Advanced Segmentation Techniques for Lung Nodules, Liver Metastases, and Enlarged Lymph Nodes in CT Scans

Jan Hendrik Moltz, Lars Bornemann, Jan-Martin Kuhnigk, Volker Dicken, Elena Peitgen, Stephan Meier, Hendrik Bolte, Michael Fabel, Hans-Christian Bauknecht, Markus Hittinger, Andreas Kießling, Michael Püsken, and Heinz-Otto Peitgen

Abstract—This article presents advanced algorithms for segmenting lung nodules, liver metastases, and enlarged lymph nodes in CT scans. Segmentation and volumetry are essential tasks of a software assistant for oncological therapy monitoring. Our methods are based on a hybrid algorithm originally developed for lung nodules that combines a threshold-based approach with model-based morphological processing. We propose extensions that deal with particular challenges of each lesion type: lung nodules that are attached to non-convex parts of the pleura, rim-enhancing and peripheral liver metastases and lymph nodes with an extensive contact to structures of similar density. We evaluated our methods on several hundred lesions in clinical datasets and the quality of segmentations was rated by radiologists. The results were classified as acceptable or better in 81% to 92% of the cases for the different algorithms and readers.

Index Terms—Biomedical image processing, computed tomography, image segmentation, liver, respiratory system, tumors.

I. INTRODUCTION

In oncological therapy monitoring, the estimation of tumor growth from consecutive CT scans is an important aspect in deciding whether the given treatment is adequate for the patient. Traditionally, this is done by measuring and comparing the largest axial diameter of each lesion manually. According to the current RECIST standard, an increase in diameter of at least 20% within three to six months indicates a progressive disease where a decrease of more than 30% is considered as partial response to therapy [1].

This approach implies several problems. First, manual examinations are always subjective, error-prone and time-consuming. Second and even more importantly, a 3-D quantity (volume) is estimated based on a 1-D measurement (diameter). This simplification would be valid if tumors were perfectly spherical and grew symmetrically but in practice it leads to inaccurate results. It should also be noted that a diameter increase of 20% means that the volume has already grown by 70% if a spherical shape is assumed. However, depending on the lesion size and the image resolution, a variation in diameter measurement of 20% may correspond to only a few voxels and is not unusual in practice. More details on the disadvantages of RECIST measurements can be found in [2].

Although volumetry has the potential to enhance the accuracy and reproducibility of growth estimation, measuring the lesion volume manually would take too much time in the workflow of a radiologist. This is the motivation for employing software assistants in oncological therapy monitoring since they are able to perform automatic volume measurements in 3-D. In order to be accepted in clinical routine, they have to work both fast and accurately. In particular, we found that the response time after marking a lesion should in general not exceed 3 s. An additional requirement is an option to change results manually since a software will never be able to guarantee correct findings for all cases.

In a completely automated workflow, three subtasks can be distinguished: detection, segmentation, and volumetry. This article exclusively deals with segmentation. The automatic detection of lung nodules has been subject to research for several years, recent publications being [3]–[5], but for liver metastases and especially lymph nodes, investigation has just begun and the clinical benefit is yet to be proved. In this work, we rely on an easy user interaction for locating lesions: drawing

1932-4553/$25.00 © 2009 IEEE
a stroke across the lesion. This is intuitive because it is similar to marking the largest diameter but it does not have to be exact. The second advantage is that it gives us information about the size and the gray value range of the tumor which can be used for segmentation.

Once a tumor has been segmented, a coarse volumetry can be achieved by simply counting voxels. More sophisticated techniques take the partial volume effect into account and weight surface voxels according to their gray values in relation to typical lesion and background values. We have already presented a solution for lung nodules [6], but transferring this principle to liver metastases and lymph nodes is not straightforward due to the more complex anatomical relations.

Lung nodules, liver metastases, and lymph nodes that are relevant for oncological therapy monitoring have some features in common: they are mostly homogeneous, roughly spherical or ellipsoid-shaped objects that are often located close to other structures that exhibit a similar density in CT images. This motivates using the same basic segmentation procedure for all three of them. We use a hybrid algorithm that was originally developed for lung nodules and combines a threshold-based approach with model-based morphological processing. It was described elaborately in [6] and will be reviewed briefly in Section II.

In order to cover the most common lesion types in oncological routine, we developed modified versions for liver metastases and enlarged lymph nodes and presented first results in [2]. In this article, we provide more details on these modifications and propose further extensions of the algorithm that deal with particular challenges of each lesion type: solid lung nodules that are attached to non-convex parts of the pleura (Section III), rim-enhancing and peripheral liver metastases (Section IV) and lymph nodes with an extensive contact to structures of similar density (Section V). In these three sections, we survey the state of the art, give in-depth descriptions of our methods, and present the results of evaluation studies with clinical datasets.

In most of these studies, we had the quality of the segmentation results ranked by two radiologists. Comparison with manual segmentation was only performed in a small study for liver metastases and presented first results in [2]. In this article, we provide more details on these modifications and propose further extensions of the algorithm that deal with particular challenges of each lesion type: solid lung nodules that are attached to non-convex parts of the pleura (Section III), rim-enhancing and peripheral liver metastases (Section IV) and lymph nodes with an extensive contact to structures of similar density (Section V). In these three sections, we survey the state of the art, give in-depth descriptions of our methods, and present the results of evaluation studies with clinical datasets.

Before the actual calculation is started a Gaussian smoothing filter is applied to the ROI if necessary. We estimate the noise in the image by computing the standard deviation in a small region outside the body. If this value exceeds a threshold, smoothing is enabled.

II. THE BASIC ALGORITHM: “SMART OPENING”

A. Initial Segmentation

The smart opening algorithm was developed by Kuhnigk et al. [6] for segmenting solid lung nodules. It proved useful for other tumor types as well because it tackles a problem that is common to all of them: the contact to vessels or other thin, elongated structures that feature a similar density in CT scans. It is obvious that a mere threshold-based method is not sufficient for segmentation in such a situation. Since the lesions are mostly homogeneous it can, however, be used as a first step to obtain a superset of voxels that may be part of the lesion. This can be implemented efficiently as a 3-D region growing starting from the center of the ROI. The thresholds can be fixed for lung nodules or determined adaptively from an analysis of the density distribution in the ROI (see Section IV-B for details).

For a typical lung nodule, the region growing result contains the complete lesion and additionally parts of the attached vasculature (Fig. 1). A morphological opening operation is an obvious choice to remove the vessels, but the challenge lies in determining the optimal erosion strength. Since tumors and their supplying vessels can differ in size significantly, an erosion with a fixed-size kernel cannot be used since this would either maintain thick vessels that should be removed or erode the lesion too strongly so that details of its boundary would be lost.

The idea of smart opening is to choose the erosion strength adaptively. To facilitate the computations, erosion is implemented by thresholding on a distance map that contains the distance of each mask voxel to the closest background voxel.
The underlying assumption is that all vessels are connected to the boundary of the ROI and that the diameter of a vessel decreases monotonically in its course. In order to disconnect the mask, all paths from the ROI center to the boundary are considered. The maximum of all minimum path diameters is the cut-off value that removes all vessels (Fig. 1).

The second step of the opening operation is a dilation that reconstructs the lesion in its original size without regrowing the vessels. The dilation is again implemented by thresholding on a distance map, but this time it shows the distances of all background voxels to the eroded mask. The threshold is chosen slightly higher than the erosion threshold so that in a final refinement step all boundary details can be reproduced by intersection with the region growing result.

It should be noted that this procedure alone is not suitable if more extensive connections to structures of similar density are present. Juxtapleural lung nodules, for instance, lie adjacent to the chest wall, the heart, or the diaphragm and are often hard to delineate. Peripheral liver metastases with contact to the intercostal musculature or conglomerates of lymph nodes pose similar problems. In cases like these, the necessary erosion threshold would delete the entire nodule or any irregularities in its shape. In order to overcome this restriction, this article presents preprocessing methods that separate broadly connected structures before the actual opening is performed.

B. Interactive Correction

An advantage of using the same basic algorithm for all lesion types is the possibility to offer a uniform concept for interactive correction of segmentation results. The smart opening algorithm contains an efficient and intuitive mechanism for this purpose. The user can change the erosion strength and thereby slightly alter the overall shape of the mask in cases where the underlying anatomical assumptions are not fulfilled or where the result is not optimal due to imaging artifacts. The possible values for the actual erosion threshold are mapped to the range from 0% to 100%. After a change, the segmentation is partly recomputed and the result is available almost immediately. This concept improves the flexibility of the algorithm while preserving reproducibility. The procedure is useful for all three lesion types considered here even though it does not cover all cases where the user might want to improve the result. This kind of interactive correction was allowed in all evaluation studies mentioned in this article.

III. SEGMENTATION OF JUXTAGLEURAL LUNG NODULES

Lung nodules are mostly located centrally within the lung parenchyma, but they can also be attached to the pleura, a thin membrane that covers the lungs. In CT scans of the thorax, the voxels can basically be divided into two classes: while the dark ones represent the lung parenchyma, bright voxels may be nodules, but also blood vessels or structures adjacent to the lungs such as the chest wall, the heart, or the diaphragm [Fig. 2(a)]. Since the pleura itself is invisible in CT images the boundary between a juxtapleural nodule and any of these structures shows little or no contrast and it is sometimes impossible even for a radiologist to determine the exact boundary of a nodule.

Fig. 2. Step-by-step illustration of the segmentation algorithm for juxtapleural lung nodules, exemplified by the central axial slice. (a) Result mask of the initial region growing and rays cast towards the boundary of the mask. (b) Valid ray endpoints on the nodule boundary. (c) Ellipsoid fitted to the boundary points. (d) Part of the inverse mask within the dilated ellipsoid. (e) Convex hull of the inverse mask within the dilated ellipsoid. (f) Difference of convex hull and inverse mask.

A. State of the Art

Several authors that worked on solid lung nodule segmentation have also proposed solutions for juxtapleural nodules. An obvious idea is to compute a lung segmentation in order to separate the nodule from structures outside the lungs as it is done by Fettita et al. [11]. However, we decided not to incorporate any global information in order to keep computation times as low as possible and to be able to integrate the algorithm into an existing workstation. Van Ginneken [3] uses a local 2-D lung field segmentation but no evaluation for juxtapleural nodules is given.

One of the first dedicated segmentation algorithms for juxtapleural nodules was presented by Shen et al. [12]. Assuming that the chest wall is physiologically smooth and that a nodule creates a “bump” with a high local curvature, the nodule can be separated by smoothing the wall surface. This is implemented by projecting the surface to a plane whose normal is the mean of all surface voxel normals. On the projection image, the nodule appears as a region of high values which are replaced by a cubic polynomial interpolation of the other values. The smoothed 3-D surface is then computed via backprojection. This idea is promising but there will be problems when the wall itself has points of high curvature as in Fig. 3(e).
In our previous article [6] we presented a preliminary solution as well. To separate a nodule from the chest wall, the algorithm makes use of the fact that the lungs are convex in most parts and that juxtapleural nodules create a concavity in this shape. The idea is to reconstruct the shape in the tumor-free state by computing the convex hull of the lung parenchyma within the ROI and cut off the nodule along the boundary of the convex hull. However, in regions where the lungs are not convex, such as the boundaries to the heart or the diaphragm, the convex hull does not remove the attached structure completely [Fig. 3(a), (c), and (e)]. In this contribution, we present an improvement to the algorithm that can handle this case as well. This work has previously been published in German [17].

B. Method

The goal of our extension of the original algorithm [6] was to improve the segmentation of solid nodules located at concave parts of the pleura while changing the original method as little as possible in order to get consistent results. The convex hull operation is obviously not suitable for reconstructing the shape of a concave object but we observed that the error decreases when the ROI is made smaller since the convex hull is basically determined by the most distant points of the lung boundary that are contained in the ROI. Therefore our approach is to make the ROI as small as possible so that the disturbing effect of the concavity is minimized. As a minimal ROI, we choose a dilated ellipsoid that is computed as an approximation of the nodule shape. Ellipsoid approximation of lung nodules has been used with different goals and methods in the literature [15], [18].

Our method is a preprocessing step for the smart opening algorithm [6] and consists of three parts which are described in the following sections and illustrated in Fig. 2:

1) identification of points on the nodule boundary by region growing and subsequent ray casting from the seed point;
2) calculation of an ellipsoid that approximates the shape of the nodule;
3) convex hull operation as in [6], but restricted to the dilated ellipsoid.

1) Region Growing and Ray Casting: Initially, region growing is performed, using the ROI center as a seed point. Since we only need to separate the solid nodule and attached high-density structures from the lung parenchyma in this first step, we can use \(-400\) HU as a fixed threshold. Non-solid nodules are not targeted by our algorithm. In order to find points on the boundary between the nodule and the parenchyma, we apply a ray casting approach. Starting from the seed point, rays are sent out through all surface voxels of a \(5 \times 5 \times 5\) cube around the seed point. This ensures a symmetric distribution of rays and an alignment to the voxel grid. These \(5^3 - 3^3 = 96\) rays are traced until they reach either the boundary of the region growing mask or leave the ROI [Fig. 2(a)]. In the former case, the endpoints are stored, otherwise discarded [Fig. 2(b)]. Since some false boundary points may be found due to noise or other structures in the outer parts of the ROI, ray endpoints above a certain distance from the seed point should also be discarded. We found the 95% quantile of the distances of all points to provide a good threshold.

An algorithm proposed by Wiemker et al. [13] starts with region growing and determines the optimal cutoff value retrospectively by means of an objective function that separates the wall at its strongest inflection. Unfortunately, no results for juxtapleural nodules are shown.

Reeves et al. [14] approximate the pleural surface by a clipping plane that is iteratively refined until a leap in volume change is observed when it actually reaches the pleura. This works for small nodules where the actual convex or concave shape of the lungs can be ignored in a local view, but in other cases a plane is not suitable and the algorithm will fail.

Okada et al. [15] presented an approach that uses morphological opening similar to ours, but applied it in an inverse way: the size of the structure element is chosen such that the nodule is removed and the adjacent structure is retained. This assumes, however, that in the ROI the wall region is significantly larger than the nodule. Therefore, the approach will fail when the nodule is large or when the wall has a concave shape. Furthermore, it is implemented such that segmenting a juxtapleural nodule takes more than twice as long as for a central one.

In a recent publication, Dehmeshki et al. [16] describe an algorithm that uses sphericity-constrained region growing on a fuzzy connectedness map. Successful results are reported for nodules “very close to lung wall or diaphragm” but the method always needs a visible contrast between the nodule and the adjacent structure.
2) Ellipsoid Approximation: Typically, the points found by the ray casting procedure cover a major part of the actual nodule surface. Assuming that the nodule has approximately an ellipsoid shape, we aim at reconstructing this shape by fitting an ellipsoid to the points [Fig. 2(c)].

A 3-D ellipsoid is defined as a conic section
\[ \{ x \in \mathbb{R}^3 | x^T A x + b^T x + c = 0 \} \]
where the symmetric matrix \( A \in \mathbb{R}^{3 \times 3} \) is positive or negative definite. Due to its symmetry, \( A \) has only six degrees of freedom, plus a total of four for \( b \in \mathbb{R}^3 \) and \( c \in \mathbb{R} \). From the valid endpoints of the 98 rays we want to determine those ellipsoid parameters which are optimal in a least squares sense. This establishes a nonlinear equation system which, however, can be reduced to a generalized eigenvalue problem and solved efficiently with a method proposed by Grammalidis and Strintzis [19]. It does not guarantee \( A \) to be definite, but our experiments showed that this is almost always the case. If the points are distributed in a way such that it is not possible to fit an ellipsoid to them—if the nodule, for example, has an irregular shape or very extensive contact to other structures—a sphere can be computed instead with the radius as the only free parameter. Although this is a coarser approximation, it can still yield acceptable results in most of these rare cases. It should be noted that the center of the ellipsoid is included in the optimization. The user-defined seed point influences only the distribution of the boundary points. Since the equation system is highly overdetermined it is robust against variations caused by different user interactions.

3) Convex Hull: For the following computations we use a dilated version of the ellipsoid as a new minimal ROI [Fig. 2(d)]. The dilation strength is chosen such that all valid ray endpoints are enclosed in order to ensure that the nodule is contained completely. At its margin, the ellipsoid contains some parenchyma voxels as well, so the convex hull operation can now be applied to reconstruct the original lung shape within this ROI [Fig. 2(e)]. This is sufficient for determining the boundary of the nodule and it works in concave parts as well because these concavities are no longer visible inside the ellipsoid. For performance reasons, the convex hull has been implemented as the union of slice-wise convex hulls in axial, sagittal, and coronal views.

Subsequently, the original algorithm [6] is continued and the adaptive opening procedure removes attached vessels. The final result is shown in Fig. 2(f).

C. Results and Discussion

For our evaluation, we used a database of 333 ROIs of juxta-pleural solid nodules from various patients, clinics and CT scanners with seed points set manually by radiologists. Since extensive studies have been conducted for the original version [6], we focused on the effects of the extensions presented above. It is often impossible to determine the exact boundary between a nodule and a structure it is attached to. For lack of a reliable ground truth, we evaluated the segmentation results visually and examined if they were consistent with our approach to reconstruct the tumor-free shape of the lung parenchyma.

While in 71% of the cases the result of the original algorithm was classified as good, our extension could increase this portion to 89%. For an additional 5%, a good result was obtained after interactive correction. It should be noted that all of the nodules in this study had contact to the pleura and constituted a sample of difficult cases. Most of the nodules that could not be segmented had complex shapes or a very extensive connection to high-density structures that made it impossible to fit ellipsoids to them. Fig. 3 shows some examples of successful segmentations and reveals a significant improvement over the previous results.

In a study recently published by Vogel et al. [8], our algorithm was evaluated on 101 lung nodules from 28 patient scans that were reconstructed with slice thicknesses of 1, 3, and 5 mm. The fraction of nodules with contact to the pleura was 40% and the material was described as challenging by the authors. The quality of the segmentation was rated on a five-point scale by a single reader. In 88% of the cases, the results were classified as acceptable or better and in 82% as good or very good, almost independently of the slice thickness. The study also showed the superiority of volumetry over diameter measurement: the 95% limit of agreement was approximately ±0.66 mm for the effective diameter of the automatically computed volume compared to ±1.3 mm for the RECIST diameter.

IV. Segmentation of Liver Metastases

With respect to segmentation, the main difference between lung nodules and liver metastases is that the latter are much more diverse in their appearance. Lung nodules have a high contrast to the surrounding parenchyma and the density of both structures is well known so that fixed thresholds can be used for the initial region growing. The gray values of parenchyma and metastases in the liver, on the other hand, depend on primary cancer, contrast agent, contrast timing, scan parameters, and patient conditions. Lesions can be brighter (hyperdense) or darker (hypodense) than the surrounding tissue or they can appear inhomogeneous if they have a contrast-enhanced rim around them or if they are partly necrotic or calcified.

A. State of the Art

The segmentation of liver metastases, unlike that of lung nodules, has not been an area of intensive research so far. Few algorithms have been designed for 3-D CT data. Lu et al. [20] use 2-D active contours that are initialized manually. This is relatively slow and not very accurate. Other authors focus on the detection of metastases in a segmented liver and apply rather simple methods for lesion segmentation. Park et al. [21] and Ciecholewski and Ogiela [22] both try to find optimal thresholds based on statistical analysis of the histogram under the liver mask, but this will fail on poorly contrasted or inhomogeneous lesions. Li et al. [23] apply a machine learning technique to classify possible boundary positions on 1-D intensity profiles. Again, poor contrasts and especially lesions with contact to a structure of the same density will pose problems.

One of the most promising methods was proposed by Li and Jolly [24]. They use a graph-theoretic approach, representing voxels by vertices and surfaces by paths along the edges. The optimal path minimizes an objective function that incorporates boundary, regional and elasticity constraints. Notably, the algorithm is able to detect multiple surfaces simultaneously, making
it possible to segment tumors with necroses and calcifications in one pass. Its rather abstract nature, however, makes it difficult to integrate an intuitive interactive correction procedure.

Another interesting approach was presented by Jolly and Grady [25]. Their algorithm is suitable for different kinds of lesions including liver metastases. It learns the gray value distribution of a tumor from user interaction and several 2-D segmentations on orthogonal planes and finally applies the random walker algorithm. Due to its general character, the method lacks a handling of typical difficult cases such as liver metastases or lymph nodes adjoining an isodense structure.

A rough description and a preliminary evaluation of our method was given in [2].

B. General Method

In order to make the smart opening algorithm applicable for liver metastases, some changes have to be made. Due to the high diversity in their appearance, the thresholds for the initial region growing have to be determined adaptively. This is based on an analysis of the density distribution in the ROI. Information given by the user is reflected in the size and center of the ROI. Inhomogeneous metastases can be detected using the stroke that the user is expected to draw across the lesion. The density profile under the stroke provides information about the "relative density" of the lesion compared to the parenchyma.

Since thresholds depending on the user input are always a limitation for reproducibility we tried to incorporate medical knowledge about CT imaging and use fixed thresholds and constants wherever possible. In particular, we assume that the difference between the typical lesion and parenchyma values should be at least 25 HU, based on the typical enhancement of 50 HU.

As a first step, a typical parenchyma value is estimated which allows us to decide whether the lesion is hypodense, hyperdense or inhomogeneous. Since the size of the ROI is chosen as a multiple of the stroke length, we can assume that the largest part of the ROI contains healthy liver parenchyma and that the maximum peak of the ROI histogram represents a typical parenchyma density $P$. There are two cases where this can lead to a wrong conclusion. First, for metastases close to the liver surface, the ROI does not only contain liver tissue and the maximum peak may correspond to a structure that is actually outside the liver. To prevent this, the histogram is restricted to voxels with at least 40 HU, so as to exclude less dense materials such as fluids or air. Second, if several metastases lie close to each other, they can contribute to the maximum peak rather than the parenchyma. This can easily be detected and is handled in the following.

Next, a histogram of the values under the stroke is computed. The stroke is dilated with a $3 \times 3 \times 3$ kernel without elongating it in order to get a larger amount of representative values [Fig. 4(a)]. From the histogram, we compute the ratio of voxels covered by the dilated stroke that have a density less than $P$. This ratio is a measure of the "hypodensity" $H$ of the lesion compared to the healthy parenchyma. A value close to 100% indicates that almost all voxels under the stroke are below $P$ and that the lesion is hypodense. If $H$ is near 0%, we have a hyperdense lesion. In order to take inaccurately drawn strokes and noise into account, we set thresholds at 20% and 80% and regard lesions with $H$ between these values as inhomogeneous [Fig. 4(b)]. This case is handled in Section IV-C. Otherwise we compute a Gaussian-weighted mean value in a small area around the center of the ROI and store it as the typical lesion value $L$.

Given $P$ and $L$, the thresholds $T_0$ and $T_1$ for region growing are determined as follows (all values in HU).

- Hypodense lesions: $T_0 = 10$, $T_1 = (P + L)/2$.
- Hyperdense lesions: $T_0 = (P + L)/2$, $T_1 = 180$.

The "inner" thresholds are set to the average of $P$ and $L$, the "outer" thresholds to constant values in order to remove materials that can be excluded a priori such as fat, air, or bones.

If $L$ and $P$ have a distance of less than 25 HU, we suspect that $P$ is not representative of the parenchyma because it is not the largest region in the ROI. In this case, we try to find a better estimate than $P$ by searching for a quantile of the ROI histogram that has a reasonable distance to $L$.

- Ambiguous cases with $L < Q_{25}$: $T_0 = L - 25$

$$T_1 = \begin{cases} 
(L + Q_{25})/2, & \text{if } L < Q_{25} - 25 \\
(L + Q_{50})/2, & \text{if } L < Q_{50} - 25 \\
(Q_{50} + Q_{75})/2, & \text{otherwise}.
\end{cases}$$

- Ambiguous cases with $L > Q_{50}$: analogously.

where $Q_n$ denotes the $n\%$ quantile of the restricted ROI histogram. An example of such a case is shown in Fig. 6(h). Here,
the only peak in the ROI histogram corresponds to the lesion value and the parenchyma is not distinguishable. Using $Q_{50}$ instead of $P$ yielded a successful segmentation.

Finally, we enlarge the threshold range so that most of the values under the stroke are covered but a safety distance to $P$ is preserved

$$T_0 = \min(T_0, \max(S_{10}, P - 10))$$
$$T_1 = \max(T_1, \min(S_{10}, P + 10))$$

where $S_{10}$ is the $10\%$ quantile of the stroke histogram. For ambiguous cases, $P$ is replaced by $Q_{50}$. This is especially useful for slightly inhomogeneous lesions such as hyperdense metastases with calcifications where the calcified parts would not be covered by the initial threshold range [Fig. 6(d)]. We are aware that this increases variability but on the other hand we want to use the information given by the user-defined stroke. Our experiments confirmed that this is a reasonable compromise that improves the overall quality of the results.

Now the smart opening procedure is started with the thresholds $T_0$ and $T_1$.

C. Segmentation of Rim-Enhancing Lesions

In most cases, metastases identified as inhomogeneous have a hypodense inner part and a hyperdense outer part. These are usually hypodense lesions with a contrast-enhanced rim but could also be hyperdense lesions with a necrotic core. In a first step, we segment the hypodense part only [Fig. 4(c)]. Since the ROI

Fig. 6. Example results of the liver metastasis segmentation algorithm. (a) Isolated hypodense lesion. (b) Isolated hyperdense lesion. (c) Hypodense lesion at the liver surface. (d) Hyperdense lesion with calcifications. (e) Hyperdense lesion with a hypodense necrotic core. (f) Hyperdense lesion with a hyperdense enhanced rim and extensive vascularization. (g) Peripheral hyperdense lesion with an incomplete hyperdense rim. (h) Hyperdense lesion surrounded by other hypodense structures; a case where the parenchyma is not well represented in the ROI histogram. (i) Irregular inhomogeneous lesion where the algorithm for rim-enhancing metastases fails. (j) Large peripheral tumor where the liver shape estimation fails and a part of the lesion is missing.
center might not be a suitable seed point for this purpose, we shift it to the closest voxel that is clearly hypodense ($< P - 25$) and change $L$ accordingly. The thresholds are then determined as above.

After the first run of smart opening, a post-processing step is added. The basic idea is to fill the segmented inner part with a typical value of the outer part so that a “virtual” hyperdense lesion is created. Due to the partial volume effect, there is a narrow zone of voxels between the originally hypodense and hyperdense parts of the lesion that have a density similar to the parenchyma. In order to analyze the outer part of the lesion but exclude this partial volume zone, the current segmentation mask is dilated twice with a $5 \times 5 \times 5$ kernel and a histogram is computed of the ring-shaped area that was added by the second dilation. This is supposed to be a representative sample of the hyperdense part. Let $R_0$ denote the $n\%$ quantile of this histogram. Then we use the fill value $F = \min(R_0, 180)$ where 180 HU is a threshold to exclude bones that are accidentally covered by the histogram (Fig. 4(d)). Using the 95% quantile we can be sure to have a representative value of the hyperdense rim because the histogram can be distorted by the partial volume zone.

If only the result mask of the first segmentation is filled, the partial volume zone remains. Obviously, it cannot be included in the threshold range since then the parenchyma would be covered as well. On the other hand, filling a dilated version of the mask can deteriorate the result because the hyperdense part does not always enclose the hypodense part completely. As a compromise, we use a $5 \times 5 \times 5$ dilated mask but remove voxels with a density less than $T_0$ of the first segmentation or greater than $F$. This is useful for metastases situated at the liver surface that have only an incomplete hyperdense rim because the histogram can be distorted by the partial volume zone.

Interactive correction was performed for 36% of the lesions. The fact that the special handling for inhomogeneous lesions is triggered by an analysis of the stroke histogram allows the user to decide whether the hyperdense part should be segmented without requiring any additional interaction. If the stroke is drawn across the hypodense part only, the postprocessing step is omitted. We found that often radiological expertise is needed to make this decision. Furthermore, it is necessary in order to make follow-up measurements consistent.

There are other forms of inhomogeneous liver metastases that do not have such a regular structure. These cases demand a different approach and cannot yet be handled by our algorithm (Fig. 6(i)).

It should be noted that this procedure is robust in cases where a rim was detected but is not actually present. Due to the sophisticated threshold determination, the second segmentation causes no significant changes to the mask. Therefore the thresholds for the “hypodensity” $H$ are not critical.

### D. Segmentation of Peripheral Lesions

Peripheral liver metastases pose a similar problem as juxtpleural lung nodules. These lesions often have the same density as intercostal muscles or the parenchyma of other abdominal organs such as the kidneys or the stomach and are sometimes hard to delineate even visually (Fig. 5(a)). In such cases, the algorithm presented so far has a tendency to create a roundish segmentation that leaks into the adjacent structure. Obviously, an exact liver mask would avoid this problem but as with lung nodules we decided to get all our information exclusively from the ROI.

Instead, we try to estimate the boundary of the liver geometrically within the ROI, making use of the fact that the liver is almost perfectly convex. If a tumor is not too large and not situated in an area where the liver contour has a very high curvature, it is possible to reconstruct the liver shape by computing a convex hull of the parenchyma contained in the ROI. A coarse parenchyma mask can be obtained by thresholding with $[P - 20, P + 20]$ where $P$ is the estimated typical parenchyma value as in Section IV-B (Fig. 5(b)). Since this mask can include other structures, notably parts of the ribs, it is opened with a $5 \times 5 \times 5$ kernel and the largest connected component is chosen (Fig. 5(c)).

Since a convex hull computation in 3-D is very expensive, we compute the 2-D convex hulls on all axial, sagittal, and coronal slices and use the union of all voxels contained in any of these convex hulls as an approximation (Fig. 5(d)). This is a subset of the actual 3-D convex hull and thus not necessarily convex. The approximation can be refined by iterating the process but for our purposes this was not necessary. The result of this procedure is used to mask the ROI for the algorithm of Section IV-B (Fig. 5(e)).

### E. Results and Discussion

We evaluated the segmentation algorithm in a study with contrast-enhanced CT datasets from 48 patients with liver metastases caused by different primary tumors. The data was acquired with two different scanners and had a reconstruction increment of 3 mm. For each patient, two scans performed within an interval of about three months were available. The lesions in both baseline and follow-up scans were segmented by two radiologists and the results were rated on a five-point scale independently by both readers. The total number of lesions was 246. At the time of the study, the automatic detection of inhomogeneous metastases had not been implemented yet. Instead, the readers could start the filling and re-segmenting procedure manually. The special handling for peripheral lesions was not included either.

The results of the study are summarized in Table I. In 92% and 81%, the results were classified as acceptable or better (0, $+\alpha$, or $++$) by the first and second reader, respectively. A good or very good rating was given in 84% (52%) of the cases. Interactive correction was performed for 36% of the lesions.

<table>
<thead>
<tr>
<th>TABLE I</th>
<th>RATINGS FOR LIVER METASTASIS SEGMENTATION</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>++</td>
</tr>
<tr>
<td>Reader 1</td>
<td>125</td>
</tr>
<tr>
<td>Reader 2</td>
<td>4</td>
</tr>
<tr>
<td></td>
<td>129</td>
</tr>
<tr>
<td></td>
<td>27%</td>
</tr>
</tbody>
</table>

Authorized licensed use limited to: IEEE Xplore. Downloaded on February 24, 2009 at 04:46 from IEEE Xplore. Restrictions apply.
Even though the ratings of the two readers differ considerably, it can be concluded that for a major part of the lesions, segmentation was successful. Fig. 6(a)–(h) shows examples of successful segmentations for various types of liver metastases.

In a second study, we evaluated the extensions for rim-enhancing and peripheral metastases. We collected 151 lesions with enhanced rims or necroses and 144 lesions with contact to structures outside the liver from our database, covering various patients, clinics and scanners. Each tumor was randomly assigned to two of the six radiologists participating in the study and segmentation quality was rated as before.

The comprehensive results for all six readers are shown in Table II. For rim-enhancing metastases, the segmentation was classified as acceptable or better in 81% of the cases and as good or very good in 66%. For peripheral lesions, the respective numbers are 87% and 77%. Many of the tumors that could not be segmented were described as very difficult or not relevant for clinical practice by the radiologists. These cases were often confluent or the delineation to an adjacent structure was ambiguous even visually. For large tumors, the convex hull is sometimes not suitable for estimating the liver shape. In these cases, parts of the tumor are cut off [Fig. 6(j)].

We participated in the MICCAI Liver Tumor Segmentation Challenge 2008 [26] with our algorithm. The detailed results on the testing data can be found in [27]. Ten tumors from six datasets were segmented and the results were compared to a manual reference segmentation. In total, we achieved a volume overlap error of 31% and an average symmetric absolute surface distance of 1.6 mm. The respective values achieved by automatic segmentation was 13% and 0.4 mm. With these results, we achieved the highest score among nine competing teams.

A preliminary version of the algorithm was used by Heußel et al. [9] in a study that compares volumetry to the standard diameter measurements. Two radiologists analyzed 198 lesions in consecutive CT scans of 82 patients. The authors conclude that a reliable tumor response evaluation is only possible with volumetric analysis. Unfortunately, no explicit evaluation of the segmentation quality is given.

V. SEGMENTATION OF ENLARGED LYMPH NODES

Even more than lung nodules and liver metastases, lymph nodes often exhibit an extensive connection to muscles, vessels or other structures of similar density. Sometimes two or more lymph nodes conglomerate but should be segmented individually. Therefore lymph node segmentation is widely considered a particular challenge and has not been an object of intensive research so far.

### Table II

<table>
<thead>
<tr>
<th></th>
<th>++</th>
<th>+</th>
<th>0</th>
<th>-</th>
<th>--</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Rim-enhancing</td>
<td>31</td>
<td>69</td>
<td>24</td>
<td>20</td>
<td>7</td>
<td>151</td>
</tr>
<tr>
<td>(all readers)</td>
<td>21%</td>
<td>46%</td>
<td>16%</td>
<td>13%</td>
<td>5%</td>
<td></td>
</tr>
<tr>
<td>Peripheral</td>
<td>49</td>
<td>62</td>
<td>14</td>
<td>12</td>
<td>7</td>
<td>144</td>
</tr>
<tr>
<td>(all readers)</td>
<td>34%</td>
<td>43%</td>
<td>10%</td>
<td>8%</td>
<td>5%</td>
<td></td>
</tr>
</tbody>
</table>

A. State of the Art

One of the first attempts was a 3-D active surface approach by Honea et al. [28], based on image gradients and shape constraints. Unfortunately, no evaluation on clinical data was conducted. Dornheim et al. [29] proposed a solution with stable 3-D mass-spring models that additionally incorporates information on the characteristic gray value range. It was specifically designed for neck lymph nodes and showed promising results, but it was not evaluated for other anatomical locations. Yan et al. [30] use a marker-based watershed transform. In contrast to our approach, they determine markers separately on each slice, apply a 2-D watershed transform, and propagate the result to the next slice in order to find new markers.

B. Method

The smart opening algorithm was designed for eliminating thin structures, but it is not suitable for handling extensive connections between tissues of similar density. Therefore, a preprocessing step was added to handle this problem. In order to combine shape knowledge with a powerful grayvalue-based segmentation method, we chose the interactive watershed transform (IWT) which was introduced by Hahn and Peitgen [31] and has already been used for a variety of medical imaging problems such as brain and bone segmentation. In IWT, markers can be set at arbitrary image positions to influence the basin-merging process. If two adjacent basins contain different markers, they will not be merged even if there is only a low ridge between them. On the other hand, basins with the same markers will never be separated. These markers can be set by the user and processed in real time. Since we are, however, interested in reducing user effort and increasing reproducibility, we decided to place the markers automatically based on a coarse estimation of the shape of the lymph node.

1) Ellipsoid Approximation: Knowing that lymph nodes have a roughly ellipsoidal shape, we can apply an ellipsoid approximation in a similar way as with lung nodules. Lymph nodes vary in their average gray value but unlike liver metastases, they are mostly homogenous. We found a threshold range from $m - 65$ to $m + 65$ to be suitable for the initial region growing where $m$ is the mean value in the vicinity of the seed point [Fig. 7(b)]. Based on the resulting mask, the ray casting and the computation of the ellipsoid parameters are analogous to the description in Section III-B [Fig. 7(c)].

2) Interactive Watershed Transform: The watershed transform is a powerful tool for segmenting images based solely on their values, without incorporating any *a priori* knowledge. In our use case, however, we do have such knowledge: First, we are given a seed point by the user and we are only interested in a binary segmentation. Second, from the density at the seed point and our homogeneity assumption we can determine a range of values that is likely to cover the lymph node. In order to combine this knowledge with the capabilities of the watershed transform, we do not apply the transform directly on the image, but on the Euclidean distance map of the region growing mask obtained with the thresholds given above. Thereby, we discard the gray value information of the image and reduce it to a binary representation. Performing a watershed transform on a distance map is a common procedure to separate connected objects in binary
images. Theoretically, local maxima of the distance map correspond to object centers. We place markers at those local maxima to guide the segmentation.

Since our segmentation problem is binary, we use only two types of markers which will be called include and exclude markers. Include markers are positioned at the seed point and at all local maxima inside the ellipsoid. Exclude markers are placed at all local maxima outside the dilated ellipsoid and at several points on the ROI boundary [Fig. 7(d)]. The output of the watershed transform is a mask that contains the lymph node and possibly vessels but no larger adjacent structures [Fig. 7(e)]. The ROI is then masked with the watershed result and smart opening is performed to get the final segmentation mask [Fig. 7(f)].

C. Results and Discussion

Our algorithm was evaluated in a study with CT datasets from 50 melanoma patients with 222 enlarged lymph nodes in the chest, abdomen and pelvis. For some of the patients, follow-up data was available as well. The reconstruction increment was 1 mm. The lymph nodes were segmented independently by two radiologists and the quality of the segmentation results was rated on a five-point scale.

![Fig. 7. Step-by-step illustration of the segmentation algorithm for lymph nodes. (a) Original ROI, showing an abdominal lymph node between two vessels. (b) Result of the initial region growing. (c) Ellipsoid approximating the shape of the lymph node. (d) Distance map of the region growing mask with include (+) and exclude (×) markers. (e) Result of the watershed transform. (f) Result of the smart opening.](image)

![Fig. 8. Example results of the lymph node segmentation algorithm at different anatomical locations: (a) axillary, (b) mediastinal, (c) abdominal, (d) inguinal.](image)

Table III summarizes the results of the study. In 87% and 86% of the cases, the result was classified as acceptable by the first and second reader, respectively. Good or very good results were obtained in 86% (68%). Interactive correction was performed for 31% of the lymph nodes. In general, the algorithm performed better on small lymph nodes (diameter <1 cm) in the axillary and inguinal regions which are often surrounded by fat and are therefore easier to separate. Larger lymph nodes (diameter >2 cm), especially in the iliac and mesenterial regions, caused more problems. This is no surprise since they tend to have more complex connections to similar structures. Results were slightly better on the baseline datasets because lymph nodes growing over time become more difficult to delineate. About half of the lesions with poor rating were classified as very difficult or unsegmentable by the readers. Fig. 8 shows examples of successful segmentations for lymph nodes at different anatomical locations.

![Figure 8](image)

**TABLE III**

| Ratings for Lymph Node Segmentation |
|-------------------------------|---|---|---|---|---|
| Reader 1 | ++ | + | 0 | - | -- |
| 213 | 124 | 2 | 22 | 29 | 390 |
| Reader 2 | ++ | + | 0 | - | -- |
| 99 | 155 | 67 | 17 | 34 | 372 |

| Reader 1 | 312 | 279 | 69 | 39 | 63 | 762 |
| Reader 2 | 41% | 37% | 9% | 5% | 8% | |

A similar study without follow-up data and with an earlier version of the algorithm was conducted by Fabel et al. [10]. This study compared automatic volumetry results to RECIST diameters and manual volumes. The 95% limits of agreement between the two observers were [−4.04 mm, 5.84 mm] (effective diameter of automatic volumetry), [−1.75 mm, 2.87 mm] (effective diameter of manual volumetry) and [−1.54 mm, 1.94 mm] (RECIST), respectively. This shows that the inter-observer variability is reduced considerably if software-assisted volumetry
is used. Furthermore, the time needed for a complete examination of a patient was comparable for RECIST and automatic volumetry.

VI. CONCLUSION AND OUTLOOK

In this article, we reviewed the smart opening algorithm for segmenting lung nodules and presented extensions to handle other tumor entities and more complex cases. We proposed a ray-casting-based ellipsoid approximation for juxtapleural lung nodules, showed a threshold determination scheme for liver metastases based on histogram analysis, presented a filling and re-segmenting procedure for rim-enhancing lesions in the liver as well as a local liver shape estimation for peripheral metastases and used a watershed transform as a preprocessing step for lymph nodes with extensive connection to structures of similar density. Thereby, we covered the three lesion types that are by far the most frequent ones in oncological routine. We performed evaluation studies for all methods and radiologists considered the segmentation results as successful in most of the cases (between 81% and 92% for the different algorithms and readers). In both the liver metastasis and the lymph node segmentation, many of the lesions with poor segmentation ratings were classified as hard to delineate even visually and as not relevant for volumetry in the clinical routine. Furthermore, the algorithms are fast (runtime 1–3 s per lesion without interactive correction) and require only a stroke across the lesion from the user.

For future work, we identified complex inhomogeneous lesions as a typical problem case. They can be partly necrotic, calcified or fattened liver metastases or lymph nodes. In these cases, at least two disjoint threshold ranges would be necessary but due to their irregular structure a two-step approach as for rim-enhancing liver metastases is not possible.

We are also planning to develop more sophisticated volumetry methods. For lung nodules, Kuhnigk et al. [6] presented a solution that counts voxels at the surface of the mask only partially in order to compute a more accurate volume. For liver metastases and lymph nodes, this is more difficult because the lesions themselves and their surroundings can be much more diverse. A clustering algorithm could find the different adjacent structures and partial volume effects could be examined at the boundary between each of these structures and the lesion.

As another research topic, we are currently working on different interactive correction techniques. Adapting the erosion threshold does not always suffice. It limits the user in order to preserve reproducibility but this implies that complex modifications such as adding or removing particular parts of the lesion mask are not possible. The interaction should be as easy and intuitive as possible, so the problem is to extrapolate the necessary information for a 3-D operation. Furthermore, the better the initial segmentation becomes, the greater is the challenge to find a general correction framework for the remaining difficult cases.

REFERENCES

MOLTZ et al.: ADVANCED SEGMENTATION TECHNIQUES FOR LUNG NODULES

Jan Hendrik Moltz received the diploma in computer science from the University of Lübeck, Lübeck, Germany, in 2006.

Since then, he has been a Research Assistant at Fraunhofer MEVIS, Institute for Medical Image Computing, Bremen, Germany, working on tumor segmentation in CT and other image processing techniques for oncological therapy monitoring.

Lars Bornemann received the diploma in computer science from the University of Lübeck, Lübeck, Germany, in 2000.

He previously worked as a Software Developer and Lead Programmer for Ascaron GmbH, Gütersloh, Germany. Since 2003, he has been a Research Assistant at Fraunhofer MEVIS, Institute for Medical Image Computing, Bremen, Germany, working on tumor segmentation and application development.

Jan-Martin Kuhnigk received the diploma in computer science from the University of Lübeck, Lübeck, Germany, and the Ph.D. degree in medical image analysis from the University of Bremen, Bremen, Germany, in 2008.

Currently, he works on research and development of software assistance in medical imaging at Fraunhofer MEVIS, Institute for Medical Image Computing, Bremen. His main interests lie in quantitative analysis of lung morphology and function in CT images.

Volker Dicken studied mathematics and physics at the University of Marburg, Marburg, Germany, and Cambridge University, Cambridge, U.K. He received the Ph.D. degree in numerical analysis from the University of Potsdam, Potsdam, NY, in 1998 for work on image reconstruction in SPECT. His dissertation was on wavelet methods for inverse problems and received the grand-challenge2008.bigr.nl/proceedings/pdfs/lts08_MeVis.pdf

Following work on image compression and imbalance detection in aircraft engine, he joined Fraunhofer MEVIS, Institute for Medical Image Computing, Bremen, Germany, in 2002, where he works mainly on analysis of thoracic and oncological medical image data with a focus on clinical evaluation.

Elena Peitgen received her medical education at the Philippi-University Marburg, Marburg, Germany. She received the Ph.D. degree in 2007 and completed her internship in the Glasgow Royal Infirmary and the Vale of Leven Hospital, Glasgow, Scotland, U.K.

She started her training in diagnostic imaging at the Department of Diagnostic and Interventional Radiology, University Hospital Mainz, and is now working at the Department of Radiology, Klinikum Bremen Ost.

Stephan Meier received his medical education at the University of Mainz, Mainz, Germany. He received the Ph.D. in 2002.

He is currently a Consultant in the Department of Diagnostic and Interventional Radiology, University of Mainz. He previously held internships at the University of Stellenbosch, South Africa, the University of Cincinnati, Cincinnati, OH, and the University of the West Indies. His scientific work is focused on abdominal imaging, in particular preoperative planning of liver transplantation and liver tumor resection.

Hendrik Bolte received his medical education in neurosurgery and radiology at the Christian-Albrechts-University of Kiel, Kiel Germany, and the University of Cape Town, South Africa.

He is a Consultant of Radiology and since July 2008 an Assistant Professor in the Department of Diagnostic Radiology, University Hospital Schleswig Holstein, Kiel, Germany. His scientific projects are quantitative CT imaging (lung nodule and lymph node volumetry), musculoskeletal imaging (low dose CT of peripheral joints), and cardiac imaging.

Michael Fabel completed his studies in medicine at the University of Trier, Trier, Germany, in 2004.

He has been a Research Assistant in the Department of Diagnostic Radiology, University Hospital Schleswig Holstein, Kiel, Germany since 2007. He was an Assistant in the Division of Radiology, German Cancer Research Center, Heidelberg, in 2005 and 2006. His research is focused on therapy monitoring and comparison of RECIST and volumetry, especially for lymph nodes.
Hans-Christian Bauknecht studied medicine at the Humboldt-University of Berlin, Berlin, Germany, and graduated in 2002. He received the Ph.D. degree from the Institute for Radiology of the Charité, University of Medicine Berlin, in 2004. He has been a Research Assistant at the Institute for Radiology of the Charité, University of Medicine Berlin, since 2002. His research interests are CT of the pars petrosa, head and neck imaging, MRI planning of middle ear implants, functional MRI infrasound, MRI of the bulbos olfactorius, and volumetry of cerebral tumors.

Markus Hittinger studied medicine at the University of Munich, Munich Germany. He has been a Medical Assistant in the Institute for Clinical Radiology, University Hospital of Munich, since 2007. His research focus is imaging for therapy monitoring in oncology.

Andreas Kießling finished his studies in medicine at the University of Jena, Thuringia, Germany, in 2003 and received the Ph.D. degree in 2004. He is now an Assistant in the Department of Diagnostic Radiology, University Hospital of Gießen and Marburg, Germany. His scientific work is focused on tumor volumetry.

Michael Püsken studied medicine at the Universities of Marburg and Zürich, Germany, until 2005 and received the Ph.D. degree in 2006. He was an Assistant in the Department of Diagnostic Radiology, University Hospital of Göttingen, from 2005 to 2007. He is currently an Assistant at the Institute for Clinical Radiology, University Hospital of Münster, Münster, Germany. His scientific focus is software-assisted evaluation of liver lesions and lymph nodes in computed tomography.

Heinz-Otto Peitgen graduated in mathematics, physics, and economy in 1971 from the University of Bonn, Bonn, Germany, in 1971, where he also received the Ph.D. degree in 1973 and the Habilitation in 1977. He worked at the Institute for Applied Mathematics, University of Bonn. He obtained a Professorship in Mathematics at the University of Bremen, Bremen, Germany, and in 1991 also at Florida Atlantic University, Boca Raton. In 1995, he founded and became President of Fraunhofer MEVIS, Bremen. His research works comprises dynamical systems, numerical analysis, computer graphics, image and data analysis and computer-assistance in image-based medical diagnostics and therapy planning. He is the author of several influential books about fractal geometry and chaos theory.