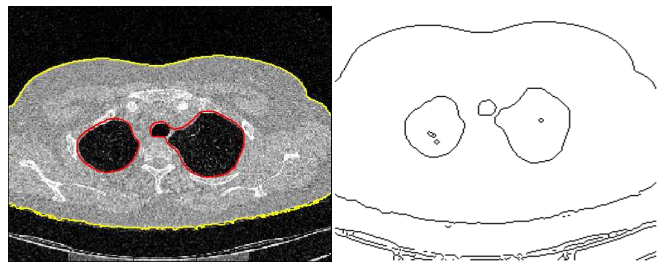
Automatic segmentation of the lungs using robust level sets

Margarida Silveira, Jacinto Nascimento and Jorge Marques
Instituto Superior Técnico - Instituto de Sistemas e Robótica
Av. Rovisco Pais, 1049-001, Lisboa, Portugal

***Abstract*—This paper presents a method for the automatic segmentation of the lungs in X-ray computed tomography (CT) images. The proposed technique is based on the use of a robust geometric active contour that is initialized around the lungs, automatically splits in two, and performs outlier rejection during the curve evolution. The technique starts by grey-level thresholding of the images followed by edge detection. Then the edge connected points are organized into strokes and classified as valid or invalid. A confidence degree (weight) is assigned to each stroke and updated during the evolution process with the valid strokes receiving a high confidence degree and the confidence degrees of the outlier strokes tending to zero. These weights depend on the distance between the stroke points and the curve and also on the stroke size. Initialization of the curve is fully automatic. Experimental results show the effectiveness of the proposed technique.**



(a) (b)

Fig. 1. Difficulties with the outlier features that prevent the contour from reaching the boundaries of the lungs; a) image with the level set segmentation superimposed b) associated edges.

I. INTRODUCTION

X-ray computed tomography (CT) is the most commonly used diagnosis technique for the analysis of the pulmonary region and the number of CT evaluations of the lungs has been steadily increasing. In most pulmonary CT image analysis applications the first step is the segmentation of the lungs. Some examples include airway analysis [7], emphysema detection [6], evaluation of lung ventilation [8], segmentation of the lobes [13] and the detection of lung nodules [10], [11].

Several algorithms have been proposed for the segmentation of the lungs. Most methods start with grey-level thresholding followed by region segmentation based on a sequence of morphological operations [4], [5], [10]. For instance in [5] the step of grey-level thresholding is performed using optimal thresholding to select the threshold automatically. Then connected components labelling is performed, the background air is eliminated by deleting regions that are connected to image borders and only the two larger regions are retained. Some methods include a priori anatomical knowledge [7], [8] which makes them more powerful but at the cost of more computational load. For instance in [8] anatomical knowledge stored in a semantic network is used to guide the low level image processing and a lung separation step based on dynamic programming is also included. Recently, an algorithm using marker based watershed transform was proposed [9] that eliminates the tasks of finding an optimal threshold and separating the attached left and right lungs. However, the identification of internal and external markers which is based on morphological operations relies on heuristics.

Active contours have also been proposed for the segmentation of the lungs. In [1] two independent ACM's based on geodesic gradient vector flow are used. Although this method works well on concavities it is unable to bridge anatomical structures such as the trachea or bronchi in case of adverse initialization. In [12] again two active contours are used that compete for the lungs boundaries using the EM algorithm.

The class of geometric active contours known as level sets has become very popular for medical image segmentation because of their independence from parametrization and the ability to automatically change topology during deformation. However, level set methods usually produce over segmented images of the lungs due to the presence of outliers features like the trachea or bronchi. The outlier features are produced by intensity transitions far from the lungs boundaries. Fig. 1 illustrates this difficulty.

In this paper we propose a level set method that overcomes this difficulty because it is robust with respect to outliers. The proposed technique is based on the use of a geometric active contour that is initialized around the lungs, automatically splits in two, and performs outlier rejection during the curve evolution. Image features are classified as valid or invalid making the curve stop only at valid features and allowing it to bridge the invalid ones. Our algorithm organizes edge points into strokes and classifies each stroke as valid or invalid. A confidence degree (weight) is assigned to each stroke and updated during the evolution process with the valid strokes receiving a high confidence degree and the confidence degrees of the outlier strokes tending to zero.

This paper is organized as follows: section II revises level set theory, section III describes the proposed algorithm for lung segmentation, section IV presents experimental results and section V concludes the paper.

II. Level Set and stopping force theory

In the level set formulation [14], the active contour is a moving front denoted by C which is represented implicitly by the zero level set $C(t) = \{(\vec{x}) | \phi(t, \vec{x}) = 0\}$ of a level set function $\phi(t, \vec{x})$. If the curve C evolves along its normal direction with speed V then the evolution of the level set function ϕ is governed by the following equation

$$\frac{\partial \phi}{\partial t} = F \left| \nabla \phi \right| \tag{1}$$

with $\phi(0, \vec{x}) = \phi_0(\vec{x})$ defining the initial contour. A particular case is the motion by mean curvature with F given by

$$F = div \left(\frac{\nabla \phi}{|\nabla \phi|} \right)$$

For the curve to stop propagating at the shape boundaries the speed function is multiplied by a stopping force c which approximates zero at the image features.

$$\frac{\partial \phi}{\partial t} = c(\vec{x}) \left(v + div \left(\frac{\nabla \phi}{|\nabla \phi|} \right) \right) |\nabla \phi| \tag{2}$$

A common choice is a stopping term based on the image gradient ∇I such as [15]

$$c(\vec{x}) = \frac{1}{1 + \left| \nabla G_\sigma(\vec{x}) * I(\vec{x}) \right|^2} \tag{3}$$

where $G_\sigma(\vec{x}) * I(\vec{x})$ is the convolution of the image with a gaussian kernel with standard deviation σ.

When the object boundary is indistinct or has gaps the contour tends to leak or bleed through the boundary. To address this problem a second type of stopping term based on edge strength has been proposed [16], [18] that pulls back the contour if it passes the boundary

$$\frac{\partial \phi}{\partial t} = c(\vec{x}) \left(v + div \left(\frac{\nabla \phi}{|\nabla \phi|} \right) \right) |\nabla \phi| + \nabla c . \nabla \phi \tag{4}$$

where v is a constant.

A third type of stopping function based on area minimization has been proposed in [17] to further prevent the boundary leaking problem

$$\begin{aligned} \frac{\partial \phi}{\partial t} = {} & c(\vec{x}) \left(v + div \left(\frac{\nabla \phi}{|\nabla \phi|} \right) \right) |\nabla \phi| + \nabla c . \nabla \phi \\ & + \left(\frac{V_0}{2} \vec{x} . \nabla c \right) |\nabla \phi| \end{aligned} \tag{5}$$

These stopping forces deal with the problem of boundary gaps but don't address the problem of spurious or outlier edges. To deal with this difficulty we propose a stopping force that classifies image features as valid (inlier) or invalid (outlier) making the curve stop only at the valid features and allowing it to bridge the invalid ones. Contrary to the above forces that are constant the proposed stopping force is adaptive and varies during the deformation process.

III. Robust level sets

In this section we introduce the proposed adaptive stopping force which allows the contour to bridge the invalid features and stop only at the valid ones. Our approach builds on the work of [19] that proposed a robust parametric active contour able to discard outlier features. The features used are connected sets of edge points, called strokes, which are more reliable than edge points.

We will start by introducing some notation. Let y be the set of all edge points detected in an image and let us assume that y is organized in connected components, called strokes, $y^j, j = 1, ..., N$ where $y^j = \{y_1^j, ..., y_n^j\}$ is the set of edge points belonging to the j-th stroke. A standard edge linking algorithm is used to compute the image strokes. Each stroke is assigned a confidence degree (weight) w^j verifying $0 \le w^j \le 1$, the valid strokes should receive a high confidence degree and the confidence degrees of the outlier strokes should tend to zero. These confidence degrees depend on the distance between the stroke points and the contour and also on the stroke size. Let x be a contour model defined by a sequence of 2D points $x_i, i = 1, \ldots, M$.

Our stopping function depends on the weights assigned to each stroke. Since we want the curve to stop at the valid strokes (with higher weights) and to bridge the invalid ones (with lower weights), the stopping function should be inversely proportional to the weights. We propose the following

$$c(x) = \frac{1}{1 + \left| G_\sigma(x) * w(x) \right|^2} \tag{6}$$

where $w(x)$ is the image of the weights calculated for all the edge points (points that are not edges have zero weight), and $w(x)$ is convolved with the gaussian kernel G_σ. This stopping function is adaptive because it depends on the weights $w(x)$ and they vary during the estimation process since they depend on the distance between the stroke points and the contour.

Therefore our algorithm is a two step iterative algorithm. In one step the contour is implicitly evolved using the evolution equation presented in eq. 4 and in the other step the stopping function is recalculated.

We will now explain how the weights are calculated. It is not easy to define a set of criteria to distinguish valid from invalid strokes. In this paper we will adopt two simple criteria, we assume that the invalid strokes are located far from the object's boundary and that they have usually a smaller length. The weights are derived in a probabilistic framework. Let k^j be a set of stroke binary labels $k^j = \{k^1, ..., k^N\}$; $k^j = 1$ if the j-th stroke is valid, $k^j = 0$ otherwise. We assume that y, x and k are random variables with a probability density function. We also assume that the strokes detected in the image are independent

$$p(y|x) = \prod_j p(y^j|x) \tag{7}$$

and that each stroke has i.i.d. edge points

$$p(y^j|x) = \prod_n p(y_n^j|x) \tag{8}$$

It is assumed that the contribution from an edge point of a valid stroke to the density is Gibbs distribution whose energy depends on a distance function between that edge point and the contour

$$p(y_n^j|k^j = 1, x) = \beta^j e^{-\sum_i d(y_n^j, x)} \tag{9}$$

where β^j is a normalization term related to the partition function. The following distance function was used [19]

$$d(y_n^j|x) = -\sum_i N(y_n^j; x_i, \sigma^2 I) \tag{10}$$

where $N(y;\mu,R)$ denotes the normal density function with mean μ and covariance R.

Substituting (9) and (10) into (8) we get

$$p(y^j|k^j = 1, x) = \beta^j e^{\sum_i \sum_n N(y_n^j; x_i, \sigma^2 I)} \tag{11}$$

In case a stroke is classified as invalid the contribution of its edge points to the density is considered a constant L^j.

$$p(y_n^j|k^j = 0, x) = \gamma^j e^{-L^j} \tag{12}$$

where γ^j is a normalization term. We used $L^j = -\frac{V}{n^j}$; since this contribution is inversely proportional to the size of the corresponding stroke, the smaller strokes will tend to be classified as outliers.

Substituting (11) into (8) we get

$$p(y^j|k^j = 0, x) = \gamma^j e^{VM} \tag{13}$$

The weights are given by the probability that a stroke is classified as valid.

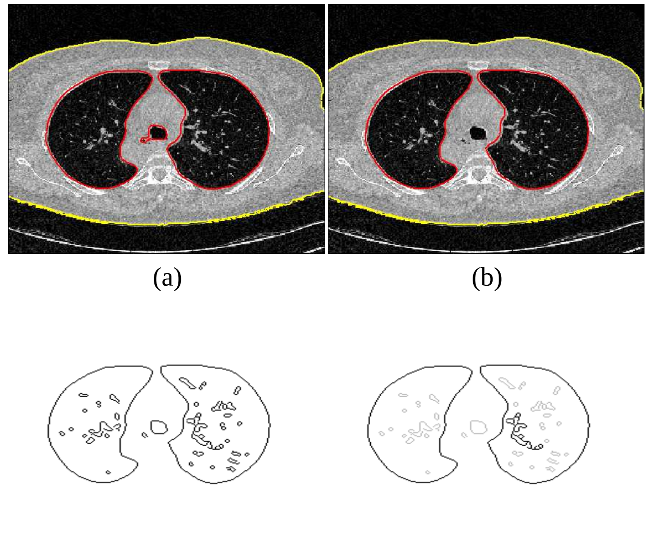
$$w^j = p(k^j = 1|y^j, x) = \frac{p(y_n^j|k^j = 1, x)p(k^j = 1)}{p(y_n^j|k^j = 1, x)p(k^j = 1) + p(y_n^j|k^j = 0, x)p(k^j = 0)} \tag{14}$$

substituting equations (11) and (13) into (14) and assuming that the a priori probabilities verify $p(k^j = 1) = \frac{\gamma^j}{\beta^j} p(k^j = 0)$ we derive the weights

$$w^j = \frac{e^{\sum_i \sum_n N(y_n^j; x_i, \sigma^2 I)}}{e^{\sum_i \sum_n N(y_n^j; x_i, \sigma^2 I)} + e^{VM}} \tag{15}$$

Contour initialization

The contour is automatically initialized at the outer chest wall, by a very simple yet effective method. The binary image obtained by gray-level thresholding using Otsu's method that was used to detect edges is processed by a morphological flood filling operation to remove holes. Then, the region with the largest area is retained and the initial contour is placed around this region. Furthermore, the edges outside this region are removed to prevent them from attracting the contour.



(a) (b)

(c) (d)

Fig. 2. Segmentation of the lungs; a) segmentation with the classical method b) segmentation with the proposed method c) detected edges and d) final inlier/ outlier classification.

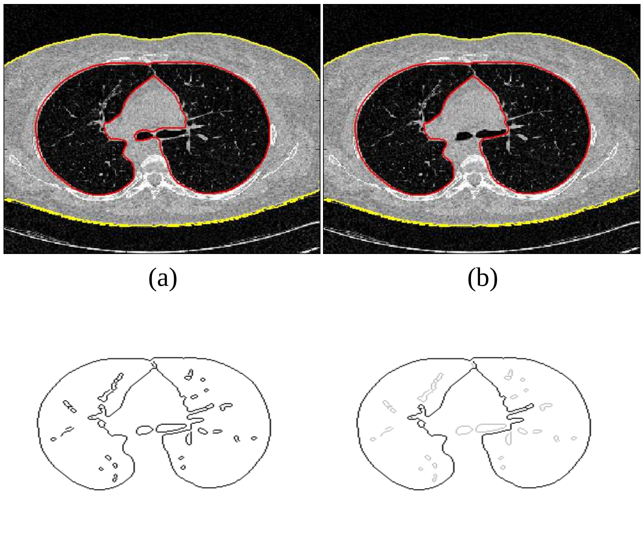
IV. Experimental results

This section presents examples to illustrate the performance of the proposed method. The input data consists of stacks of chest CT slices with X-ray attenuation ranging from -1024 to 3071 Hounsfield units, corresponding to a 12 bit quantization. Images are 512x512 with slice thickness of 1.0 millimeters. Edges were obtained with the Canny edge detector and strokes were obtained with a connected components labelling algorithm.

The first example illustrates performance of the algorithm in the middle pulmonary region and shows the robustness of the method with respect to the initialization.

Figures 2 and 3 compare the results of the classical edge based level set segmentation method with the results of the proposed method. On the top row of each of the figures the segmentation results are displayed, the initial contours superimposed in yellow and the final contours in red. The bottom row shows the edge points detected and the final classification of the strokes obtained by the proposed method, outliers in gray and inliers in black. It can be seen that in both situations the proposed algorithm was able to split the contour in two and to bridge the smaller strokes corresponding to other structures in the chest region until it reached the borders of the lungs. The classical method originated an extra region in the example of Fig. 2 and failed to converge to the lungs boundaries in the example of Fig. 3 because it was stopped by the outlier edges.

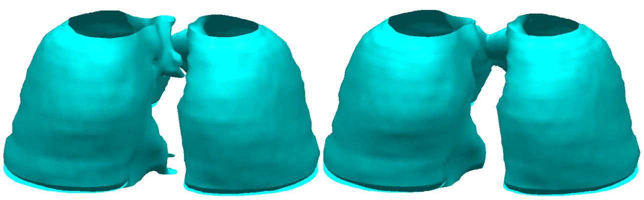
Fig. 4 compares the 3D reconstructions from the segmentation of the complete stack of 2D slices obtained using the classical edge based level set method with the ones obtained using the proposed method. The oversegmentation produced in the classical method by outlier features like the trachea or bronchi jeopardizes the 3D reconstruction of the lungs



(a) (b)

(c) (d)

Fig. 3. Segmentation of the lungs; a) segmentation with the classical method b) segmentation with the proposed method c) detected edges and d) final inlier/ outlier classification.



(a) (b)

Fig. 4. Reconstruction obtained from the Segmentation of a complete data set. a) Results of the classical method b) segmentation with the proposed method.

surface. This difficulty is overcome by the proposed method.

V. Conclusions

This paper proposes a method for the automatic segmentation of the lungs in X-ray computed tomography (CT) images using level set segmentation. Level set methods usually produce over segmented images of the lungs due to the presence of outliers features. The proposed method overcomes this difficulty by performing outlier rejection during the curve evolution. Image features are classified as valid (inlier) or invalid (outlier) making the curve stop only at the valid features and allowing it to bridge the invalid ones. It is shown that level set segmentation with the proposed stopping force is robust and able to segment the lungs bridging other anatomical structures such as the trachea or bronchi. No user intervention in required to initialize the contour since the initialization is fully automatic.

Future work will focus on lung separation and lobe identification and also on a quantitative comparison between the results obtained with the proposed method and those obtained by other methods in the literature, using for instance the metrics proposed in [4].

VI. Acknowledgemnts

This work was supported by Fundação para a Ciência e a Tecnologia (ISR/IST plurianual funding) through the POS_Conhecimento Program that includes FEDER funds. The authors would like to thank Beatriz Sousa Santos for providing the images.

References

[1] Paragios, N. Mellina-Gottardo, O. Ramesh,V.,Gradient Vector Flow Fast Geometric Active Contours, IEEE Transactions on Pattern Analysis and Machine Intelligence, 26, 3, (2004), 402-407

[2] Park, H., Schoepflin, Y.: Active contour model with gradient directional information: Directional snake. IEEE Transactions on Circuits and Systems for Video Technology, 11, 2, (2001), 252-256

[3] Otsu, N.: A threshold selection method from gray-level histograms. IEEE Transactions on Systems, Man and Cybernetics, SMC-9, (1979), 62-66

[4] Silva, A., Silvestre Silva, J., Santos, B., Ferreira, C.: Fast Pulmonary Contour Extraction in X-ray CT images: A Methodology and Quality Assessment. Medical Imaging 2001: Physiology and Function from Multidimensional Images, 4321, (2001), 216-224

[5] Hu S., Hoffman E. A., Reinhardt, J. M.: Automatic lung segmentation for accurate quantitation of volumetric X-ray CT images. IEEE Transactions on Medical Imaging, 20, 6, (2001) 490-498

[6] Blechschmidt R. A., Werthschutzky R, Lorcher: Automated CT image evaluation of the lung: a morphology-based concept. IEEE Transactions on Medical Imaging, 20, 5, (2001), 434-442

[7] Park W., Hoffman E., A. , Sonka M.: Segmentation of intrathoracic airway trees: a fuzzy logic approach. IEEE Transactions on Medical Imaging, 17, 4, (1998), 489-497

[8] Brown MS, McNitt-Gray MF, Mankovich NJ, Goldin JG, Hiller J, Wilson LS, Aberle DR.: Method for segmenting chest CT image data using an anatomical model: preliminary results. IEEE Transactions on Medical Imaging, 15, 6, (1997), 828-839

[9] Shojaii, R., Alirezaie, J., Babyn R.: Automatic lung segmentation in CT images using Watershed Transform. IEEE International Conference on Image Processing, II, (2005), 1270-1273

[10] Dajnowiec, M., Alirezaie, J.: Computer simulation for segmentation of lung nodules in CT images. IEEE International Conference on Systems Man and Cybernetics, 5, (2004), 4491-4496

[11] Yongbum, L., Hara, T., Fujita, H.,Itoh, S., Ishigaki, T.: Automated detection of pulmonary nodules in helical CT images based on an improved template-matching technique. IEEE Transactions on Medical Imaging, 20, 7, (2001), 595-604

[12] Margarida Silveira, Jorge S. Marques, ”Automatic segmentation of the lungs using multiple active contours and outlier model”, 28th IEEE Engineering in Medicine and Biology Society Annual International Conference, pp. 3122-3125, 2006.

[13] Wei, Q. Hu, Y. Macgregor, J.H. Gelfand. Segmentation of Lung Lobes in Volumetric CT images for Surgical Planning of Treating Lung Cancer. 28th IEEE Engineering in Medicine and Biology Society Annual International Conference, pp. 4869-4872 , 2006.

[14] S. Osher and J. Sethian. Fronts propagating with curvature-dependent speed: Algorithms based on Hamilton-Jacobi equations. Jour. Comp. Phys, 79:12-49, 1988.

[15] R. Malladi, J. A. Sethian, and B. C. Vemuri. Shape modeling with front propagation: A level set approach. IEEE Trans. PAMI, 17(2): 158-175, 1995.

[16] V. Caselles, R. Kimmel, and G. Sapiro. Geodesic active contours. International Journal of Computer Vision. 22: 61-79, 1995.

[17] K. Siddiqi, Y. B. Lauzi‘ere, A. Tannenbaum, and S. W. Zucker. Area and length-minimizing flows for shape segmentation. IEEE Transactions on Image Processing. 7: 433-443, 1998.

[18] A. Yezzi, S. Kichenassamy, A. Kumar, P. Olver, and A. Tannenbaum. Geometric snakes for edge detection and segmentation of medical imagery. IEEE Trans. Medical Imaging, 16: 199-209, 1997.

[19] J. C. Nascimento and J. S. Marques. Adaptive snakes using the EM algorithm. IEEE Trans. Image Processing, vol. 14, no. 11, pp. 1678-1686, 2005.