Int J CARS (2015) 10:1505–1514
DOI 10.1007/s11548-015-1150-0

ORIGINAL ARTICLE

Automatic lung tumor segmentation with leaks removal in follow-up CT studies

R. Vivanti · L. Joskowicz · O. A. Karaaslan · J. Sosna

Received: 18 November 2014 / Accepted: 31 December 2014 / Published online: 22 January 2015

Abstract
Purpose In modern oncology, disease progression and response to treatment are routinely evaluated with a series of volumetric scans. The number of tumors and their volume (mass) over time provides a quantitative measure for the evaluation. Thus, many of the scans are follow-up scans. We present a new, fully automatic algorithm for lung tumors segmentation in follow-up CT studies that takes advantage of the baseline delineation.
Methods The inputs are a baseline CT scan and a delineation of the tumors in it and a follow-up scan; the output is the tumor delineations in the follow-up CT scan; the output is the tumor delineations in the follow-up CT scan. The algorithm consists of four steps: (1) deformable registration of the baseline scan and tumor's delineations to the follow-up CT scan; (2) segmentation of these tumors in the follow-up CT scan with the baseline CT and the tumor's delineations as priors; (3) detection and correction of follow-up tumors segmentation leaks based on the geometry of both the foreground and the background; and (4) tumor boundary regularization to account for the partial volume effects.
Results Our experimental results on 80 pairs of CT scans from 40 patients with ground-truth segmentations by a radiologist yield an average DICE overlap error of 14.5 % (std = 5.6), a significant improvement from the 30 % (std = 13.3) result of stand-alone level-set segmentation.
Conclusion The key advantage of our method is that it automatically builds a patient-specific prior to the tumor. Using this prior in the segmentation process, we developed an algorithm that increases segmentation accuracy and robustness and reduces observer variability.

Keywords Lung tumor · Tumor segmentation · Follow-up CT scans · Longitudinal studies

R. Vivanti (✉) · L. Joskowicz
The Rachel and Selim Benin School of Computer Science and Engineering, The Hebrew University of Jerusalem, Givat Ram Campus, 91904 Jerusalem, Israel
e-mail: refael.vivanti@mail.huji.ac.il

O. A. Karaaslan · J. Sosna
Department of Radiology, Hadassah Hebrew University Medical Center, Jerusalem, Israel

Introduction

Radiological follow-up of tumors is the cornerstone of modern oncology. Disease progression and response to treatment are routinely evaluated with a series of volumetric scans. The number of tumors and their volume (mass) over time provides a quantitative measure for the evaluation. Today, most radiologists rely on standards such as WHO and RECIST to estimate the tumor mass based on the length and width of tumor masses on a single scan slice. It is well known that this estimate can be off by as much as 50 %, especially for tumors with multiple internal components and irregular shapes. Previous research shows that true volumetric measurements are the most accurate information for tumor monitoring [1].

Tumor delineation is the main bottleneck of tumor volume computation. Manual delineation is time-consuming, user dependent, and requires expert knowledge. Semi-automatic segmentation methods, e.g. live wire and region growing, also require user interaction and may lead to significant intra- and inter-observer variability. Automatic tumor segmentation poses significant challenges and is used in the clinic for only a handful of tumor types. Model-based methods rely on tumor priors to guide the segmentation [2–4]. They are also limited, as they require the construction of a generic tumor prior, which may not be available and is by itself

a research problem. Moreover, most methods process each scan independently without considering that it is from the same patient. Recent works incorporate the tumor delineation of the first (baseline) scan as a patient-specific prior for the automatic tumor segmentation of the next (follow-up) scan [5]. These works show that the robustness and the tumor volume and tumor volume difference accuracy may be significantly improved when the patient-specific tumor delineation from the baseline is used.

Radiological lung tumors follow-up is of great practical importance, as lung cancer leads the mortality statistics. Moreover, many other cancer types develop lungs metastases. The quantification of the tumors volumes and their changes in size and shape plays a key role in the treatment decision-making process. Consequently, a reliable automatic tool for accurate follow-up assessment may contribute to the treatment success.

In this paper, we present a new, fully automatic algorithm for lung tumor segmentation in follow-up CT studies. The inputs are the baseline scan, the tumor's delineations, and the follow-up CT scan; the outputs are the tumor's delineations in the follow-up CT scan. The delineation of the tumor in the baseline scan can be obtained by automatic or semi-automatic segmentation methods. This additional one-time effort is compensated by the time and accuracy gains on the analysis of the follow-up series. The algorithm consists of four steps: (1) deformable registration of the baseline scan and tumor's delineations to the follow-up CT scan; (2) segmentation of these tumors in the follow-up CT scan using the baseline tumor's delineations and statistical intensity models of the background and the foreground of the follow-up CT; (3) detection and removal of tumors segmentation leaks using geometry-based methods that exploit the geometry of both the tumor (foreground) and the adjacent anatomy (background); and (4) tumor boundary regularization to correct the partial volume effects.

Our method has the following advantages over existing ones: (1) it is fully automatic; (2) it builds a strong patient-specific prior from the baseline tumor delineation and statistical model of the foreground/background in the follow-up image that help improving segmentation robustness and accuracy; (3) it performs local deformable registration to model more accurately the tumor transformation; (4) it corrects for tumor segmentation leaks caused by weak/nonexistent tumor boundaries in the images with a new method based on the pulmonary surface geometry; and (5) it accounts for segmentation errors resulting from the partial volume effect. Our experimental results on CT scans from 40 patients with ground-truth segmentations generated by a radiologist yields an average overlap error of 14.5 % (std = 4.1), a significant improvement of the 30 % (std = 13.3) of stand-alone level-set segmentation.

Related work

The three most relevant research areas to our work are as follows: (1) lung follow-up studies; (2) lung tumor segmentation, and (3) lung scans registration. We briefly discuss each next.

Lung follow-up studies

Hollensen et al. [6] address the task of follow-up studies of lung tumors. Their method starts with manual rough positioning followed by rigid registration between the baseline and the follow- up scans. The baseline delineation is then used as the initialization of the follow-up segmentation with the electric flow lines and min graph-cut methods. This method, which is the closest to ours, is demonstrated on a small database of ten cases and does not handle segmentation leaks to the pleural surface or blood vessels. Opfer et al. [2] describe a semi-automatic method for lung tumors delineation in follow-up studies in which the user marks points on the baseline image. With multi-resolution rigid registration, the method then transforms the points to the follow-up scan, where the user can use them for model fitting. The main drawback is the rigid registration which is not accurate for lungs. Also, this method requires user delineation on both images.

Lung nodules segmentation and their radiological follow-up have received significant attention. Unlike lung tumors, lung nodules usually have a known diameter and a spherical shape, so their stand-alone segmentation is easier than that of lung tumors. For example, Brown et al. [7] present a method in which the baseline and follow-up lung nodules segmentation are performed independently on the scans without using a patient-specific prior. Kuhnigk et al. [8] generate lung nodule segmentations by region growing with a fixed lower threshold followed by morphological operations to remove leaks to vessels. They then use the nodule convex hull to remove leaks to the pleural surface. Reeves et al. [9] describe a method that relies on baseline lung nodule segmentation to determine the region of interest in the follow-up scan. Adaptive thresholding is then performed, followed by the removal of leaks to pleural surface using an oriented plane.

The VOLCANO'09 lung nodules follow-up challenge [10] comprises 13 groups and 50 datasets with no baseline delineation. Among the participating groups, Kostis et al. [11] present a method based on thresholding and connected component analysis. Segmentation leaks to vessels are corrected with morphological opening adjusted by the user. Pleural surface attachments are removed with a separating plane. This is adequate for specific areas of the pleural surface and for small nodules, but not for larger tumors. Jirapatnakul et al. [12] model the pleural surface with a parabola and use it to complete the lung nodules fuzzy and/or missing

boundaries. The parabolic surface is estimated using nodule-specific heuristics and is used to remove outliers. While this leak removal method is similar to ours, their heuristics may not hold for lung tumors due to their irregular shape and internal components, and location variability. Recently, Chen et al. [13] present a method for the segmentation and identification of solitary pulmonary nodules on follow-up CT scans. This method performs affine and non-rigid co-registration of follow-up scans and then matches the detected nodules in the follow-up scan with a generic similarity measure.

Lung tumor segmentation

Methods for individual stand-alone pulmonary tumors segmentation include thresholding, region growing, and level sets, often combined with advanced image-processing methods. The first challenge of these methods is the detection of the tumors. Some methods require the user to produce an initial delineation or require the known tumor location from the baseline scan. Other methods rely on the PET scan, in which automatic tumor detection, e.g., by thresholding, is much simpler. For PET/CT scans, Gribben et al. [14] propose to use the PET scan for tumor detection, followed by unsupervised Maximum A Posterior Markov Random Field on the registered CT scan values. Kanakatte et al. [16] also use the PET scan for tumor detection, but combine thresholding and components analysis to produce the final segmentation. Plajer et al. [3] classify lung tumors in stand-alone CT scans into five categories according to the nearby structures and apply mixed internal/external force segmentation and clustering. Awad et al. [4] use multi-parameter level set with a sphere shape prior. The method requires the manual adjustment of up to 20 internal parameters. Their validation on 21 tumors yields a volume overlap error of 30 %, which may be excessive for disease progression evaluation. Gu et al. [16] propose a method that uses region growing from multiple seeds followed by voting. Their method does not correct segmentation leaks. Moreover, their validation study on a small database of 15 scans focuses only on robustness to seed location.

Lung scans registration

Lung CT scans registration is challenging because the lung deformations are non-rigid and because of the sliding conditions on the lungs surface [17]. A variety of methods that rely on intensity and/or geometrical information have been proposed [18]. In the EMPIRE10 lungs registration challenge, 24 methods were tested on 30 datasets [19]. The highest scoring method by Song et al. [20] uses topology-preserving diffeomorphic transformations. Modat et al. [21] describe a block-matching approach with free-form deformation. Lung registration methods usually produce good registration results between the lungs, but may incur in large errors when used for tumor registration. In many cases, an additional fine-tuning step is required to achieve good registration in the tumor area.

Methods

The basic premise of our method is that the radiologist-validated tumor delineation in the baseline scan is a high-quality prior for the follow-up scan. The algorithm consists of four steps, which we describe in detail next.

Deformable registration of the baseline and follow-up

The initial step is to register the baseline CT scan and tumor's delineation with the follow-up scan. The resulting transformation defines the approximate location of the tumor mass in the follow-up and obviates the need for a separate stand-alone tumor detection algorithm. It is also used to build a tumor intensity prior and background intensity prior to delineate the corresponding tumor in the follow-up scan.

We start by performing a deformable registration between the baseline and follow-up scans in the lungs region of interest (ROI), which is automatically detected in the baseline scan. This lung ROI global deformable registration consists of a rigid affine registration followed by a deformable registration with B-splines. This stage usually registers the lungs properly, although the tumor itself may be poorly registered, especially for large changes in the tumor volume. In this case, the new tumor center of mass might be outside of the registered tumor.

To overcome this challenge, we perform a separate local deformable registration for *each tumor*. The baseline tumor delineation is enclosed in a bounding box that defines the local tumor ROI. The follow-up ROI is determined from the baseline tumor by projecting it to the follow-up scan using the global transformation, enclosed in a bounding box and with an added margin to account for tumor changes and residual registration errors. This allows modeling more accurately the different tumors change rates. This local registration is performed for each baseline tumor in three stages: (1) translation-only registration to account for large changes in the tumor volume; (2) rigid affine registration; and (3) deformable registration by mutual information.

Note that although this registration step yields acceptable priors for the tumor location and the tumor shape, it is not always sufficiently accurate for specific cases and clinical applications as shown by the detailed evaluation of these intermediate results described in "Experimental results" section. To obtain the desired coverage, robustness, and accuracy, we perform follow-up tumor segmentation as described next.

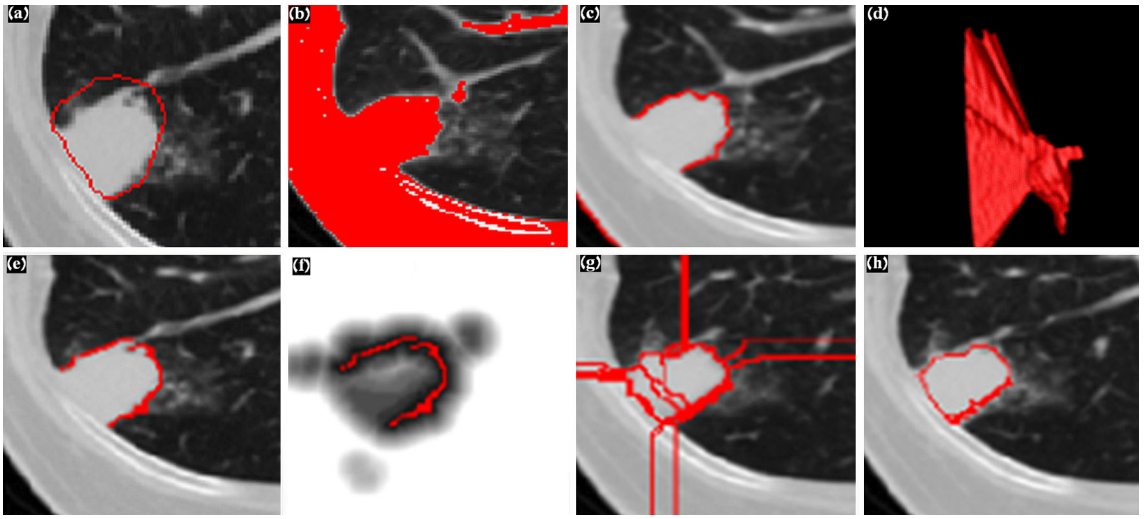

Fig. 1 Illustration of the segmentation process stages: **a** the baseline tumor delineation (*red*) overlaid on the corresponding follow-scan CT slice after registration; **b** maximum likelihood follow-up tumor segmentation (*red*) with leaks; **c** segmentation boundary B (*red*) as seen from the center point; **d** 3D view of the ray casting result; **e** follow-up tumor boundary (*red*) after leaks removal; **f** distance map from follow-up tumor boundary (*red*); **g** watershed regions, and; **h** final follow-up tumor segmentation

Initial follow-up tumor segmentation

The initial follow-up tumor segmentation inside the ROI consists of two steps: initial segmentation followed by leaks correction. The segmentation is performed using a statistical model of the foreground (tumor) and the background (other structures). The foreground parameters are computed from the prior voxels in the follow-up scan; the background parameters are computed from the prior neighborhood voxels. Note that we do not use voxels from the baseline image, as their intensity values may differ considerably from those of the follow-up scan, e.g., due to differences in scanning protocols and contrast agent variability.

Since the registration step in "Deformable registration of the baseline and follow-up" section is not always accurate (Fig. 1a), the foreground voxels may contain background voxels and vice versa. To remove them, we model the intensity distribution of the tumors and the other structures as a mixture of two Gaussian distributions, classify the voxels according to these two Gaussians using the k-means algorithm, estimate the parameters of each Gaussian, and remove the Gaussian that corresponds to the registration error. For the foreground, we select from the foreground Gaussians the one with the highest mean. Note that the eventual presence in the ROI of small regions that include parts of bones does not significantly influence its parameters. For the background, we select from the two background Gaussians the one with the lowest mean. Next, we use these two representative Gaussians of the foreground and the background to compute the initial tumor segmentation by maximum likelihood estimation. Finally, we choose the connected component that is the closest to the prior center of mass and remove all other components.

Segmentation leaks removal: bottleneck-shaped leaks

Since the tumors may be attached to neighboring structures with similar intensities, any intensity-based segmentation method, including ours, will include parts of these structures in the segmentation (Fig. 1b). Our goal is to automatically detect for each tumor these erroneous regions, called *segmentation leaks* [22], and automatically correct them. For this purpose, we use geometric boundary models for the tumor and the background structures.

We correct segmentation leaks to neighboring vessels, to the pleural wall, and/or to the diaphragm with the background geometric model. Note that we cannot use the baseline shape as a prior, as it can change dramatically during the tumor growth/shrinkage. We automatically detect and correct the segmentation leaks in two stages. The first stage handles bottleneck-shaped leaks; the second stage handles leaks caused by missing boundaries. We describe bottleneck-shaped leaks correction in this section and missing boundary leaks in the next one.

We model the tumor as a star-shaped structure. In a star-shaped structure, the kernel is the set of points that can be connected by a straight line to each point on the tumor surface without leaving the tumor volume. We observe that the majority of lung tumors are star-shaped. This empirical observation is based on our datasets, on the experience of our senior radiologist, and on the observation in Awad et al. [4].

We use our previous method for the detection and removal of bottleneck-shaped leaks [22]. For tumors that are not star-shaped, our method segments out the parts of the tumor that cannot be seen from the kernel.

The input to this step is the initial tumor segmentation with leaks and a point $c = (c_x, c_y, c_z)$ in the segmented tumor kernel. We choose this point as the center of mass of the segmentation prior which is the baseline tumor registered and projected onto the follow-up scan. If this center point turns to be outside the segmentation prior, we choose the closest point to that is inside the segmentation prior. We find the boundary of the tumor that can be seen from the center c by dense ray casting by projecting a ray p from c in all 3D directions and recording the location of the last voxel before the ray leaves the segmentation volume S as part of the segmentation boundary B:

$$\begin{aligned} &p_i \in B \; if \; f \; p_i \in S, p_{i+1} \notin S \\ &s.t. \\ &s = \max\left(|r_x - c_x|, |r_y - c_y|, |r_z - c_z|\right) \\ &d = \frac{r - c}{s} \\ &p_i = c + id \end{aligned} \tag{1}$$

where p_i is a voxel location on the ray p, $r = (r_x, r_y, r_z)$ is a voxel on the image boundary, s is the step size, and d is the smallest step size in each direction. Note that we choose a small step size that depends only on the image boundaries to ensure that the ray will not miss the tumor boundary. The result is the segmentation boundary B as seen from c (Fig. 1c, d).

Small segmentation leaks are detected as follows. When casting rays from the tumor center outwards, the rays will stop at sharp boundary segment. However, when the boundary segments are fuzzy or missing altogether, they will continue to adjacent, further away and disconnected boundaries, thus causing segmentation leaks. Consequently, the boundaries that result from the leaks will not be connected to the tumor boundary, resulting in a discontinuity. In 3D, the actual tumor boundary will form a single connected component (possibly with missing boundary segments), regardless of the number of leaks. To remove these leaks, we perform a connected components analysis on the set B (Eq. 1) and select the largest connected component that surrounds point c (Eq. 1) to be the segmentation known boundary (Fig. 1e). The missing boundaries of this known boundary are the holes that caused the segmentation leaks. We remove them by filling the boundary holes as follows.

We first compute the voxels Euclidean distance map from the known boundary using the MATLAB function *bwdist* (Fig. 1f). Next, we identify the watershed regions in this distance map (Fig. 1g) using the MATLAB function *watershed* [23]. Finally, we select the regions whose intersection with the known segmentation is greater than a threshold of 10 %, chosen empirically after testing several thresholds on all the datasets (Fig. 1h). A voxel belongs to the known segmentation if it is on a ray connecting the center point c and a point in the known boundary. Figure 1 illustrates this process.

Segmentation leaks removal: missing boundaries

A second type of segmentation leaks is caused by missing tumor boundaries. In this case, there is no prior indication where the actual tumor boundary is, regardless of what its prior shape is, as computed in the previous step. The only indication about the missing tumor boundary segments locations comes from the adjacent tissue, usually the chest wall. For example, when the connection of the tumor to the chest wall is at an obtuse angle (Fig. 2), the first leaks removal method (section "Segmentation leaks removal: bottleneck-shaped leaks") will fail to correctly complete the missing boundary. Thus, we use a local geometric prior of the adjacent structures to determine the missing boundary and remove the leak.

We model the local boundary of these adjacent structures as a parabolic surface. This boundary shape approximation holds for the pleural surface, the diaphragm, and parts of the heart walls. In the absence of clear boundary between the tumor and the leak, any tumor boundary might be inaccurate. We choose a parabolic surface to remove the leak since it is similar to the way a human delineator would complete the missing boundary segment. We empirically found that this is better than using the shape prior defined by the baseline tumor delineation, e.g., as described in [15].

To create this model, we use the tumor segmentation of the previous stage and find its boundary points (Fig. 2a, b). These boundary points can be inliers from the adjacent structures borders or outliers from the tumor or vessels.

We apply the RANSAC outlier detection framework with a parabolic surface model to simultaneously find the parabolic surface parameters and the inliers (Fig. 2c). A parabolic surface is the set of points $\{v_i = (x_i, y_i, z_i)\}$ such that:

$$z = f\ (x, y) = p_1 x^2 + p_2 xy + \ p_3 y^2 + p_4 x + p_5 y + p_6 \tag{2}$$

We also require the cloud of points to be monotonic in the z-axis direction. To achieve this, we first rotate the cloud of points so that its axis of smallest variance, obtained from principal component analysis (PCA), is aligned with the z-axis. Next, we fit a parabolic model to the resulting points cloud. The parabolic model P minimizes the squared error of

$$A_{N\times 6} P_{6\times 1} = Z_{N\times 1}$$

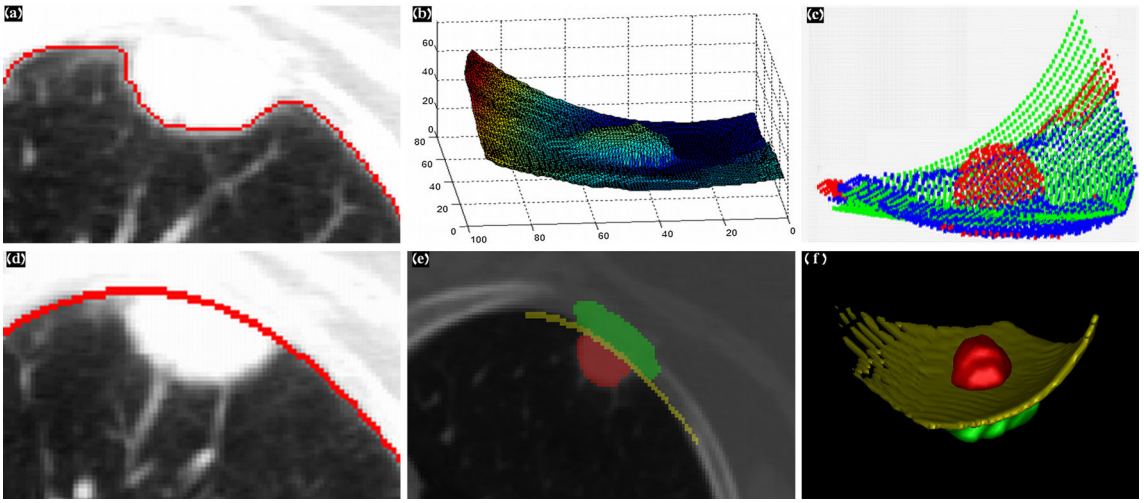

Fig. 2 Illustration of the background geometric modeling stages: **a** background and foreground boundary points (*red*) on a slice; **b** 3D model of the boundary points; **c** RANSAC result: *green* parabolic surface, *blue* inliers, *red* outliers; **d** one slice with parabolic surface; **e** segmentation results: *red* tumor, *yellow* parabolic surface, *green* leak; **f** 3D visualization

where

$$P = \left(A^{\mathrm{T}}A\right)^{-1} A^{\mathrm{T}}Z \quad (3)$$

$$A_{N\times 6} = \begin{bmatrix} x_1^2 & x_1y_1 & y_1^2 & x_1 & y_1 & 1 \\ \vdots & & & & & \\ x_n^2 & x_ny_n & y_n^2 & x_n & y_n & 1 \end{bmatrix} \quad (3)$$

and $Z_{N\times 1} = (z_1, \ldots, z_n)^T$. For each 3D point (x, y, z), we substitute the resulting matrix P and (x, y) in Eq. (2) and obtain z', the projection of the point on the parabolic surface in the (rotated) z-axis direction. Thus, $|z' - z|$ is the distance of the reconstructed point from the parabolic surface in the direction of the z-axis. Every point whose distance exceeds a predefined threshold is considered an outlier.

We repeatedly randomly choose six points with the RANSAC method. We estimate the parabolic surface parameters P from the points using Eq. (3) and use the resulting P to find which points are inliers. Finally, we choose the parabolic surface with the largest set of inliers. The final refinement step iteratively estimates the parabolic surface parameters using the inliers points from the previous iteration. It then computes a new superset of inlier points that agree with the new parabolic surface. The iterations stop when the inliers set size remains the same. Finally, we remove the part of the tumor that is separated from the center c by the parabolic surface (Fig. 2d–f).

Tumor boundary refinement

In the last step, we address the small segmentation errors caused by the partial volume effect (PVE), which blurs tumors boundaries. Since the dynamic range of a CT scan is wider than in regular images, when a radiologist delineates a tumor, he/she usually determines a lower and an upper threshold to enhance the tumor boundaries. Choosing slightly different thresholds may change the tumor boundary and thus its volume (Fig. 3a). Therefore, the PVE may cause variability in the tumor delineation by different radiologists and segmentation algorithms.

To reduce the tumor boundaries variability resulting from the PVE, we generate several possible segmentations and choose the best one as follows. First, we compute the variance map of the image from the variance of a small window around each voxel (Fig. 3b). Then, we compute several segmentations by incrementally dilating or eroding by one voxel the tumor segmentation (Fig. 3c). For each of the alternative segmentations, we compute the mean variance of its boundary voxels using the variance map and choose the one with the highest mean variance (Fig. 3d). This reduces the variability between different segmentations of the same tumor in different scans and/or by different observers. Specifically, since the radiologist co-author approved our ground-truth delineation after the PVE was treated using the above-mentioned method, it also reduced the PVE-related variability in our data, thereby increasing the overall accuracy of our results.

Experimental results

We evaluate our method on a database of clinical CT scans from 40 patients from the Hadassah Medical Center. The scans were acquired on a 64-row CT Phillips Brilliance 64 scanner and are of size $512 \times 512 \times 350$–$500$ voxels, with spatial resolution of 0.6–1.0×0.6–1.0×0.7–3 mm, with

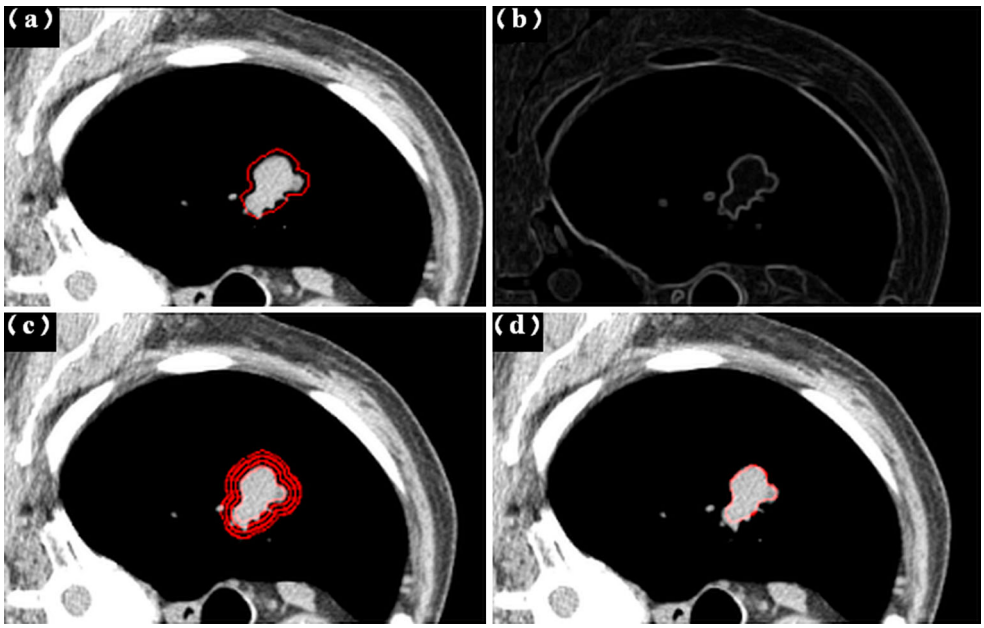

Fig. 3 Illustration of the partial volume correction process: **a** original scan with tumor segmentation so far; **b** variance map; **c** nine possible segmentations; **d** the final segmentation

contrast agent administration. The cases were carefully chosen from the hospital archive by the radiologist co-author to represent the variety of patient ages, conditions, and pathologies. The mean time between the baseline and the follow-up scans is 4.9 (std = 2.4) months. The mean tumor volume is 43.8 (std = 49.9) ml, and the mean volumetric change is 17.8 (std 29.7) ml. Of the 40 scans, 32 scans include tumors adhered to the lung wall and eight show isolated tumors. An expert radiologist produced ground-truth delineations of the tumors in both the baseline and the follow-up CT scans.

For the evaluation, we use each pair of CT scans twice: forward (from baseline to follow-up) and backwards (from follow-up to baseline). Although the backwards direction is not a real clinical case and is correlated to the pair in the forward direction, it provides additional data and attests the robustness and accuracy of our method. To prevent bias, we present the results for the forward and backwards pairs separately.

We compare the results of our method with the ground truth by computing the standard DICE volumetric overlap error (VOE) and the average symmetric surface distance (ASSD), which are defined as follows:

$$\mathrm{DICE}\,(X, Y) = \frac{2\,|X \cap Y|}{|X| + |Y|}$$

$$\mathrm{ASSD}\,(X, Y) = \frac{1}{|X|} \sum_{p \in X} \min_{p_2 \in Y} (|p - p_2|)$$

All segmentation steps were done with Elastix [24]. For the B-spline, we chose a grid spacing of 12 mm. For the watershed region, we used the MATLAB functions 'bwdist' and 'watershed' For the watershed region selection, we set the threshold to 10 %, chosen ampirically. We used our own version of the RANSAC algorithm. We set the RANSAC iterations bound to 10,000, with a threshold of three voxels. For VOE regularization (section "Tumor boundary refinement"), we compute the variance map of the image on a local $5 \times 5 \times 5$ window around each pixel. We then chose the best from nine possible segmentations as the final result.

Processing times range between 5–10 min (depending on the scan data size and the tumor size) on a PC with Intel Xenon 3.07 GHz Processor running 64-bit Windows 7. Note that the computation complexity of all of the method stages is linear in the size of the input.

We compare the results of our method to the fast marching segmentation method [25]. Note that despite the numerous publications in this field, we could not find a publicly available state-of-the-art lung tumors segmentation method and datasets for quantitative evaluation and comparison.

Fast marching requires a seed that serves as the origin of the propagation. We chose the center of mass of the ground-truth tumor segmentation as the seed. For the propagation speed function, we chose the inverted (minus) gradients map values. Since each iteration may yield a different segmentation, we stop the propagation when 90 % of the ground truth was segmented. Note that although we use the fast marching method without a shape prior, we obtain similar or better results than other state-of-the-art methods that use fast marching or level sets with shape prior. For example, Awad et al. [4] report similar error measures to those we obtained

Table 1 Experimental results of 40 forward cases and 40 backwards cases, with a comparison to Fast Marching method

	Forward VOE (%)		Forward ASSD (mm)		Backwards VOE (%)		Reversed ASSD (mm)	
	Ours	FM	Ours	FM	Ours	FM	Ours	FM
Mean	14.47	26.84	1.03	2.73	15.37	33.28	1.03	4.32
Std	4.14	7.63	0.55	1.56	6.79	16.54	0.62	3.57
Min	6.32	14.86	0.17	0.69	6.35	12.72	0.33	0.65
Max	23.25	44.2	2.46	8.54	38.37	82.1	3.07	16.09

VOE, volume overlap error in %; ASSD, average symmetric surface distance in mm; Ours, our method; FM, fast marching method

Table 2 Contribution of the registration step to the final segmentation results

	Forward VOE (%)		Forward ASSD (mm)		Backwards VOE (%)		Reversed ASSD (mm)	
	Final	Registration only	Final	Registration only	Final	Registration only	Final	Registration only
Mean	14.47	24.53	1.03	1.34	15.37	31.87	1.03	1.47
Std	4.14	12.95	0.55	2.02	6.79	16.15	0.62	1.24
Min	6.32	9.64	0.17	0.29	6.35	8.96	0.33	0.26
Max	23.25	76.86	2.46	11.18	38.37	70.9	3.07	5.88

The table shows the results after the registration step only and the results after both the registration and the segmentation step for all 40 forward cases and 40 backwards cases.VOE, volume overlap error in %; ASSD, average symmetric surface distance in mm; Final, results after registration and segmentation; Registration only, results after the registration step only

with the generic fast marching method. Note that we "help" the fast marching method by using the ground truth for both seeding and for the termination criterion—without it, the fast marching algorithm results most likely be worse.

Table 1 summarizes the results. Our method reduced the VOE and standard deviation from 30 % (std = 13.3) for the level-set method to 14.9 % (std = 5.6), an improvement of 50.4 % (std = 57.5 %). It reduced the ASSD and standard deviation from 3.5 mm (std = 2.88) for the level-set method to 1 mm (std = 0.59), an improvement of 71.4 % (std = 79.55). The minimum and maximum values were also improved considerably. The minimum VOE was reduced from 14.86 to 6.32 %, an improvement of 57 %. The minimum ASSD was reduced from 0.69 to 0.17 mm, an improvement of 75 %. The maximum VOE was reduced from 44.2 to 23.25 %, an improvement of 47 %. The maximum ASSD was reduced from 8.54 to 2.46 mm, an improvement of 71 %. When we manually selected the best stopping threshold for each case, by comparing the result after each iteration to the ground truth, the fast marching method results were VOE of 26.2 % (std = 8.4) and ASSD of 1.4 mm (std = 1.6).

To quantify the contribution of the segmentation step, we compute the accuracy of the patient-specific prior (Table 2). The individual contribution of the segmentation step can be interpreted as the registration error in terms of the volume overlap error. The VOE and ASSD after step 1 are 35.8 % (std = 17.6 %) and 4.3 mm (std = 6.6) respectively. This is a good registration result but cannot serve as the final segmentation result since it is more than twice the error of the final segmentation.

To quantify the contribution of the patient-specific prior, we left out the baseline scan and tumor's delineations and performed segmentation alone with the prior as a sphere of radius 30 voxels centered at the center of mass of the tumor ground truth. The segmentation failed in eight out of 80 cases and yielded VOE and ASSD errors of 18.6 % (std = 7.3 %) and 1.27 mm (std = 0.9) respectively for the other 72 cases. This shows that our segmentation method can be used as a stand-alone application for lung tumor delineation, without the need for a baseline scan and/or baseline tumor delineation, possibly with reduced robustness and accuracy. For example, it can be used for the tumor delineation of the baseline scan. Figure 4 presents examples for the results on three tumors. For each tumor, the inputs are a baseline scan with radiologist-approved delineation (Fig. 4a–c) and a follow-up scan (Fig. 4d–f). The results are tumor segmentation on the follow-up scan (Fig. 4h–m). Notice the effect of the parabolic surface on the final tumor shape.

The main conclusion of our experimental results is that both the patient-specific baseline prior and the local deformable baseline tumor registration are key to achieving accuracy and robustness.

Conclusion

We have presented a new automatic lung tumor segmentation method for follow-up CT studies. The inputs to the method are baseline CT scan of the lungs with delineation of the tumor and a follow-up scan. A cascade of registration steps

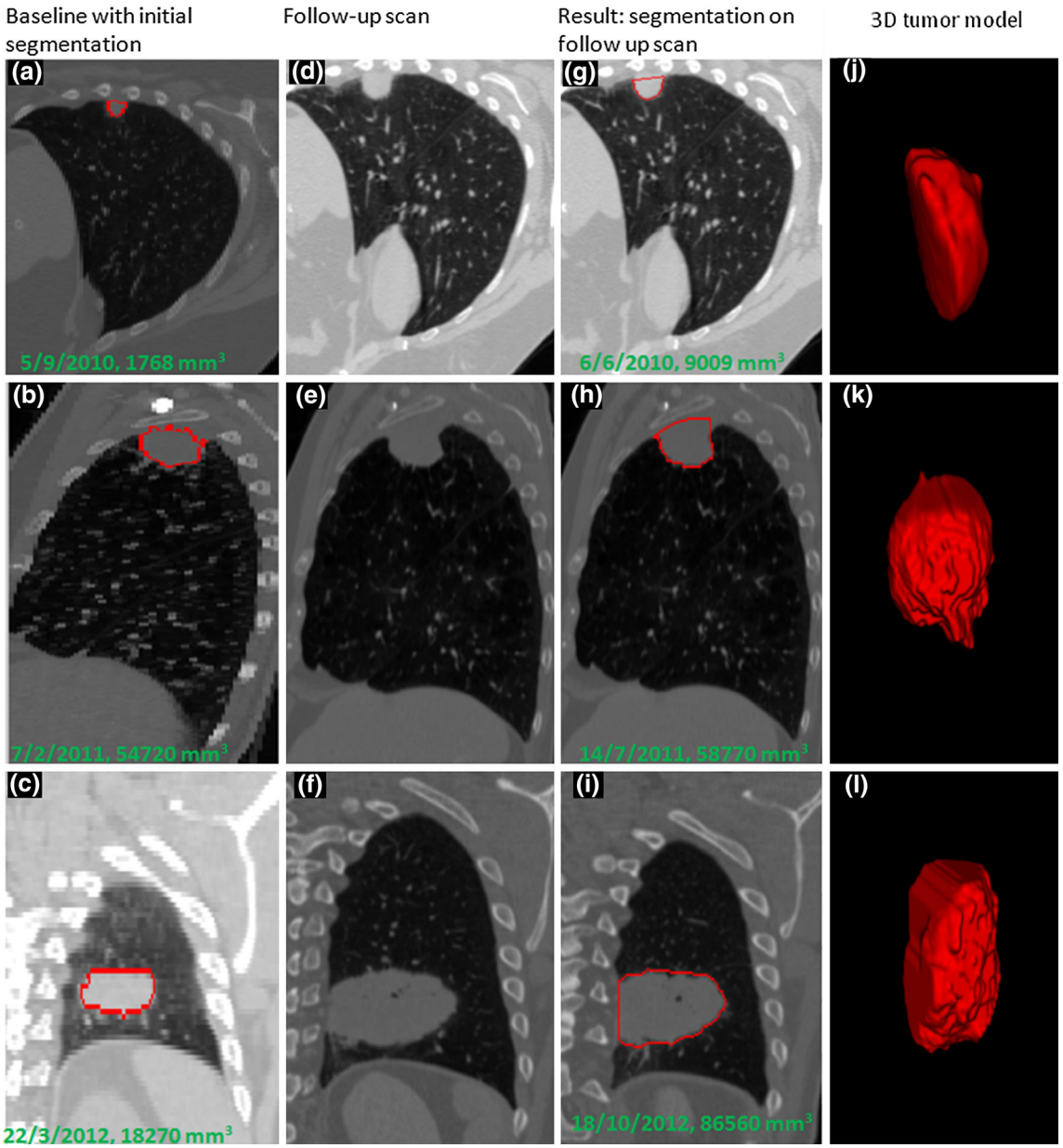

Fig. 4 Three examples of datasets with results. **a**–**c** Baseline scan showing the radiologist delineation; **d**–**f** follow-up scan; **g**, **h** follow-up scan showing the results of our method; **i**–**l** 3D model of the tumor

is used to transform the delineation into the follow-up scan, thereby creating a strong tumor prior. The initial follow-up tumor segmentation is then produced based on this prior by a statistical method. A two-stage automatic segmentation leaks detection and removal use geometrical models of the foreground and background. The final step reduces the tumor boundaries variability caused by the partial volume effect by variation analysis.

The key contribution of our work is in the use of a patient-specific model for the tumor segmentation prior. This improves the segmentation robustness by creating patient-specific statistical models of the tumors and the background. This observation is supported by the experimental study in which the segmentation step was used with a weak sphere shape prior. The failure rate was 10 % instead of 0 %, and the error rate was 25 % higher. Our registration method includes an additional tumor-specific local deformable registration step that refines the model prior. The segmentation leaks removal step relies on anatomic geometric constraints on the tumor and the adjacent structures. Note that this type of background geometric knowledge is difficult to integrate into classic active contours methods. The modeling of the adjacent structures as a parabolic surface simulates the way a human would complete the missing boundary between the

structures. The final step addresses the delineation variability caused by the PVE.

Our results on 40 pairs of CT scans, each used forward and backwards, show a significant improvement over the fast marching method and may provide relevant clinical measurements for lung tumors. We plan to apply the proposed method to other organ segmentations from various imaging modalities.

Acknowledgments This work was partially supported by KAMIN Grant 46217 from the Israeli Ministry of Trade and Industry.

Conflict of interest None of the authors has any conflict of interest. The authors have no personal financial or institutional interest in any of the materials, software, or devices described in this article.

References

1. Tuma SR (2006) Sometimes size does not matter: reevaluating RECIST and tumor response rate endpoints. J Natl Cancer Inst 98:1272–1274
2. Opfer R, Kabus S, Schneider T, Carlsen IC, Renisch S, Sabczynski J (2009) Follow-up segmentation of lung tumors in PET and CT data. SPIE, 72600X
3. Plajer IC, Richter D (2010) A new approach to model based active contours in lung tumor segmentation in 3D CT image data. In: Information Technology and Applications in Biomedicine (ITAB), 10th IEEE International Conference. IEEE, pp 1–4
4. Awad J, Owrangi A, Villemaire L, O'Riordan E, Parraga G (2012) A three-dimensional lung tumor segmentation from X-ray computed tomography using sparse field active models. Med Phys 39(2):851–865
5. Weizman L, Ben-Sira L, Joskowicz L, Precel R, Constantini S, Ben-Bashat D (2010) Automatic segmentation and components classification of optic pathway gliomas in MRI. In: Medical image computing and computer assisted intervention. Springer, Heidelberg, pp 103–111
6. Hollensen C, Cannon G, Cannon D, Bentzen S (2012) Lung tumor segmentation using electric flow lines for graph cuts. In: Image analysis recognition. Springer, Heidelberg, pp 206–213
7. Brown MS, McNitt-Gray MF, Goldin JG, Suh RD, Sayre JW, Aberle DR (2001) Patient-specific models for lung nodule detection and surveillance in CT images. Trans Med Imaging 20(12):1242–1250
8. Kuhnigk JM, Dicken V, Bornemann L, Wormanns D, Krass S, Peitgen HO (2004) Fast automated segmentation and reproducible volumetry of pulmonary metastases in CT-scans for therapy monitoring. In: Medical image computing and computer assisted intervention. Springer, Heidelberg, pp 933–941
9. Reeves A, Chan AB, Yankelevitz DF, Henschke CI, Kressler B, Kostis WJ (2006) On measuring the change in size of pulmonary nodules. Trans Med Imaging 25:435–450
10. Reeves AP, Jirapatnakul AC, Biancardi AM et al. (2009) The VOLCANO'09 Challenge: Preliminary Results. In: Brown M (ed) Proceedings of the 2nd International workshop on pulmonary image-analysis. CreateSpace Independent Publishing Platform, London, pp 353–364. http://www.lungworkshop.org/2009/proc2009/353.pdf
11. Kostis WJ, Reeves AP, Yankelevitz DF, Henschke CI (2003) Three-dimensional segmentation and growth-rate estimation of small pulmonary nodules in helical CT images. Trans Med Imaging 22(10):1259–1274
12. Jirapatnakul AC, Mulman YD, Reeves AP, Yankelevitz DF, Henschke CI (2011) Segmentation of juxtapleural pulmonary nodules using a robust surface estimate. J Biomed Imaging 1–14
13. Chen B, Hideto N, Yoshihiko N, Takayuki K, Daniel R, Hiroshi H, Hirotsugu T, Masaki M, Hiroshi N, Kensaku M (2011) Automatic segmentation and identification of solitary pulmonary nodules on follow-up CT Scans based on local intensity structure analysis and non-rigid image registration. SPIE, p 79630B
14. Gribben H, Miller P, Hanna GG, Carson KJ, Hounsell AR (2009) MAP-MRF segmentation of lung tumours in PET/CT images. Biomedical imaging: IEEE International Symposium, p 290–293
15. Kanakatte A, Gubbi J, Mani N, Kron T, Binns D (2007) A pilot study of automatic lung tumor segmentation from positron emission tomography images using standard uptake values. Computational intelligence in imaging signal processing, p 363–368
16. Gu Y, Kumar V, Hall LO, Goldgof DB, Li CY, Korn R, Bendtsen C, Velazquez ER, Dekker A, Aerts H, Lambin P, Li X, Tian J, Gatenby R, Gillies RJ (2013) Automated delineation of lung tumors from CT images using a single click ensemble segmentation approach. Pattern Recognit 46(3):692–702
17. Risser L, Baluwala H, Schnabel JA (2011) Diffeomorphic registration with sliding conditions: application to the registration of lungs CT images. MICCAI, 4th international workshop pulmonary image analysis, p 79–90
18. Gorbunova V, Durrleman S, Lo P, Pennec X, De Bruijne M (2010) Lung CT registration combining intensity, curves and surfaces. International symposium on biomedical imaging: From Nano to Macro, p 340–343
19. Murphy K, van Ginneken B, Reinhardt J, Kabus S, Ding K, Deng X, Cao K, Du K, Christensen G, Garcia V, Vercauteren T, Ayache N, Commowick O, Malandain G, Glocker B, Paragios N, Navab N (2011) Evaluation of registration methods on thoracic CT: the EMPIRE10. Chall Trans Med Imaging 30(11):1901–1920
20. Song G, Tustison N, Avants B, Gee JC (2010) Lung CT image registration using diffeomorphic transformation models. Medical image analysis for the clinic: a grand challenge, pp 23–32
21. Modat M, McClelland J, Ourselin S (2010) Lung registration using the NiftyReg package. Med Image Anal Clin:33–42
22. Kronman A, Joskowicz L, Sosna J (2012) Anatomical structures segmentation by spherical 3d ray casting and gradient domain editing. In: Medical image computing and computer assisted intervention:363–70
23. Tan Y, Schwartz LH, Zhao B (2013) Segmentation of lung lesions on CT scans using watershed, active contours, and Markov random field. Med Phys 40(4):043502
24. Klein S, Staring M, Murphy K, Viergever MA, Pluim JPW (2010) Elastix: a toolbox for intensity-based medical image registration. Trans Med Imaging 29(1):196–205
25. Bærentzen JA (2001) On the implementation of fast marching methods for 3D lattices. Math Model 13:1–19