Graph-based skin lesion segmentation of multispectral dermoscopic images
Olivier Lézoray, Marinette Revenu, Michel Desvignes

To cite this version:
Olivier Lézoray, Marinette Revenu, Michel Desvignes. Graph-based skin lesion segmentation of multispectral dermoscopic images. 21st IEEE International Conference on Image Processing (ICIP 2014), Oct 2014, Paris, France. hal-01080036

HAL Id: hal-01080036
https://hal.archives-ouvertes.fr/hal-01080036
Submitted on 4 Nov 2014

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L’archive ouverte pluridisciplinaire HAL, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d’enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.
Graph-based skin lesion segmentation of multispectral dermoscopic images

O. Lezoray, M. Revenu *
GREYC CNRS UMR 6072, Caen, France

M. Desvignes
GIPSA-Lab CNRS UMR 5216, Grenoble, France

Abstract

Accurate skin lesion segmentation is critical for automated early skin cancer detection and diagnosis. We present a novel method to detect skin lesion borders in multispectral dermoscopy images. First, hairs are detected on infrared images and removed by inpainting visible spectrum images. Second, skin lesion is pre-segmented using a clustering of a superpixel partition. Finally, the pre-segmentation is globally regularized at the superpixel level and locally regularized in a narrow band at the pixel level.

Index Terms — Skin cancer, dermoscopy, hair detection, hair removal, graphs, segmentation.

1. Introduction

The incidence of both non-melanoma and melanoma skin cancers has been increasing over the past decades, especially melanoma for fair-skinned populations in industrialized countries [1]. Early diagnosis is particularly important since melanoma can be cured with a simple excision if detected early [2]. Dermoscopy is a non-invasive skin imaging technique used in dermatology for early diagnosis of skin lesions [3]. The visual diagnosis from dermoscopy images is difficult, subjective and lacks reproducibility [2]. Therefore, the development of computerized image analysis techniques that can assist dermatologists is of paramount importance [4]. An essential step is the automatic segmentation of the lesions. The segmentation is very challenging due to factors such as illumination variations, irregular structural and color variations, presence of hair and reflections. In this paper we propose a method for the segmentation of multispectral dermoscopic images composed of 6 spectral bands (3 in visible light and 3 in infrared -IR- light).

2. Image Pre-processing

Presence of hairs in dermoscopic images is a challenge in the design of an automated segmentation of skin lesions. Indeed, hair pixels occlude some of the information of the lesion such as its boundary and texture. Hair detection is complicated due to thin and thick hair, hairs with different colors and similar color with the lesion [5]. The interest of multispectral imaging in dermoscopy [6] relies on the principle that light of the visible and infrared spectrums penetrates the skin in different depths and this enables to reveal different features of the skin lesion. We can make the most of the information provided with infrared light since, in addition to reveal the melanin present in the deeper layers of the skin [7], it also better exhibit the hairs than under visible light (see the two first columns in Figure 1). Therefore, we have based our hair detection method on the IR images. Assuming that in IR images, hairs are darker that the surrounding zones, a morphological closing top-hat filtering is applied [8]. Considering that hairs are thin linear structures, the structuring element is a straight line. Since the hair direction is not known, a top-hat filtering operation is performed four times using a straight line with four different orientations and the maximum of the four top-hat responses is calculated. This is done for each component of the IR image and is expressed by

\[
\max_{i \in \{0, 1, 2, 3\}} \max_{\theta \in \{45, 90, 135\}} \left( \phi_{y \in B^\theta(x)}(f_i^{IR}(y)) - f_i^{IR}(x) \right)
\]

where \(B^\theta\) is a directional structuring element of length 9 and \(f_i^{IR}\) is the \(i\)-th component of the IR image. Since the detection can be noisy, the result is smoothed using a non-local means algorithm [9] with patches of size 3 \times 3 in a window of size 5 \times 5. This filter enables to filter the residual noise while well preserving the edges (i.e., detected hairs). The denoised image is then automatically binarized and structures longer than a given length are selected. This length is checked in different orientations to include hairs with different positions and orientations [10]. We have fixed the length to 33 and considered 24 different possible orientations. Third and fourth columns of Figure 1 shows sample results. Once hairs have been extracted, we have to fill-in the areas of detected hairs to repair hair-occluded information. Hair removal is performed on visible spectrum images. Given the hair mask, we use a recent inpainting method we developed in [11]. This method allows reducing artefacts produced by any patch-based inpainting algorithm [12] and enables to lower possible inter-patch seams and inconsistencies, rendering the image perceptually more visually coherent which is very important for dermoscopy images. Finally a last artifact has to be removed from the images: light reflection. Reflection areas are

*Work supported by grant ANR-10-TECS-018 MELASCAN
detected by comparing if the pixel blue intensity is high and well above the average blue intensity in its neighborhood. Detected reflection areas are inpainted using again our spatial patch blending inpainting [11]. The two last columns of Figure 1 shows sample results of hair and reflection removal. As it can be seen, our method works well for many different configurations.

3. GRAPH-BASED IMAGE SEGMENTATION

Once the artifacts have been removed from the images in the visible spectrum, we propose to perform the segmentation using a graph-based image processing approach [13]. Our approach uses graphs either at the superpixel or at the pixel level and is very different to usual methods used to skin lesion border detection [14, 15]. We detail each of these steps in the sequel but start with needed graph theoretical notions and algorithms.

3.1. Notations and Preliminaries

A weighted graph \( G = (\mathcal{V}, \mathcal{E}, w) \) consists of a finite set \( \mathcal{V} = \{v_1, \ldots, v_N\} \) in vertices and a finite set \( \mathcal{E} \subset \mathcal{V} \times \mathcal{V} \) of weighted edges. Let \( \langle v_i, v_j \rangle \) be the edge of \( \mathcal{E} \) that connects two vertices \( v_i \) and \( v_j \) of \( \mathcal{V} \). Its weight, denoted by \( w(v_i, v_j) \), represents a similarity measure and is computed with a positive symmetric function \( w : \mathcal{V} \times \mathcal{V} \rightarrow \mathbb{R}^+ \) satisfying \( w(v_i, v_j) = 0 \) if \( (v_i, v_j) \notin \mathcal{E} \) and the notation \( v_i \sim v_j \) is used to denote two adjacent vertices. Let \( \mathcal{H}(\mathcal{V}) \) be the Hilbert space of real-valued functions defined on the vertices of a graph (i.e., vectors associated to vertices).

3.2. Difference Operators on Weighted Graphs

Let \( G = (\mathcal{V}, \mathcal{E}, w) \) be a weighted graph. The difference operator [16], denoted \( d_u : \mathcal{H}(\mathcal{V}) \rightarrow \mathcal{H}(\mathcal{E}) \), is defined for all \( f \in \mathcal{H}(\mathcal{V}) \) and \( \langle v_i, v_j \rangle \in \mathcal{E} \) by:

\[
(d_u f)(v_i) = \sqrt{w(v_i, v_j)}(f(v_j) - f(v_i))
\]

The weighted gradient operator of a function \( f \in \mathcal{H}(\mathcal{V}) \), at a vertex \( v_i \in \mathcal{V} \), is the vector defined by:

\[
(\nabla w f)(v_i) = (d_u f(v_i)_{v_i})^T
\]

The \( \ell_p \) norm of this vector is defined, for \( p \geq 1 \), by:

\[
\| (\nabla w f)(v_i) \|_p = \left( \sum_{v_j \sim v_i} w(v_i, v_j)^{p/2} |f(v_j) - f(v_i)|^p \right)^{1/p}
\]

3.3. Regularization on graphs

In this section, we consider a general function \( f^0 : \mathcal{V} \rightarrow \mathbb{R}^m \) defined on graphs of the arbitrary topologies and we want to regularize this function. The regularization of such a function corresponds to an optimization problem that can be formalized by the minimization of an energy as a weighted sum of two regularization and data fitting terms:

\[
\mathcal{E}_{w,p}(f, f^0, \lambda) = \frac{1}{p} \sum_{v_i \in \mathcal{V}} \| (\nabla w f)(v_i) \|_p^2 + \frac{1}{2} \| f - f^0 \|_2^2
\]

In previous works [16], we have shown that solutions minimizing such an energy on weighted graphs can be obtained by the following iterative algorithm \( \forall u \in \mathcal{V} \), with \( (\gamma_{w,p} f)(v_i, v_j) = w_{ij} \| (\nabla w f)(v_i) \|_2^2 + \| (\nabla w f)(v_i) \|_2^2 \):

\[
\begin{cases}
    f^{(0)} = f^0 \\
    f^{(n+1)}(v_i) = \frac{\lambda f^{(n)}(v_i) + \sum_{v_j \sim v_i} (\gamma_{w,p} f^{(n)})(v_i, v_j) f^{(n)}(v_j)}{\lambda + \sum_{v_j \sim v_i} (\gamma_{w,p} f^{(n)})(v_i, v_j)}
\end{cases}
\]

This formulation can be used to denoise a function, but also for interpolation. In particular it can be used for semi-supervised clustering [17]. Let \( \mathcal{E}_l \) be a set of labeled vertices, these latter belonging to the \( \ell \)-th class. Let \( \mathcal{V}_0 = \bigcup \{ \mathcal{E}_l \}_{l=1,...,k} \) be the set of initial labeled vertices and let \( \mathcal{V} \setminus \mathcal{V}_0 \) be the initial unlabeled vertices. Each vertex of \( v_i \in \mathcal{V} \) is then described by a vector of labels \( F_\ell(v_i) = (f^{(l)}(v_i))_{l=1,...,k} \) with \( f^{(l)}(v_i) = +1 \) if \( v_i \in \mathcal{E}_l \) and \( f^{(l)}(v_i) = 0 \), \( \forall v_i \in \mathcal{V} \setminus \mathcal{V}_0 \). The semi-supervised clustering problem then reduces to interpolate the labels of the unlabeled vertices \( (\mathcal{V} \setminus \mathcal{V}_0) \) from the labeled ones \( (\mathcal{V}_0) \). To solve this interpolation problem, we consider the variational problem (1) and estimate class membership probabilities with Eq. (2). Final classification can be obtained for a given vertex \( v_i \in \mathcal{V} \) by the maximum membership

\[
l(v_i) = \arg \max_{\ell} \left\{ \frac{f^{(\ell)}(v_i)}{\sum_{\ell} f^{(\ell)}(v_i)} \right\}
\]

3.4. Superpixel segmentation

Given these notions, our method starts by computing a segmentation of the image into superpixels. In this paper, we have considered the Eikonal-based Region Growing Clustering algorithm [18]. ERGC is a recent superpixel algorithm formulated as the solution of an Eikonal equation. We consider here a more general formulation of ERGC on arbitrary graphs. The Eikonal equation describes the evolution of a propagation front \( \mathcal{V}_0 \):

\[
\begin{cases}
    \| (\nabla w f)(v_i) \|_p = P(v_i) & \forall v_i \in \mathcal{V} \\
    f(v_i) = 0 & v_i \in \mathcal{V}_0
\end{cases}
\]

where \( \mathcal{V} \) corresponds to the set of vertices of a 8-grid graph associated to an image, \( \mathcal{V}_0 \) is an initial set of seed vertices, \( P \) is a positive potential function, and \( f(v_i) \) is the distance of \( v_i \) from source \( \mathcal{V}_0 \). This problem can be efficiently solved on graphs using a generalization of the fast marching method on graphs [19]. In ERGC, the potential map \( P \) evolves during diffusion process and is defined as

\[
P(v_i) = \sqrt{\| f_{c}(v_i) - \bar{f}_c(v) \|^2 + \frac{m \| f_{c}(v_i) - f_{c}(v) \|^2}{S}}
\]

where \( \bar{v}_c \) denotes the superpixel region \( v_i \) belongs to, \( S = \sqrt{|\mathcal{V}|/|\mathcal{V}_0|} \) is the initial sampling step of the seeds, \( f_{c} \) denotes the coordinates of a vertex or of the initial seed of a superpixel, \( f_{c} \) and \( \bar{f}_c \) denote the color in the CIELAB color space of a vertex or a superpixel (mean color in this case). The first term of
the potential favors the creation of homogeneous and close to edges superpixels and the second term favors regular-shaped superpixels with the compactness factor \( m \). The initialization of the algorithm consists in sampling the set of seeds vertices \( \mathcal{V}_0 \) on the vertices \( \mathcal{V} \) of a regular 8-grid graph. First line of Figure 2 shows sample results of the obtained superpixel partition.

### 3.5. Superpixel clustering and regularization

To classify the different areas in a dermoscopy image, each superpixel is described by a vector of 6 values providing the average and standard deviation of the superpixel for each CIELAB color channel. These feature vectors are computed on a simplified version of the image obtained by a regularization (Eq. (2)) of the image on a 8-grid graph \( \mathcal{S} = (\mathcal{V}, \mathcal{E}, \mathcal{w}_1) \) with no data-term (\( \lambda = 0 \), \( p = 2 \) and bilateral weights

\[
w_1(v_i, v_j) = G_\sigma_c(\|f_p(v_i) - f_p(v_j)\|_2^2) \cdot G_\sigma_n(\|f_c(v_i) - f_c(v_j)\|_2^2)
\]

where \( G_\sigma \) is a Gaussian kernel. Then we use a 3-means clustering to classify each superpixel in three classes. We consider the first and second terms of the \( k \)-means algorithm [20] that is not sensitive to initialization. Second column of Figure 2 presents the results of the obtained clustering. This result is a good initial guess but does not provide a good localization of the skin lesion. To enhance this result, we first perform a regularization of the labels associated to each superpixel. We first associate a graph \( \mathcal{G} = (\mathcal{V}, \mathcal{E}, \mathcal{w}_2) \) to the superpixel partition where each superpixel \( \tilde{v}_i \) is identified with a node of \( \mathcal{V} \) and \( \mathcal{E} \) is filled with edges that connect adjacent superpixels. In our case, after the 3-means clustering, each superpixel has a label, therefore there are no unlabeled vertices. Our algorithm can however be used to spatially regularize the classification to enhance its coherence. We apply Eq. (2) until convergence with \( p = 1 \) and \( \lambda = 1 \) (low values of \( \lambda \) enable a modification of the initial labels). However, this method is efficient only if the graphs weights well reflect the similarity between the superpixels. Since dermoscopic images are highly textured, we have used a similarity between superpixels that takes into account both color and texture by

\[
w_2(\tilde{v}_i, \tilde{v}_j) = G_\sigma_c(\|f_c(\tilde{v}_i) - f_c(\tilde{v}_j)\|_2^2) \cdot G_\sigma_n(\|h(\tilde{v}_i) - h(\tilde{v}_j)\|_2^2)
\]

where \( h(\tilde{v}_i) \) is the histogram of uniform gray-scale and rotation invariant Local Binary Patterns [21] of the superpixel \( \tilde{v}_i \) and \( \| \cdot \|_\chi \) is the \( \chi \)-distance between two histograms. Third column of Figure 2 presents results of regularized clustering. At the end of the regularization process, three classes are still available. Therefore we post-process the final classification to determine if the second class (light green in Figure 2) belongs to the lesion or the surrounding skin using criteria based on area, color and spatial position relative to the detected lesion (dark-green in Figure 2). Fourth column of Figure 2 presents results of merging.

### 3.6. Pixel boundary spatial refinement

The obtained segmentation has been obtained on the superpixel graph, and the boundaries are not very precise. In a last step we want to refine the obtained superpixel labeling at the pixel level, starting from the result obtained from the superpixels. First we assign to each pixel \( v_i \) of a 8-grid graph (i.e., the pixel level) the label of the superpixel \( \tilde{v}_i \) it belongs to: \( l(v_i) = l(\tilde{v}_j) \). Each vertex is then associated with a cluster label \( l : \mathcal{V} \rightarrow \mathbb{Z} \). The set of vertices associated to one cluster \( j \) is: \( \mathcal{V}_j = \{ v_i \in \mathcal{V} : l(v_i) = j \} \) where \( j = 1, \ldots, C \) with \( C \) the number of classes. The aim of the clustering spatial refinement is to modify the labels assigned to pixel vertices in order to have a clustering that is better delineated along its boundaries. Therefore, we use again discrete label regularization. Since the spatial refinement has to be performed only around the boundaries of objects, we consider a specific grid-graph that is a subset of the whole grid-graph. The set of vertices that corresponds to the boundaries between two different clusters is defined by: \( \partial \mathcal{V} = \{ v_i \in \mathcal{V} : \exists v_j \in \mathcal{V} \text{ with } (v_i, v_j) \in E \text{ and } l(v_i) \neq l(v_j) \} \). The set of vertices that belongs to a narrow band of size \( 2\delta + 1 \) around the set \( \partial \mathcal{V} \) is defined by: \( \partial \mathcal{V}^+ = \{ v_i \in \mathcal{V} : \exists v_j \in \partial \mathcal{V} \text{ with } d(v_i, v_j) \leq \delta \} \) where \( d(v_i, v_j) \) is the length of the path \( \{ v_i = v_1, v_2, \ldots, v_n = v_j \} \) with \( (v_j, v_{j+1}) \in E \). The set of edges \( \mathcal{E}^+ \) is defined as the subset of edges in \( E \) that connects two vertices of \( \partial \mathcal{V}^+ \): \( \mathcal{E}^+ = \{(v_i, v_j) \in E : v_i, v_j \in \partial \mathcal{V}^+ \} \). Then, \( \mathcal{E}^+ \) is reduced to \( \mathcal{E}^{++} \) such that each vertex \( v_i \) is connect to only its 8-nearest neighbors (in terms of \( 3 \times 3 \) patch \( L_2 \) distance). The clustering spatial refinement is then accomplished by label regularization processes on the graph \( \mathcal{G}^+ = (\partial \mathcal{V}^+, \mathcal{E}^{++}, \mathcal{w}_3) \). We use \( p = 1 \), \( \lambda = 1 \), and weights are defined as \( w_3(v_i, v_j) = G_\sigma(||F(f_0, v_i) - F(f_0, v_j)||_2^2) \) where \( F(f_0, v_i) \) is a \( 3 \times 3 \) patch at \( v_i \). Last column of Figure 2 presents the results of spatial boundary refinement superimposed on the inpainted visible spectrum image.

### 4. RESULTS AND CONCLUSION

Our images were acquired with a multispectral camera under development by the Intuiskin company. We have 32 images of size 800 \( \times \) 600 with 6 spectral bands (3 under visible light and 3 under IR light) from the Grenoble Hospital. To evaluate the segmentation results, we have used image ground truths that have been manually created and we have considered the quantitative error metrics defined in [14]. Our method gives the following results: Sensitivity=93.62\%, Specificity=98.54\% and Error Prob.=2.79\%. This shows how well our method is able to extract the lesion from the multispectral image. In conclusion, we have presented a whole strategy for multispectral dermoscopy images that enables to segment images with a graph-based approach after the removal of hairs and reflections. In future works, we plan to optimize important parameters of the method.
Fig. 1. Image pre-processing steps illustration. From left to right: Visible Light image, Infra-Red Light image, Filtered Top-Hat, Hair detection, Hair removal, Reflection removal. See text for details.

Fig. 2. Image Segmentation steps illustration. From left to right: Superpixel segmentation, Superpixel clustering, Superpixel regularized clustering, Superpixel merged clustering, final segmentation after boundary spatial refinement, overlay on the image. See text for details.
5. REFERENCES


