CELL-BASED GRAPH CUT FOR SEGMENTATION OF 2D/3D SONOGRAPHIC BREAST IMAGES

Hsin-Hung Chiang, Jie-Zhi Cheng, Pei-Kai Hung, Chun-You Liu, Cheng-Hong Chung, Chung-Ming Chen

Institute of Biomedical Engineering, National Taiwan University, Taipei, Taiwan.

ABSTRACT

Boundary delineation is the fundamental basis of many sonographic image analyses. In sonographic breast lesion images, it’s complicated and time consuming for physicians to delineate the lesion boundaries. When it comes to three dimensional sonographic breast lesions image, delineation of lesion boundary becomes much more complicated. Taking advantage of cell competition algorithm along with its good region structure, generated cells can be served as elegant nodes for graph cut. Further, a similar weight function plays an important role in the estimation of lesion boundary to avoid visible weak edge and isolated node in graph cut.

The integration of cell competition and graph cut can be intuitively implemented in three dimensional images, in addition to the reduction of computational time. With efficiency and accuracy of lesion detection, a computer aided system was therefore developed to fulfill clinical applications.

Index Terms— Sonographic Image, Cell Competition, Graph Cut

1. INTRODUCTION

Morphological features from boundaries are informative clues in the clinical diagnosis of breast lesions in sonograms. The morphological features can be utilized for further quantitative analysis [1], or even applied to computer-aided diagnosis systems [2, 3], to assist the diagnosis of malignancy/benignancy. Important as it is, the task of boundary delineation is not trivial at all. Manual delineations are quite inefficient and the correctness highly depends on the experience of medical doctors. Particularly, when it comes to the lesion analysis in three dimensional breast ultrasound images, manual delineations are even far from practical. It is because extra efforts need to be made to assure the consistence among the consecutive slices. Accordingly, many segmentation methods were devised to save the boundary delineation efforts for sonographic imaged breast lesions from tedious manual workings.

In the literature, many studies attempted to address the problem of demarcating breast lesions in sonography by applying conventional approaches. Methods may range from thresholding [4, 5], deformable models [6], level-set methods [7], and graph-cut approaches [8]. Although satisfactory performances had been attained in the specific image sets used, each algorithm remains fall short in dealing with the complex shapes and texture patterns within and surrounding of various kinds of lesions.

By and large, most methods in the literature sought the best solution in pixel-by-pixel fashion. In [9], pixel-based methods for breast sonography were shown to be less statistically powerful against speckle noises. While pixel-based approaches may be improved by considering a specific range of local neighborhood for one pixel, determining the neighborhood size remains a crucial issue. On the other hand, cell-based scheme [9-10] was shown to be effective in breast sonography, where ‘cell’ is a small region with homogeneous intensity distribution. The cell structure is a natural tessellation of original sonography by a two-pass watershed transformation [9, 11]. The advantage of tessellating sonography into cell structure can be mainly twofold. Firstly, the cell structure can provide sufficient information of local echogenisity distribution to resist the speckle noise. Secondly, the search space spanned by cell structure is extremely less than pixel space. As a result, it will be more efficient to find the optimal segmentation solution by cells than pixels.

The key point of the cell-based scheme to achieve promising performance consists in how to group the cell structure in meaningful regions. Although the cell competition algorithm [9] were able to partition the sonograms into perceptual prominent regions, lesions with complicated echogenicity, especially malignant lesions, may still be fragmented into several prominent regions. It is because the cell competition algorithm only incorporated local features in its cost function. Moreover, most lesions with complicated echogenicity are constituted of more than one prominent region in sonograms. As a result, a higher level framework considering relatively global features is needed to piece the prominent regions into a whole lesion. To achieve this end, a graph-cut scheme based on the prominent region structure is devised to delineate integral breast lesions in sonograms. The benefit of utilizing graph-cut scheme is that the foreground/background priors provided from user can be easily incorporated. Since it
remains far from practical to automatically locate breast lesions in sonograms, the user priors will be extremely helpful to guide the assembly of prominent regions into a whole lesion.

In this paper, a hierarchical segmentation framework is proposed by tessellating the image/ROI from low to high level structures. The primitive tessellation, i.e., the cell structure, is generated first by the two-pass watershed transformation as the fundamental basis structure. The cell competition is then performed to re-organize the cell structure into a middle level structure, which is constituted of prominent regions. The prominent region is defined as a homogeneous region circumscribed with salient boundary. The tessellation of prominent region as middle level structure is motivated by human perception behavior that people intuitively notice the prominent regions at first glance when they are looking at an image. On the technical perspective, the searching space on the prominent region tessellation can be significantly reduced while the structural information of objects in the image is still preserved.

With the prominent region structure, a higher level process can be applied to assemble those similar prominent regions to demarcate the lesion with relatively global features. The demarcation of breast lesion is realized by formulating a graph-cut scheme on the prominent region structure. The prominent region tessellation can be naturally transformed into a graph by taking each prominent region as a graph node. The linkage among graph nodes can be established from the neighboring relation of prominent regions. With the given foreground/background priors, the optimal bi-partition of prominent region graph is sought by the graph-cut scheme with an associated similarity cost function.

The contribution of this paper consists in developing a framework that seeks the boundaries of breast lesions from the hierarchical structures of cells, prominent regions, and lesions. Comparing to pixel-based graph-cut approaches, the prominent region structure offers more chance from being trapped into undesired local optima. Moreover, the computational search space is significantly less in prominent region structure and hence the efficiency can be improved. The remaining parts of this paper will be structured as follows. First, the cell competition algorithm will be reviewed and the concept of cells and prominent regions will be detailed. Following that, the graph-cut scheme based on the prominent region structure and the devised similarity function will be discussed. Afterward, the experimental results on 2D breast sonograms will be presented. To demonstrate the capability of generalization to N-dimensional image data, the extension to 3D breast ultrasound volume is also discussed.

2. MATERIAL AND METHOD

This paper proposes a hierarchical segmentation framework to delineate the lesion boundaries in breast sonograms.

Fig. 1. Flow chart of proposed algorithm.

In the first phase, cell competition algorithm is performed to tessellate the image/ROI into prominent region structure. Based on the prominent region structure, a graph-cut scheme is applied to demarcate the breast lesion with relatively global features. The flow chart of the proposed framework is summarized in Figure 1.

Cell Structure and Cell Competition

The cell structure is generated by two-pass watershed transformation [9]. In the first pass, immersive watershed transformation is inundated on the landscape of gradient map of the original image. The second pass then eliminates the inferior watersheds with very small gradient magnitude and results in cell tessellation.

With the primitive cell tessellation, the cell competition algorithm is applied to find the best partition of cell structure into prominent region structure. The main goal of cell competition algorithm is to seek the homogeneous regions with salient boundaries. Accordingly, the cost function to be minimized is composed of two major parts, i.e., region and edge costs. The region cost characterizes the region homogeneity by the variance measure of the intensity distribution within one region. The ‘region’ here is constituted of a few connected cells at one time point of competition process. The edge cost described the statistical boundary strength of one region. The measurement of the edge cost is obtained by performing three passes of statistical tests on the profiles of inner and outer vicinities and boundary.

During the competition process, the cells are allowed to merge and split with each other whenever the cost function is able to be minimized. The minimum of the cost function is reached in iteratively gradient descending fashion. More details of the cell competition algorithm can be found in [9]. Because some weak but important edges can’t be preserved with the cost re-initialization strategy [9] in some cases, the re-initialization step of cell competition is not conducted in this study.

Demarcation of Breast Lesion by Graph-Cut

Cell Structure and Cell Competition

The cell structure is generated by two-pass watershed transformation [9]. In the first pass, immersive watershed transformation is inundated on the landscape of gradient map of the original image. The second pass then eliminates the inferior watersheds with very small gradient magnitude and results in cell tessellation.

With the primitive cell tessellation, the cell competition algorithm is applied to find the best partition of cell structure into prominent region structure. The main goal of cell competition algorithm is to seek the homogeneous regions with salient boundaries. Accordingly, the cost function to be minimized is composed of two major parts, i.e., region and edge costs. The region cost characterizes the region homogeneity by the variance measure of the intensity distribution within one region. The ‘region’ here is constituted of a few connected cells at one time point of competition process. The edge cost described the statistical boundary strength of one region. The measurement of the edge cost is obtained by performing three passes of statistical tests on the profiles of inner and outer vicinities and boundary.

During the competition process, the cells are allowed to merge and split with each other whenever the cost function is able to be minimized. The minimum of the cost function is reached in iteratively gradient descending fashion. More details of the cell competition algorithm can be found in [9]. Because some weak but important edges can’t be preserved with the cost re-initialization strategy [9] in some cases, the re-initialization step of cell competition is not conducted in this study.

Demarcation of Breast Lesion by Graph-Cut

Cell Structure and Cell Competition

The cell structure is generated by two-pass watershed transformation [9]. In the first pass, immersive watershed transformation is inundated on the landscape of gradient map of the original image. The second pass then eliminates the inferior watersheds with very small gradient magnitude and results in cell tessellation.

With the primitive cell tessellation, the cell competition algorithm is applied to find the best partition of cell structure into prominent region structure. The main goal of cell competition algorithm is to seek the homogeneous regions with salient boundaries. Accordingly, the cost function to be minimized is composed of two major parts, i.e., region and edge costs. The region cost characterizes the region homogeneity by the variance measure of the intensity distribution within one region. The ‘region’ here is constituted of a few connected cells at one time point of competition process. The edge cost described the statistical boundary strength of one region. The measurement of the edge cost is obtained by performing three passes of statistical tests on the profiles of inner and outer vicinities and boundary.

During the competition process, the cells are allowed to merge and split with each other whenever the cost function is able to be minimized. The minimum of the cost function is reached in iteratively gradient descending fashion. More details of the cell competition algorithm can be found in [9]. Because some weak but important edges can’t be preserved with the cost re-initialization strategy [9] in some cases, the re-initialization step of cell competition is not conducted in this study.

Demarcation of Breast Lesion by Graph-Cut

Cell Structure and Cell Competition

The cell structure is generated by two-pass watershed transformation [9]. In the first pass, immersive watershed transformation is inundated on the landscape of gradient map of the original image. The second pass then eliminates the inferior watersheds with very small gradient magnitude and results in cell tessellation.

With the primitive cell tessellation, the cell competition algorithm is applied to find the best partition of cell structure into prominent region structure. The main goal of cell competition algorithm is to seek the homogeneous regions with salient boundaries. Accordingly, the cost function to be minimized is composed of two major parts, i.e., region and edge costs. The region cost characterizes the region homogeneity by the variance measure of the intensity distribution within one region. The ‘region’ here is constituted of a few connected cells at one time point of competition process. The edge cost described the statistical boundary strength of one region. The measurement of the edge cost is obtained by performing three passes of statistical tests on the profiles of inner and outer vicinities and boundary.

During the competition process, the cells are allowed to merge and split with each other whenever the cost function is able to be minimized. The minimum of the cost function is reached in iteratively gradient descending fashion. More details of the cell competition algorithm can be found in [9]. Because some weak but important edges can’t be preserved with the cost re-initialization strategy [9] in some cases, the re-initialization step of cell competition is not conducted in this study.
With the prominent region tessellation from the cell competition algorithm, relatively global features can be obtained. The problem of demarcating the breast lesion here is formulated as the problem of bi-partitioning the graph.

The prominent region structure is transformed into a graph, denoted as PR-graph, by taking each prominent region as a graph node. The graph links are then established from the neighboring relation of prominent regions. The transformation of prominent region as a PR-graph can be illustrated in Figure 2.

Given the high level priors of foreground (lesion) and background from user, the best bi-partition of the prominent region graph is sought in a graph-cut scheme. The major issue of a graph-cut scheme is the design of the energy function. Given a P-R graph, \( \chi = [V, E] \), the energy function is defined as:

\[
\begin{align*}
    c(u, s) &= 0, \forall u \in V - \{F \cup B\} \\
    c(u, t) &= 0, \forall u \in V - \{F \cup B\} \\
    c(u, s) &= \infty, \forall u \in \{F\} \\
    c(u, t) &= \infty, \forall u \in \{B\} \\
    c(u, v) &= e^{-\frac{\sigma^2}{\sigma^2} + \frac{\sigma^2}{\sigma^2}}, \{u, v\} \in Nb \\
    c(u, v) &= 0, \{u, v\} \notin Nb
\end{align*}
\]

where \( F \) and \( B \) represent the foreground and background respectively. \( u \) and \( s \) represent the source and sink nodes in the graph. \( c(\alpha, \beta) \) is the flow capacity of the graph link \( \alpha, \beta \) that connect nodes \( \alpha \) and \( \beta \) (\( \forall \alpha, \beta \in V \)). In the energy function, the capacity \( c(u, v) \) between node \( u \) and \( v \) is characterized by edge strength \( \delta_{uv} \), as well as the region intensity difference \( \delta_{uv} \). \( \sigma \) and \( \sigma^2 \) are the standard deviation of overall measures of edge strength and region intensity difference. \( Nb \) is the graph link subset of \( E \) which includes all graph edges except those connecting to source and sink nodes. The obtained sets of the edges which cut PR-graph into two isolated graphs will then be transformed back into the breast lesion and background.

3. RESULT AND DISCUSSION

Sixteen patients with malignant or benign breast lesions were involved in this study. The segmentation result of one benign breast lesion and the cell structure are demonstrated in Figure 3. Accordingly, the proposed algorithm is able to delineate complicated shapes.

The delineation of a solid lesion in three dimensional breast ultrasound images can be much more complicated. The topology of one breast lesion may vary in different slices. More specifically, one lesion may appear as one lesion component in some slices but sometimes may become several disconnected lesion components in other slices, especially in the ending parts.

![Fig. 2. Transformation of prominent region structure into a graph. (A) A PR-graph of (B). (B) A prominent region structure. The dotted line in (A) represents a graph cut and corresponds to the border line in (B).](image)

![Fig. 3. Illustration of three segmentation results (ABC/DEF/GHI). (A), (D), and (G) are Region of Interests from three different breast sonograms. (B, E, H) represent the cell tessellation and (C, F, I) final result of (A), (D), and (G), respectively.](image)

<table>
<thead>
<tr>
<th>Group</th>
<th>COD</th>
<th>IOD</th>
<th>COD-IOD</th>
<th>95% CI</th>
</tr>
</thead>
<tbody>
<tr>
<td>Group A</td>
<td>3.178</td>
<td>3.978</td>
<td>-0.8</td>
<td>(-1.306, -0.671)</td>
</tr>
<tr>
<td>Group B</td>
<td>2.937</td>
<td>3.92</td>
<td>-0.983</td>
<td>(-1.228, -0.739)</td>
</tr>
<tr>
<td>Group C</td>
<td>3.364</td>
<td>4.144</td>
<td>-0.78</td>
<td>(-1.021, -0.616)</td>
</tr>
<tr>
<td>Group D</td>
<td>2.71</td>
<td>3.705</td>
<td>-0.995</td>
<td>(-1.306, -0.671)</td>
</tr>
</tbody>
</table>

We also extend our proposed algorithm to three dimensional sonographic data. Figure 4 shows segmentation results of 3D sonographic breast images. It can be observed that the topological changes in different sonographic slices can be well addressed in Figure 4. The total computational time of 2D/3D sonographic breast images highly depends on the result of cell competition algorithm.
We evaluate 2D results with four sets of manual delineations. In this study, the assessment method is from [13], Table 1 summarizes the comparison of COD and IOD. From column three and four, we may find that the computer-to-observer distances are smaller than inter-observer distances. It can be suggested that the computer results are within the variation of four sets of manual delineations.

4. CONCLUSION

In this study, facilitating from the regional structure of cell competition algorithm, the prominent regions can be served as elegant nodes for graph cut. Cell-based graph cut is able to circumvent the disadvantages on conventional pixel-based methods, and proper similar weight function can address the problem of weak edge. Comparing to manual delineation, reliable and satisfying segmentation results can be achieved by our cell-based graph cut algorithm.

5. REFERENCES