

## Research Article

# Automatic Graph Cut Segmentation of Lesions in CT Using Mean Shift Superpixels

**Xujiong Ye, Gareth Beddoe, and Greg Slabaugh**

*R&D Department, Medicsight PLC, 66 Hammersmith Road, London W14 8UD, UK*

Correspondence should be addressed to Xujiong Ye, xujiong.ye@medicsight.com

Received 29 April 2010; Revised 18 August 2010; Accepted 5 September 2010

Academic Editor: Kenji Suzuki

Copyright © 2010 Xujiong Ye 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.

This paper presents a new, automatic method of accurately extracting lesions from CT data. It first determines, at each voxel, a five-dimensional (5D) feature vector that contains intensity, shape index, and 3D spatial location. Then, nonparametric mean shift clustering forms superpixels from these 5D features, resulting in an oversegmentation of the image. Finally, a graph cut algorithm groups the superpixels using a novel energy formulation that incorporates shape, intensity, and spatial features. The mean shift superpixels increase the robustness of the result while reducing the computation time. We assume that the lesion is part spherical, resulting in high shape index values in a part of the lesion. From these spherical subregions, foreground and background seeds for the graph cut segmentation can be automatically obtained. The proposed method has been evaluated on a clinical CT dataset. Visual inspection on different types of lesions (lung nodules and colonic polyps), as well as a quantitative evaluation on 101 solid and 80 GGO nodules, both demonstrate the potential of the proposed method. The joint spatial-intensity-shape features provide a powerful cue for successful segmentation of lesions adjacent to structures of similar intensity but different shape, as well as lesions exhibiting partial volume effect.

## 1. Introduction

Accurate and automatic segmentation of medical images is an essential component of a computer-aided diagnosis (CADx) system. However, medical image segmentation is typically a difficult task due to noise resulting from the image acquisition process, irregular shape and variable size of anatomical objects, as well as the characteristics of the object's neighborhood. For example, a lung nodule or a colon polyp is usually embedded in a complex surrounding region. In CT imaging, the intensities of such lesions (e.g., juxtavascular nodules, juxtapleural nodules, or colonic polyps) are usually very similar to their adjacent tissues. In this case, traditional intensity-based or morphological methods [1–5] may fail to accurately segment the lesion.

Energy minimization techniques for image segmentation have shown much promise in medical image computing. In particular, graph cut methods [6–12], which construct an image-based graph and achieve a global minimum of energy functions are often found in image segmentation. Two key issues in the graph cut technique are as follows: (1) how to

define an energy function so its minimization leads to the desired result; (2) how to represent the energy function in terms of the graph construction.

In most graph cut methods, the graph vertices are constructed at the image pixels, and the segmentation energy is composed of intensity terms only. For example, Zheng et al. [7] proposed a framework to simultaneously segment and register the lungs and nodules in CT data. For segmentation, a 2D pixel-based graph cut algorithm was applied to 3D lung and nodule datasets. It is noted that, when a graph vertex is placed at each image pixel, the number of nodes in the graph increases exponentially with the image size, dramatically increasing the computation time. To improve the efficiency, Li et al. [8] introduced a graph built on a presegmented image using a watershed algorithm. However, their graph cut formulation is solely based on the image intensities. It is known that pixel intensity can be locally erroneous due to noise and other image acquisition issues, such as Partial Volume Effect (PVE). Also, in CT imaging, different nearby anatomies may have similar intensities. Xu et al. [9] presented a graph

TABLE 1: Comparison of pixel-based and superpixel-based graph cut algorithms on three different types nodules.

Nodule type	Pixel-based method		superpixel-based method	
	Number of vertices	Time(s)	Number of vertices	Time(s)
GGO (a)	46195	16	946	1.2
part-solid (b)	22740	6.6	409	0.6
Solid (c)	17254	3.4	208	0.5

TABLE 2: Summary of Dice coefficients for different types of solid nodules based on two different methods.

Type	Number nodules	4D-based method		Proposed method	
		Mean coefficient	Std ( $\sigma$ )	Mean coefficient	Std ( $\sigma$ )
Isolated	28	0.80	0.05	0.81	0.04
Vascular	53	0.68	0.1	0.80	0.06
Pleural	20	0.67	0.09	0.73	0.06

cuts-based active contour approach to object segmentation, which Slabaugh and Unal [10] extended by incorporating a shape prior. These methods show promising results, but require manual initialization and quantitative evaluation on different types of lesions which are not available. Liu et al. [11] applied ordering constraints into an energy smoothness term based on an initial labeling. A simple geometric shape prior was also incorporated in a graph cut segmentation. Zheng et al. [12] constructed a graph Laplacian matrix and solved linear equations for the estimation of Ground Glass Opacity (GGO) nodules in CT, with assistance from some manually drawn scribbles for which the opacity values are easy to determine manually. More general-purpose graph cut methods show much utility in image segmentation, however, are ideally suited to nonmedical images. Other medical imaging approaches apply to nodule segmentation but are evaluated only on very small CT datasets. Further, most available methods require user interaction to identify initial foreground and background seeds.

The goal of this paper is to develop an automatic and robust superpixel-based graph cut method for accurate segmentation of different types of lesions in CT imaging including solid and GGO nodules, as well as colonic polyps.

One of the original contributions of this paper is that our graph is built on mean-shift superpixels. In this paper, a *superpixel* is a connected set of pixels sharing similar properties. We obtain superpixels by using five-dimensional mean shift clustering that incorporates joint spatial, intensity, and shape (JSIS) features. The superpixels express the local structure of the data in the five-dimensional feature space. Mean shift clustering reduces the data variation, thereby helping to produce accurate segmentations. By connecting superpixels instead of pixels, the algorithm produces better results and improves speed significantly.

A second original contribution is our energy function, which incorporates the image intensity and the shape feature into a Markov Random Field (MRF) minimized with graph

cuts. In particular, the intensity and shape information appear in both our unary and pairwise terms. Compared to our previous work in [13], we have improved our unary term in the energy function by considering both image intensity and shape features and investigated the method on a relatively large dataset that contains different types of nodules, including juxtapleural, juxtapleural, as well as GGO. Moreover, the method has also been applied to colonic polyp segmentation. The experimental results, evaluated both visually and quantitatively, demonstrate high performance for CT lung nodule segmentation. Initial results on colon polyps demonstrate the generalization of the method to other anatomies.

The paper is organized as follows: Section 2 describes the proposed algorithm, which covers five-dimensional mean shift clustering based on JSIS features (Section 2.1) and automatic graph cut segmentation of the mean shift JSIS superpixels (Section 2.2). Section 3 presents our experimental results, followed by a discussion and future work in Section 4 and conclusion in Section 5.

## 2. Methodology

Our approach is a combination of five-dimensional mean shift clustering followed by energy minimization based on graph cuts. The method is also guided by prior knowledge about the lesion (e.g., nodule/polyp). The flow chart of our method is illustrated in Figure 1. In the following sections, each stage is described in detail.

**2.1. Mean Shift Clustering of JSIS Features.** The method first computes the JSIS features, which are then clustered in a five-dimensional space using mean shift. In this section, we review the volumetric shape index feature and our five-dimensional mean shift approach.

**2.1.1. Volumetric Shape Index: A 3D Geometric Feature.** The volumetric shape index (SI) at voxel  $p$  can be defined as follows [14, 15]:

$$SI(p) = \frac{1}{2} - \frac{1}{\pi} \arctan \frac{k_1(p) + k_2(p)}{k_1(p) - k_2(p)}, \quad (1)$$

where  $k_1(p)$  and  $k_2(p)$  are the principal curvatures, defined as

$$\begin{aligned} k_1(p) &= H(p) + \sqrt{H^2(p) - K(p)}, \\ k_2(p) &= H(p) - \sqrt{H^2(p) - K(p)}, \end{aligned} \quad (2)$$

where  $K(p)$  and  $H(p)$  are the Gaussian and mean curvatures, respectively.

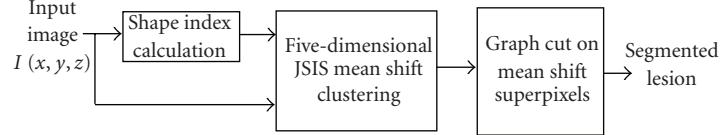


FIGURE 1: Flow diagram of the proposed graph cut-based method.

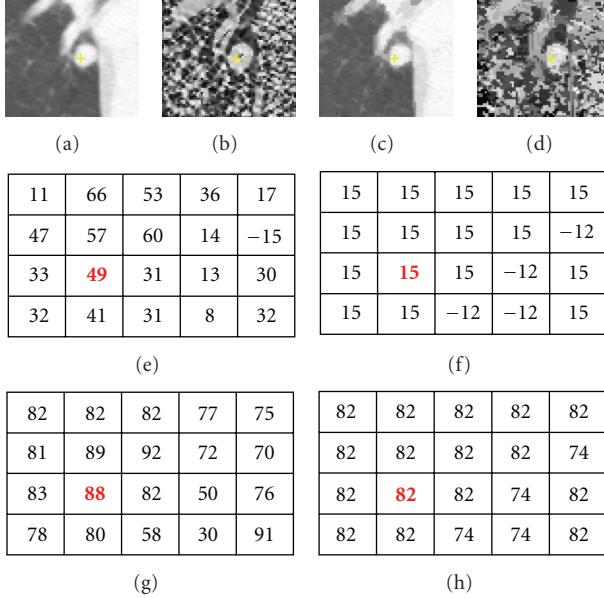


FIGURE 2: An attached nodule with its intensity and shape mode maps. (a) Original CT subimage; (b) shape index map based on (1); (c) and (d) intensity mode and shape index mode maps; (e)–(h) intensity values, intensity mode values, shape index values, and shape index mode values at the same voxel in the nodule. (It is noted that shape index are values multiplied by 100.)

The calculation of the Gaussian and mean curvatures is based on the first and second fundamental forms of differential geometry. A practical approach for computing these forms is to use the smoothed first and second partial derivatives of the image as suggested in [16]. In this paper, prior to shape index calculation, a single-scale 3D Gaussian filter, with standard deviation of 1.5, is applied to obtain a smoothed image.

The shape index provides a local shape feature at each voxel. Every distinct shape, except for the plane, corresponds to a unique shape index. For example, a shape index value of 1.0 indicates a sphere-like shape (e.g., lung nodule or colonic polyp), while 0.75 indicates a cylinder-like shape (e.g., vessel or colonic fold). Based on the definition, the volumetric shape index directly characterizes the topological shape of an isosurface in the vicinity of each voxel without explicitly calculating the isosurface. This feature provides rich information that complements the image intensity and is useful for automated segmentation. In particular, lesions in a CT image may appear within an area of complicated anatomy (such as a lung nodule neighboring a blood vessel or colonic polyp attached to the colon wall)

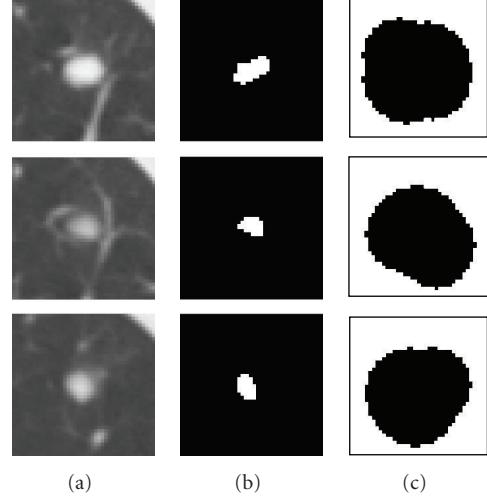


FIGURE 3: An example of the initialization based on spherical concentration on an attached nodule. (a) 3D nodule in three contiguous CT slices; (b) initial foreground based on high spherical concentration; (c) initial background region.

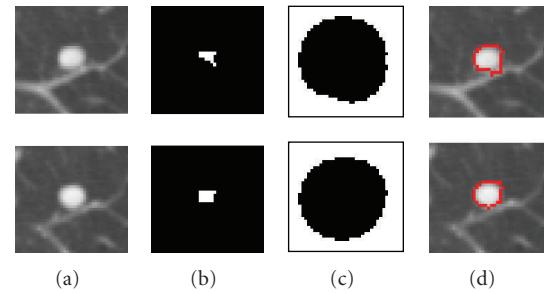


FIGURE 4: An example of the segmentation of an attached solid nodule based on the automatic calculation of the initial foreground and background. (a) 3D nodule in two contiguous CT slices; (b) initial foreground based on high spherical concentration; (c) initial background; (d) segmentation results by the proposed graph cut based method.

where adjacent structures have similar image intensities but different shapes.

Note that we use the term *spherical concentration* throughout the paper to describe a spatially connected set of voxels that have a high shape index value.

**2.1.2. Inclusion of the Shape Index Feature in the Mean Shift Framework.** For each voxel, the 3D spatial location, intensity, and volumetric shape index features are concatenated in the

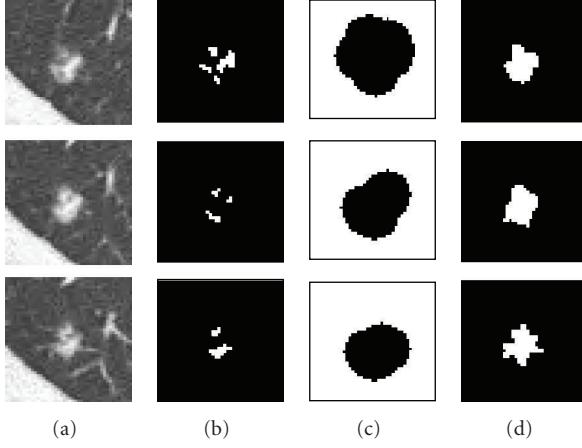


FIGURE 5: An example of the segmentation of a part-solid nodule based on the automatic calculation of the initialization (a) 3D nodule in three contiguous CT slices; (b) initial foreground based on high spherical concentration; (c) initial background; (d) segmentation results by the proposed graph cut based method.

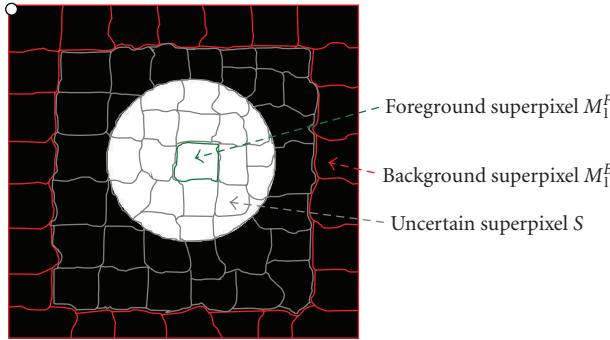


FIGURE 6: Schematic diagram of different types of superpixels: foreground (green) superpixels, background (red) superpixels, and uncertain (gray) superpixels.

joint spatial-intensity-shape domain of dimension  $d = 5$ . Given  $n$  data points  $p_i$ ,  $i=1,\dots,n$  in a five-dimensional space, the multivariate kernel is defined as the product of three radially symmetric kernels

$$K_{h_s h_r h_{si}}(p) = c_{k,5} k\left(\left\|\frac{p^s}{h_s}\right\|^2\right) \cdot k\left(\left\|\frac{p^r}{h_r}\right\|^2\right) \cdot k\left(\left\|\frac{p^{si}}{h_{si}}\right\|^2\right), \quad (3)$$

where  $c_{k,5}$  is a normalization constant which assures that  $k(p)$  integrates to 1.  $p^s$ ,  $p^r$ , and  $p^{si}$  are the spatial location, intensity, and shape index features, associated with each voxel  $p$ ;  $h_s$ ,  $h_r$ , and  $h_{si}$  are the kernel bandwidths for spatial, intensity, and shape index kernel functions. The normal kernel is used in this paper, where  $k(p) = (2\pi)^{-d/2} \exp(-1/2\|p\|^2)$ .

In the mean shift framework [17], the shape index feature can be combined with the intensity and spatial features

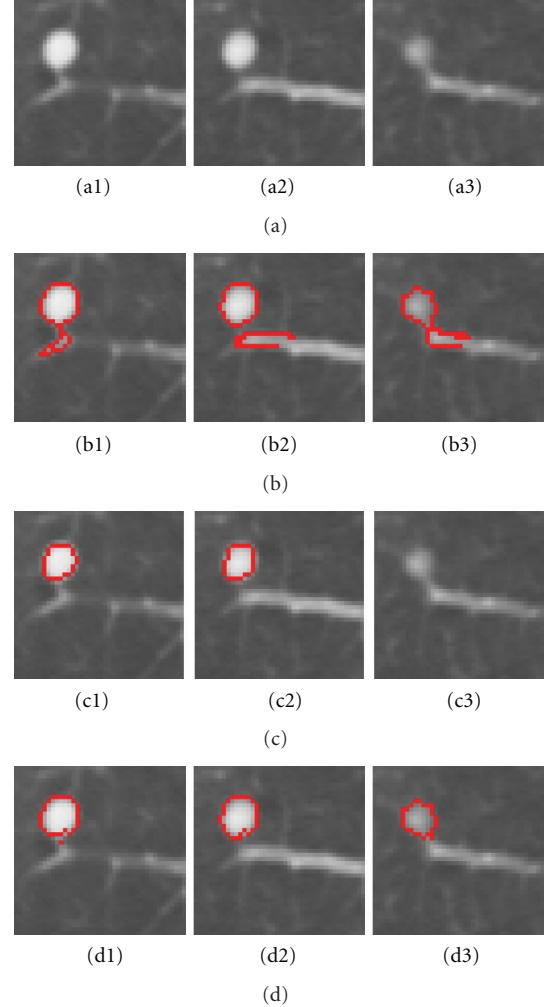


FIGURE 7: A example of an attached nodule segmentation using different pairwise smoothing energies. (a1-a3) 3D nodule in three contiguous slices in CT; (b1-b3) nodule segmentation using the first term only in (9); (c1-c3) results by using the second term only in (9); (d1-d3) nodule segmentation by using the complete energy defined in (9).

for clustering. The mean shift vector with three kernel bandwidths (spatial, intensity, and shape index) is defined as

$$m_{h_s h_r h_{si}}(p) = \frac{\sum_{i=1}^n p_i g\left(\left\|p_i^s/h_s\right\|^2\right) \cdot g\left(\left\|p_i^r/h_r\right\|^2\right) \cdot g\left(\left\|p_i^{si}/h_{si}\right\|^2\right)}{\sum_{i=1}^n g\left(\left\|p_i^s/h_s\right\|^2\right) \cdot g\left(\left\|p_i^r/h_r\right\|^2\right) \cdot g\left(\left\|p_i^{si}/h_{si}\right\|^2\right)} - p, \quad (4)$$

where  $g(p) = -k'(p)$ . The mean shift vector always points in the direction of the maximum increase in the density function.

It is noted that the mean shift algorithm estimates the *modes* (the densest regions) of the multivariate distribution underlying the feature space. The set of points that converge to the same mode is defined as the attraction basin. Mean

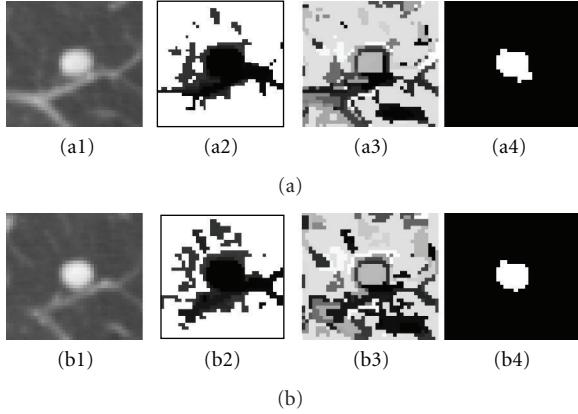


FIGURE 8: Example of cost functions. (a1, b1) 3D attached nodule in two contiguous slices in CT; (a2, b2) unary cost; (a3, b3) pairwise smoothing cost; (a4, b4) minimization of both energy terms based on (5).

shift maps all the data samples to the local maxima of their corresponding attraction basin based on five-dimensional features. *Superpixels* are formed for the set of pixels in each attraction basin. Each superpixel has a constant shape index and intensity, represented with *mode maps*, namely, an intensity mode map ( $M_I$ ) and a shape index mode map ( $M_{SI}$ ). Additionally, there is a spatial mode map ( $M_S$ ) that is not incorporated directly in our energy function; however, spatial information is utilized when defining neighbors in our graph. Figure 2 shows a nodule attached to a vessel and its corresponding intensity and shape index mode maps. It can be noted that the mode maps ((c) and (d)) from JSIS mean shift clustering can be seen as “filtered” images and are less contaminated by outliers. In the following section, graph cut-based segmentation is applied to these superpixels.

**2.2. Automatic Graph Cut Segmentation Using JSIS Superpixels.** The mode map obtained by the above JSIS mean shift technique expresses the local structure of the data in a given region of the feature space. The number of resulting superpixels depends on the kernel bandwidth and the data itself; a key advantage of the mean shift clustering is its ability to locally group regions of similar intensity and shape, reducing the data variance in superpixel.

The aim of this section is to group superpixels into two classes: foreground (lesion) and background (nonlesion). It is known that the image labeling problem can be formulated using an energy function in a Bayesian framework in the context of maximizing a posteriori (MAP) and MRF theory and can be solved by energy minimization. The energy function includes both unary and pairwise terms, the latter providing smoothing by modeling the interaction between neighboring superpixels. In this section, the MAP-MRF is transformed to a graph cut problem. A novel energy formulation that incorporates shape and intensity in both unary and pairwise terms is defined on superpixels and minimized with graph cuts using the maxflow/mincut method [6].

Two key issues are addressed in the following subsections: (1) how to initialize the segmentation and (3) how to define the energy function for minimization.

**2.2.1. Initialization Based on Spherical Concentration.** In this paper, we assume that a nodule (or polyp) is generally either spherical or has a local spherical concentration, while a blood vessel (or colonic fold) is usually oblong. The initial seeds for the graph cut are computed automatically based on a spherical concentration.

A spherical concentration  $\mathfrak{R}_s$  is defined, using hysteresis thresholding [16], as a 3D connected region that satisfies the following criteria.

- (i) All voxels in the region have a shape index greater than or equal to a low threshold  $T_l$ .
- (ii) At least one voxel in the region has a shape index greater than or equal to a high threshold  $T_h$ .

Typical thresholds are in the range  $T_l \in [0.8, 0.9]$  and  $T_h \in [0.9, 1]$ .

The spherical concentration defines the foreground seeds in the graph cut initialization. The background seeds are determined using adaptive enlargement of the spherical concentration. A distance transform [18] of  $\mathfrak{R}_s$  is computed and then adaptively enlarged (e.g.,  $\mu_{\text{factor}} * f\text{DistanceMax}$ , where  $\mu_{\text{factor}}$  is the enlargement factor and  $f\text{DistanceMax}$  is the maximum distance to the region boundary). The background region can then be obtained by inverting this enlarged foreground region. The enlargement of the foreground region is designed so that the background region does not cover the foreground object. Figure 3 shows an example of the initialization for a juxtavascular nodule using the above method, where  $T_l$  and  $T_h$  were set to 0.82 and 0.92, respectively. Figures 3(b) and 3(c) are the initial foreground and background regions, respectively, where  $\mu_{\text{factor}}$  is set to be 10.

The thresholds of  $T_l$  and  $T_h$  are fixed throughout the paper and are based on the assumption that a nodule which may not be entirely spherical has at least spherical elements concentrated within the nodule. In fact, using shape index to extract the spherical concentration is the first step in our automatic lung nodule detection algorithm [19]. Very high detection sensitivity can be achieved at this stage. In most cases (e.g., high contrast solid nodule), the sphericity concentration covers the core part of the nodule. However, in the case of some part-solid or GGO nodules, the initial foreground region might spread in the nodule object. Figures 4 and 5 show examples of two different segmented nodules using automatic calculation of foreground and background regions, where, in Figure 4(b), the initial foreground contains the core part of the solid nodule object, while in Figure 5(b), the foreground region scatters in the part-solid nodule (it is noted, the foreground region is a 3D connected region but it may appear as several small regions in each 2D slice as shown in Figure 5(b)). In both cases, the proposed method provides better initialization. Figures 4(d) and 5(d) show the segmentation results by using the proposed graph cut-based method based on the initial foreground (b) and background

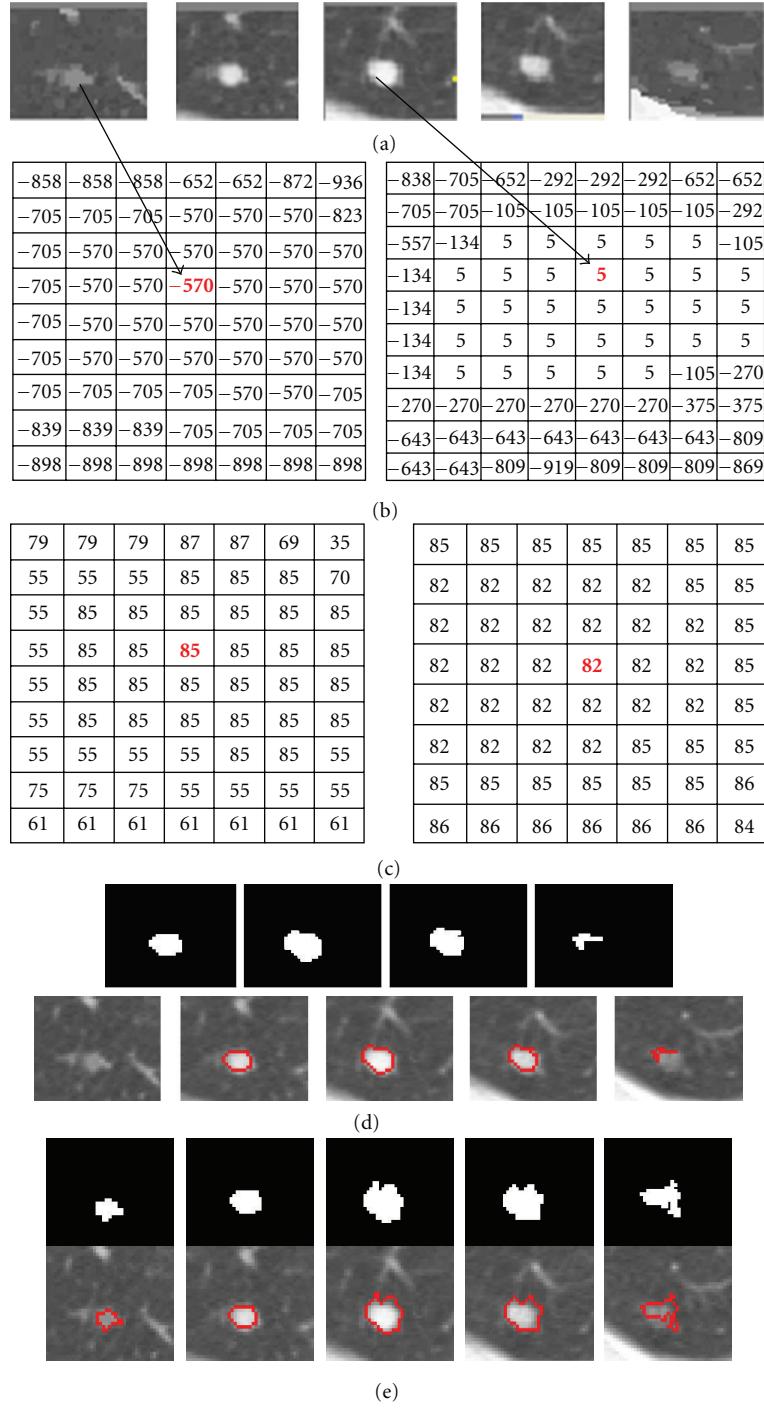


FIGURE 9: An example of a solid nodule (attached to a small vessel) segmentation. (a) Original 3D nodule on 5 contiguous slices; (b) and (c) intensity and shape index (multiplied by 100) mode maps from 5D mean shift clustering; (d) segmentation results without considering the shape feature; (e) results based on the proposed method.

seeds (c). However, occasionally the above initialization may fail to give good foreground and background seeds, as the spherical concentration may contain part of the nodule as well as lung tissue. An example is provided in Figure 21. More discussion for the initial foreground seeds will be given in Section 4.

**2.2.2. MAP-MRF Energy.** In this section, we employ a graph  $G = (V, E)$ , defined with vertices  $v \in V$  representing superpixels determined from five-dimensional mean shift clustering, and edges  $e \in E$  connecting adjacent superpixels. A key difference from the usual graph construction is that we connect superpixels instead of the original pixels. As a result,

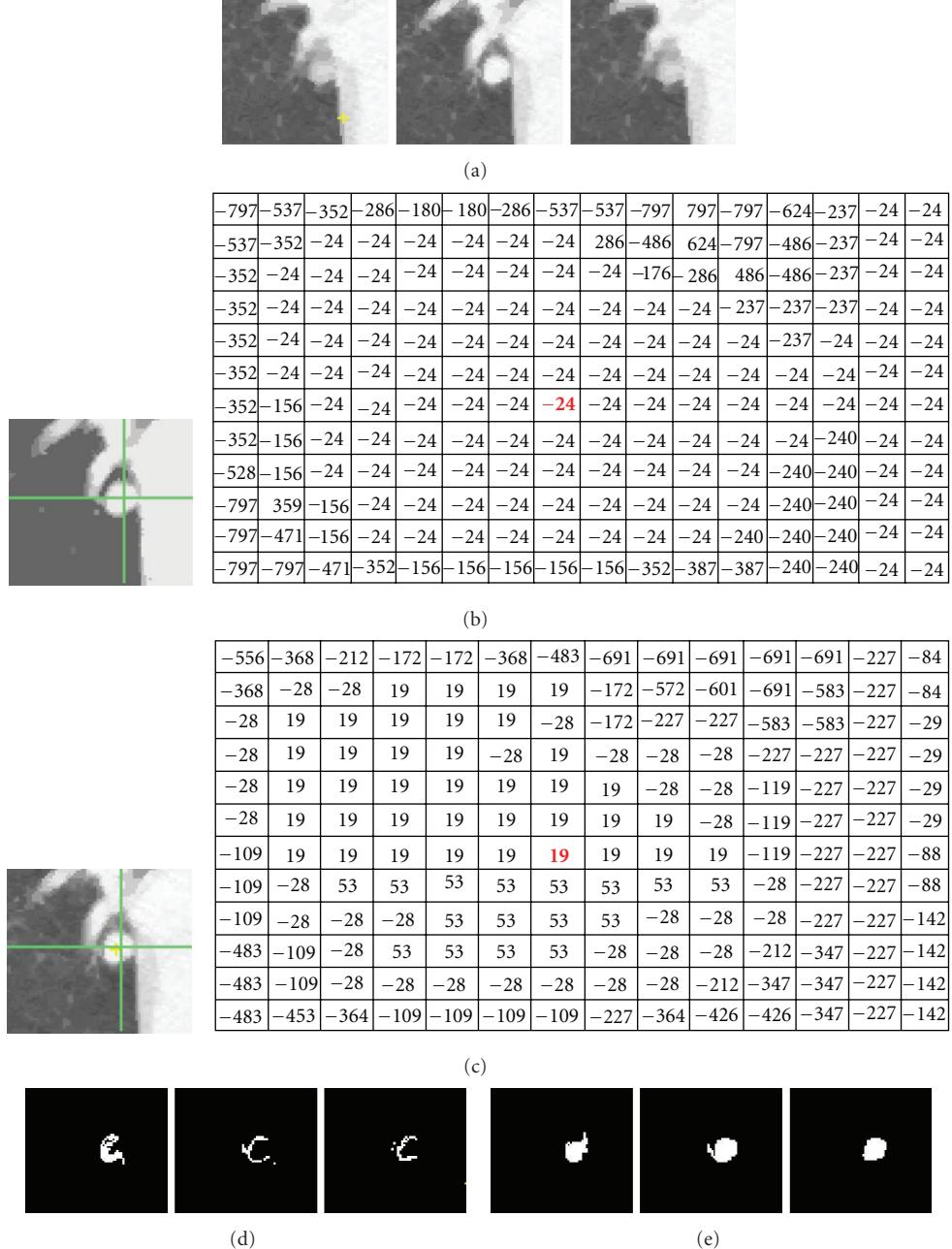


FIGURE 10: An example of segmentation of an attached nodule. (a) Original 3D nodule on three contiguous slices; (b) intensity mode map and its corresponding intensity mode values based on 4D mean shift clustering (without shape); (c) intensity mode map and its corresponding intensity mode values based on 5D mean shift clustering (with shape). (d) and (e) segmentation results without and with the shape feature (proposed method), respectively.

the number of vertices in  $G$  is greatly reduced compared to the original number of pixels in the image.

The lesion segmentation problem is formulated as a binary labeling problem, so the goal is to assign a unique label  $l_i \in \{0, 1\}$  (where 0 is background and 1 represents foreground (lesion)) to each superpixel  $i$  by minimizing the following energy function  $E(L)$  [20]:

$$E(L) = \sum_{i \in V} E_1(l_i) + \lambda \sum_{(i,j) \in \varepsilon} E_2(l_i, l_j), \quad (5)$$

where  $E_1(l_i)$  is the unary data term representing the cost to assign the label  $l_i$  to the superpixel  $i$ ,  $E_2(l_i, l_j)$  is the pairwise smoothness term representing the cost of assigning the labels  $l_i$  and  $l_j$  to adjacent superpixels  $i$  and  $j$ , and  $\lambda$  is a weighting factor. The details of energy minimization via the graph cut algorithm for binary labeling can be found in [6]. Below we focus on how to define the two energy terms.

*Unary Data Term.* We are given the initial foreground  $\{M_m^F\}$  and background regions  $\{M_t^B\}$ , automatically calculated from

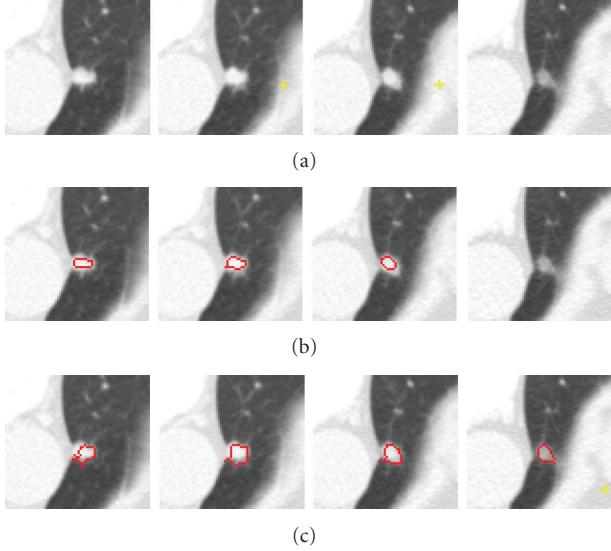


FIGURE 11: An example of a pleural nodule segmentation. (a) Original 3D nodule on four contiguous slices; (b) and (c) segmentation results without and with the shape feature (proposed method), respectively.

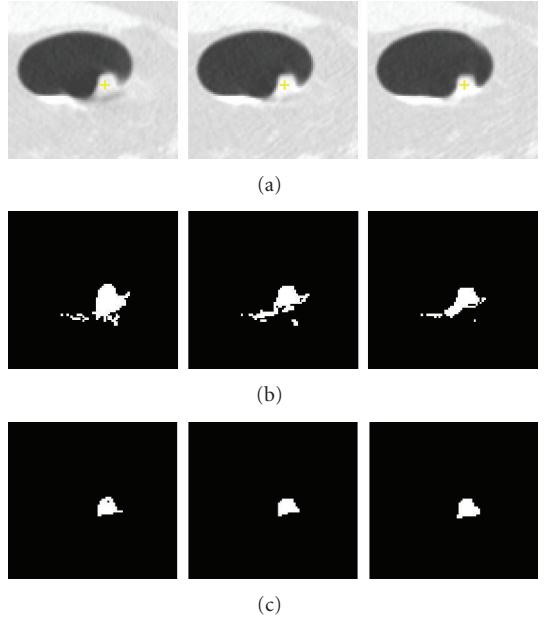


FIGURE 13: Segmentation of a colonic polyp. In (a), we show a 3D polyp in three contiguous slices, in (b) segmentation results without the shape feature, and in (c) segmentation results based on the proposed method.

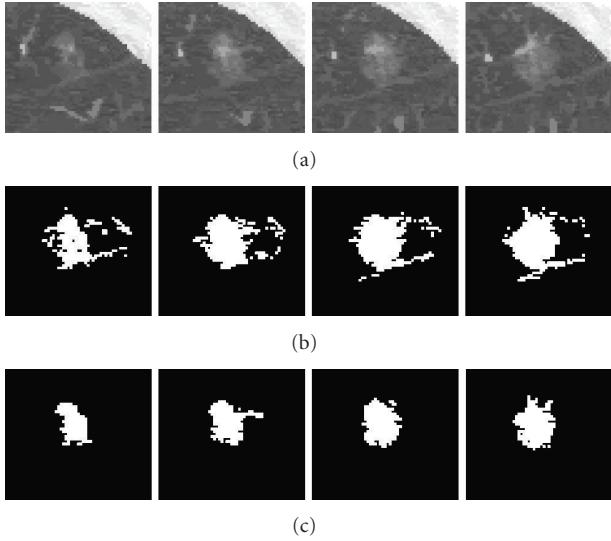


FIGURE 12: An example of a large GGO nodule segmentation. (a) Original 3D nodule on eight contiguous slices; (b) and (c) segmentation results without and with shape feature (proposed method), respectively.

the previous section, where  $m$  and  $t$  are the superpixel indices for initial foreground and background, respectively. Figure 6 shows the schematic diagram for different types of superpixels, for example, foreground, background, and uncertain superpixels. For each superpixel  $i$ , the unary term is defined as

$$E_1(l_i = 1) = \frac{c_i^F}{c_i^F + c_i^B}, \quad E_1(l_i = 0) = \frac{c_i^B}{c_i^F + c_i^B}, \quad (6)$$

where  $c_i^F$  and  $c_i^B$  are the costs of a superpixel  $i$  to the foreground superpixels and background superpixels, determined from the initialization.

The foreground cost ( $c_i^F$ ) is calculated as

$$c_i^F = \min_m \left\{ 1 - \exp \left( -\frac{(S_I(i) - M_m^F)^2}{\sigma_i^2} \right) \right\} \times \left( 1 - \exp \left( -\frac{(S_{SI}(i) - 1.0)^2}{\sigma_{si}^2} \right) \right), \quad (7)$$

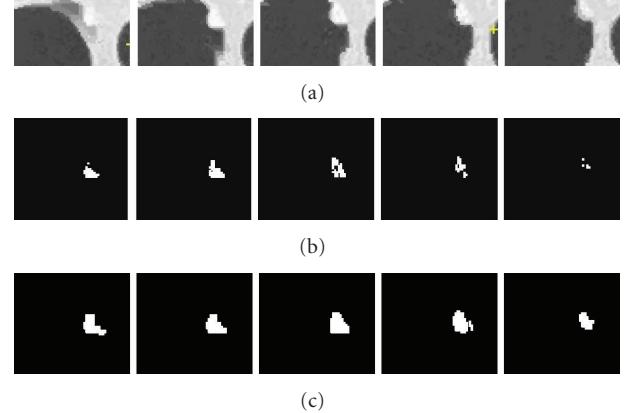


FIGURE 14: Another example of polyp segmentation. (a) 3D polyp in five contiguous slices; (b) segmentation results without shape feature; (c) segmentation results using the proposed method.

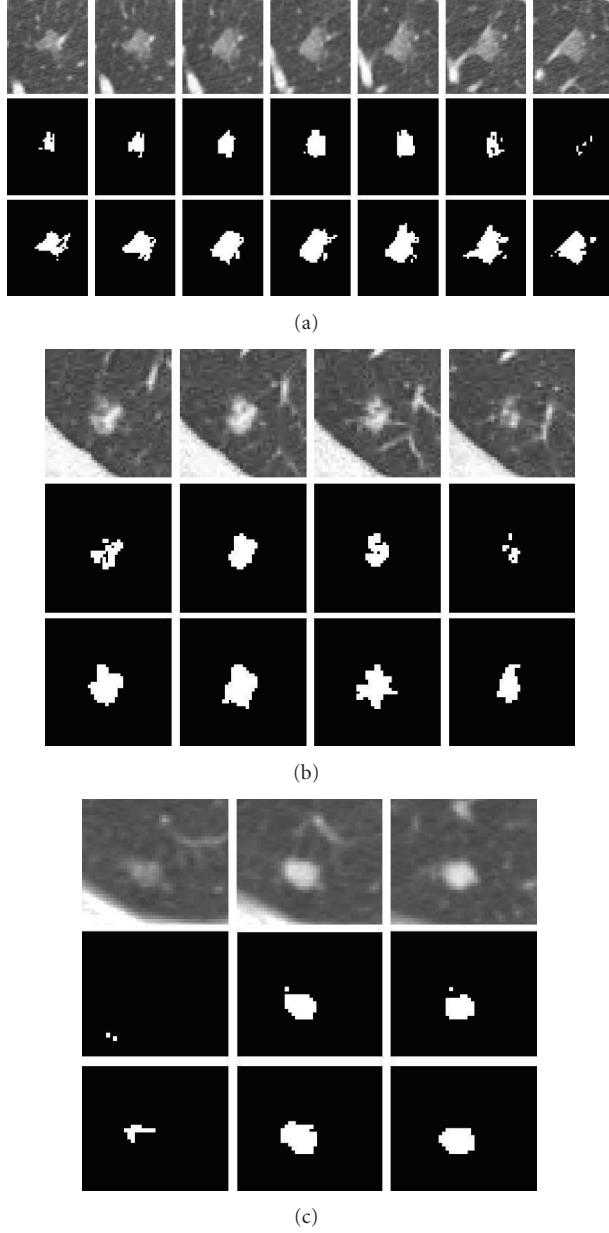


FIGURE 15: Comparison of pixel-based and superpixel-based graph cut algorithms on three different types of nodules. For each nodule in (a), (b), and (c), the first row shows the 3D nodule in the original CT subimages; the second row shows segmentation results based on the pixel-based method, and the last row shows results using the proposed superpixel based method.

where  $S_I(i)$  and  $S_{SI}(i)$  are the intensity value and shape index value at superpixel  $i$ , determined from mean shift clustering. It is noted that the intensity and shape features are normalized to the same scale. Also, in this paper, the parameters  $\sigma_i$  and  $\sigma_{si}$  are determined experimentally (e.g.,  $\sigma_i$  is chosen to be 0.6, while,  $\sigma_{si}$  is 0.3). It can be seen that  $c_i^F$  encourages a superpixel to have the same label as the initial foreground superpixels if the superpixel has similar intensity to the foreground superpixels and also its shape feature is closer to one.

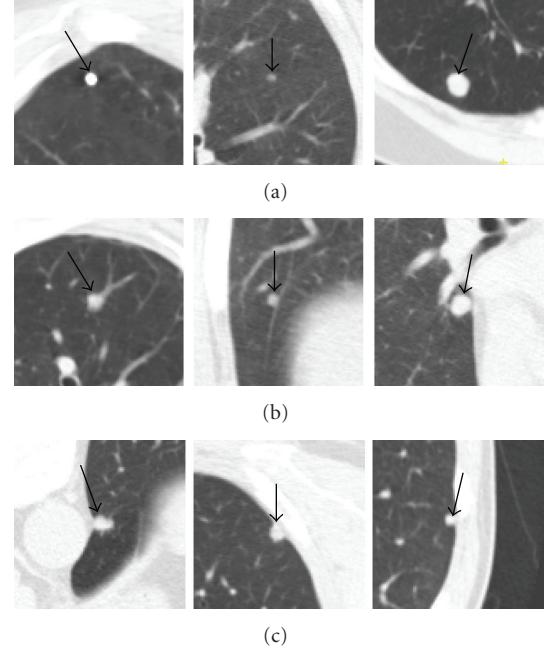


FIGURE 16: Example of different types of solid nodules. (a) Isolated nodules; (b) vascular nodules; (c) pleural nodules.

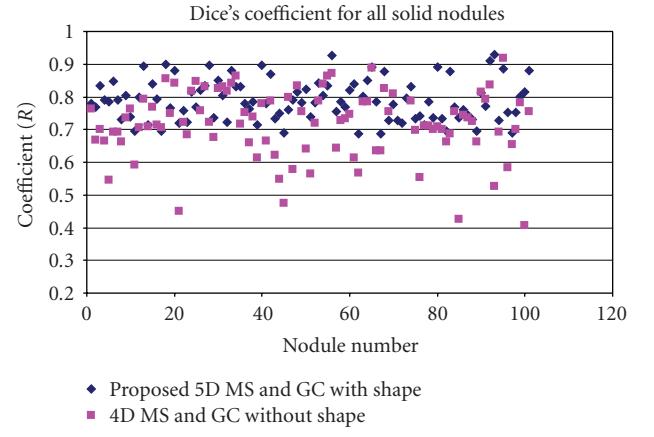


FIGURE 17: Dice coefficients based on the two different methods for all the solid nodules.

Similarly, the background cost  $c_i^B$  is defined as

$$c_i^B = \min_t \left\{ 1 - \exp \left( - \frac{(S_I(i) - M_t^B)^2}{\sigma_i^2} \right) \right\} \times \exp \left( - \frac{(S_{SI}(i) - 1.0)^2}{\sigma_{si}^2} \right). \quad (8)$$

*Pairwise Smoothing Term.* The second term  $E_2(l_i, l_j)$  in (5) is defined as

$$E_2(l_i, l_j) = |l_i - l_j| \cdot (w_{si}(E_I + E_{SI}) + (1 - w_{si}) \cdot E_I \cdot E_{SI}), \quad (9)$$

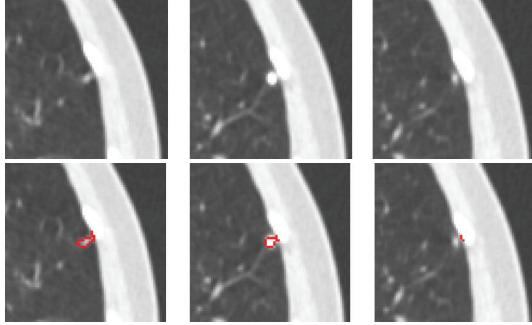


FIGURE 18: Example of a pleural nodule segmentation. Top row: original 3D nodule on three contiguous slices; bottom row: segmentation results based on the proposed method for which a small part of nodule on the last slice was incorrectly segmented.

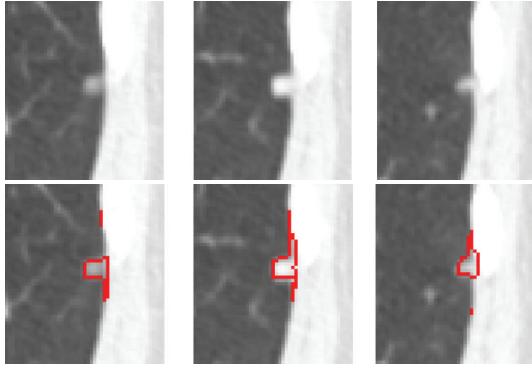


FIGURE 19: Another example of a pleural nodule segmentation. Top row: original 3D nodule on three contiguous slices; bottom row: segmentation results based on the proposed method in which the segmented nodule includes a small amount of lung wall.

where  $E_I$  is an intensity energy term denoting the intensity difference between two adjacent superpixels  $i$  and  $j$ , defined as  $E_I(l_i, l_j) = 1/(\|S_I(i) - S_I(j)\| + 1)$ . This means that superpixels with similar intensities have a larger  $E_I$ , which leads to assigning them the same labels.

$E_{SI}$  is the shape energy term denoting the shape difference between two adjacent superpixels, defined as  $E_{SI}(l_i, l_j) = 1/(\|S_{SI}(i) - S_{SI}(j)\| + 1)$ . Similar to the intensity term, the shape energy term captures the shape features for the two adjacent superpixels. If adjacent superpixels have similar shape,  $E_{SI}$  is larger, with high probability, both superpixels have the same label.

As can be seen from (9), the intensity term and shape term are combined through a weighting function  $w_{si}$  which is defined as

$$w_{si}(S_{SI}(i), S_{SI}(j)) = \exp\left(-\left(\frac{S_{SI}(i) - S_{SI}(j)}{\sigma_{si}}\right)^2\right). \quad (10)$$

It is noted that when two adjacent superpixels have the same shape, namely,  $w_{si} = 1$ , the energy depends on the first term of (9). However, when two adjacent superpixels have very different shapes,  $w_{si}$  is small, so the (9) depends on the second term. Figure 7 shows a juxtavascular nodule

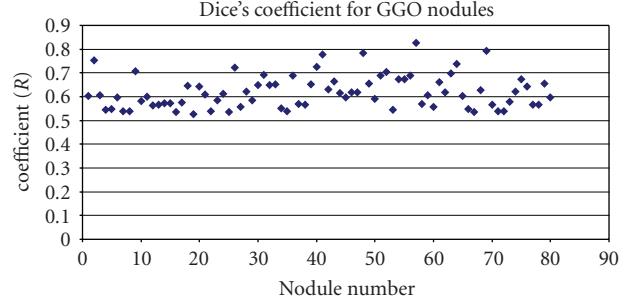


FIGURE 20: Dice coefficients for GGO nodules based on the proposed method.

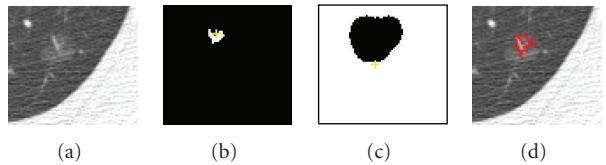


FIGURE 21: Segmentation results on a GGO nodule; (a) one slice of the GGO nodule; (b) initial foreground seeds located on the top border of the nodule; (c) initial background; (d) segmented nodule which is missing parts of the opacified component.

segmentation that uses different pairwise smoothing terms. Due to PVE in CT imaging, part of the nodule's pixels (e.g., Figure 7(a3)) has relatively low intensities, compared to those on other slices. With just the first term of (9), the pixels with low intensity (but similar shape) can still be correctly identified as being part of the nodule object, as seen in Figure 7(b3). However, some small amounts of vessels (which has similar intensity but a different shape feature value) are also included in the segmentation, as seen in Figures 7(b1–b3). This is because the first term equally considers the similarity for both of the intensity and shape features. Figures 7(c1–c3) are the results using the second term only in (9). It can be seen that the shape feature is only used as a weighting to the intensity feature. Superpixels with different shapes result in a lower weighting; this is why part of the nodule can be properly separated from the adjoining vessel despite the similar intensity, as shown in Figures 7(c1–c2). However, the superpixels with lower intensity due to PVE are wrongly identified as background due to the different intensities, compared to that on the other slices, as shown in Figure 7 (c3). Figures 7(d1–d3) are the results by combining both of terms as in (9), in which the nodule boundary can be properly delineated despite the PVE and the presence of vessels with similar intensities.

Figure 8 shows a step-by-step example of using different cost functions for the energy minimization. Figures 8(a1) and 8(b1) are two contiguous slices for one 3D vascular nodule. Figures 8(a2) and 8(b2) are the unary cost (7) for each slice, respectively. The dark area in the image indicates lower cost for being foreground, while the bright area means high cost. Figures 8(a3) and 8(b3) are the pairwise smoothing term (9) for each slice, respectively. It can be seen

that there are low costs for the boundary, while high costs are usually within the regions; Figures 8(a4) and 8(b4) are the minimization results in which the energy function uses both of unary term (6) and pairwise smoothing term (9).

### 3. Experimental Results

In this section, we present results demonstrating the effectiveness of the proposed algorithm applied to clinical CT scans. The evaluation of the proposed segmentation method has been conducted in two parts. The first, visual analysis is performed on several different types of CT pulmonary nodules and colonic polyps to provide insight into the method's performance for any visual gross missegmentation, such as a failure in separating nodules from vasculature. The second, quantitative experiments evaluate the segmentation results on large nodule datasets, containing 101 solid nodules and 80 GGO nodules. It is noted that, all the nodules in our database are confirmed by three experienced thoracic radiologists. However, each nodule boundary was manually delineated by one qualified radiologist.

In all experiments, the three kernel bandwidths (spatial  $h_s$ , intensity  $h_r$ , and shape index  $h_{si}$  in (4)) in the five-dimensional JSIS mean shift clustering were set to 3.0, 6.5, and 3.0, respectively. The proposed graph cut algorithm was applied to the mean-shift superpixels using the energy formulation based on (5)–(10), where the weighting factor  $\lambda$  was set to be 1. The MRF parameters were selected manually based to maximize the overlap ratios with the ground truth data.

#### 3.1. Comparison with Four-Dimensional Mean Shift and Graph Cut without the Shape Feature

##### 3.1.1. Lung Nodule Segmentation

**Attached Nodules.** Figure 9 demonstrates a detailed example of the segmentation of a solid nodule attached to a small vessel. The figure shows (a) the original CT lung nodule on five contiguous slices, (b), and (c) corresponding intensity mode and shape index mode values on two of the slices of the superpixels resulting from the five-dimensional JSIS mean shift clustering. Since each superpixel is an attraction basin in the 5D feature space, each superpixel has a different pair of intensity and shape index mode values. In Figure 9(b), after five-dimensional JSIS mean shift clustering, most of the voxels within the nodule on the same slice have the same mode (such as the value of “−570” in intensity mode map of the first slice and “5” of the third slice). It can be seen that, although the intensity modes within nodule are very different, their shape modes are quite similar.

The nodule superpixels can be merged using the proposed graph cut method. Figure 9(d) shows the segmented nodule resulting from a graph cut without the shape feature, for which the intensity energy term  $E_I$  was only used in the pairwise term (9). The algorithm correctly segments most of the nodule (specifically, the part of nodule in the middle three slices). However, the other parts of the nodule with

lower intensity (e.g., pixels on the first and the last slice) are wrongly identified as background due to the different intensities. Figure 9(e) shows the segmented nodule using the proposed method. It can be seen that, by combining both the intensity energy  $E_I$  and shape energy  $E_{SI}$  into the pairwise energy term (9), the parts of the nodule with low intensity can be successfully segmented due to the similar shapes on the first and last slices.

Figure 10 shows another example of a segmentation of an attached nodule. The intensity of the nodule is very similar to that of the adjoining vessel and heart wall. Figure 10(a) shows the nodule on three continuous CT slices; Figures 10(b), and 10(c) are two intensity mode maps corresponding to the same nodule slice produced by four-dimensional mean shift clustering (without the shape index feature) and five-dimensional mean shift clustering (with the shape index feature), respectively. It can be seen that, without considering the shape feature, a large amount of nonnodule voxels (e.g., heart) have a same superpixel as that of nodule voxels (e.g., “−24” in Figure 10(b)). However, using the proposed five-dimensional JSIS feature, most of the nonnodule object is separated from the nodule. This is due to the different local shapes of the nodule and the surrounding anatomy (such as the adjoining vessel and heart wall). The proposed five-dimensional mean shift clustering not only considers spatial and intensity information, but also shape features. Figure 10(d) shows the final segmented nodule using graph cut on the above four-dimensional mean shift superpixels, for which only the intensity term is used in the energy function. It is noted that the superpixel with intensity mode value of “−24” in Figure 10(b) produced from four-dimensional mean shift without taking into account the shape feature contains part of the nodule but also nonnodule objects (such as heart). By representing each superpixel as a node in the graph construction and minimizing the energy function defined in (5), this superpixel is identified as nonnodule object. That is why most of the nodule voxels within this super pixel were wrongly identified as background (nonnodule object). For comparison, the segmentation results of the proposed method are shown in Figure 10(e). It is noted that, by using the proposed method, the misclassified nodule part in Figure 10(c) can be successfully identified as being part of nodule object.

**Pleural Nodule.** A pleural nodule is defined as a nodule attached to the pleural wall. Figure 11 shows an example of a pleural nodule segmentation. Figure 11(a) is a 3D nodule in four contiguous CT slices. Similar to the above two examples, Figures 11(b) and 11(c) compare results based on the 4D approach (no shape feature) to the 5D approach (using the shape feature). It can be seen that, without considering the shape feature, the algorithm is unable to segment the nodule on the last slice, while this missing part can be identified as part of the nodule by using our method.

**GGO Nodule.** A GGO nodule appears as a hazy increased CT attenuation of the lung. It is challenging to properly segment GGO boundaries due to the faint contrast, irregular shape, and fuzzy margins of GGO nodules. Figure 12 shows an

example of the five-dimensional method on a large GGO nodule (containing four slices as shown in Figure 12(a)), attached to small vessels with similar intensity. The segmentation results shown in Figure 12(c) demonstrate good performance of the method. For comparison, the results without the shape feature are also given in Figure 12(b). It can be seen that, without considering the shape feature, a large amount of small adjoining vessels are wrongly identified as part of the nodule object. However, using the shape index feature in both the mean shift clustering and the graph cut energy function, the nodule boundary can be properly delineated from the background despite the low intensity of the nodule and the presence of other nontarget structures with similar intensity.

Quantitative evaluation of GGO nodule segmentation and further discussion will be given in next subsection.

**3.1.2. Colonic Polyp Segmentation.** The segmentation of colonic polyps in CT images is a complex task due to the irregularity of polyp morphology and the complexity of surrounding regions. The boundaries between polyp tissues and nonpolyp tissues are much less obvious. Results showing the proposed method applied to colonic polyp segmentation appear in Figures 13 and 14. Similar to the nodule segmentation mentioned above, for comparison, Figures 13(b) and 14(b) show results without the shape feature, while Figures 13(c) and 14(c) are the results based on the proposed method. It can be seen that, by considering the shape index feature in both the mean shift clustering and the graph cut energy, both polyp boundaries can be properly delineated and separated from nonpolyp tissues. Both results demonstrate good performance of the proposed method for colonic polyp segmentation.

**3.2. Comparison with Pixel-Based Graph Cut without Mean Shift Clustering.** The performance of our superpixel-based graph cut algorithm was also compared with that of a standard pixel-based method, where the graph was constructed on each pixel and only the intensity was considered in the energy function.

Figure 15 shows segmentation results based on the two methods on three different nodules: (a) one large vascular GGO nodule, (b) one part-solid nodule, and (c) one solid nodule. Testing was performed on a system with a 2.39GHz CPU and 2GB memory. Table 1 shows the total number of vertices used in the graph cut algorithm and the computation time for both of pixel-based and superpixel-based methods. It is noted that the computation time in Table 1 includes construction of the graph and energy minimization. The majority of the computation time is in the graph construction, which includes the calculation of the both energy terms for each vertex. For the superpixel-based method, this also includes time for mean shift calculation. From Table 1, it can be seen that, on the superpixel-based graph, the fewer vertices result in a much faster run-time. Furthermore, from Figure 15, It is noted that on the pixel-based graph, only the core part of the nodule object is segmented while the outside part is wrongly labeled as a nonnodule object as shown in the middle row for each

nodule Figures 15(a)–15(c). However, the proposed method shown in the third row for each nodule can correctly identify the nodules, separating nodules from the adjoining vessels.

Since the superpixels from the five-dimensional JSIS mean shift algorithm express the local structure of the data, it produces better results and improves the speed significantly.

**3.3. Quantitative Evaluation.** In this section, the proposed algorithm has been evaluated on a large set of 130 thoracic CT scans with a slice thickness ranging from 0.5 mm to 2.0 mm. The tube current ranged from 30 mA to 250 mA. Each scan was read individually by three experienced thoracic radiologists to produce a gold standard of 181 nodules (101 solid nodules and 80 GGO nodules). Among those solid nodules, 28 nodules are isolated nodule (in this paper, *isolated* nodule is a nodule not located near distracting structures of similar intensity, like large blood vessels, the pleural wall, etc.), 53 are vascular nodules, and another 20 are pleural nodules. Examples of different nodule types are shown in Figure 16. The size of the nodules ranged between 5 mm to 20 mm in diameter. To produce the ground truth, each nodule boundary was manually delineated by one qualified radiologist.

Dice's coefficient ( $R$ ) between each segmented nodule and the ground truth nodule is calculated as follows:

$$R = \frac{2\|V_s \cap V_g\|}{\|V_s\| + \|V_g\|}, \quad (11)$$

where  $V_s$  and  $V_g$  are the segmented nodule object and the ground truth object. The notation  $\|\bullet\|$  gives the total number of voxels in the region, and the operator  $\cap$  is an intersection.

**3.3.1. Evaluation on a Solid Nodule Dataset.** Figure 17 shows the overall distribution of the coefficients ( $R$ ) calculated based on (11) for all solid nodules (101 nodules) using the proposed method and the method without the shape feature. It is noted that, without the shape feature, the mean Dice coefficient for the whole dataset is 0.71 with standard deviation ( $\sigma$ ) of 0.1. However, using the proposed method, the mean Dice coefficient has been increased to 0.79 with the  $\sigma$  decreasing to 0.06.

For comparison, we split all the solid nodules into three different types (isolated, vascular and pleural nodules) and evaluated the segmentation performance for each type separately. Table 2 is the summary of Dice coefficients for each of the types of nodules, where the mean and standard deviation are calculated for the two different methods.

For the isolated nodules, both methods give good results. For these nodules, the mean Dice coefficient for the 4D method is 0.80, and 0.81 for the 5D method. As can be seen in Figure 16(a), in general, isolated nodules have high contrast and well-defined boundaries, so the image intensity is the predominant feature, hence both methods obtain similarly good results. However, the proposed 5D method provides better results in the presence of PVE. In such cases, by considering the shape feature, the parts of the nodule with low intensity but similar shape can be successfully segmented.

For both vascular nodules and pleural nodules, the proposed method gives a much better mean Dice coefficient. Especially for vascular nodules, the mean Dice coefficient increased from 0.68 to 0.8 with  $\sigma$  decreasing from 0.1 to 0.06 using the proposed method. In general, vascular nodule is embedded in a complex surrounding anatomy with similar intensities, and considering the shape feature provides rich information for the accurate segmentation. For pleural nodules, the mean Dice coefficient is 0.73 for the proposed method and 0.67 without the shape feature. Among these 20 pleural nodules, the best Dice coefficient is 0.88 by using the proposed method. Example of good segmentation of pleural nodule can be seen in Figure 11. However, Figure 18 shows an example of a small pleural nodule segmentation based on the proposed method, where the segmented nodule missed a small part of the nodule on the last slice, and Figure 19 shows a case where the segmented nodule boundary includes a small amount of lung wall tissue. More discussion for the pleural nodule segmentation will be given in next section.

**3.3.2. Evaluation on a GGO Nodule Dataset.** The proposed method has also been evaluated on 80 GGO nodules (with pure and part opacified components). Most GGO nodules exhibit low contrast, irregular shapes and are often attached to vessels with very similar intensities. Figure 20 shows the overall distribution of the Dice coefficients computed between each segmented GGO nodule and the ground truth, based on the proposed method. The mean overlap ratio is 0.63, with  $std$  of 0.07. The best Dice coefficient is 0.83, while the worst is 0.53.

GGO nodules pose a more difficult segmentation challenge than solid nodules. The proposed method provides a robust approach that combines both of local shape features and intensity. More discussion for GGO nodule segmentation will be provided in the next section.

#### 4. Further Discussion and Future work

Experimental results presented in the previous section demonstrate the promise and generalization of the proposed method to different types of lung nodules in CT. Most of nodules can be properly delineated from the lung parenchyma despite the presence of other nontarget structures such as vessels or the lung wall.

Based on the quantitative evaluation shown in Table 2, the mean Dice coefficient of the proposed method for both isolated nodules (0.81) and vascular nodules (0.8) is on average higher than that of pleural nodules. Moreover, for the vascular and pleural nodules, the mean Dice coefficient has been greatly improved when considering the shape feature, compared to without shape feature. Figures 18 and 19 show two small pleural nodules for which the proposed method either missed a small amount of nodule or included a small amount of adjoining tissues. We further investigated these cases for probable causes. In the mean shift clustering step, different fixed bandwidths are used in (3) for both the intensity and shape features, respectively. However, a fixed bandwidth might not be ideal, as a bandwidth that is too

large might create superpixels that are too large, grouping small differences in intensity or shape that should otherwise be ungrouped. This can lead extraneous parts included in the final segmentation. To improve the performance, variable bandwidths for mean shift clustering will be further investigated.

Figures 12 and 15(a) show our visual results on two GGO nodules, and Figure 20 provides a quantitative measure on 80 GGO nodules. The mean Dice coefficient of GGO nodules is, as expected, relatively lower compared to that of solid nodules. However, it is more difficult to segment GGO nodules due to the large variations in the intensities and vague boundaries. As an initial study, both in terms of a visual analysis and a quantitative evaluation on a relatively large dataset, the results are encouraging for accurate segmentation of GGO nodules in CT images. In general, GGO nodule has shown certain opacity pattern as shown in Figures 12 or 15(a); texture features such as grey level cooccurrence (GLCM) can be incorporated into mean shift clustering for the further improvement of the GGO segmentation.

Also, in this paper, the initial foreground seeds for graph cut segmentation are automatically obtained based on spherical concentration. For our data, the initialization, described in Section 3, never failed completely, that is, the foreground initialization always returned some seeds within the nodule object for subsequent processing. In most cases, the initialization provided good foreground and background seeds for the graph construction and cut. The concentration of the high shape index values can either cover the core part of the object (e.g., solid nodules) or scatter within the nodule (e.g., some part-solid or GGO nodules). However, in the case of some part-solid or GGO nodules, the spherical concentration might be on the nodule boundary, which contains part of the nodule as well as the lung tissue. An example can be seen in Figure 21, where the initial foreground is on the top of the nodule border that contains some nonnodule parts, while the automatically obtained initial background also contains part of the nodule object. That can lead to the final segmentation missing part of the lesion. To resolve this issue, we are investigating a robust method which not only considers the shape feature but also intensity for the automatic generation initial seeds.

The proposed method has also been applied to colonic polyp segmentation. Visual analysis (Figures 13 and 14) shows the method's promise. The polyp boundaries can be accurately delineated from the adjoining nonpolyp tissues with very similar intensity. However, further analysis on a large dataset with different types of colonic polyps is required to test the generalization of the proposed method on this anatomy. This is one of the directions we will investigate in the near future.

#### 5. Conclusion

We have presented a new, automatic method of extracting lesions from CT data. A five-dimensional JSIS mean shift clustering is firstly used to produce superpixels comprised of intensity and shape index mode maps. A graph cut

algorithm is then applied using a novel energy formulation that considers not only image intensity but also shape. The initial foreground and background can be automatically obtained based on shape index concentration. The JSIS features provide rich information for lesion segmentation. Both by visual inspection on different types of lesions (lung nodules and colonic polyps), as well as using a quantitative evaluation on 101 solid nodules and 80 GGO nodules, demonstrate the potential of the proposed method. The method can not only successfully segment lesions adjacent to structures of similar intensity but different shape, but also can correctly identify some part of lesions with different intensity (due to PVE in CT imaging) but similar shape.

## References

- [1] S. A. Hojjatoleslami and J. Kittler, "Region growing: a new approach," *IEEE Transactions on Image Processing*, vol. 7, no. 7, pp. 1079–1084, 1998.
- [2] J. Dehmeshki, H. Amin, M. Valdivieso, and X. Ye, "Segmentation of pulmonary nodules in thoracic CT scans: a region growing approach," *IEEE Transactions on Medical Imaging*, vol. 27, no. 4, pp. 467–480, 2008.
- [3] J. Yao, M. Miller, M. Franaszek, and R. M. Summers, "Colonic polyp segmentation in CT colonography-based on fuzzy clustering and deformable models," *IEEE Transactions on Medical Imaging*, vol. 23, no. 11, pp. 1344–1352, 2004.
- [4] J. J. Dijkers, C. Van Wijk, F. M. Vos et al., "Segmentation and size measurement of polyps in CT colonography," in *Proceedings of the 8th International Conference on Medical Image Computing and Computer-Assisted Intervention (MICCAI '05)*, vol. 3749, pp. 712–719, 2005.
- [5] L. Lu, A. Barbu, M. Wolf, J. Liang, M. Salganicoff, and D. Comaniciu, "Accurate polyp segmentation for 3D CT colonography using multi-staged probabilistic binary learning and compositional model," in *Proceedings of the 26th IEEE Conference on Computer Vision and Pattern Recognition (CVPR '08)*, June 2008.
- [6] Y. Boykov, O. Veksler, and R. Zabih, "Fast approximate energy minimization via graph cuts," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 23, no. 11, pp. 1222–1239, 2001.
- [7] Y. Zheng, K. Steiner, T. Bauer, J. Yu, D. Shen, and C. Kambhamettu, "Lung nodule growth analysis from 3D CT data with a coupled segmentation and registration framework," in *Proceedings of the 11th IEEE International Conference on Computer Vision (ICCV '07)*, October 2007.
- [8] Y. Li, J. Sun, C.-K. Tang, and H.-Y. Shum, "Lazy snapping," *ACM Transactions on Graphics*, vol. 23, no. 3, pp. 303–308, 2004.
- [9] N. Xu, N. Ahuja, and R. Bansal, "Object segmentation using graph cuts based active contours," *Computer Vision and Image Understanding*, vol. 107, no. 3, pp. 210–224, 2007.
- [10] G. Slabaugh and G. Unal, "Graph cuts segmentation using an elliptical shape prior," in *Proceedings of the IEEE International Conference on Image Processing (ICIP '05)*, pp. 1222–1225, September 2005.
- [11] X. Liu, O. Veksler, and J. Samarabandu, "Graph cut with ordering constraints on labels and its applications," in *Proceedings of the 26th IEEE Conference on Computer Vision and Pattern Recognition (CVPR '08)*, Anchorage, Alaska, USA, June 2008.
- [12] Y. Zheng, C. Kambhamettu, T. Bauer, and K. Steiner, "Estimation of ground-glass opacity measurement in CT lung images," in *Proceedings of the International Conference on Medical Image Computing and Computer-Assisted Intervention (MICCAI '08)*, pp. 238–245, 2008.
- [13] X. Ye, M. Siddique, A. Douiri, G. Beddoe, and G. Slabaugh, "Graph-cut based automatic segmentation of lung nodules using shape, intensity and spatial features," in *Proceedings of the 2nd International Workshop on Pulmonary Image Analysis, Held in Conjunction with the 12th International Conference on Medical Image Computing and Computer Assisted Intervention (MICCAI '09)*, 2009.
- [14] O. Faugeras, *Three-Dimensional Computer Vision: A Geometric View-Point*, MIT Press, Cambridge, Mass, USA, 1993.
- [15] O. Monga and S. Benayoun, "Using partial derivatives of 3D images to extract typical surface features," *Computer Vision and Image Understanding*, vol. 61, no. 2, pp. 171–189, 1995.
- [16] H. Yoshida and J. Näppi, "Three-dimensional computer-aided diagnosis scheme for detection of colonic polyps," *IEEE Transactions on Medical Imaging*, vol. 20, no. 12, pp. 1261–1274, 2001.
- [17] D. Comaniciu and P. Meer, "Mean shift: a robust approach toward feature space analysis," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 24, no. 5, pp. 603–619, 2002.
- [18] R. C. Gonzalez and R. E. Woods, *Digital Image Processing*, Addison Wesley, San Francisco, Calif, USA, 2002.
- [19] X. Ye, X. Lin, J. Dehmeshki, G. Slabaugh, and G. Beddoe, "Shape-based computer-aided detection of lung nodules in thoracic CT images," *IEEE Transactions on Biomedical Engineering*, vol. 56, no. 7, pp. 1810–1820, 2009.
- [20] S. Geman and D. Geman, "Stochastic relaxation, gibbs distributions, and the bayesian restoration of images," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 6, no. 6, pp. 721–741, 1984.