Automated Image Analysis of Lung Branching Morphogenesis from Microscopic Images of Fetal Rat Explants

Pedro L. Rodrigues,1,2,3 Nuno F. Rodrigues,1,2 Duarte Duque,2 Sara Granja,1 Jorge Correia-Pinto,1 and João L. Vilaça1,2

1 ICVS/3B’s, PT Government Associate Laboratory, Guimarães, Braga, Portugal
2 DIGARC, Polytechnic Institute of Câvado and Ave, Barcelos, Portugal
3 Life and Health Sciences Research Institute (ICVS), School of Health Sciences, University of Minho, 4710-057 Braga, Portugal

Correspondence should be addressed to Pedro L. Rodrigues; pedrorodrigues@ecsaude.uminho.pt

Received 28 May 2014; Accepted 21 July 2014; Published 28 August 2014

Copyright © 2014 Pedro L. Rodrigues et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

1. Introduction

Branching morphogenesis results on the creation of branched structures in the body and is a key and fundamental feature of several organs development and growth, such as lungs, pancreas, salivary glands, mammary glands, kidney, and prostate [1–3]. The lung branching morphogenesis (LBM) of fetal rat explants, grown in vitro, has been an essential tool in research of molecular and cellular development mechanisms. This methodology has been widely studied, at different gestational ages in vivo and in vitro, in many research centers due to its stability and versatility [4–7].

Usually, LBM analysis involves a morphometric analysis of lung explants differentiation and growth, during a 5-day period, using stereo microscope images acquired at 24-hour intervals. For each day, a comprehensive study of the branching pattern of embryonic lungs by quantifying the branches perimeter, area, outer contour, and number of peripheral airway buds is performed.

Although, over the past decade, there have been significant advances in understanding of the genetic control of lung development, to the best of our knowledge, all LBM studies are still performed by manual image quantification using generic 2D curves software. LBM analysis remains a time-consuming procedure, dependent on researcher expertise and error-prone. Often, it prevents the biological result comparison among different researchers, to deduce biological validations and theorems, due to ambiguous and inaccurate measurements [8]. Therefore, besides the different image processing approaches proposed in the biological research domain [9–18], none of these works are applied to LBM analysis preventing us from discussing the state-of-the-art technology to deal with the same problem.
Considering the literature pitfalls, we propose a new methodology capable of automatically performing the LBM morphometric analysis in order to reduce or even eliminate the researcher dependence, providing fast, robust, user-independent, and accurate results.

2. Methods

All methods were developed using C++, ITK (Segmentation & Registration Toolkit), and VTK (Visualization Toolkit). RGB images were acquired at the Life and Health Sciences Research Institute (ICVS) of School of Health Sciences, University of Minho, Portugal, using a stereo microscope (Olympus SZX16). Each CT slice has $768 \times 576$ pixels, with a pixel resolution of 40 $\mu$m. All images were acquired at the $\times 20$ magnification.

Depending on the biological laboratory trials, LBM images can be acquired with different color conditions (examples in Figures 1(a) and 1(b)). The computer application presented in this work was tested and validated for both kinds of images. The segmentation process assumes that the region of interest to be segmented is the union of one or more small primitive regions previously calculated from the input image. To this extent, different seeds were automatically calculated and placed within the lung epithelia allowing a multithreading cluster growth.

2.1. Preprocessing Filtering. The RGB input images were first converted to grayscale values by averaging and normalizing the 3-color components. Then, the grayscale image was input to an anisotropic diffusion algorithm [19] which reduces the noise spots corrupting the image.

This algorithm depends on three parameters (empirically defined): the number of iterations, edge parameter ($\sigma$), and an edge-stopping diffusivity function $g(x, \sigma)$, according to Tukey’s function: $g(x, \sigma) = (1/2)[1 - (X/\sigma)^2]$, if $|x| \leq \sigma$ or $g(x, \sigma) = 0$, if $|x| > \sigma$. The anisotropic diffusion algorithm worked as a low-pass filter for noise reduction but preserving sharper boundaries and image contours, producing uniformity in the output image intensity (Figures 2(a) and 2(b)).

This outcome was used to automatically segment (1) the outer contour of the lung explant object and also (2) the inner epithelia by merging different clusters according to an image partitioning.

2.2. Lung Explant Outer Contour. An adaptive and multiscale thresholding technique was used to accomplish the segmentation of the whole lung explant object from the background.

Initially, a global threshold that maximizes the image entropy between a segmented object (the lung explant) and its background was automatically calculated. For that, consider an image $I$ with $N$ pixels, $I(i)$ as the image intensity at position $i$ ($i = 1, 2, \ldots, N$), and IMin and IMax equal to 0 and 255. Moreover, consider $T(i)$ as the result of a threshold algorithm with a threshold level at $I(i)$, $E_O(i)$ as the entropy of the objects of $T(i)$, $E_B(i)$ as the entropy of the background of $T(i)$, and $E_T$ as the total entropy ($E_B(i) + E_O(i)$). Ranging from IMin to IMax, the value of $T$ was determined by the intensity $I(i)$ that maximized $E_T$.

Although this global threshold produces suitable contours in all lung objects with image properties as Figure I(a), it fails for Figure I(b) due to the irregularities, small contrast variability, and outer contour ambiguities.

Consequently, for this kind of images, the initial outer contour was redefined with the following steps:

(a) calculation of an initial binary object and its centroid ($C$ in Figure 3(b)) is done;

(b) determination of different lines with origin at $C$ (slopes incremented from $0^\circ$ to $360^\circ$ with $45^\circ$ of step) is done;

(c) let $N$ be the total number of lines ($N = 8$); calculate the distances ($D_i, i = 1, 2, \ldots, N$) between the origin point and the one that intersects the initial contour ($C_i, i = 1, 2, \ldots, N$);
2.3. Epithelial Segmentation

2.3.1. Image Partitioning. The image gradient magnitude of the filtered image was input to an algorithm that divides the input image into small regions. Figure 4(a) shows how the different regions were labelled by a starting point (red, Figure 4(a)) and follow the flow line, whose direction was the gradient of minimum local intensity. The minimum gradient (green, Figure 4(a)) path of each pixel \( p \) of the input image was tracked by recursively selecting a pixel \( q \) in the 8-connected neighborhood (yellow, Figure 4(a)). If more than one pixel \( q \) exist, the last pixel found was taken, considering \( p \) as a reference pixel. Every pixel \( q \) along the path is marked as a local minimum of the gradient magnitude and assigned a distinct label.

In the end, the whole image was divided into primitive regions. Each region shares the same statistical properties and the boundaries coincide with the ridges of the gradient magnitude surface (Figures 4(a) and 4(b)).

2.3.2. Clustering Regions

Creating Seeds. The image partitioning contains a set of nonoverlapping regions. Although the probability of having region boundaries corresponding to boundaries of important objects increases with oversegmentation, it can also create
Figure 4: (a) Labeling process using the image gradient magnitude: (left) image gradient magnitude; (center) zooming image area; (right) the numbers are the pixel intensities, red is the starting pixel, green is minimum pixel intensity in the 8-connected neighborhood (in yellow), and the arrow is the searching direction. (b) and (c) are image partitioning in rat lung explants for Figures 1(a) and 1(b).

many insignificant boundaries. This stage describes how one dealt with this problem and the inner lung epithelium was automatically determined. Briefly, the procedure consists of the identification and clustering of similar primitive regions.

Each cluster starts growing from different seeds, initialized within the lung epithelium, along different lines $L_i$ (with $i = 1, 2, \ldots, 8$ (total number of lines), Figure 5(a)).

First, a neighborhood $N_i$ (with $i = 1, 2, \ldots, 8$, white circle in Figure 5) defines a kernel with center at centroid $C$, circle shape, and radius of 8 pixels. An iterative process transverses each line $L_i$ with $N_i$ (white arrow in Figure 5), while it calculates the average distribution of the kernel neighborhood originating different candidate seeds. The final seed for clustering grown will be the one where the average distribution within the kernel was minimum (black circles in Figure 5 defining a seed $S_i$).

The regions belonging to each seed $S_i$ were used to calculate initial statistical properties of the epithelia (centroid, average distribution, minimum and maximum values, region edges, region neighbors, and entropy) that were used for clustering growth.

**Clustering Growth.** The merging procedure was based on the similarity between regions formulated mathematically as a local optimization problem using the minimum description length principle [21]. If any primitive region is neighbor of a cluster, initialized at a seed $S_i$, a decision rule will state if it should be included or not.

Let $f(x, y)$ be a two-dimensional function that denotes the 2D input image with $k$ constant regions and $I_i$ ($i = 1, 2, \ldots, k$) the original image intensity in the $i$th region. Let $Q(x, y)$ be a region of $f(x, y)$ with circular shape that includes a maximum number of regions ($n = 100$, experimentally calculated), located within the outer contour and center at the new query region $R_n$ (tested whether its inclusion in the cluster $C_j$ is adequate) (Figure 6(a)). Using a region $Q(x, y)$ with a limited number of primitive regions reduces the probability to merge statistically outlier regions, providing a truer picture of the inner epithelia.

According to the minimum description length, the image data was coded in order to determine the total number of bits necessary to encode the region $Q(x, y)$ given by $B_R = \sum B_i(R_i) + B_B(R)$, where

1. $B_i(R_i)$ is the total number of bits needed to describe the image intensity for each region given: $B_i(R_i) = nR_iH(R_i)$ (with $nR_i$ being the number of pixels and image intensity entropy within $R_i$);

2. $B_B(R)$ is the number of bits needed to code the region boundary information given by $B_B(R) = N_B(R) \cdot b_1 + N_B(R) \cdot b_2$ (with $N_B(R)$ being the number of regions in $R$, $N_B(R)$ the total boundary length of the
Figure 5: Overview of the clustering procedure. The white circle represents a kernel at center $C$; the black circles represent the initial seeds for clustering growth.

Figure 6: Overview of the clustering growth procedure: (a) shows how a local region $Q(x, y)$ is calculated in a cluster $C_i$ with seed $S_i$ with center at $R_n$ (query region tested whether its inclusion is suitable using the minimum description length principle); (b) and (c) show a multicenter growth simultaneously.

As a local optimization problem, the inclusion of new region $R_n$ in the cluster $C_i$ aims at providing the largest positive description length gain $G$ in $Q(x, y)$ at each step with local optimization.

If a new region is merged in the cluster, the cluster will have more pixels. Hence, more bits are needed to encode the cluster image intensities. However, the common boundary segment between $R_n$ and $C_i$ disappears and the total description length might decrease, since the number of bits needed to encode the new region boundary information decreases.

With $nb(R_n, C_i)$ being the number of common boundary elements of $R_n$ and $C_i$ and $C_{new}$ the new cluster resulting from the merging procedure, the value of description length gain $G$ associated with this merging is given by

\[ G = nR_n H(R_n) + nC_i H(C_i) - nC_{new} H(C_{new}) + b_1 + nb(R_n, C_i) \cdot b_2. \]  

If $G > 0$, the image intensities between $R_n$ and $C_i$ are quite similar, whereby these regions belong to the same object. Thus the entropy increase is compensated by the elimination of the common boundary, and then these two regions should be merged in order to minimize $B(R)$.

2.4. Buds Counting. The number of peripheral buds was counted based on the lung inner epithelial skeletonizing [22]. The skeletonizing is performed by removing the boundary and corner points of the epithelial object, until only the skeleton remains as white pixels.

The number of peripheral buds was determined as the number of parents (circles in blue, Figure 7(c)) of the skeleton branched ends (circles in red, Figure 7(c)). Only the branched ends located near the outer contour (distancing less than 25%
of the axis that transverses the lung object vertically) were considered.

3. Results

The performance of this new computer algorithm was tested and validated in a total of 210 images: 90 (18 images for each day of culture) with image conditions as Figure 1(a) and 120 (24 images for each day of culture) with image conditions as Figure 1(b). All images were previously segmented by three experienced researchers: each user manually segmented the inner epithelia and outer contour and counted the number of peripheral buds.

The performance of the manual quantification was accessed by comparing the results among different researchers. The inner epithelia and outer contour were evaluated using the dice similarity coefficient (DSC) that quantifies the spatial overlap between different segmentations as follows:

\[ \text{DSC}(A, B) = \frac{2|A \cap B|}{|A| + |B|} = \frac{2\text{TP}}{2\text{TP} + \text{FP} + \text{FN}}, \]

where TP is the number of true positives, FP the false positives, and FN the false negatives. The DSC score ranges from 0%, indicating no spatial overlap between sets of binary segmentation results, to 100%, indicating complete overlap.

Table 1 shows the mean DSC score and error difference for the LBM morphometric analysis among three researchers. Table 2 shows the mean differences when the automatic method was compared to each manual result. The peripheral buds error is the difference average error between users in Table 1 and difference average error between manual and automatic methods in Table 2.

Best results were obtained in the first two days for both types of images with less standard deviations. High similarities between the manual and automatic procedure were not easily obtained for the last two days of culture, due to the increased number of peripheral airway buds and lung architecture complexity.

Statistical analysis using ANOVA test (under SPSS Windows version 17.0 software where \( P \) values lower than 0.05 were accepted as significant) shows nonsignificant differences between the automatic and manual results (inner epithelia and outer counter) in all culture days. On the other hand, the error of peripheral buds was nonsignificant only on the first three days of culture. Nonsignificant differences were found between researchers. The time for LBM morphometric analysis was reduced to 1-2 seconds/image in contrast to the manual one (2-3 minutes/image).

Results from different images regarding images with different lighting conditions are shown in Figure 8.

A user interface was also developed that allows the user to spatially understand the microscope image and rapidly produce an automatic segmentation. If necessary, manual editing may be used to correct automatic segmentation results. Although this manual editing (used to eliminate false merges) improves DSC scores (>98%) and decreases the standard deviations, it also adds user dependence and slows the morphometric analysis (depending on the editing degree).

This strategy has been used at the ICVS research lab to study and analyze the effect of different concentrations of a specific inhibitor in lung branching morphogenesis.

4. Discussion

A computer application was developed, providing an automatic procedure to enable a fast LBM analysis. The morphometric analysis efficiency and robustness were increased while time consumption, user dependence, and subjectivity were decreased or even eliminated. The observer variability was eliminated since all regions were computed and merged automatically.

The segmentation rate depends on the number of regions needed to be merged to select the entire lung epithelial region. However, the processing time was always about 98% less than the manual one.

The merging procedure was essential to achieve a good segmentation, since a significant amount of regions were initially created by the partitioning algorithm. The automatic seed selection was suitable to segment the inner epithelia using the minimum description length criterion that
Figure 8: LBM morphometric analysis for different type of images: outer counter in yellow, inner epithelia in red, and branched skeleton in white.

Table 1: DSC scores of LBM analysis comparing different researchers results when the study was made in different culture days and image of types.

<table>
<thead>
<tr>
<th>Days of culture</th>
<th>Inner epithelial DSC %</th>
<th>Outer counter DSC %</th>
<th>Peripheral buds error</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>A images</td>
<td>B images</td>
<td>A images</td>
</tr>
<tr>
<td>1</td>
<td>91.77 ± 5.13</td>
<td>86.91 ± 5.83</td>
<td>97.87 ± 4.65</td>
</tr>
<tr>
<td>2</td>
<td>88.99 ± 2.03</td>
<td>86.45 ± 4.23</td>
<td>97.32 ± 5.84</td>
</tr>
<tr>
<td>3</td>
<td>87.93 ± 5.44</td>
<td>87.26 ± 4.72</td>
<td>97.73 ± 4.89</td>
</tr>
<tr>
<td>4</td>
<td>87.97 ± 6.76</td>
<td>85.53 ± 8.23</td>
<td>96.19 ± 7.00</td>
</tr>
<tr>
<td>5</td>
<td>87.60 ± 7.94</td>
<td>85.78 ± 9.69</td>
<td>97.15 ± 5.46</td>
</tr>
</tbody>
</table>

selectively cluster regions based on their image intensity distribution similarity.

It was seldom observed that clusters merged dissimilar regions due to ambiguity and lack of definition of the inner lung explants contours. The worst results were obtained in the last two days of culture (Table 2), whose images present low details and perceptibility, and the branched ramifications increase drastically. However, the usage of a local image region, centered at the query region, tested whether its inclusion is suitable using the minimum description length principle. This allowed a reliable cluster to grow along epithelial object that changes gradually over space.

The local threshold was efficient to automatically delineate the outer contour with high DSC scores in images with lighting characteristics presented in Figure 1(b).

The main difficulties found in the segmentation procedure were the contrast ambiguity and variances of inner epithelia, complexity of the branched shape, and size, and the presence of neighborhood regions with the same density. Moreover, response to drugs and biological markers and different culture medium can induce image intensity variations and shadows near the outer contour.

These difficulties were more evident at the last two days of culture where DSCs are lower, even when comparing the results from different researchers. However, the inner epithelial segmentation was always segmented with scores around 90%, even at the last two culture days, indicating successful segmentation.

However one has to recognize some limitations within this work. The proposed algorithm overpredicts the number of peripheral buds with statistical significance with manual counting. However, peripheral buds counting is a controversial issue, without any stated method. As presented in Table 1, the coefficient of variation (defined as the ratio between the standard deviation and mean) among researchers is around 1, showing that there is no coherency between them. With this methodology, one aimed to broaden and generalize this procedure among different researchers. However, further
5 Conclusions

This work presents an automatic segmentation procedure and implementation, providing a technique for LBM morphometric analysis. The proposed method introduces a series of advantages related to its intuitive use and accuracy, making the technique suitable to images with different lighting characteristics. The total number of human decisions, time consumption, and user dependence were significantly decreased.

Due to its automatization nature, this application allows a reliable comparison of different researchers’ results and the possibility for more than one researcher to perform the same LBM study.

Results show that further work is needed regarding the merging procedure and the development of image enhancement techniques to improve inner lung epithelial contrast, mainly in the last days of culture. Moreover, a new strategy must be developed for counting the number of peripheral airway buds. Finally, this method has the potential to be used in different research environments to improve research outcomes.

Conflict of Interests

The authors declare that there is no conflict of interests regarding the publication of this paper.

Acknowledgment

The authors acknowledge Foundation for Science and Technology (FCT), Portugal, for the fellowship with the references: SFRH/BD/74276/2010 and SFRH/BPD/46851/2008.

References


Submit your manuscripts at
http://www.hindawi.com