OPPA European Social Fund
Prague & EU: We invest in your future.

Epidermis/dermis boundary segmentation in OCT skin images

Hynek Urban, Tomáš Sixta
December 31, 2010

1 Introduction

With its 1.5 - 2 m^2 the skin is one of the largest organ in human body. It protects the body from various exogenous threats like UV radiation, mechanical trauma, chemical toxins or infectious agents, it plays an important role in maintaining homeostasis (maintaining fluid balance, regulation of body temperature, synthesis of vitamines etc.), it also contains variety of receptors that allow humans to sense external stimuli and it has lot of other functions.
The skin is composed of three main layers. The outermost layer is epidermis. It is the thinnest on the eyelids at 0.05 mm and the thickest on the palms and soles at 1.5 mm. From biological point of view it is a stratified squamous epithelium mainly composed of cells called keratinocytes in various stages of development. Their main function is the formation of keratin layer, which protects the underlying tissues from exogenous threats. The layer beneath epidermis is called dermis. It is a connective tissue composed of loosely arranged (upper part, papillary dermis) and densely packed (lower part, reticular dermis) collagen fibers which provides nutrition for the epidermis. Its thickness varies from 0.3 mm to 3 mm. Epidermis and dermis are connected by a basement membrane called dermoepidermal junction. The lowermost layer is hypodermis (subcutaneous tissue) and its main purpose is storaging fat.

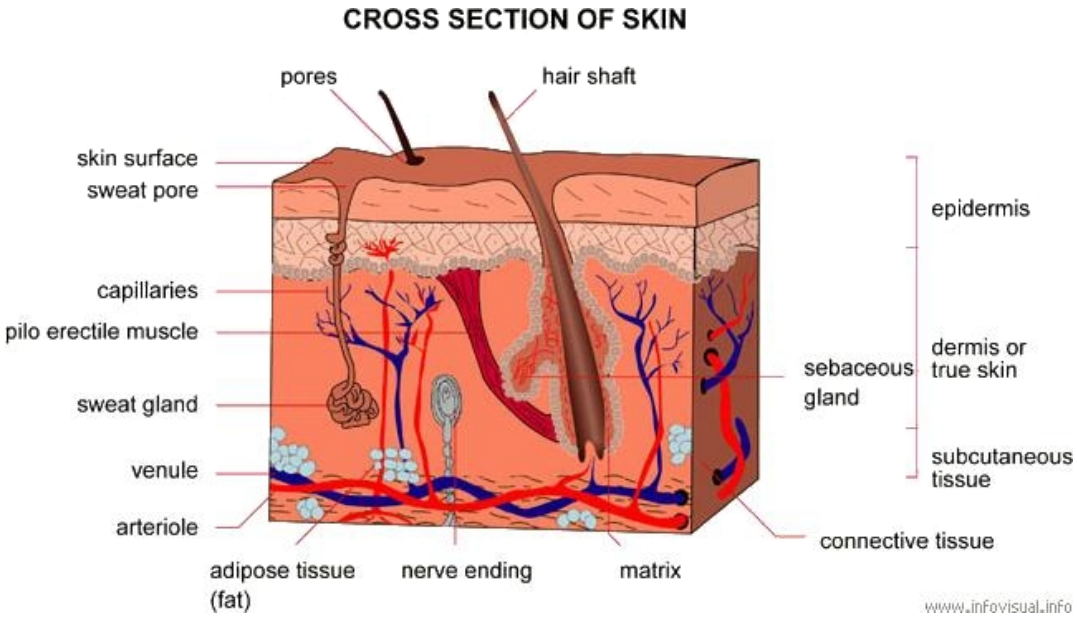


Figure 1.1 Cross section of skin (image taken from http://www.infovisual.info/03/036_en.html)

Thickness of the epidermis and the local shape of the dermoepidermal junction are important factors in monitoring skin health, aging and photodamage [4]. The classical approach for examination is performing histological analysis and there is a long-term effort to find a non-invasive method of biopsy. One candidate for such method is the optical coherence tomography (OCT) [2]. It allows to examine the skin as deeply as 1 mm and it is able to reveal some of the skin structures including change in contrast around dermoepidermal junction. Examples of OCT scans are provided in figure **1.2**.
In this work we describe an automatic method for finding the surface of the epidermis (boundary 1) and the dermoepidermal junction (boundary 2) in OCT images. Each boundary is described by a hidden markov chain as a sequence of height values. The boundaries are found nearly-realtime (less than 20 s per image). Our algorithm is also robust to noise, although it does not require any noise filtering.

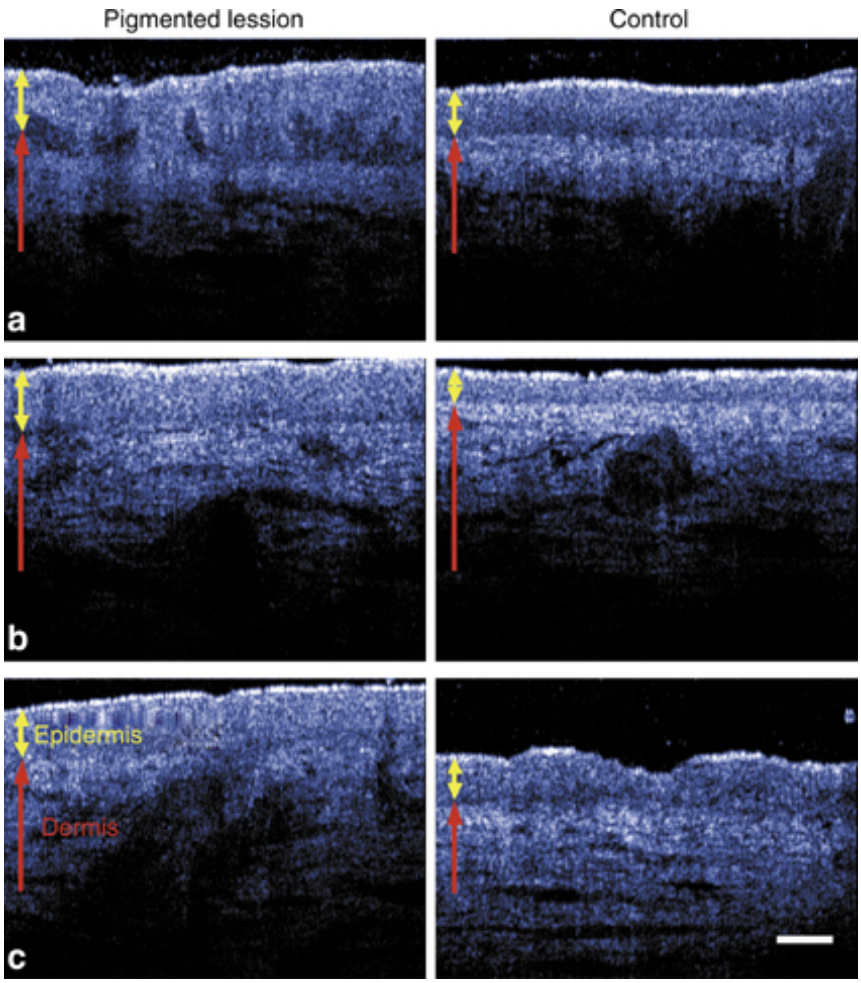


Figure 1.2 OCT imaging of the epidermis of solar lentigines [3], white bar is 100 μm

2 Model description

2.1 Hidden states

Each state represents one possible distance of a (single) boundary from the top edge of the image. Therefore the boundary is directly encoded by sequence $\mathbf{s} = (s_1, s_2, \ldots, s_n)$[1], where s_i is distance of a boundary from the top edge of the image in its i-th column. Actual set of states is constructed during learning process: it contains only such states (distances) that are found at least once in the training data.

2.2 Observations

Observation for a single boundary point is an unordered set of pairs (x, r), where $x \in \mathbb{N}$ is an index corresponding to possible position of a boundary and $r \in (0, 1]$ is "relevance" of that estimation. Observation set may contain zero to three pairs (x, r) (the first boundary only 2 at maximum).

2.2.1 Boundary 1 (background/epidermis)

Boundary between background and epidermis is characterized by rapid change of brightness. The background is dark whereas the uppermost part of epidermis is very bright on available images. From this reason we propose to robustly compute the change of brightness in vertical direction by convolving the image I with double-triangle-like kernel (figure **2.1**) so we obtain modified image I' with the same size

[1] Actual value of n is 600 for available data.

as I. Each column of I' is viewed as 1D function $f(x)$ and observations for given column are such points x (their indices) that fulfill following requirements:

1. $f(x)$ is local maximum
2. $f(x) > 0$
3. $f(x) \geq 0.1 \max f(x)$
4. $|x - x_{max}| > 10$, where $x_{max} = \arg\max f(x)$ and $x \neq x_{max}$
5. $x \geq x_{max}$, where $x_{max} = \arg\max f(x)$
6. If x is not a global maximum of $f(x)$, for all points $y \neq x, y \neq x_{max}$ fulfilling requirements 1. - 5. hold that $f(y) \leq f(x)$

Requirements 1. and 2. mean that we are looking for points, where the brightness is changing from dark to bright and the speed of change is highest at some neigbourhood.
Available images contain two types of noise. The first one is caused by properties of used CCD camera. We believe, that this type of noise is mostly additive and is thus filtered out thanks to the size of convolution kernel **2.1**. The second type are small impurities of possibly pieces of skin above the actual epidermis which are also recorded by OCT scanner (see figure **2.2**). Their width is often greater than 15 px, so they could easily cause improper identification of the first boundary. Requirements 3., 4. and 5. should increase probability, that the boundary is found everywhere at right position.
Requirement 6. only means that two points are chosen at maximum.
Although it is not explicitely stated at previous requirements, instead of x we chose point $x + h_1$ (h_1 is a constant common for all images). Parameter h_1 can be estimated from training data and we used $h_1 = 4$.

It suprisingly turned out that there is no need for sophisticated computation of relevances r, so all observations x has relevance $r = 1$.

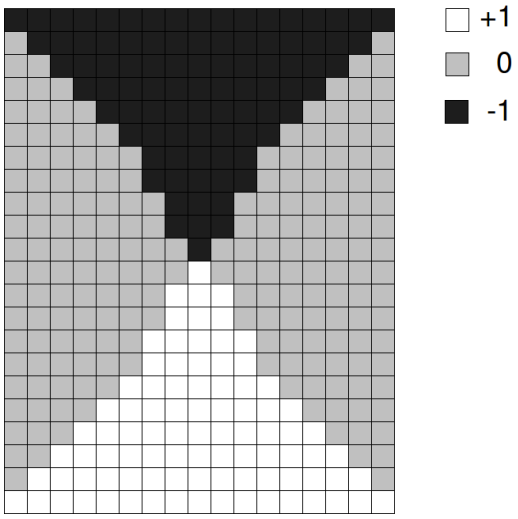


Figure 2.1 Convolution kernel for computing vertical change of brightness

The double-triangle-like kernel was chosen instead of rectangular one, because the boundary is not everywhere horizontal and at these places the rectangular kernel would be less precise in finding position of the boundary.

2.2.2 Boundary 2 (dermoepidermal junction)

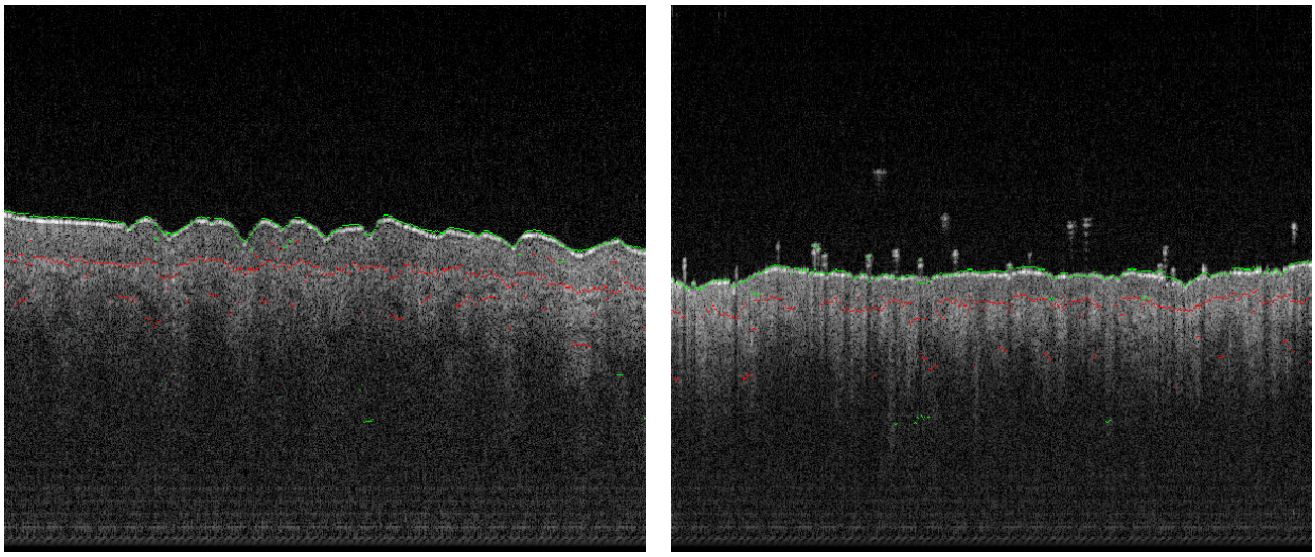
Observations for the second boundary are sought in similar way like for the first one. The input image is also convolved with the kernel from picture **2.1**, each column of the result is again viewed as 1D function $f(x)$ and observations for given column are such points x (their indices) that fulfill following requirements:

1. $f(x)$ is local maximum
2. $f(x) > 0$
3. $x - t_m >$ distance of the first boundary at this column from the top edge of the image (t_m is the minimum thickness of epidermis found in training data)

4. x minus distance of the first boundary at this column from the top edge of the image is lower than maximum thickness of epidermis found in training data
5. $f(x) \geq 0.1 f(x_{chmax})$, where x_{chmax} fulfills requirements 1. - 4. and $f(x_{chmax})$ is maximum among these points
6. $|x - x_{chmax}| > 10$ for all x satisfying requirements 1. - 5., $x \neq x_{chmax}$
7. If $x \neq x_{chmax}$, there is at most one $y \neq x, y \neq x_{chmax}$ fulfilling requirements 1. - 6. such that $f(y) \geq f(x)$

Requirements 1. and 2. again mean that we are looking for points, where the brightness is changing from dark to bright and the speed of change is highest at some neigbourhood. However, in case of the first boundary there are always some points, for which $f(x) > 0$. This requirement is more restrictive for the second boundary and it is more likely that no x is chosen.
Requirements 3. and 4. ensure that the chosen estimation x is always in reasonable distance from the first boundary.
Requirement 7. only means that three points are chosen at maximum.
Like in the case of the first boundary a point $x + h_2$ is chosen instead of x (h_2 is again constant common for all images). Considering the training data we used $h_2 = -10$.

We tested two ways of computing relevances r. In the first we computed relevance by dividing $f(x)$ (where x is chosen point the relevance is computed for) by $f(x_{chmax})$ (where x_{chmax} is such chosen point, for which the $f(x_{chmax})$ is maximum). The second approach was as straightforward as in section **2.2.1**: all relevances were simply set to 1. As the first approach proved us slightly better performance, we decided to use it at our experiments.



a) b)

Figure 2.2 Observations at two sample images

2.3 Probability models

Both boundaries use the same probability model. Prior and transition probability functions are simply defined by enumeration and estimated during learning process (see section **3.1**).
Let x_i be a set of all chosen observations (x_{ij}, r_{ij}) at i-th column. The emission probability function $p(x_i|s_i)$ is than defined as

$$p(x_i|s_i = k) = \begin{cases} \frac{1}{z} & x_i = \emptyset \\ \frac{1}{z} \max_j \{r_{ij} \mathcal{N}(k; x_{ij}, \sigma^2)\} & \text{else} \end{cases} \tag{2.1}$$

where

$$\mathcal{N}(k; \mu, \sigma^2) = \frac{1}{\sqrt{2\pi}\sigma} e^{\frac{(k-\mu)^2}{2\sigma^2}} \text{ and}$$

z is a normalizating constant such that

$$\sum_{x_i \in X} p(x_i|s_i = k) = 1 \tag{2.2}$$

Parameter σ controls "smoothness" of the inferred boundary (the lower σ the less smooth the boundary is). Because it is unclear, how smooth the boundary should be (it would apparently depend on given medical application), we used $\sigma^2 = 10$ for the first boundary and $\sigma^2 = 40$ for the second one.

3 Method description

Both boundaries are sought in following steps:

1. Load training data.
2. Estimate $p(s_1)$ and $p(s_i|s_{i-1})$ for both boundaries (section **3.1**).
3. Load testing data.
4. Compute observations $x_i, i = 1..n$ for the first boundary (section **2.2.1**).
5. For all $i = 1..n$, $s_i \in K$ and given x_i enumerate emission probability $p(x_i|s_i)$ for the first boundary (section **2.3**).
6. Find the first boundary (section **3.2**).
7. Compute observations $x_i, i = 1..n$ for the second boundary (section **2.2.2**).
8. For all $i = 1..n$, $s_i \in K$ and given x_i enumerate emission probability $p(x_i|s_i)$ for the second boundary (section **2.3**).
9. Find the second boundary (section **3.2**).
10. Show results.

3.1 Estimation of $p(s_1)$ and $p(s_i|s_{i-1})$

It turned out, that actual values of $p(s_1)$ have only local influence on the shape of inferred boundary. Therefore we found all possible distances of given boundary in training data (at all columns) and $p(s_1)$ is than defined as

$$p(s_1 = k) = \begin{cases} \frac{1}{p_s} & \text{k found in the training data} \\ 0 & \text{else} \end{cases} \tag{3.1}$$

where p_s is a normalizating constant such that

$$\sum_{s_1 \in K} p(s_1) = 1 \tag{3.2}$$

Our model of the boundary is a homogenous markov chain, which means, that

$$p(s_i|s_{i-1}) = p(s_j|s_{j-1}) \quad \forall i, j = 2..n \tag{3.3}$$

We also found useful to make it invariant to vertical translation, so we learned transition probabilities in such way, that

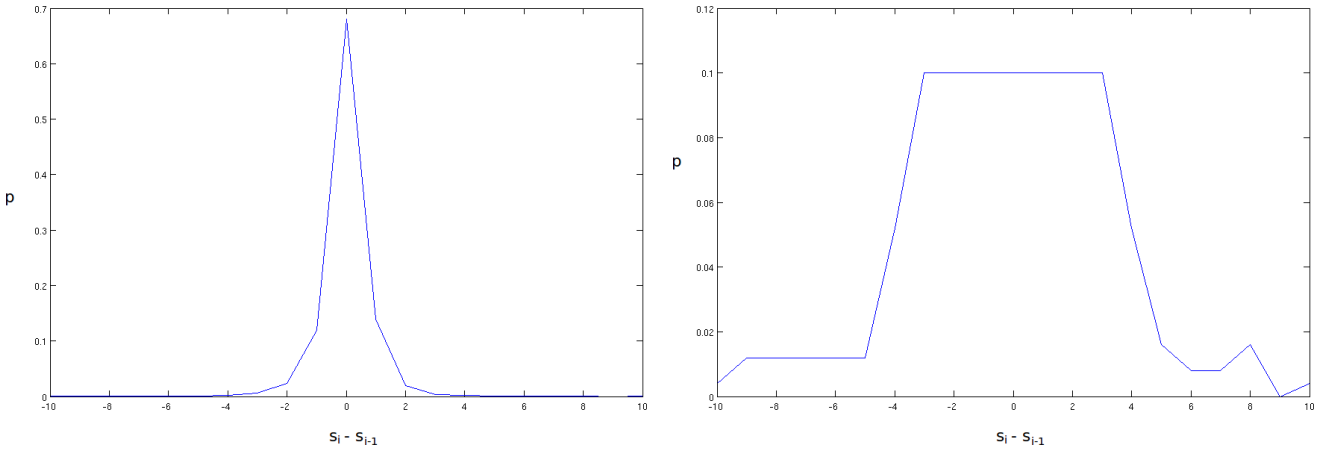
$$p(s_i = k|s_{i-1} = l) = p(s_i = k + a|s_{i-1} = l + a) \quad \forall a \tag{3.4}$$

where $k, l \in K$ and $a \in \mathbb{Z}$ is some feasible shift. When the learning process is over the transition probabilities $p(s_i|s_{i-1})$ are iteratively modified so that for all $i = 2..n$ holds

$$\sum_{s_i \in K} p(s_i|s_{i-1}) = 1 \tag{3.5}$$

$$\max p(s_i|s_{i-1}) \leq t \tag{3.6}$$

where $t \in (0, 1]$ is a chosen threshold. Without this modification the inference algorithm would strongly prefer horizontal boundary and would be unable to handle its local shape.



Without iterative modification | With iterative modification

Figure 3.1 Transition probability $p(s_i|s_{i-1})$ for the first boundary

Threshold $t = 0.1$ gave us reasonable results for both boundaries (figure **3.1**).

3.2 Inference

Both boundaries are inferred as sequences of the most probable hidden states by forward-backward algorithm. In order to prevent potential numerical instability (underflowing) we implemented scaling technique described in [6].

4 Results

The described approach was implemented in Matlab environment. We were given 401 partially anotated images in four folders (Lernstichprobe, Proband 139, Proband 141 and Proband 142). 14 images from Lernstichprobe were used as training data and the rest were used for testing. Some of the results are shown in image **4.1**
Processing of one image takes about 15 seconds on AMD Turion 64 X2/4 GB RAM laptop.
Similarity of our solution and the ground truth is described in table **4.1** by two statistics: average distance of two appropriate boundaries and bias of our solution.
Average distance:

$$d_{avg} = \frac{1}{600} \sum_{i=1}^{600} |s_i - g_i|, \tag{4.1}$$

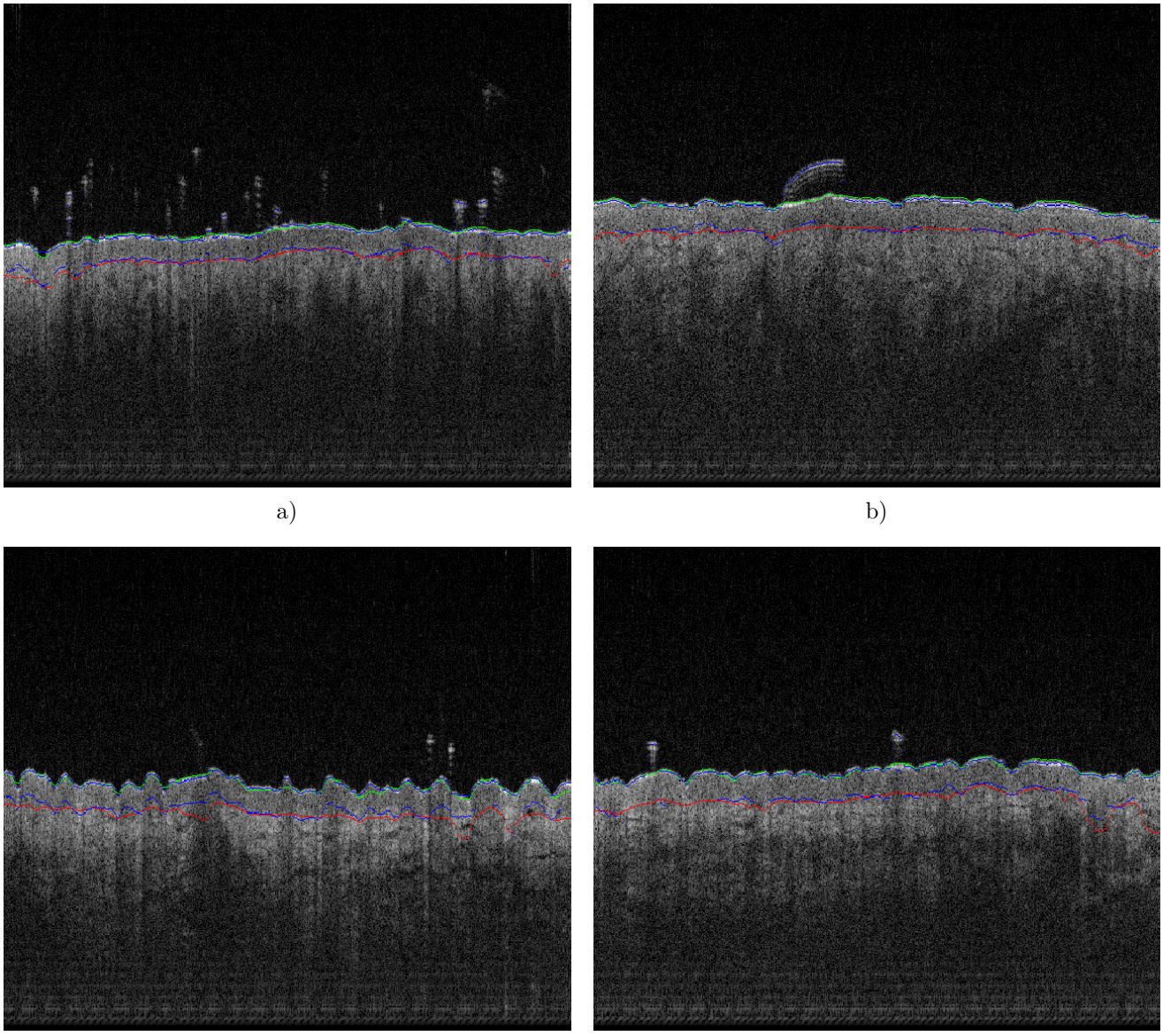
where s_i is inferred position of given boundary at i-th column,
g_i is position of the same boundary given by annotation and
600 is width of the image.
Bias:

$$b = \frac{1}{600} \sum_{i=1}^{600} s_i - g_i, \tag{4.2}$$

Length of available annotation was greater than 380 for all four images.

5 Conclusions

We implemented an automated method for recognizing boundary between background and epidermis and dermoepidermal junction in OCT images. Each boundary is encoded by a hidden markov chain.



a) b)

c) d)

Figure 4.1 Boundaries at 4 sample images. Green points denote inferred boundary background/epidermis, red denote inferred dermoepidermal junction and blue points represent ground truth

	b (boundary 1)	d_{avg} (boundary 1)	b (boundary 2)	d_{avg} (boundary 2)
Image **4.1** a)	0.644	3.693	3.601	4.373
Image **4.1** b)	1.721	4.900	2.005	3.645
Image **4.1** c)	1.111	2.815	3.239	6.738
Image **4.1** d)	-0.002	2.556	2.132	3.927

Table 4.1 Average distance and bias of inferred boundaries

Our method is robust to various types of noise although it does not require any noise filtering during preprocession.

5.1 Possible improvements

1. Although our method is able to ignore small pieces of skin above the true surface of epidermis, it rarely happens, that it fails in localizing the true surface. Therefore it would be convenient to perform watershed segmentation to find small isolated bright areas and remove them from the image.

2. If the direction of a boundary is nearly vertical, even kernel **2.1** may suffer from poor localization. One possible solution could be rotating the kernel **2.1** to certain angles (e.g. -45°, 0°, 45°), for each rotation perform the convolution and choose the highest response.

6 References

[1] Graham-Brown R., Burns T.. *Lecture notes: Dermatology*. Wiley-Blackwell, 224 pages, ISBN: 978-1-4051-3977-9, 2006.

[2] Huang D., Swanson E.A., Lin C.P., Schuman J.S., Stinson W.G., Chang W., Hee M.R., Flotte T. et al. *Optical coherence tomography*. Science 254 (5035): 1178–81, 1991.

[3] Yamashita T., Negishi K., Hariya T., Kunizawa N., Ikuta K., Yanai M., Wakamatsu S.. *Intense Pulsed Light Therapy for Superficial Pigmented Lesions Evaluated by Reflectance-Mode Confocal Microscopy and Optical Coherence Tomography*. Journal of Investigative Dermatology 126, 2281–2286. doi:10.1038/sj.jid.5700414; published online, 2006.

[4] Weissman J., Hancewicz T., Kaplan P.. *Optical coherence tomography of skin for measurement of epidermal thickness by shapelet-based image analysis*. Optics Express, Vol. 12, Issue 23, pp. 5760-5769, 2004.

[5] Schwabe G.. *Erkennung der Grenze zwischen Dermis und Epidermis in OCT Aufnahmen (Master's thesis)*. Technische Universitat Dresden, 2006.

[6] Khreich W., Granger E., Miri A., Sabourin R.. *On the memory complexity of the forward-backward algorithm*. Pattern Recogn. Lett. 31, 2, 91-99. DOI=10.1016/j.patrec.2009.09.023 http://dx.doi.org/10.1016/j.patrec.2009.09.023, January 2010.

OPPA European Social Fund
Prague & EU: We invest in your future.