Appearance Models for Robust Segmentation of Pulmonary Nodules in 3D LDCT Chest Images

Aly A. Farag[1], Ayman El-Baz[1], Georgy Gimel'farb[2], Robert Falk[3],
Mohamed A. El-Ghar[4], Tarek Eldiasty[4], and Salwa Elshazly[1]

[1] Computer Vision and Image Processing Laboratory,
University of Louisville, Louisville, KY 40292
{farag, elbaz}@cvip.Louisville.edu
http://www.cvip.louisville.edu
[2] Department of Computer Science, Tamaki Campus,
University of Auckland, Auckland, New Zealand
[3] Director, Medical Imaging Division, Jewish Hospital
[4] Urology and Nephrology Department,
University of Mansoura, Mansoura, Egypt

Abstract. To more accurately separate each pulmonary nodule from
its background in a low dose computer tomography (LDCT) chest im-
age, two new adaptive probability models of visual appearance of small
2D and large 3D pulmonary nodules are used to control evolution of de-
formable boundaries. The appearance prior is modeled with a translation
and rotation invariant Markov-Gibbs random field of voxel intensities
with pairwise interaction analytically identified from a set of training
nodules. Appearance of the nodules and their background in a current
multi-modal chest image is also represented with a marginal probability
distribution of voxel intensities. The nodule appearance model is isolated
from the mixed distribution using its close approximation with a linear
combination of discrete Gaussians. Experiments with real LDCT chest
images confirm high accuracy of the proposed approach.

1 Introduction

Because lung cancer is the most common cause of cancer deaths, fast and
accurate analysis of pulmonary nodules is of major importance for medical
computer-aided diagnostic systems (CAD). In [1] we introduced a fully auto-
matic nodule detection algorithm showing the accuracy up to 93.3% on the
experimental database containing 200 real LDCT chest data sets with 36,000
2D slices. Below we focus in the next CAD stage, namely, on accurate segmenta-
tion of the detected nodules for subsequent volumetric measurements to monitor
how the nodules change in time.

We use a two-step procedure to separate the nodules from their background:
(*i*) an initial LDCT slice is segmented with algorithms introduced in [2] to isolate
lung tissues from surrounding structures in the chest cavity as shown in Fig. 1
and (*ii*) the nodules in the isolated lung regions are segmented by evolving
deformable boundaries under forces that depend on the learned current and prior
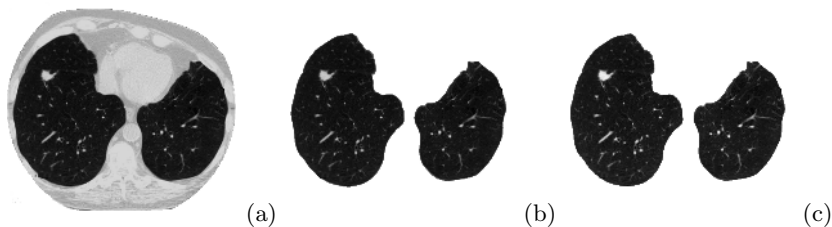
Fig. 1. Step 1 of our segmentation approach (e.g., [2]): an LDCT slice (a) with isolated lungs (b), and (c) the normalized segmented lung image

appearance models. At Step 1 each LDCT slice is modelled as a bi-modal sample from a simple Markov–Gibbs random field of interdependent region labels and conditionally independent voxel intensities (gray levels). This step is necessary for more accurate separation of nodules from the lung tissues at Step 2 because voxels of both the nodules and other chest structures around the lungs are of quite similar intensity.

Previous work. At present, segmentation of pulmonary nodules is under extensive studies. Typical conventional techniques are based on fitting a Gaussian model to empirical data [3] but this approach becomes a challenge when initial measurements are corrupted with outliers and margin-truncation due to neighboring structures. Okada et al. [4] proposed an anisotropic intensity model fitting with analytical parameter estimation. Zhao et al. [5] and Kostis et al. [6] proposed to segment 2D and 3D nodules based on thresholding the voxel intensity. Their algorithms accurately segment well-defined solid nodules with similar average intensities but become unreliable on the cavity or non-solid nodules. ***Our segmentation overcomes these drawbacks due to using at Step 2 deformable boundary models such that their evolution is controlled by both a learned prior probability model of the visual nodule appearance and an adaptive appearance model of the nodules in a current image to be segmented.***

Basic notation. Let (x, y, z) denote Cartesian coordinates of points in a finite arithmetic lattice $\mathbf{R} = [(x, y, z) : x = 0, \ldots, X - 1; y = 0, \ldots, Y - 1, z = 1, \ldots, Z - 1]$. It supports a given 3D grayscale image $\mathbf{g} = [g_{x,y,z} : (x, y, z) \in \mathbf{R}; g_{x,y,z} \in \mathbf{Q}]$ with gray levels from a finite set $\mathbf{Q} = \{0, \ldots, Q - 1\}$ and its region map $\mathbf{m} = [m_{x,y,z} : (x, y, z) \in \mathbf{R}; m_{x,y,z} \in \mathbf{L}]$ with region labels from a finite set $\mathbf{L} = \{\mathsf{nd}, \mathsf{bg}\}$. Each label $m_{x,y,z}$ indicates whether the pixel (x, y, z) in the corresponding data set $\mathbf{g}$ belongs to the goal object (pulmonary nodule), $m_{x,y,z} = \mathsf{nd}$, or to the background, $m_{x,y,z} = \mathsf{bg}$. Let $\mathbf{b} = [\mathbf{P}_k : k = 1, \ldots, K]$ be a deformable piecewise-linear boundary with K control points $\mathbf{P}_k = (x_k, y_k, z_k)$. The index k can be considered as a real number in the interval $\mathbf{K}$ indicating continuous positions around the boundary, e.g. $\mathbf{K} = [1, K]$ for the positions from $\mathbf{P}_1$ to $\mathbf{P}_K$.

Conventional deformable model. moves in the direction that minimizes a boundary energy E such as e.g. in [7]:

$$E = E_{\text{int}} + E_{\text{ext}} = \int_{k \in \mathbf{K}} \left(\xi_{\text{int}} \left(\mathbf{b}(\mathbf{P}_k) \right) + \xi_{\text{ext}} \left(\mathbf{b}(\mathbf{P}_k) \right) \right) dk \qquad (1)$$

where $\xi_{\text{int}} \left(\mathbf{b}(\mathbf{P}_k) \right)$ and $\xi_{\text{ext}} \left(\mathbf{b}(\mathbf{P}_k) \right)$ are internal and external forces, respectively.

In this paper we present a new class of the the external energy that guided the evolution of deformable model based on two new probability models that roughly describe the prior and current visual appearance of the nodules.

2   Data Normalization

To account for monotone (order-preserving) changes of signals (e.g. due to different illumination or sensor characteristics), for each segmented data set, we will calculate the occurrence histogram, then we normalize the segmented data set to make $q_{max} = 255$ for each segmented data set, (e.g see Fig. 1(c)).

3   MGRF-Based Prior Appearance Model

To exclude an alignment stage before segmentation, the appearance of both small 2D and large 3D nodules is modeled with a translation and rotation invariant generic MGRF with voxel-wise and central-symmetric pairwise voxel interaction specified by a set $\mathbf{N}$ of characteristic central-symmetric voxel neighborhoods $\{\mathbf{n}_\nu : \nu \in \mathbf{N}\}$ on $\mathbf{R}$ and a corresponding set $\mathbf{V}$ of Gibbs potentials, one potential per neighborhood.

(a)   (b)
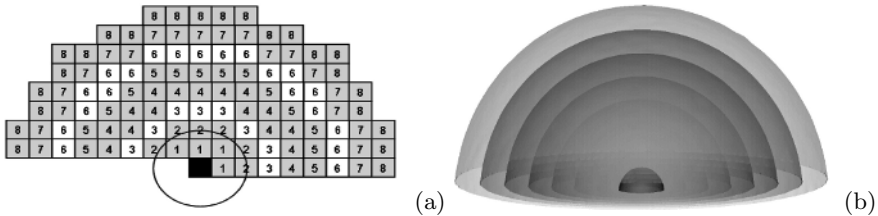
Fig. 2. Central-symmetric 2D (a) and 3D (b) neighborhoods for the eight distance ranges $[d_{\nu,\min} = \nu - 0.5, d_{\nu,\max} = \nu + 0.5)$; $\nu \in \mathbf{N} = \{1, \dots, 8\}$ on the lattice $\mathbf{R}$

A central-symmetric voxel neighborhood $\mathbf{n}_\nu$ embraces all voxel pairs such that (x, y, z)-coordinate offsets between a voxel (x, y, z) and its neighbor (x', y', z') belong to an indexed semi-open interval $[d_{\nu,\min}, d_{\nu,\max})$; $\nu \in \mathbf{N} \subset \{1, 2, 3, \dots\}$ of the inter-voxel distances: $d_{\nu,\min} \leq \sqrt{(x - x')^2 + (y - y')^2 + (z - z')^2} < d_{\nu,\max}$. Figure 2 illustrates the neighborhoods $\mathbf{n}_\nu$ for the uniform distance ranges $[\nu - 0.5, \nu + 0.5)$; $\nu \in \mathbf{N} = \{1, \dots, 8\}$.

The interactions in each neighborhood $\mathbf{n}_\nu$ have the same Gibbs potential function $\mathbf{V}_\nu$ of gray level co-occurrences in the neighboring voxel pairs, and the voxel-wise interaction is given with the potential function $\mathbf{V}_{\text{vox}}$ of gray levels in the voxels:

$$\mathbf{V} = \left[\mathbf{V}_{\text{vox}} = [V_{\text{vox}}(q) : q \in \mathbf{Q}] ; \left\{ \mathbf{V}_\nu = \left[V_\nu(q, q') : (q, q') \in \mathbf{Q}^2 \right] : \nu \in \mathbf{N} \right\} \right]$$

Model identification. Let $\mathbf{R}_t = \{(x, y, z) : (x, y, z) \in \mathbf{R} \ \wedge \ m_{t;x,y,z} = \text{nd}\}$ and $\mathbf{C}_{\nu,t}$ denote the part of the 3D lattice $\mathbf{R}$ supporting the training nodules in the image–map pair $(\mathbf{g}_t, \mathbf{m}_t) \in \mathbf{S}$ and the family of voxel pairs in $\mathbf{R}_t^2$ with the co-ordinate offsets $(\xi, \eta, \gamma) \in \mathbf{n}_\nu$, respectively. Let $\mathbf{F}_{\text{vox},t}$ and $\mathbf{F}_{\nu,t}$ be a joint empirical probability distribution of gray levels and of gray level co-occurrences in the training nodules from the image $\mathbf{g}_t$, respectively: $\mathbf{F}_{\text{vox},t} = [f_{\text{vox},t}(q) = \frac{|\mathbf{R}_{t,q}|}{|\mathbf{R}_t|}; \ \sum_{q \in \mathbf{Q}} f_{\text{vox},t}(q) = 1]$ and $\mathbf{F}_{\nu,t} = [f_{\nu,t}(q,q') = \frac{|\mathbf{C}_{\nu,t;q,q'}|}{|\mathbf{C}_{\nu,t}|}; \sum_{(q,q') \in \mathbf{Q}^2} f_{\nu,t}(q,q') = 1]$ where $\mathbf{R}_{t,q} = \{(x,y,z) : (x,y,z) \in \mathbf{R}_t \ \wedge \ g_{x,y,z} = q\}$ is a subset of voxels supporting the gray level q in the training nodules from the image $\mathbf{g}_t$ and $\mathbf{C}_{\nu,t;q,q'}$ is a subfamily of the voxel pairs $\mathbf{c}_{\xi,\eta,\gamma}(x,y,z) = ((x,y,z),(x+\xi,y+\eta,z+\gamma)) \in \mathbf{R}_t^2$ supporting the gray level co-occurrence (q,q') in the same nodules, respectively.

The MGRF model of the t-th object is specified by the joint Gibbs probability distribution on the sublattice $\mathbf{R}_t$:

$$P_t = \tfrac{1}{Z_t} \exp\left(|\mathbf{R}_t| \left(\mathbf{V}_{\text{vox}}^\mathsf{T} \mathbf{F}_{\text{vox},t} + \sum_{\nu \in \mathbf{N}} \rho_{\nu,t} \mathbf{V}_{\nu,t}^\mathsf{T} \mathbf{F}_{\nu,t} \right) \right) \tag{2}$$

where $\rho_{\nu,t} = |\mathbf{C}_{\nu,t}|/|\mathbf{R}_t|$ is the average cardinality of the neighborhood $\mathbf{n}_\nu$ with respect to the sublattice $\mathbf{R}_t$.

To simplify notation, let areas of the training nodules be similar, so that $|\mathbf{R}_t| \approx R_{\text{nd}}$ and $|\mathbf{C}_{\nu,t}| \approx C_{\nu,\text{nd}}$ for $t = 1, \ldots, T$, where R_{nd} and $C_{\nu,\text{nd}}$ are the average cardinalities over the training set $\mathbf{S}$. Assuming the independent samples, the joint probability distribution of gray values for all the training nodules is as follows:

$$P_\mathbf{S} = \frac{1}{Z} \exp\left(T R_{\text{nd}} \left(\mathbf{V}_{\text{vox}}^\mathsf{T} \mathbf{F}_{\text{vox}} + \sum_{\nu \in \mathbf{N}} \rho_\nu \mathbf{V}_\nu^\mathsf{T} \mathbf{F}_\nu \right) \right)$$

where $\rho_\nu = C_{\nu,\text{nd}}/R_{\text{nd}}$, and the marginal empirical distributions of gray levels $\mathbf{F}_{\text{vox},\text{nd}}$ and gray level co-occurrences $\mathbf{F}_{\nu,\text{nd}}$ describe now all the nodules from the training set. Zero empirical probabilities caused by a relatively small volume of the training data available to identify the above model are eliminated if fractions defining the empirical probabilities in terms of cardinalities of the related sublattices or subfamilies are modified as follows: $(\langle\text{nominator}\rangle + \varepsilon)/(\langle\text{denominator}\rangle + S\varepsilon)$. With the Bayesian quadratic loss estimate, $\varepsilon = 1$ and $S = Q$ for the first-order or $S = Q^2$ for the second-order interactions.

Using the analytical approach similar to that in [2], the potentials are approximated with the scaled centered empirical probabilities:

$$\begin{aligned} V_{\text{vox},\text{nd}}(q) &= \lambda \left(f_{\text{vox},\text{nd}}(q) - \tfrac{1}{Q} \right); \quad (q) \in \mathbf{Q}; \\ V_{\nu,\text{nd}}(q,q') &= \lambda \left(f_{\nu,\text{nd}}(q,q') - \tfrac{1}{Q^2} \right); \ (q,q') \in \mathbf{Q}^2; \nu \in \mathbf{N} \end{aligned} \tag{3}$$

where the common factor λ is also computed analytically. It can be omitted ($\lambda = 1$) if only relative potential values are used for computing relative energies $E_{\nu,\text{rel}}$ of the central-symmetric pairwise voxel interactions in the training data. The energies that are equal to the variances of the co-occurrence distributions:

$$E_{\nu,\mathrm{rel}} = \sum_{q,q' \in \mathbf{Q}^2} f_{\nu,\mathsf{nd}}(q,q') \left(f_{\nu,\mathsf{nd}}(q,q') - \frac{1}{Q^2} \right)$$

allow for ranking all the central-symmetric neighborhoods $\mathbf{n}_\nu$ and selecting the top-rank, i.e. most characteristic ones $\mathbf{N}' \subset \mathbf{N}$ to include to the prior appearance model of Eq. (3). Under the model, any grayscale pattern within a deformable boundary $\mathbf{b}$ in an image $\mathbf{g}$ is described by its Gibbs energy

$$E(\mathbf{g}, \mathbf{b}) = \mathbf{V}_{\mathrm{vox},\mathsf{nd}}^{\mathsf{T}} \mathbf{F}_{\mathrm{vox},\mathsf{nd}}(\mathbf{g}; \mathbf{b}) + \sum_{\nu \in \mathbf{N}'} \mathbf{V}_{\nu,\mathsf{nd}}^{\mathsf{T}} \mathbf{F}_{\nu,\mathsf{nd}}(\mathbf{g}, \mathbf{b}) \qquad (4)$$

where $\mathbf{N}'$ is an index subset of the selected top-rank neighborhoods, and the empirical probability distributions are collected within the boundary $\mathbf{b}$ in $\mathbf{g}$.

4   LCDG-Based Current Appearance Model

The visual appearance of nodules in each current data set $\mathbf{g}$ to be segmented typically differ from the appearance of the training nodules due to non-linear intensity variations from different data acquisition systems and changes in patient tissue characteristics, radiation dose, scanner type, and scanning parameters. This is why, in addition to the appearance prior learned from the normalized training nodules, we model the marginal gray level distribution within an evolving boundary $\mathbf{b}$ in $\mathbf{g}$ with a dynamic mixture of two distributions for current candidates for nodules and their background, respectively. The mixture is closely approximated with a bi-modal linear combination of discrete Gaussians (LCDG) and then partitioned into the nodule and background LCDGs. The approximation is performed with the modified EM-based approach in [2].

5   Boundary Evolution Using Two Appearance Models

The following external energy term in Eq. (1) combines the learned prior and current appearance models to guide an evolving boundary in a way such that maximizes the energy within the boundary:

$$\xi_{ext}\left(\mathbf{b}(\mathbf{P}_k = (x, y, z))\right) = -p_{\mathrm{vox},\mathsf{nd}}(g_{x,y,z}) \pi_{\mathbf{p}}(g_{x,y,z}|\mathbf{S}) \qquad (5)$$

where $p_{\mathrm{vox},\mathsf{nd}}(q)$ is the marginal probability of the gray level q in the LCDG model for the nodules, arteries, and veins (see Section 4) and $\pi_{\mathbf{p}}(q|\mathbf{S})$ is the prior conditional probability of the gray level q, given the current gray values in the characteristic central-symmetric neighborhoods of $\mathbf{P}_k$, for the MGRF prior model in Section 3:

$$\pi_{\mathbf{P}}(g_{x,y,z}|\mathbf{S}) = \frac{\exp\left(E_{\mathbf{P}}(g_{x,y,z}|\mathbf{S})\right)}{\sum_{q \in \mathbf{Q}} \exp\left(E_{\mathbf{P}}(q|\mathbf{S})\right)}$$

where $E_{\mathbf{P}}(q|\mathbf{S})$ is the voxel-wise Gibbs energy for a gray level q assigned to $\mathbf{P}$ and the current fixed gray levels in all neighbors of $\mathbf{P}$ in the characteristic neighborhoods $\mathbf{n}_{\nu}$; $\nu \in \mathbf{N}$:

$$E_{\mathbf{P}}(q|\mathbf{S}) = V_{\mathrm{vox,nd}}(q) + \sum_{\nu \in \mathbf{N}} \sum_{(\xi,\eta,\gamma) \in \mathbf{n}_{\nu}} (V_{\nu,\mathrm{nd}}(g_{x-\xi,y-\eta,z-\gamma}, q) + V_{\nu,\mathrm{nd}}(q, g_{x+\xi,y+\eta,z+\gamma}))$$

The boundary evolution in each 2D section with the fixed z-coordinate termi-
nates after the total energy $E_{\mathbf{r}}$ of the region $\mathbf{r} \subset \mathbf{R}$ inside the boundary $\mathbf{b}$ does
not change:

$$E_{\mathbf{r}} = \sum_{\forall \mathbf{P} = (x,y,z) \in \mathbf{r}} E_{\mathbf{P}}(g_{x,y,z}|\mathbf{S}) \tag{6}$$

The deformable boundary $\mathbf{b}$ evolves in discrete time, $\tau = 0, 1, \ldots, T$, as follows:

1. Initialization ($\tau = 0$):
   (a) Initialize a boundary inside a nodule (e.g. automatically as in [2]).
   (b) Using voxels within and outside the initial boundary, estimate the cur-
       rent "nodule" and "background" LCDGs $\mathbf{p}_{\mathrm{vox,nd}}$ and $\mathbf{p}_{\mathrm{vox,bg}}$.
2. Evolution ($\tau \leftarrow \tau + 1$):
   (a) Calculate the total energy of Eq. (6) within the current boundary $\mathbf{b}_{\tau}$.
   (b) For each control point $\mathbf{P}_k$ on the current boundary, indicate the exterior
       $(-)$ and interior $(+)$ nearest neighbors with respect to the boundary.
   (c) For each $(+)$-point, calculate the total energy of Eq. (1) for each new
       candidate for the current control point.
   (d) Select the minimum-energy new candidate.
   (e) Calculate the total energy of Eq. (6) within the boundary that could
       have appeared if the current control point has been moved to the selected
       candidate position.
   (f) If the total energy increases, accept this new position of the current
       control point, otherwise for each $(-)$-point, calculate the total energy of
       Eq. (1) for each new candidate for the current control point.
   (g) Select the minimum-energy new candidate.
   (h) Calculate the total energy of Eq. (6) within the boundary that could
       have appeared if the current control point has been moved to the selected
       candidate position.
   (i) If the total energy increases, accept this new position of the current
       control point.
   (j) Otherwise do not move the current control point because it is already
       located on the edge of the desired nodule.
   (k) Mark each voxel visited by the deformable boundary.
   (l) If the current control point moves to the voxel visited earlier, then find
       the edge formed by the already visited voxels and use the edge points as
       the new control points of the deformable boundary.
   (m) If the new control points appear, interpolate the whole boundary using
       cubic splines and then smooth its control points with a low pass filter.
   (n) If the total energy within the boundary does not change, terminate the
   process; otherwise return to Step 2b.

6   Experimental Results and Conclusions

The proposed segmentation algorithm was tested on a database of clinical multislice 3D chest LDCT scans of 29 patients with $0.7 \times 0.7 \times 2.5$ mm^3 voxels that contains 350 nodules, in particular, 150 solid nodules of larger than 5 mm in diameter, 40 small solid nodules of less than 5 mm diameter, 10 cavity nodules, 61 nodules attached to the pleural surface, and 89 largely non-spherical nodules. The diameters of the nodules range from 3mm to 30mm.
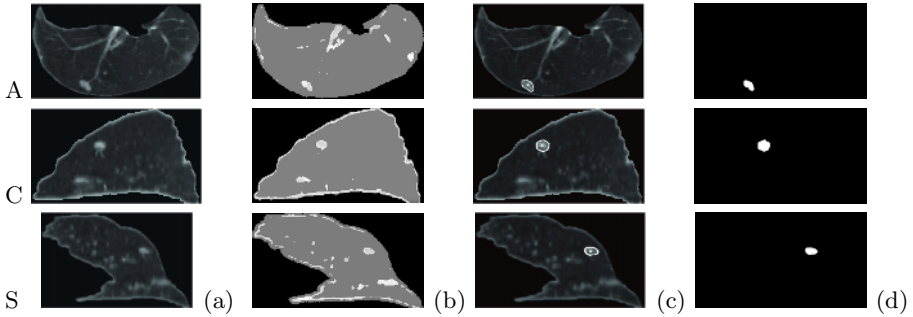
Fig. 3. 3D segmentation of pleural attached nodules; results are projected onto 2D axial (A), coronal (C), and saggital (S) planes for visualization: 2D profile of the original nodule (a), pixel-wise Gibbs energies (b) for $\nu \leq 11$, our segmentation (c), and (d) the radiologist's segmentation.

Figure 3 illustrates results of segmenting pleural attached nodules shown by axial, sagittal, and coronal cross sections. The pixel-wise Gibbs energies in each cross section are higher for the nodules than for any other lung voxels including the attached artery. Therefore, our approach separates accurately the pulmonary nodules from any part of the attached artery. The evolution terminates after 50 iterations because the changes in the total energy become close to zero. The error of our segmentation with respect to the radiologist "ground truth" is 1.86%.
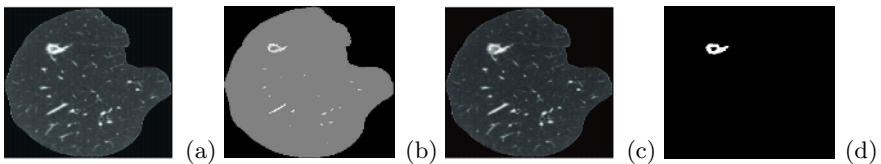
Fig. 4. 2D segmentation of cavity nodules: (a) 2D profile of the original nodule, (b) pixel-wise Gibbs energies for $\nu \leq 11$, (c) our segmentation, and (d) the radiologist's segmentation

The main advantage of our approach over the existing algorithms is in the more accurate segmentation of thin cavity nodules, i.e. the nodules that appear

only in a single slice. Experimental results in Fig. 4 show that the error of our segmentation with respect to the radiologist is 2.1%. It is worthy to note that all the existing approaches fail to segment this cavity nodule because it is totally inhomogeneous. Figure 5 presents more segmentation results obtained by our algorithm. In total, our segmentation of the 350 nodules has an error range of 0.4% – 2.35% with a mean error of 0.96%, and a standard error deviation of 1.1%.
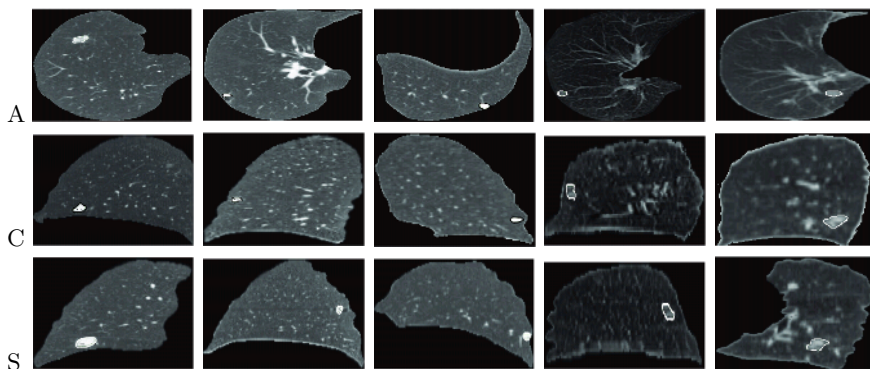
Fig. 5. Our segmentation for five more patients

We introduced a new method to accurately segment small 2D and large 3D pulmonary nodules on LDCT chest images. In our approach, the evolution of a deformable model is controlled with probability models of visual appearance of pulmonary nodules. The prior MGRF model is identified from a given training set of nodules. The current appearance model adapts the control to each bi-modal image of lung regions to be segmented. Both the models are learned using simple analytical and numerical techniques. Experiments with real LDCT chest images confirm high accuracy of our segmentation with respect to the radiologist's ground truth. Our segmentation outperforms other existing approaches for all types of nodules, in particular, for cavity nodules where other existing approaches fail.

References

1. A. A. Farag, A. El-Baz, and G. Gimelfarb, "Quantitative Nodule Detection in Low Dose Chest CT Scans: New Template Modeling and Evaluation for CAD System Design," *Proc. of International Conference on Medical Image Computing and Computer-Assisted Intervention (MICCAI)*, Palm Springs, California, USA, October 26-29, 2005, pp. 720-728.
2. A. A. Farag, A. El-Baz, and G. Gimel'farb, "Precise Segmentation of Multi-modal Images," *IEEE Transactions on Image Processing*, vol. 15, no. 4, pp. 952-968, April 2006.

3. K. Fukunaga, *Statistical Pattern Recognition*. San Diego: Academic, 1990.
4. K. Okada, D. Comaniciu, and A. Krishnan, "Robust Anisotropic Gaussian Fitting for Volumetric Characterization of Pulmonary Nodules in Multislice CT," *IEEE Trans. on Medical Imaging*, vol. 24, no. 3, pp.409–423, March 2005.
5. B. Zhao, D. Yankelevitz, A. Reeves, and C. Henschke, "Two-dimensional multicriterion segmentation of pulmonary nodules on helical CT images," *IEEE Trans. on Medical Imaging*, vol. 22, pp. 1259–1274, 2003.
6. W. J. Kostis, A. P. Reeves, D. F. Yankelevitz, and C. I. Henschke, "Three-dimensional segmentation and growth-rate estimation of small pulmonary nodules in helical CT images," *IEEE Trans. on Medical Imaging*, vol. 22, pp. 1259–1274, 2003
7. M. Kass, A. Witkin, and D. Terzopoulos, "Snakes: Active contour models," *International Journal of Computer Vision*, vol. 1, pp. 321–331, 1987.