuan, J., Bae, E., Tai, X-C. A study on continuous max-flow and min-cut approaches; Computer Vision and Pattern Recognition (CVPR), 2010 IEEE Conference on; 2010 Jun. p. 2217-2224.
26. Meng Y, He Y, Wu J, Chen S, Li A, Gong H. Automatic detection and quantitative analysis of cells in the mouse primary motor cortex. Proceedings of SPIE 9230, Twelfth International Conference on Photonics and Imaging in Biology and Medicine (PIBM). 2014; 9230:92 301E–14.
27. Cootes TF, Edwards GJ, Taylor CJ. Active appearance models. IEEE Transactions On Pattern Analysis And Machine Intelligence. 2001 Jun; 23(6):681–685.

Biographies

Sindhu Ghanta received her MS from Texas Tech University in 2010 and PhD in Electrical and Computer Engineering at Northeastern University, Boston, USA in 2014. She was a postdoctoral fellow at BIDMC and Harvard Medical School in the department of pathology after that, where she worked on detection and classification of features from histopathological (breast cancer) images. She is currently a research scientist at Parallel Machines. Her research interests are in Bayesian nonparametric models and parallelization of machine learning algorithms in distributed systems.

Michael I. Jordan is the Pehong Chen Distinguished Professor in the Department of Electrical Engineering and Computer Science and the Department of Statistics at the University of California, Berkeley. He received his Masters in Mathematics from Arizona State University, and earned his PhD in Cognitive Science in 1985 from the University of California, San Diego. He was a professor at MIT from 1988 to 1998. His research interests bridge the computational, statistical, cognitive and biological sciences. Prof. Jordan is a

member of the National Academy of Sciences, a member of the National Academy of Engineering and a member of the American Academy of Arts and Sciences. He is a Fellow of the American Association for the Advancement of Science. He has been named a Neyman Lecturer and a Medallion Lecturer by the Institute of Mathematical Statistics. He received the IJCAI Research Excellence Award in 2016, the David E. Rumelhart Prize in 2015 and the ACM/AAAI Allen Newell Award in 2009. He is a Fellow of the AAAI, ACM, ASA, CSS, IEEE, IMS, ISBA and SIAM.

Kivanc Kose Kivanc Kose received his M.Sc. and Ph.D. degrees from Electrical an Electronics Engineering Department at Bilkent University. He is currently working at Dermatology Service, Memorial Sloan Kettering Cancer Center as a post-doctoral research scholar. His research interests include, image and video processing, machine learning in image processing, computer vision and their applications in medical imaging. His current focus is in developing imaging-based diagnostic methods for confocal microscopy of tissue, especially skin. He has (co)authored several research articles in various engineering and medical journals (JID, IEEE Signal Processing Magazine), and refereed conferences (ICIP, ICAASP). He has been serving as an associate editor for “Signal, Image and Video Processing” journal. He has also been serving as a reviewer for several IEEE journals and conferences.

Dana H. Brooks is Professor of Electrical and Computer Engineering, Northeastern University, cofounder of the Biomedical Signal Processing, Imaging, Reasoning, and Learning (B-SPIRAL) group there, and co-Director of the Simulation and Estimation Core of the Center for Integrative Biomedical Computing headquartered at University of Utah. His research includes application of signal and image processing and machine learning to medical and biological imaging, modeling, inverse problems, and optimization.

Milind Rajadhyaksha is a mechanical/optical engineer, an Associate Member of the faculty in Dermatology, and directs a research program in confocal microscopy at Memorial Sloan Kettering Cancer Center. His work spans the entire spectrum from laboratory research

through commercialization through clinical studies to clinical implementation. He develops and translates confocal microscopes to noninvasively guide diagnosis and therapy of skin cancers, and enjoys working in the so-called “valley of death” (and living through near-death experiences) between laboratory and clinic and between academia and industry. Two of his microscopes have been commercialized (VivaScopes, Caliber Imaging and Diagnostics) and are now being implemented in clinics, with promising early impact on patient care. He is a co-founder of another company (ColdSteel Laser, LLC), to advance endoscopic laser surgery for head-and-neck cancers, and he wildly fantasizes about a future in which multimodal optical imaging may routinely guide diagnosis and treatment directly on patients, at the bedside, in real time and at low cost.

Jennifer G. Dy is a professor at the Department of Electrical and Computer Engineering, Northeastern University, Boston, MA, where she first joined the faculty in 2002. She received her M.S. and Ph.D. in 1997 and 2001 respectively from the School of Electrical and Computer Engineering, Purdue University, West Lafayette, IN, and her B.S. degree (Magna Cum Laude) from the Department of Electrical Engineering, University of the Philippines, in 1993. Her research spans both fundamental research in machine learning and data mining and their application to biomedical imaging, health, science and engineering, with research contributions in clustering, dimensionality reduction, feature selection and sparse methods, learning from complex data, learning from crowds and Bayesian nonparametric models. She received an NSF Career award in 2004. She has served as action editor/editorial board member for Machine Learning and the Journal of Machine Learning Research, organizing and or technical program committee member for premier conferences in machine learning and data mining (ICML, ACM SIGKDD, AAAI, IJCAI, AISTATS, UAI, SIAM SDM), and program chair for SIAM SDM 2013.

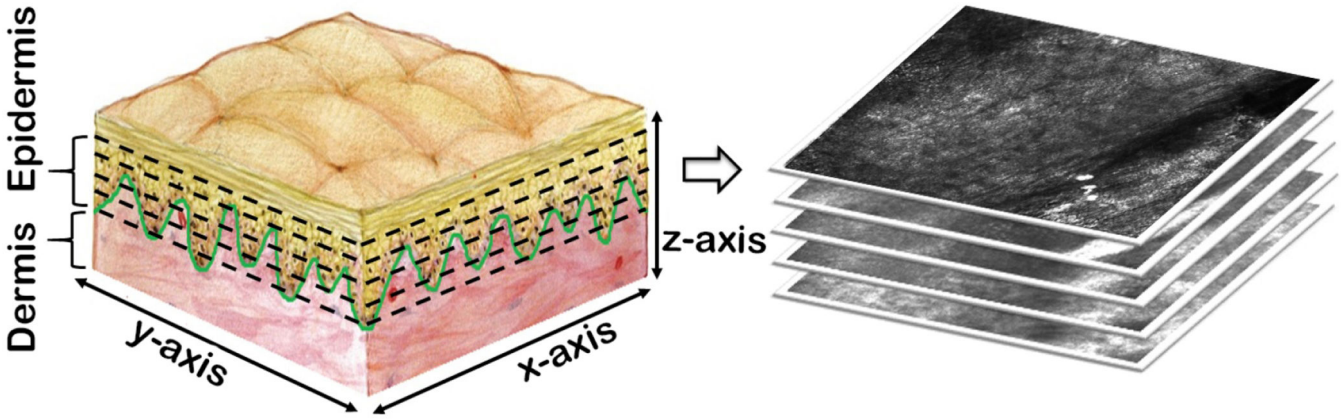


Fig. 1.

An illustration of the dermal-epidermal junction (DEJ) of human skin is shown in the left hand side figure. Green color indicates the DEJ boundary. Black dashed lines represent the section where RCM imaging captures slices. Right hand side figure shows a stack of these slices.

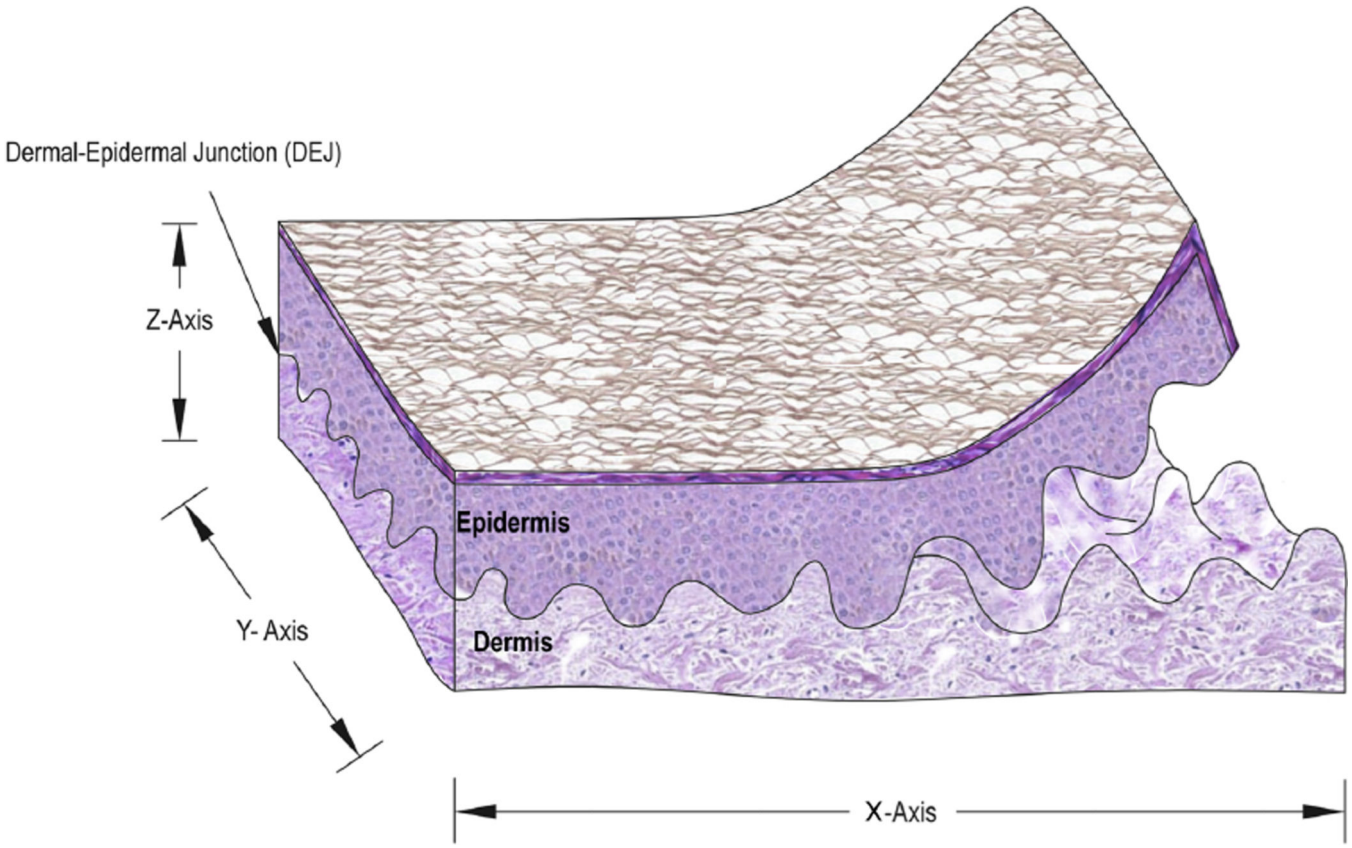


Fig. 2.
Hill and valley like structure of dermal-epidermal junction (DEJ) of human skin.

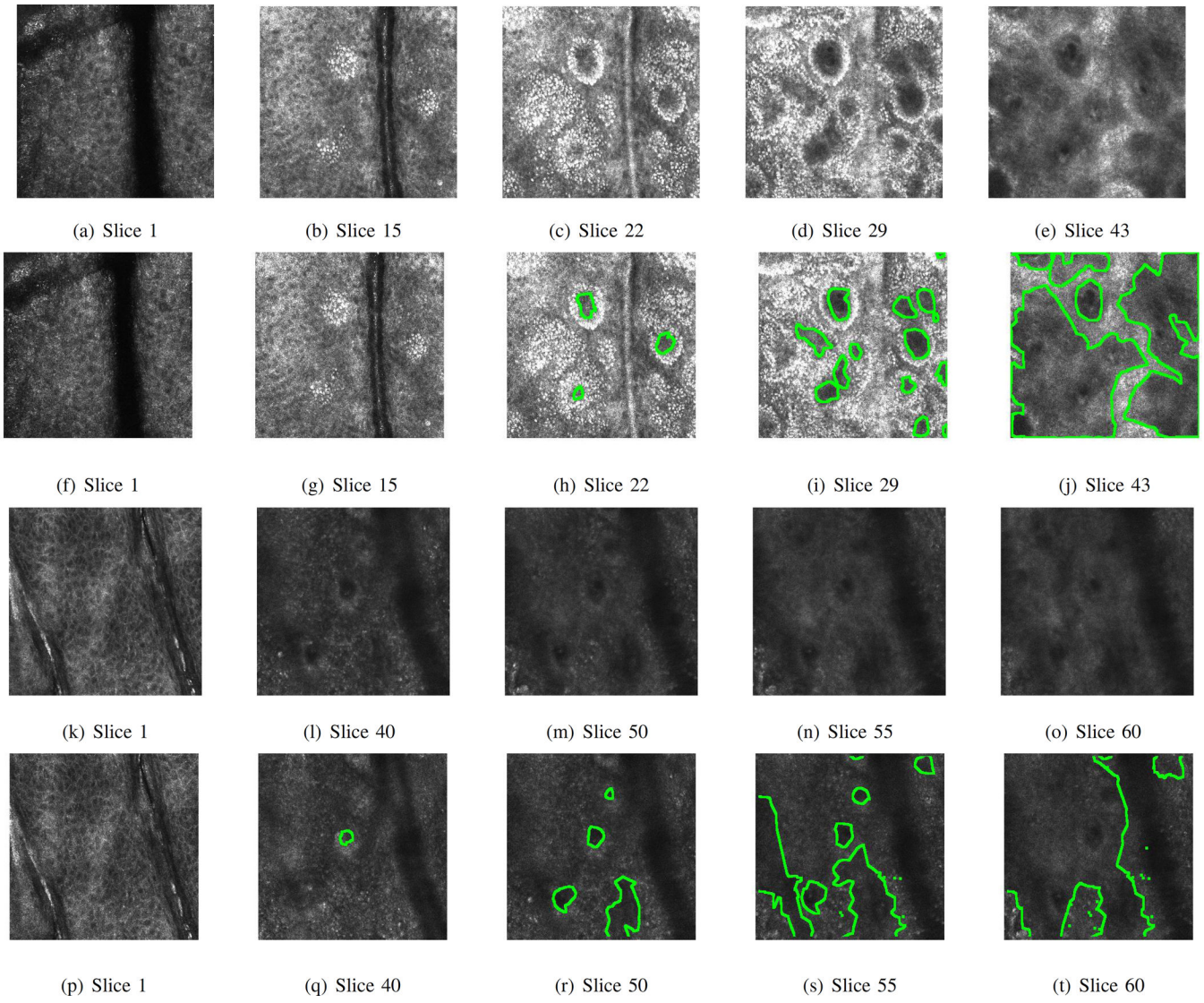
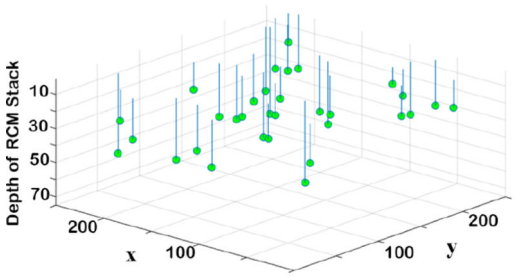
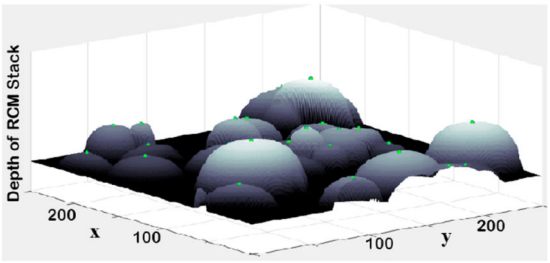


Fig. 3.

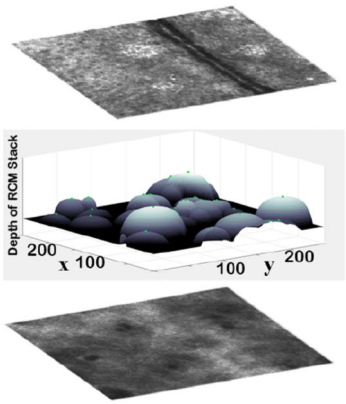
(a)–(e) and (k)–(o) show original unsegmented example *en-face* slices from a dark and fair skin stack respectively where the slice numbers go from top (epidermis) to bottom (dermis). (f)–(j) and (p)–(t) show the corresponding expert labels (shown in green) for dark and fair skin stacks respectively.



(a) Green dots represent points generated from a Poisson process.



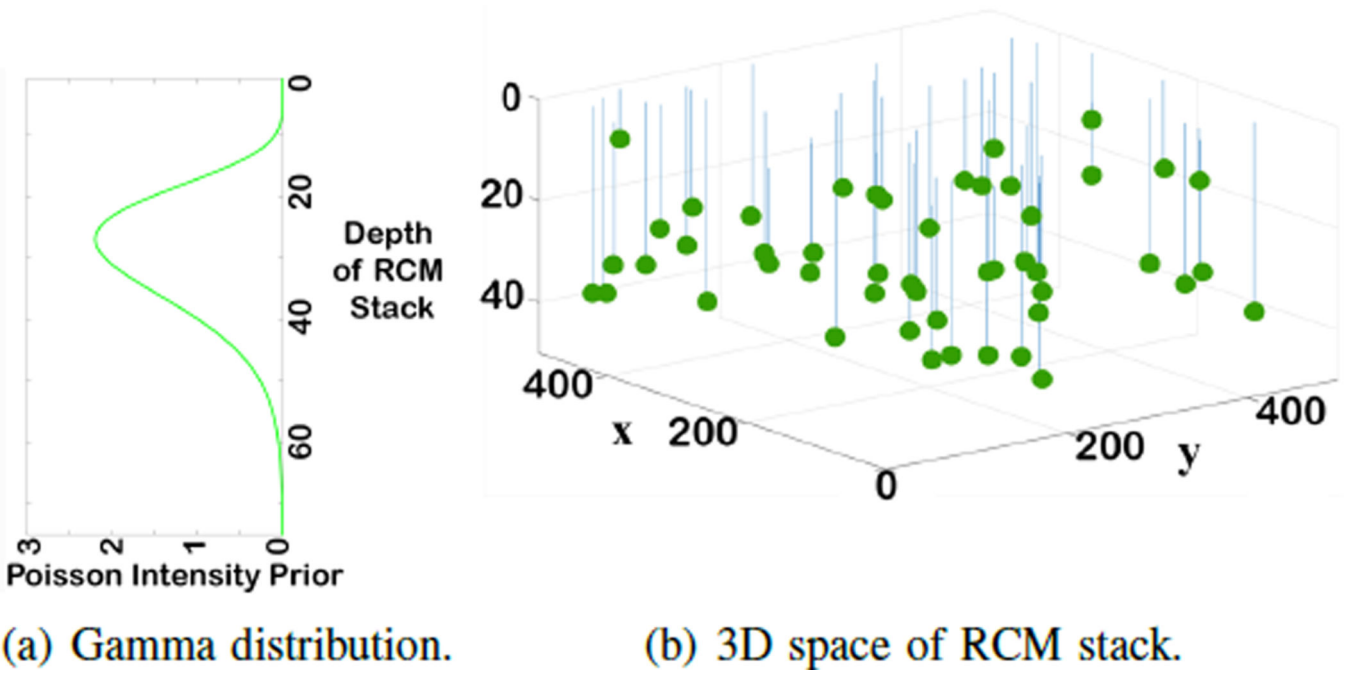
(b) Shapes are generated from the shape prior model at each location of the points generated in the previous step.



(c) Pixels belonging to dermis and epidermis are generated from an appearance model.

Fig. 4.

Generative model for dermis-epidermis segmentation. First, M points at random location are generated from a spatial Poisson process. The green circles represent these points. Lines in 4(a) that connect to these points are drawn for clarity in visualization of depth. An object boundary involving hills is generated from a shape model at each location followed by modeling of appearance of voxels above and below the shape boundary.



(a) Gamma distribution. (b) 3D space of RCM stack.

Fig. 5.

Poisson intensity along the z-axis varies according to a Gamma distribution$\Gamma(10, 3)$ shown in 5(a). Uniform Poisson intensity prior across the x–y plane is assumed with the integral Poisson intensity value $\sum_{\rho}\beta = 50$. Green circles in 5(b) represent points generated from a random outcome of this Poisson model. Lines hat connect to these points are drawn for clarity in visualization of depth.

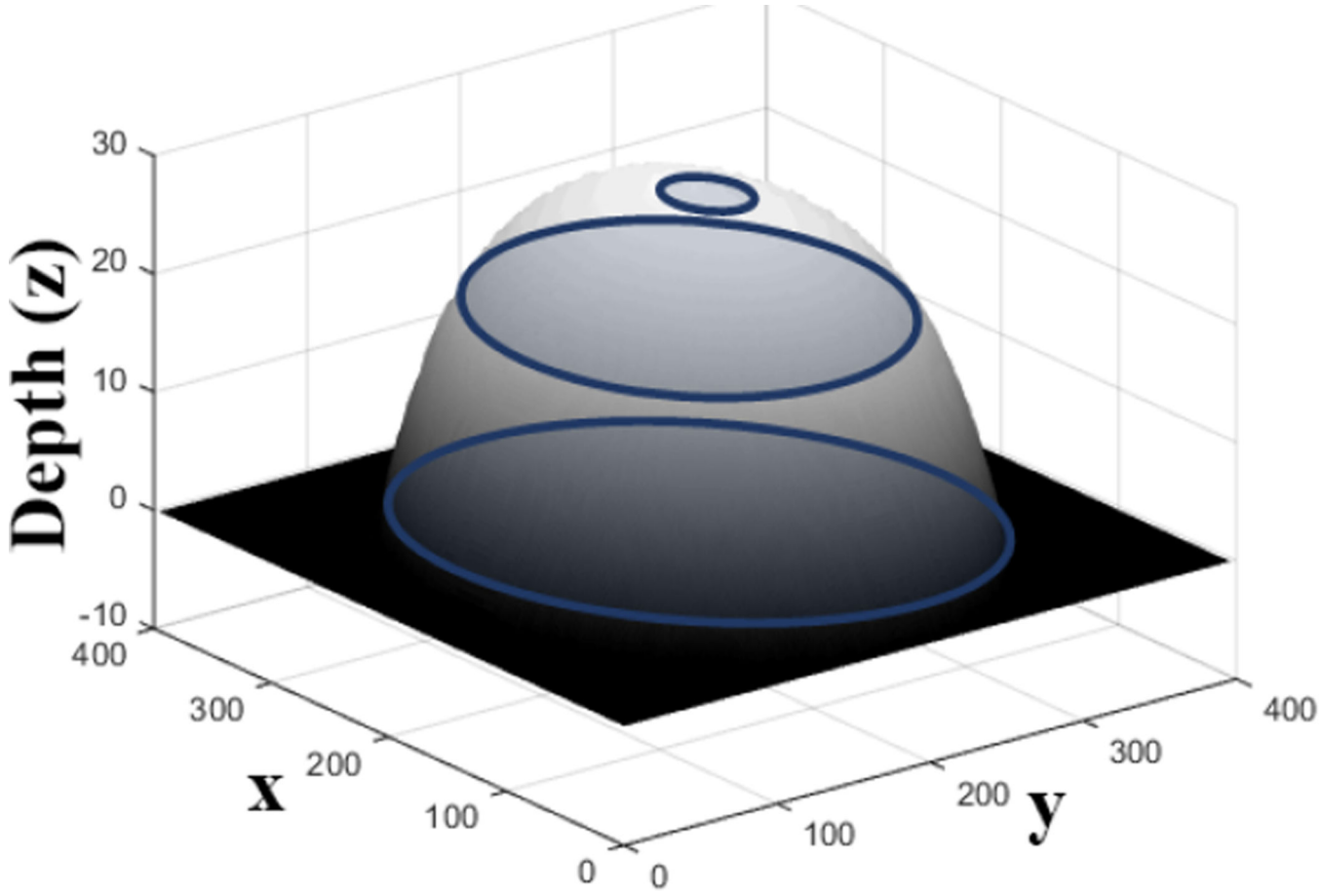


Fig. 6.
A random sample from the proposed shape model.

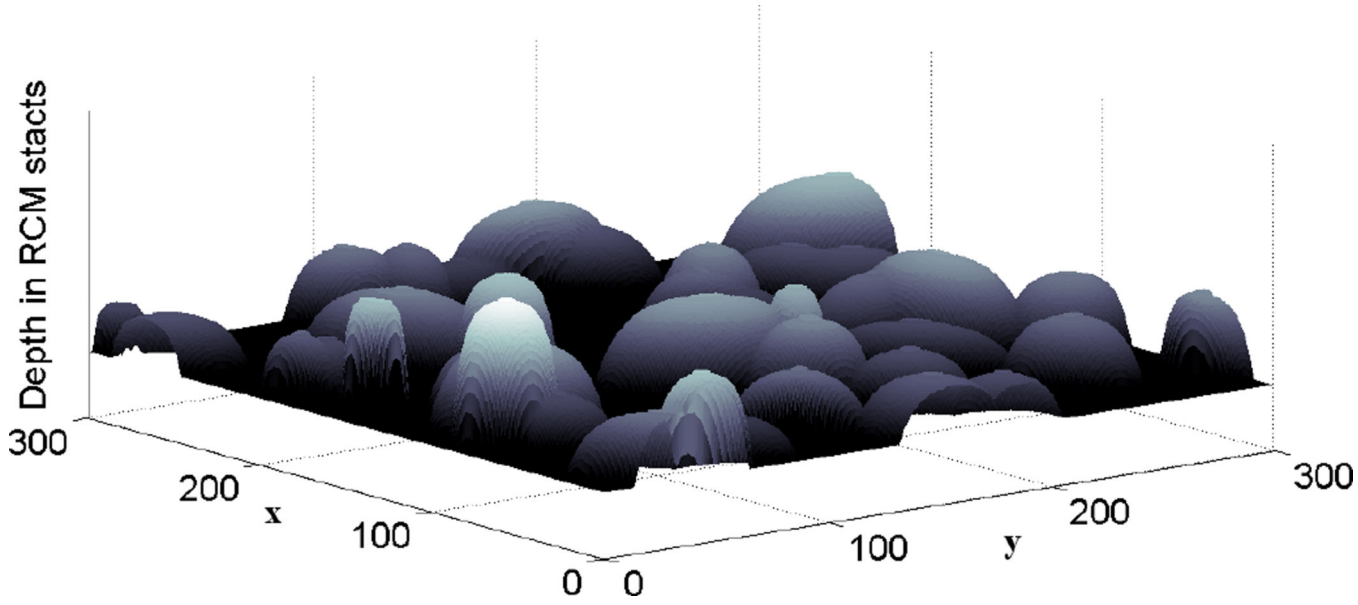


Fig. 7.

A random sample from the latent shape marked spatial Poisson process. This shape/boundary is generated by sampling the hill peaks from a spatial Poisson process, followed by sampling the hill shape parameters from a shape prior model. Refer to text for more details.

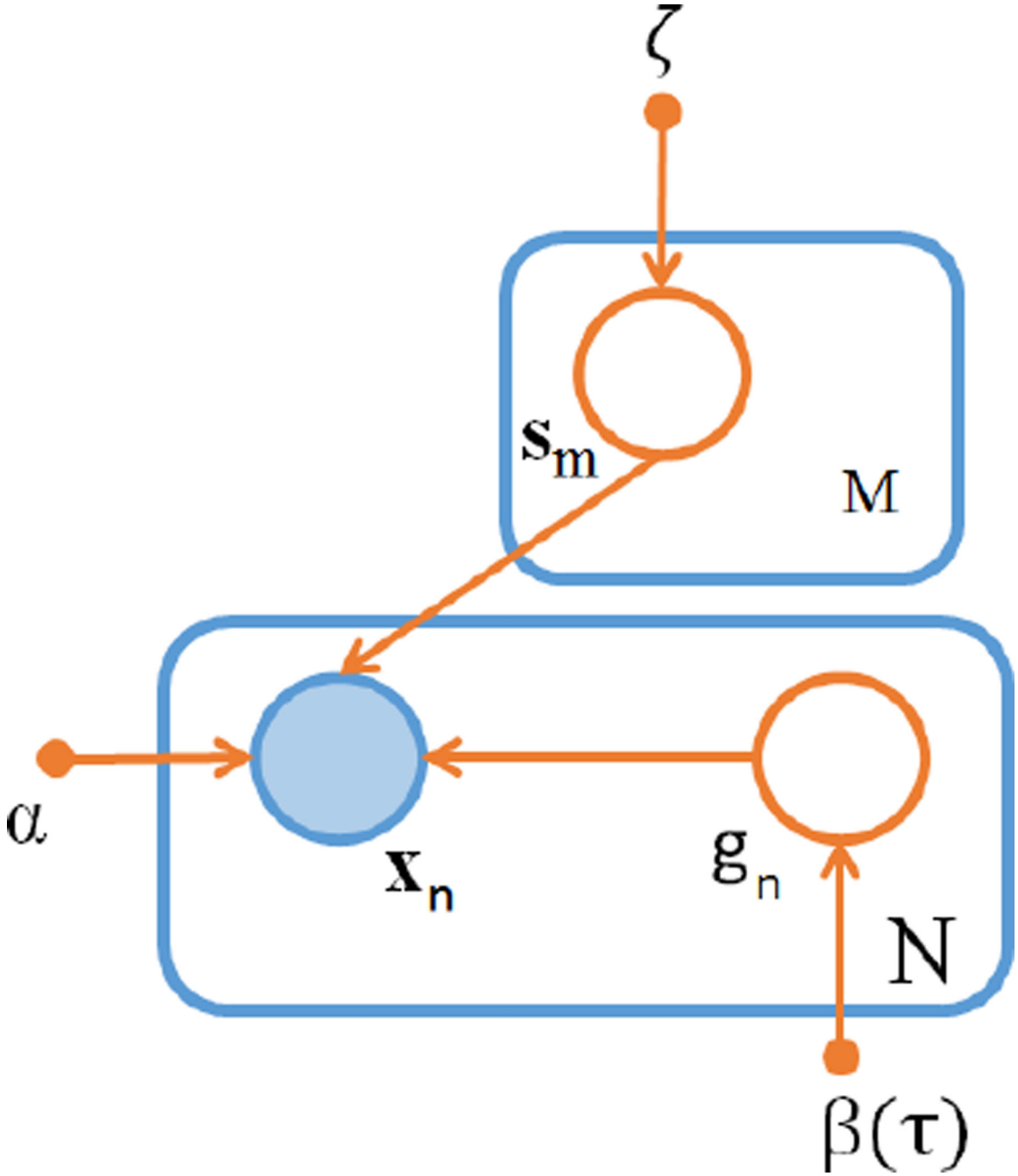


Fig. 8.

Graphical representation of the MPP-skin model. Larger circles show random variables (random variables, shaded if observed, latent otherwise), while the much small circles show deterministic ones. Arrows show statistical dependency of the variables. Boxes around random variables, denoted by an upper case letter M/N, indicates that there are M/N such random variables.**x** are voxels and **s** the shape parameters. The hyperparameters α (appearance), β (position) and ζ (shape) determine the distribution of **x**.

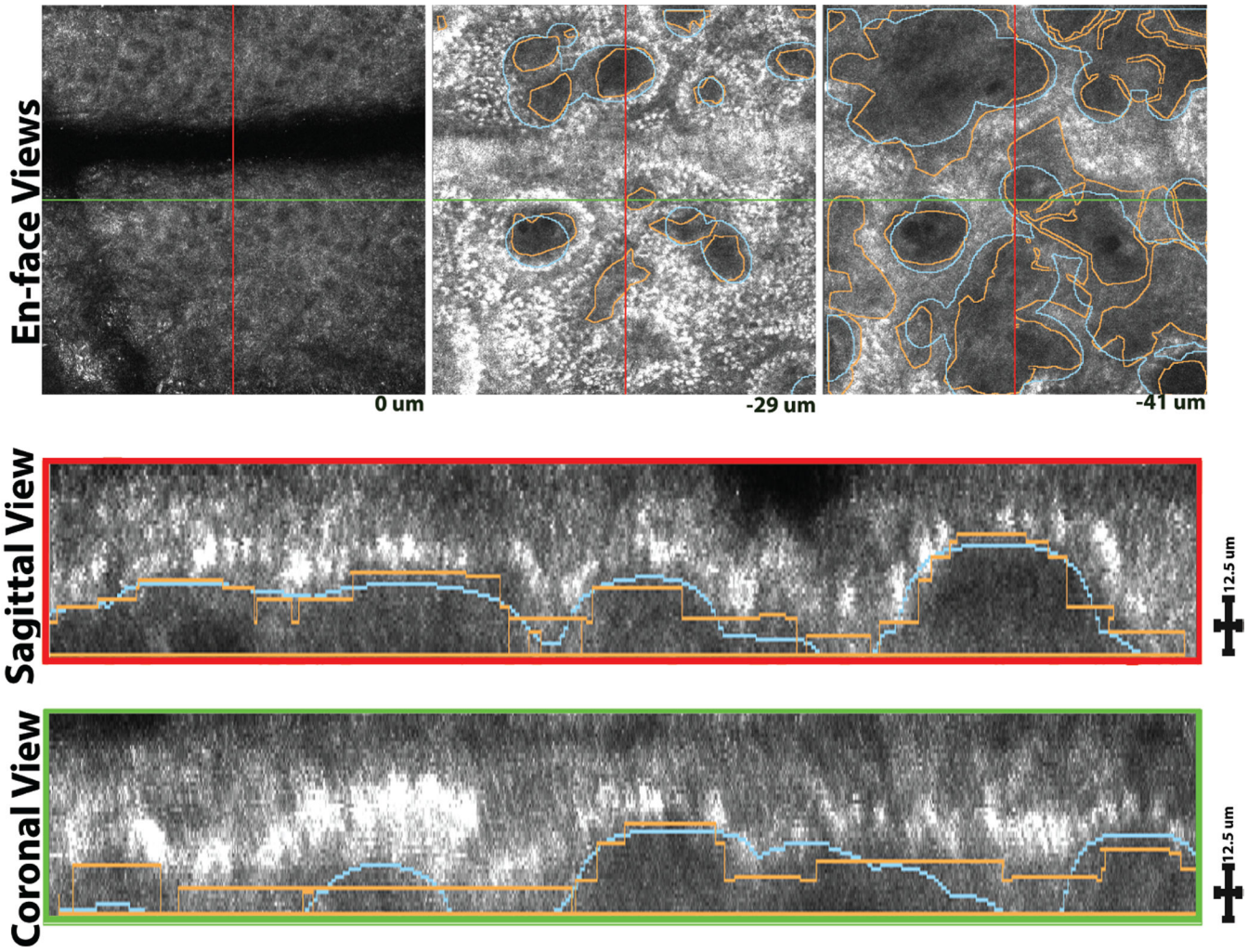


Fig. 9.

Example images from an RCM stack of 45 images (1μ*m* depth spacing) of dark skin, showing layers from lower epidermis, DEJ and papillary dermis. The topmost row shows images in depth (axial views), from left to right, collected at 0μ*m* (epidermis), 29μ*m* (DEJ), and 41μ*m* (papillary dermis), with respect to the initial imaging level at the stratum corneumgranular layer boundary. Sagittal and coronal sections, oriented perpendicular to the plane of this page and located at the red and green lines in axial views, are shown in the second and third row, respectively. In all the views, DEJ boundary drawn by expert clinicians and the algorithmic delineation are shown in orange and light blue, respectively. Scale bars in sagittal and coronal views show 12.5μ*m*.

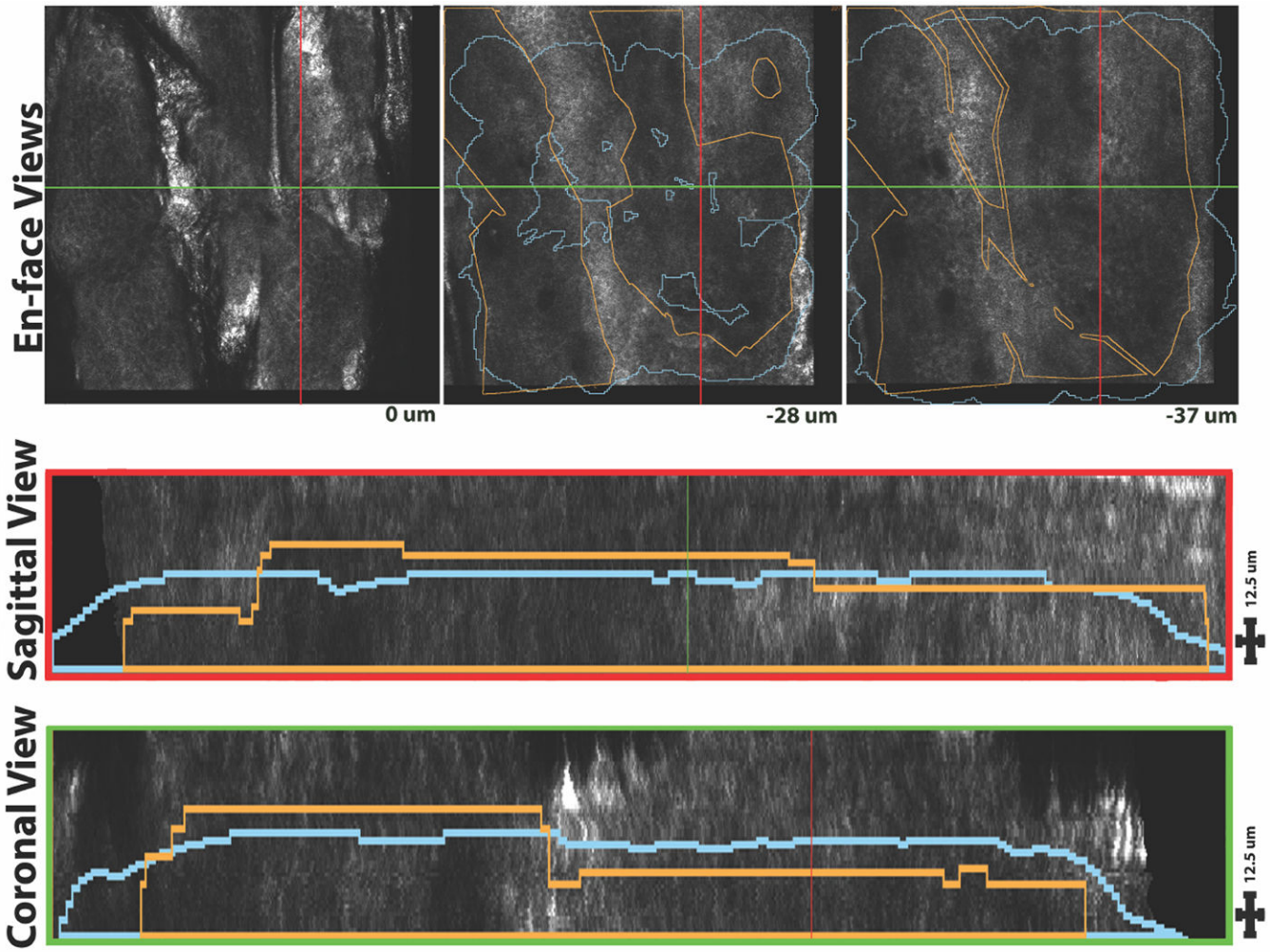
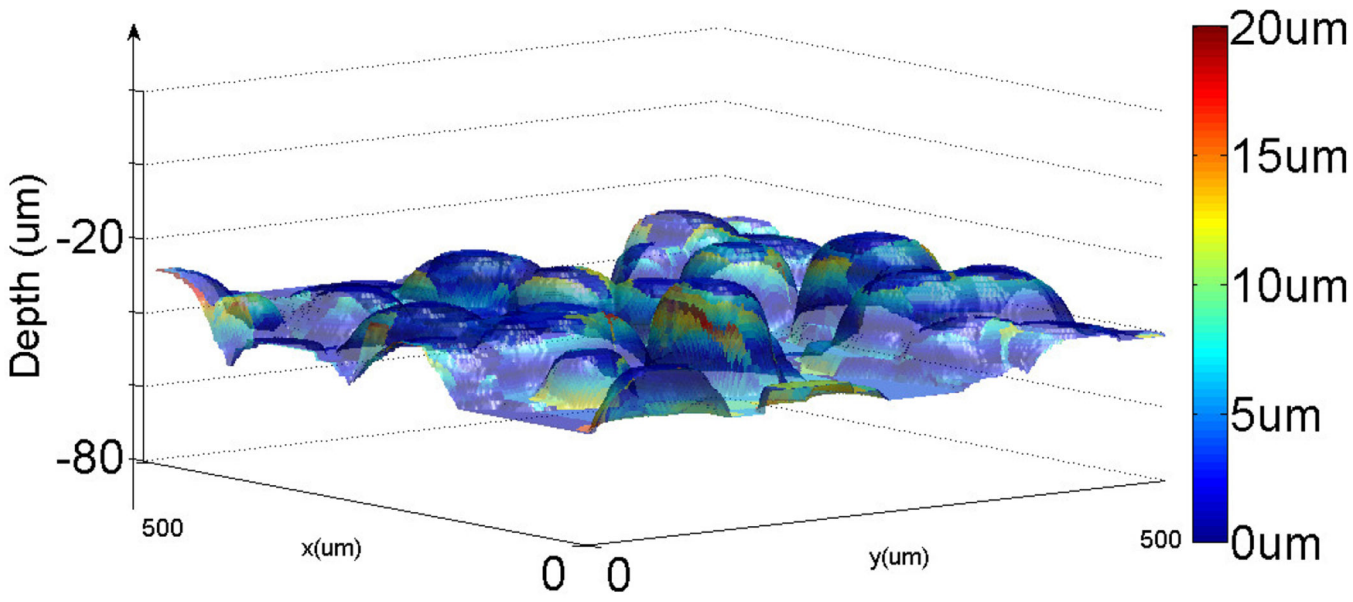
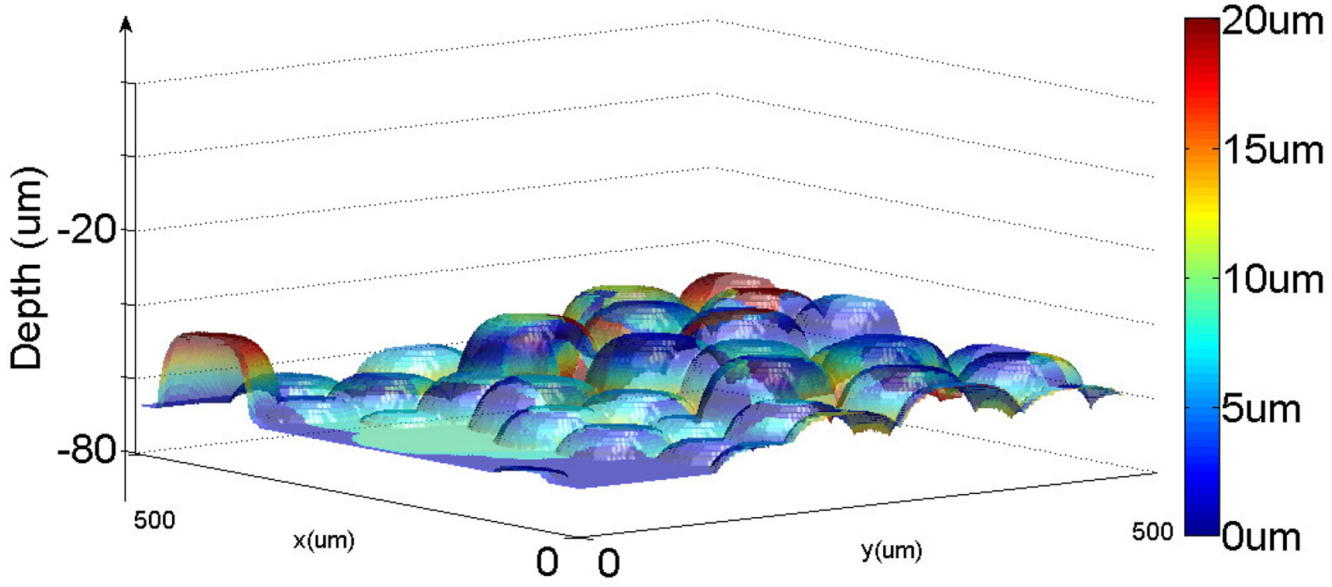


Fig. 10.

Example images from an RCM stack of 60 images (1μ*m* depth spacing) of fair skin, showing layers from lower epidermis, DEJ and papillary dermis. The topmost row shows images in depth (axial views), from left to right, collected at 0μ*m* (epidermis), 28μ*m* (DEJ), and 37μ*m* (papillary dermis), with respect to the initial imaging level at the stratum corneumgranular layer boundary. Sagittal and coronal sections, oriented perpendicular to the plane of this page and located at the red and green lines in axial views, are shown in the second and third row, respectively. In all the views, DEJ boundary drawn by expert clinicians and the algorithmic delineation are shown in orange and light blue, respectively. Scale bars in sagittal and coronal views show 12.5μ*m*.



(a) Dark skin type.



(b) Fair skin type.

Fig. 11.

Dermal-epidermal junction detected by our MPP-skin algorithm. Color indicates the error in DEJ detection compared to expert labels.

TABLE I

Hyper-parameter values used in the experiments.

Model Component	Prior distribution	Parameter values
Poisson: ($\beta(\tau)$)	Gamma	(10, 3)
Shape major and minor axis: (ρ_{ma}, ρ_{mi})	Gamma	(5, 5)
Shape orientation: (ρ_o)	Uniform	(0, π)
Shape rate of increase: (w_{ma}, w_{mi})	Uniform	(2, 20)

TABLE II

Evaluation of the performance of our MPP-skin algorithm compared to our previous approach [4] for dark skin type. We report the mean and standard deviation of error in estimation of DEJ along the z-axis. All units are in μm.

Test set	Graphcuts [25] ↓	Kurugol et al. [4], [16] ↓	MPP-skin ↓
Stack 1	17.71 ± 12.3	3.15 ± 3.26	3.6 ± 3.4
Stack 2	3.6 ± 14.0	7.13 ± 6.36	10.1 ± 5.3
Stack 3	8.7 ± 11.1	5.89 ± 4.58	4.2 ± 3.1
Stack 4	47.6 ± 5.3	3.49 ± 3.28	6.0 ± 5.5
Stack 5	6.7 ± 7.9	14.45 ± 7.45	4.9 ± 3.6
Stack 6	34.18 ± 6.2	2.99 ± 2.65	2.1 ± 3.0
Stack 7	22.3 ± 14.5	7.73 ± 6.7	6.0 ± 5.4
Stack 8	17.4 ± 8.1	7.04 ± 5.98	7.0 ± 5.3
Stack 9	18 ± 13	9.95 ± 7.61	4.8 ± 4.8
Average	19.5 ± 9.24	6.86 ± 4.78	5.41 ± 3.94

TABLE III

Evaluation of the performance of our MPP-skin algorithm for fair skin type. We report the mean and standard deviation of error in estimation of DEJ along the z-axis. All units are in μm.

Test set	Graphcuts [25]↓	Kurugol et al. [4], [16] ↓		MPP-skin ↓
		Epidermis - Transition	Dermis - Transition	
Stack 1	20.94 ± 13.62	4.96 ± 8.27	4.87 ± 8.23	9.5 ± 7.8
Stack 2	16.08 ± 3.81	5.18 ± 8.61	5.24 ± 8.69	12.4 ± 8.31
Stack 3	11.86 ± 8.03	9.99 ± 8.59	8.36 ± 8.55	5.8 ± 5.53
Stack 4	41.99 ± 19.5	15.06 ± 12.02	14.13 ± 11.45	14.6 ± 5.9
Stack 5	15.87 ± 7.04	11.42 ± 5.22	12.02 ± 6.11	14.9 ± 5.2
Stack 6	13.36 ± 11.5	9.65 ± 10.43	9.31 ± 11.14	15.2 ± 10.1
Stack 7	15.75 ± 11.66	24.09 ± 15.65	22.22 ± 14.82	7.9 ± 7.4
Stack 8	13.2 ± 9.1	9.96 ± 10.29	10.50 ± 10.31	13.0 ± 7.8
Stack 9	22.81 ± 13.11	6.84 ± 11.83	6.94 ± 11.52	10.2 ± 9.0
Stack 10	9.5 ± 10.76	10.88 ± 9.38	11.49 ± 9.71	17.9 ± 8.9
Stack 11	21.44 ± 9.99	6.45 ± 11.21	6.76 ± 10.81	6.8 ± 4.1
Stack 12	8.98 ± 8.29	7.16 ± 5.53	7.37 ± 6.34	10.2 ± 6.8
Stack 13	23.28 ± 4.68	4.56 ± 4.38	4.48 ± 4.32	19.0 ± 5.6
Stack 14	18.54 ± 6.9	13.09 ± 13.72	15.27 ± 12.7	12.1 ± 7.0
Stack 15	11.01 ± 4.59	6.38 ± 10.06	6.04 ± 9.44	13.6 ± 6.83
Average	17.64 ± 9.5	9.71 ± 9.68	9.67 ± 9.62	12.1 ± 7.0