Active Contour Models: Application to Oral Lesion Detection in Color Images

Ghassan Hamarneh, Artur Chodorowski, and Tomas Gustavsson
Image Analysis Group, Chalmers University of Technology
Göteborg SE-412 96, Sweden

Abstract

This paper presents the application of active contour models (Snakes) for the segmentation of oral lesions in medical color images acquired from the visual part of the light spectrum. The aim is to assist the clinical expert in locating potentially cancerous cases for further analysis (e.g. classification of cancerous vs. non-cancerous lesions). In order to apply the conventional snake formulation, color images were converted into single-band images. A number of different single-bands were evaluated including those resulting from the original and normalized RGB, perceptual HSI space, I_1I_2I_3, and the Fisher discriminant function. Examples of Snake segmentation results of oral lesions are presented.

Index terms: active contour model, snakes, oral lesions, cancer, true color images.

I. INTRODUCTION

The human oral mucosa is a site of a variety of disorders. Numerous diseases or lesions have been clinically classified. In particular, there exist lesions that have a potential to develop into oral cancer. The American Cancer Society estimates 30,200 new cases and 7,800 deaths in the US in the year 2000 of oral cancer. The preliminary diagnosis of oral disease is based on ocular inspection and registration of the patient’s oral cavity as true-color digital images. Although complementary techniques exist, based e.g. on infrared or fluorescence spectroscopy, in clinical practice the decision about further treatment of the patient is predominantly based on lesion appearance from the visual part of light spectrum. The automatic detection (segmentation) of color images of the oral mucosa is thus an important part of computer-aided oral lesion diagnosis systems (CADx). It is of great interest for the medical community working with oral lesions to have an automatic (or semi-automatic) method for segmenting the lesions in true-color images, since by doing that the next step of extracting the different features and the consequent classification (examining the potentiality of a malignant cancerous lesion) can be immediately performed and evaluated. A previous study evaluated the classification of lesions based on different color features with the lesions being manually segmented by medical experts. The oral specialists usually agree on the position of the lesion boundaries in the recorded images. However, this is still a challenging computer vision problem due to the shape and appearance variability of oral lesions. On the other hand, the machine is usually more efficient, after supervised learning, than humans in discrimination of different oral diseases. The automatic segmentation algorithm will simplify analysis of oral lesions and can be used in clinical practice to detect potentially cancerous lesions.

Currently our image database includes cases of two common oral lesions, the potentially cancerous lesions called leukoplakia and the usually harmless lesions called lichenoid reactions. Furthermore, the lichenoid reactions can be divided into atrophic, plaque-formed and reticular lesions. Thus the subsequent classification problem can be studied as a 2-class problem (cancerous vs. non-cancerous) or a 4-class problem (complete classification), see Figure 1. Both of the lesion types appear reddish/whitish to the human observer and are not easily differentiated. From a clinical viewpoint the boundaries of the lesions form a closed contour with no gaps. Most of the research in this field arises from dermatology and skin cancer detection. In contrast to skin lesions, the oral lesions are predominantly reddish and occupy a narrow band of hue-spectrume.
case the images are vector-valued and color seems to be an important part in human-based detection we have implemented and evaluated a snake formulation suitable for multi-band images. Essentially, different single bands are derived from the original RGB bands and their suitability for snake segmentation is compared. Snake segmentation results for a variety of typical cases are presented.

II. BACKGROUND

Active contour models belonging to the class of deformable models, have gained large acceptance as a segmentation tool. This is due to a collection of factors including the way Snakes consider the boundary as a single, inherently connected, and smooth structure. Snakes also support intuitive interactive mechanisms for guiding the segmentation deformations. Many variations to the original Snakes formulation have been proposed to improve their performance. For example, incorporating an inflation force changing topology using dynamic programming simulated annealing (in Brownian strings) genetic algorithms for snake energy minimization dual Snakes for dealing with initialization problems an adaptive Spline model, and many others. Another useful boundary detection method to use in this application could be the “live-wire” technique, which has been recently incorporated into united Snakes. A recent review on deformable models in medical image analysis can be found in Most of the previous work on deformable models was directed towards scalar-valued (intensity or gray level) images. Nevertheless, attempts have been made to modify, extend, and/or apply Snakes to detecting and segmenting objects in multi-band images.

III. METHODS

In this section we present the snake formulation and some implementation details, followed by a description of the single band generation and error calculation procedures.

A. Active Contour Model (Snake) Formulation

In active contour models, a contour is initiated on the image and is left to deform in a way that, firstly, moves it toward features of interest in the image and, secondly, maintains a certain degree of smoothness and continuity in the contour. In order to favor this type of contour deformation, an energy term is associated with the contour and is designed to be inversely proportional to the contour’s smoothness and the fit to desired image features. The deformation of the contour in the image plane will change its energy, thus one can imagine an energy (potential) surface on top of which the contour moves (in a way that resembles the slithering of a snake and hence the name) seeking valleys of low energy. This can also be formulated using a force field that causes this energy change (analogous to physical systems). Moreover, certain forces can be designed (or derived from energy terms) in a way that the resulting contour deformations will reduce its energy, thus yielding a smooth contour located along desired image features such as edges.

In our implementation a polygonal discrete active contour model is used and is represented by a set of nodes or vertices: \( v_n(t) = (x_n(t), y_n(t)) \) where \( n = 1, 2, \ldots, N \) is the node number and \( t \) denotes time or iteration number. The equation used for updating a single snake node is

\[
v_i(t + \Delta t) = v_i(t) - \frac{\Delta t}{\gamma} \left( \alpha F_{\text{tensile}}^i + \beta F_{\text{flexural}}^i \right) - F_t^i(t) - F_{\text{inflation}}^i(t)
\]

(1)

where \( \alpha \) and \( \beta \) are weighting factors, \( \Delta t \) is the time step, and \( \gamma \) is a damping coefficient. Equation (1) is derived from \( \mu \ddot{v}_i + \gamma \dot{v}_i + \alpha F_{\text{tensile}} + \beta F_{\text{flexural}} = F_{\text{external}} + F_{\text{inflation}} \), by setting the mass \( \mu \) to zero and using

\[
\dot{v}_i = \frac{(v_i(t + \Delta t) - v_i(t))/\Delta t}
\]

(2)

\( F_{\text{tensile}}^i(t) \) is a tensile force (resisting stretching) acting on node \( i \) at time \( t \) and is given by

\[
F_{\text{tensile}}^i(t) = 2v_i(t) - v_{i-1}(t) - v_{i+1}(t)
\]

(3)

\( F_{\text{flexural}}^i(t) \) is a flexural force (resisting bending) and is given by

\[
F_{\text{flexural}}^i(t) = 2F_{\text{tensile}}^i(t) - F_{\text{tensile}}^{i-1}(t) - F_{\text{tensile}}^{i+1}(t)
\]

(4)

\( F_{\text{external}}^i(t) = \nabla P(x_i(t), y_i(t)) \)

(5)

where \( P(x,y) \propto -\|I_s(x,y)\| \) and \( I_s(x,y) \) is the intensity of a pixel \( (x,y) \) in a smoothed version of the original image. \( F_{\text{inflation}}^i(t) \) is an inflation force that enables us to initialize the snake farther away from the target boundary and is given by

\[
F_{\text{inflation}}^i(t) = q_i F(I(x_i(t), y_i(t))) n_i(t)
\]

(6)

where \( n_i(t) \) is the unit vector in the direction normal to the contour at node \( i \) and the binary function

\[
F(I(x,y)) = \begin{cases} +1 & \text{if } I(x,y) \geq T \\ -1 & \text{otherwise} \end{cases}
\]

links the inflation force to the image data, and \( T \) is an image intensity threshold. In order to dampen the inflation force when the snake nodes reach the target boundary we associate a node-specific inflation weight \( q_i \). Having only a single value for the inflation force for all the nodes proved insufficient in our experiments, since this caused the nodes that reach the target boundary earlier than others to pass over it and cause the snake to “leak”. If a node reached the target boundary the inflation direction is
reversed (inflation becomes deflation and vice versa) and if a certain number of inflation reversals occurred within a limited number of past iterations then the inflation force is dampened for this particular node only. We also applied an adaptive subdivision scheme, where the number of snake nodes is re-sampled based on the distance between nodes and the curvature along the snake. Our Snakes’ implementation was also equipped with a facility that allows the user to place certain forced nodes on the target boundary through which the snake must pass.

B. Single Band Generation From Color Images

In order to apply the discussed ACM formulation on color images without deriving more complex multi-band forces, we derived a number of single bands from the original three-band (RGB) images. We have investigated the use of the single bands shown in Table 1 seeking the band in which the detected edges are most pronounced and coincide with the true lesion boundaries.

<table>
<thead>
<tr>
<th>Color space</th>
<th>Single bands</th>
</tr>
</thead>
<tbody>
<tr>
<td>RGB R: red G: green B: blue</td>
<td></td>
</tr>
<tr>
<td>HSI H: hue S: saturation I: intensity</td>
<td></td>
</tr>
<tr>
<td>Normalized RGB</td>
<td></td>
</tr>
<tr>
<td>R = R/(R+G+B) G = G/(R+G+B) B = B/(R+G+B)</td>
<td></td>
</tr>
<tr>
<td>Other</td>
<td></td>
</tr>
<tr>
<td>I1 = I/(I+I+I) I2 = (I-I)/2 I3 = (I+I-I)/2</td>
<td></td>
</tr>
<tr>
<td>Other</td>
<td></td>
</tr>
<tr>
<td>F: Fisher projection M: modified Fisher projection</td>
<td></td>
</tr>
</tbody>
</table>

Table 1. Single bands

The different single bands generated from one example color image are shown in Figure 2. All the single bands used are either linear or nonlinear transformations of the original RGB values to other color coordinates. However, the Fisher (F) and the modified Fisher (M) single bands require training. In order to generate the F and M single bands, manually segmented lesions (by clinical experts) in true-color images were supplied. For a single image two classes were formed representing the two regions near the boundary; inside the lesion (in) and outside (out), see Figure 3.

The Fisher single band image, $I_F(x,y)$, is calculated as (26)

$$I_F(x,y) = f^T I_{RGB}(x,y)$$

where

$$I_{RGB}(x,y) = [I_R(x,y) I_G(x,y) I_B(x,y)]^T$$

are the original RGB values,

$$f = (S^{in} + S^{out})^{-1} (\bar{T}^{out} - \bar{T}^{in})$$

$$S^{in} = \sum_{(x,y) \in in} (I_{RGB}(x,y) - \bar{T}^{in})(I_{RGB}(x,y) - \bar{T}^{in})^T$$

$$\bar{T}^{in} = \frac{1}{N_{in}} \sum_{(x,y) \in in} I_{RGB}(x,y)$$

and similarly for $S^{out}$ and $\bar{T}^{out}$.

In the modified fisher the within class scatter matrices ($S^{in}, S^{out}$) used in the original Fisher formulation were ignored (i.e. $S$ replaced by an identity matrix) which firstly simplifies calculations and secondly gives complete emphasis on generating a single band that possesses high contrast along the boundary edge, i.e. giving maximum separation of the means of the two classes with no regards to their respective variance.

![Figure 2. Different single bands derived from an RGB image of an oral lesion.](a) (b)

Figure 3. (a) Original image with expert delineation. (b) Inner (dark) and outer (white) samples used for the Fisher training.

C. Comparing Single bands

In the snake implementation there are different weighting factors and parameters to be set. Many of these depend mainly on the shape of the lesion, for example the tensile and flexural weights and the subdivision parameters. The threshold value $T$ (see equation (6)), on the other hand, is directly linked to the intensity of the image and hence to the single band under investigation. In order to compare the performance of Snakes using the different single bands, we fixed the values of all the parameters except for the threshold value. For each single band generated we performed a fully automated snake segmentation (without any manual corrections) over a feasible range of threshold values. To quantify the difference between the manually delineated boundary, $M$, and the snake-segmented boundary, $S$, we defined the following error measure...
\[
\varepsilon = \frac{A(S) \cup A(M) - A(S) \cap A(M)}{A(M)} \quad (11)
\]
where \(A(M)\) and \(A(S)\) are the areas enclosed within \(M\) and \(S\), respectively. This error measure was then used to determine which bands perform well to be used in the segmentation procedure. Figure 4 illustrates the various error measures for a selection of single bands obtained as described above for an example color image.

**Figure 4.** Error measures for various single bands vs. the threshold value calculated for one cancerous lesion.

IV. RESULTS

In this section we present some visual results of the snake segmentation of the oral lesions performed on single band images generated from the true color digital images. Figure 5(a) shows a single band image with only five snake nodes used for initialization and placed inside the target lesion region. Figure 5(b) shows the final result of the snake segmentation with the forced points used to constrain the snake shown as circles. The expert manual tracings of the oral lesion are shown in Figure 5(c). Figure 6 shows a similar snake segmentation result on a different lesion image using only four initial snake nodes. Figure 7 depicts the deformation of the snake and the progress of the segmentation. The four snake nodes used for initialization are shown in Figure 7(a) while Figure 7(b) shows the snake after deformations and subdivisions. Figure 7(c) shows the snake in a later stage of the deformation where it is stuck on an erroneous edge, not being able to reach the correct left side of the lesion boundary. The placement of a single forced point at the correct lesion boundary (the part that the snake couldn’t latch to) and how this improves the segmentation is depicted in Figure 7(d). The final segmentation result is shown in Figure 7(e) and the expert delineation of the oral lesion is shown in Figure 7(f). Notice how the number of snake nodes is adaptively increased to accommodate the complexity of the lesion boundary. Figure 8 illustrates the calculation of the error term in equation (11). The binary image in Figure 8(a) shows the area \(A(S)\) of the snake-segmented oral lesion of Figure 7(e) and Figure 8(b) shows the area \(A(M)\) of the manually segmented lesion. Figure 8(c) shows the area described by the numerator of the error term in (11).

![Figure 5](image_url)

**Figure 5.** Segmentation example using the Green band: (a) Initial snake nodes. (b) Final segmentation result (snake nodes shown as white dots and forced points as white circles). (c) The manual expert delineation of the oral lesion overlaid on the original color image.
Figure 6. Segmentation example using the Blue band: (a) Initial snake nodes. (b) Final segmentation result (snake nodes shown as white dots and forced points as white circles). (c) The manual expert delineation of the oral lesion overlaid on the original color image.

Figure 7. Segmentation example using the original Fisher projection band: (a) Initial snake nodes. (b) Progress of snake deformation (snake nodes shown as white dots). (c) Snake stuck on erroneous edge (left-side of lesion). (d) Addition of a forced point (white circle). (e) Final segmentation result. (f) The manual expert delineation of the oral lesion overlaid on the original color image.

Figure 8. Error calculation: (a) Area of snake-segmented lesion. (b) Area of manual delineation of the same lesion. (c) The erroneous area, $\varepsilon = 0.0953$ (9.53%).

V. DISCUSSION

Snakes proved to be a valuable method for the segmentation of oral lesions by guaranteeing continuous and smooth lesion boundaries. However, it is important that a user (or an add-hoc method) assists in the segmentation procedure by pointing to a rough region in the image where the target lesion exists (done here by specifying a few initial snake nodes). The user should also be ready to intervene by placing constrained (forced) points to assist the snake if it clings to erroneous edges.

Choosing the weights and parameters of the snake model is an important and often tedious task. For object images of similar modalities it is usually required to be performed only once for a specific application. However, in this case we encounter great variability between different lesions, which means that the same parameters will not necessarily be optimal for all cases. A potential solution to the optimal estimation of the internal snake parameters, with respect to all images, could lie in the use of reinforcement learning[27]. An alternative approach would be to design different Snakes for different medical cases. In particular, we have found it
more difficult to detect boundaries of reticular lichenoid reactions, since their shapes are usually more complex. We are currently investigating several enhancements to our current snake segmentation tool, most importantly, better utilization of the expert delineated results for choosing the weights and parameters of the snake model (possibly varying adaptively along the snake contour) and for defining energy terms that are suitable for all the images in our database.

VI. CONCLUSION

We have applied Snakes for semi-automatic segmentation of oral lesions in color images of the human oral cavity. Snakes reduced the need for edge linking compared to traditional edge based segmentation and led to small segmentation errors. However some operator interaction was still needed due to the large variability of the objects and images in this application. To further automatize and improve segmentation, additional or enhanced energy terms and more human knowledge should be incorporated into the Snakes design.

VII. ACKNOWLEDGEMENTS

Ghassan Hamarneh is funded by the Visual Information Technology (VISIT) program supported by the Swedish Foundation for Strategic Research (SSF). This work was also supported in part by KK-foundation, grant no. 1997/1716.

VIII. REFERENCES