Sketch-Based Interactive Segmentation and Segmentation Editing for Oncological Therapy Monitoring
– PhD Thesis Summary –

Frank Heckel
Fraunhofer Institute for Medical Image Computing MEVIS, Bremen

1 Introduction

Cancer is one of the most common causes of death. Worldwide about 14 million new cancer cases and 8.2 million cancer deaths were reported for 2012 [1]. Even though minimally invasive treatment methods like radiofrequency ablation (RFA) become more and more popular, chemotherapy remains an important systemic treatment method, especially for metastatic cancer. An important criterion for assessing a tumor’s response to chemotherapy is its change in size. The current gold standard for measuring the size of solid tumors and their progression over time is the maximum diameter as defined in the RECIST guideline (Response Evaluation Criteria In Solid Tumors) [2]. A diameter is not always appropriate, though, since most tumors grow and shrink irregularly in 3D [3]. Therefore, the volume of a tumor has shown to be a more accurate and reproducible measure for tumor size in various studies (e.g., [4]). To measure a tumor’s volume based on computed tomography (CT), it needs to be delineated (segmented) in the 3D image. Segmentation in medical imaging is an intensively studied field and many algorithms have been developed during the past decades.

For a long time, research has aimed at automatic segmentation methods. In recent years, however, interactive methods have grown in interest due to limitations of automatic methods in clinical routine. In real-world applications, control over the segmentation result is an essential requirement [5,6]. However, as CT images suffer from partial volume effects caused by their limited spatial resolution, the delineation of a tumor is an ill-defined problem. Particularly for interactive methods, this can result in significant differences in the volume measured by different readers or at different points of time. This thesis tackled the problems of interactive segmentation, the editing of a given segmentation result, and the accurate and reproducible measurement of a tumor’s volume based on a given segmentation mask in the context of oncological therapy response assessment. Each contribution is summarized in the following sections.

2 Contribution 1: Interactive Segmentation with Energy-Minimizing Implicit Functions

Most segmentation algorithms are image-data-driven and have been designed for specific use cases, i.e., they make assumptions on the objects of interest, in our case tumors, for example how they look with a certain imaging technique. Interactive algorithms overcome this limitation at the expense of an increased time required by the user for segmentation. However, not many general 3D segmentation algorithms exist that can be used for any tumor in any imaging modality even if the image quality is poor. This thesis proposed an image-independent interactive segmentation method to fill this gap. The object is segmented by the user on a few 2D slices of the 3D image by drawing contours along its border (cf. Fig. 1(b)), which is a common way to manually delineate objects in medical images. From these contours, a smooth 3D surface is generated. The algorithm also allows the contours to be drawn in any MPR.

2.1 Methods

The proposed segmentation algorithm uses an object reconstruction approach called variational interpolation [7]. Variational interpolation defines an implicit function $f(x)$ that fulfills all constraints given by a set of unordered points (a point cloud) while it minimizes an energy function $E$ that measures the smoothness of $f(x)$. The object’s surface is defined by $f(x) \equiv 0$. Using an appropriate radial basis function (RBF) $\phi(x)$, $f(x)$ can be written as

$$f(x) = P(x) + \sum_{j=1}^{k} w_j \phi(x - c_j),$$

Fig. 1: Example for segmentation of a liver metastasis in CT using 17 contours: (a) A user-drawn contour in 2D, (b) Interpolation result in 3D. Also notice the cap on top of the segmentation in (b), where the segmentation is smoothly closed.
where \( c_j = (c_j^x, c_j^y, c_j^z) \) are the points in 3D space where the function is constrained to have a specific value, \( w_j \) are the weights of each RBF and \( P(x) = p_0 + p_1 c^x + p_2 c^y + p_3 c^z \) is a degree-one polynomial that accounts for the linear and constant portions of \( f(x) \). We use the triharmonic spline \( \phi(x) = \|x\|^3 \) as RBF, which results in a \( C^2 \)-continuous surface. \( f(x) \) must fulfill the constraints \( c_i \) whose values are given by \( h_i \). The constraints are generated from the user-drawn contours. If a constraint \( c_i \) is located on the surface of the object, \( h_i \) equals 0, which is called a surface or boundary constraint. In addition, normal constraints are used to unambiguously define \( f(x) \). Normal constraints also allow a more accurate reproduction of the contours. They are defined as either \( h_i = 1 \) or \( h_i = -1 \), depending on whether they point inside or outside of the object. This results in the following linear system, with \( \phi_{ij} = \phi(c_i - c_j) \):

\[
\begin{bmatrix}
\phi_{11} & \phi_{12} & \cdots & \phi_{1k} & 1 & c_1^x & c_1^y & c_1^z \\
\phi_{21} & \phi_{22} & \cdots & \phi_{2k} & 1 & c_2^x & c_2^y & c_2^z \\
\vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \vdots & \vdots \\
\phi_{k1} & \phi_{k2} & \cdots & \phi_{kk} & 1 & c_k^x & c_k^y & c_k^z \\
1 & 1 & \cdots & 1 & 0 & 0 & 0 & 0 \\
\end{bmatrix}
\begin{bmatrix}
w_1 \\
w_2 \\
\vdots \\
w_k \\
p_0 \\
p_1 \\
p_2 \\
p_3 \\
\end{bmatrix}
= 
\begin{bmatrix}
h_1 \\
h_2 \\
\vdots \\
h_k \\
0 \\
0 \\
0 \\
0 \\
\end{bmatrix}
\]

Solving Eq. 2 gives the weights \( w_j \). The linear system is symmetric and positive semi-definite, so it can be efficiently solved using an algorithm by Bunch and Kaufman [8].

As the utilized object reconstruction approach is a global interpolation, the main challenge is the reduction of the computation time such that interactive rates are possible. To achieve this, a shape-preserving constraint reduction is parallelized to utilize the processing power of modern multi-core CPUs.

\[
\omega_i^\alpha = \frac{1}{2} \left( 1 - \frac{(c_i - c_{i-1}) \cdot (c_{i+1} - c_i)}{||c_i - c_{i-1}|| \cdot ||c_{i+1} - c_i||} \right) \quad \omega_i^\beta = \frac{||c_i - c_{i+1}|| \cdot ||c_{i+1} - c_i||}{||c_i - c_{i-1}|| + ||c_{i+1} - c_i||}
\]

The total weight \( \omega_i \) of a contour point is calculated by \( \omega_i = \omega_i^\alpha \omega_i^\beta \). Points that have the least influence on the contours’ geometry, i.e., with the lowest weight \( \omega_i \), are iteratively removed until a specific number of points is reached. This number is given by a quality factor \( q \in [0, 1] \) multiplied by the initial number of points. Finally, all computations are parallelized to utilize the processing power of modern multi-core CPUs.

### 2.2 Results and Discussion

Without constraint reduction, the computation time for the tumor example shown in Fig. 1 was 5.34s (17 contours, 6246 constraints). With \( q = 0.2 \), the number of constraints is reduced to 1390, resulting in a computation time of 0.69s, which is fast enough to allow interactive modifications. Compared to the result with \( q = 1 \), the Jaccard coefficient was 99.32% with a Hausdorff distance of 1.09 mm, i.e., there is almost no loss in segmentation quality. The algorithm was evaluated by two experienced radiographers on 15 liver metastases of different size (ranging from 0.05 ml to 53.26 ml) and complexity and 1 liver. For contouring, a touch-screen display with a digitizing pen was used. The quality factor for constraint reduction was set to \( q = 0.2 \). Tab. 1 gives an overview on the liver metastases results. Both the quality and the time were rated on a 5-point scale with 1 being unacceptable and 5 being perfect.

As variational interpolation does not use image information but only considers the object geometry given by the user-drawn contours, the algorithm is able to properly segment objects also if there is no contrast to the background, i.e., if the object does not exhibit distinct boundaries. This is a common problem for many objects in medical images, at least at certain parts of the object, which is one reason why automatic algorithms frequently fail. Because of its interpolation character, the reconstruction algorithm guarantees that all contours given by the user are part of the surface. Moreover, it plausibly extrapolates beyond the contours (cf. Fig. 1(b)). If the user gives contradictory inputs, the interpolation property is a drawback, as it could result in unexpected topologies (e.g., holes). In addition, irregular surfaces cannot be reconstructed accurately. Both can be slightly compensated by an approximation strategy that

**Table 1:** Results of the evaluation of the image-independent segmentation algorithm on 15 liver metastases. All results are given as average over all cases and participants. The overlap and distance measures are given relative to the manual segmentation.

<table>
<thead>
<tr>
<th></th>
<th>Manual</th>
<th>Variational interpolation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Required time</td>
<td>111.4s</td>
<td>64.2s</td>
</tr>
<tr>
<td>Number of contours</td>
<td>20.6</td>
<td>7.1</td>
</tr>
<tr>
<td>Quality rating</td>
<td>4.9</td>
<td>3.8</td>
</tr>
<tr>
<td>Time rating</td>
<td>3.8</td>
<td>4.0</td>
</tr>
<tr>
<td>Overlap to manual</td>
<td>100%</td>
<td>75.77%</td>
</tr>
<tr>
<td>Maximum surface distance to manual</td>
<td>0mm</td>
<td>2.94mm</td>
</tr>
</tbody>
</table>

2 F. Heckel
As an image-independent algorithm were suggested that solve this problem in 3D based on an intuitive sketching the most probably corrected error generated based on the current and a given reference segmentation. The simulation consists of three steps: 1. finding such contradictions, we ignore all constraints from previous inputs in a specific range around each constraint. Sometimes the user gives contradictory information, particularly if the correction is done in different views. To resolve these contradictions, we ignore all constraints from previous inputs in a specific range around each constraint.

To simulate the user input on the next (previous) slice $s \pm 1$, $C_{\text{user}}$ is replaced. This can be interpreted as adding some part, cutting away some part, combinations of both or replacing the segmentation. To extrapolate the 2D correction given by the user into 3D, we first estimate the extent of the edited region in z-direction (the correction depth $d(C_{\text{user}})$) based on its geometrical properties in the edited slice $s$, which defines the first and the last slice $s_{\text{start}}$ and $s_{\text{end}}$ between which the segmentation is modified. For tumor segmentations, the z-extent of the edited region often corresponds to its thickness in the current slice.

To simulate the user input on the next (previous) slice $s \pm 1$, $C_{\text{user}}$ is replaced. This can be interpreted as adding some part, cutting away some part, combinations of both or replacing the segmentation. To extrapolate the 2D correction given by the user into 3D, we first estimate the extent of the edited region in z-direction (the correction depth $d(C_{\text{user}})$) based on its geometrical properties in the edited slice $s$, which defines the first and the last slice $s_{\text{start}}$ and $s_{\text{end}}$ between which the segmentation is modified. For tumor segmentations, the z-extent of the edited region often corresponds to its thickness in the current slice.

To simulate the user input on the next (previous) slice $s \pm 1$, $C_{\text{user}}$ is replaced. This can be interpreted as adding some part, cutting away some part, combinations of both or replacing the segmentation. To extrapolate the 2D correction given by the user into 3D, we first estimate the extent of the edited region in z-direction (the correction depth $d(C_{\text{user}})$) based on its geometrical properties in the edited slice $s$, which defines the first and the last slice $s_{\text{start}}$ and $s_{\text{end}}$ between which the segmentation is modified. For tumor segmentations, the z-extent of the edited region often corresponds to its thickness in the current slice.

To simulate the user input on the next (previous) slice $s \pm 1$, $C_{\text{user}}$ is replaced. This can be interpreted as adding some part, cutting away some part, combinations of both or replacing the segmentation. To extrapolate the 2D correction given by the user into 3D, we first estimate the extent of the edited region in z-direction (the correction depth $d(C_{\text{user}})$) based on its geometrical properties in the edited slice $s$, which defines the first and the last slice $s_{\text{start}}$ and $s_{\text{end}}$ between which the segmentation is modified. For tumor segmentations, the z-extent of the edited region often corresponds to its thickness in the current slice.

To simulate the user input on the next (previous) slice $s \pm 1$, $C_{\text{user}}$ is replaced. This can be interpreted as adding some part, cutting away some part, combinations of both or replacing the segmentation. To extrapolate the 2D correction given by the user into 3D, we first estimate the extent of the edited region in z-direction (the correction depth $d(C_{\text{user}})$) based on its geometrical properties in the edited slice $s$, which defines the first and the last slice $s_{\text{start}}$ and $s_{\text{end}}$ between which the segmentation is modified. For tumor segmentations, the z-extent of the edited region often corresponds to its thickness in the current slice.

To simulate the user input on the next (previous) slice $s \pm 1$, $C_{\text{user}}$ is replaced. This can be interpreted as adding some part, cutting away some part, combinations of both or replacing the segmentation. To extrapolate the 2D correction given by the user into 3D, we first estimate the extent of the edited region in z-direction (the correction depth $d(C_{\text{user}})$) based on its geometrical properties in the edited slice $s$, which defines the first and the last slice $s_{\text{start}}$ and $s_{\text{end}}$ between which the segmentation is modified. For tumor segmentations, the z-extent of the edited region often corresponds to its thickness in the current slice.

To simulate the user input on the next (previous) slice $s \pm 1$, $C_{\text{user}}$ is replaced. This can be interpreted as adding some part, cutting away some part, combinations of both or replacing the segmentation. To extrapolate the 2D correction given by the user into 3D, we first estimate the extent of the edited region in z-direction (the correction depth $d(C_{\text{user}})$) based on its geometrical properties in the edited slice $s$, which defines the first and the last slice $s_{\text{start}}$ and $s_{\text{end}}$ between which the segmentation is modified. For tumor segmentations, the z-extent of the edited region often corresponds to its thickness in the current slice.
The proposed algorithms were evaluated on 131 insufficient tumor segmentations (lung nodules, liver metastases, lymph nodes) by 5 radiologists with a different level of experience. The initial segmentations were generated using the algorithms by Moltz et al. [11]. The editing algorithms were rated on a 5-point scale from unacceptable (−−) to perfect (++). The ratings were averaged after mapping them onto the interval [0 ... 1], giving the editing rating score \( r_{\text{edit}} \).

Next, the MICCAI score [12] in each step was analyzed and accumulated into the editing quality score \( m_{\text{edit},S_{\text{max}}} \):

\[
m_{\text{edit},S_{\text{max}}} = \frac{1}{S_{\text{max}}} \left( \min(S,S_{\text{max}}) \sum_{j=1}^{m} m_j + \hat{S} \cdot m_S \right)
\]

\[
\hat{S} = \begin{cases} 
S_{\text{max}} - S & \text{if } S_{\text{max}} > S \\
0 & \text{if } S_{\text{max}} \leq S
\end{cases}
\]

The results are given in Fig. 4. They show that both algorithms allow an effective and efficient editing of tumor segmentations. The image-based method has been rated better, as the image-independent method sometimes suffers from the issue of contradictory inputs and irregular shapes. For the image-based method, the median computation time was 0.28 s with an overall editing time of 52 s. Using the image-independent method editing took 63 s with a computation time of 0.09 s per step.

Concerning the evaluation, user studies are the most important instrument. The simulation-based approach complements this during development, where frequent user studies are impossible and suboptimal due to their high effort and limited reproducibility. In this context, a simulation-based evaluation can provide crucial support during algorithm improvement (e.g., bug-fixing and parameter tuning). It also allows a more objective comparison of different algorithms. However, the simulation results currently do not correlate with the results by real humans (cf. Fig. 5(b)).
Due to the limited spatial resolution of CT, voxels at the border of an object do not only cover the object itself but also adjacent structures. Thus, the CT value of such voxels is a mixture of different tissues. As a consequence, the border between the object and its background is ill-defined. This intrinsic uncertainty could cause a significant variability in size measurements by different readers or at different points of time, particularly for manual and interactive segmentation tools. To make volumetric measurements more reliable even for inhomogeneous tumors, this thesis proposed a partial volume correction approach that utilizes the information of a given segmentation result. The proposed algorithm is an extension of the algorithm for lung nodules by Kuhnigk et al. [13].

4 Contribution 3: Segmentation-Based Partial Volume Correction

The influencing tissues for a certain partial volume voxel are defined by a spatial subdivision of the lesion into spherical sectors (segments $S$) (cf. Fig. 6). The segments are calculated by an angular subdivision such that each segment has a similar volume with respect to a sphere. The center of this sphere is located at the center of gravity of the segmentation mask $M$. Based on $M$, inner and outer partial volume regions $P_i$ and $P_o$ as well as tissue regions $T_i$ and $T_o$ are defined as sets of all voxels within a specific distance to the surface of $M$. For each region of each segment references values $t_i$, $p_i$, $t_o$, and $p_o$ are computed as the median of all voxel-values within $T_i$, $P_i$, $T_o$, and $T_o$, respectively. Finally, the lesion core is computed by $C = M \setminus (P_i \cup T_i)$. Given the influencing tissues and their reference values, their fraction in each partial volume voxel can be calculated, i.e., the weight $w(V)$. For each voxel of each partial volume region, $w(V)$ is estimated by a linear combination of the reference values of the associated inner and outer tissue regions of the specific segment $s$:

$$w(V) = \frac{t_o - v}{t_o - t_i}, V \in P_i \cup P_o.$$  

For all voxels within $C$ and $T_i$, $w(V)$ is set to 1 (cf. Fig. 6(b)). Given the weight of each voxel we can calculate the lesions volume $Vol_L$, with $C = C \cup T_i \cup P_i \cup P_o$ being the lesion and $Vol_V$ being the volume of a single voxel:

$$Vol_L = \sum_{V \in C} w(V)Vol_V.$$  

4.2 Results and Discussion

The algorithm has been evaluated on a software phantom, hardware phantoms as well as a total number of 1516 lesion segmentation pairs. The data contained lung nodules, liver metastases, and lymph nodes. The segmentations were generated using the algorithms by Moltz et al. [11] and optionally edited using the methods proposed in Sec. 3. Fig. 7 shows some of results from these evaluations. Overall, the variability (with respect to the interquartile range) for the phantom data was reduced by 49% ($p<0.001$) and the variability between different readers was reduced by 28% ($p<0.001$). The average computation time is 0.2 s. The algorithm has also shown to improve the reproducibility on 1851 lung nodules from the LIDC-IDRI database that have been segmented manually by at least two readers.

The proposed algorithm has shown a more accurate estimation of the real volume and its ability to reduce inter- and intra-observer variability significantly, hence, allowing more reliable volumetric measurements. Moreover, it is fast enough to be combined with interactive segmentation methods. Its main limitation is that one segment might contain more than one tissue class inside or outside of the lesion, resulting in incorrect estimations of $t_i$, $t_o$, and thus of $w(V)$. If the main shape of a lesion is not convex, the implicitly done ellipsoidal shape assumption might not represent the lesion appropriately, which could as well result in suboptimal estimations of partial volume and tissue regions.

5 Conclusion

The goal of this thesis was the development of algorithms that support the radiologist in efficiently and reliably measuring a tumor’s volume in CT. It could be shown that an object reconstruction method can be used as a general approach to estimate the volume of a lesion from its segmentation. The proposed algorithm has shown a more accurate estimation of the real volume and its ability to reduce inter- and intra-observer variability significantly, hence, allowing more reliable volumetric measurements. Moreover, it is fast enough to be combined with interactive segmentation methods. Its main limitation is that one segment might contain more than one tissue class inside or outside of the lesion, resulting in incorrect estimations of $t_i$, $t_o$, and thus of $w(V)$. If the main shape of a lesion is not convex, the implicitly done ellipsoidal shape assumption might not represent the lesion appropriately, which could as well result in suboptimal estimations of partial volume and tissue regions.
image-independent segmentation tool. Interactive computation times can be achieved by a quality preserving constraint reduction and the utilization of multi-core CPUs. The main contribution of this thesis is a comprehensive discussion of the segmentation editing problem. It could be shown that sketching provides a simple and intuitive interface for segmentation editing in 2D. For tumor segmentation, sketch-based modifications can be effectively and efficiently extrapolated into 3D using image information (in terms of gradients) as well as purely geometrically by utilizing the object reconstruction approach. In order to assess the quality of editing tools, two measures have been proposed, which summarize qualitative ratings as well as the quality of intermediate segmentation results. This has been complemented by a reproducible evaluation without the need for a user, where plausible interactions are simulated. Finally, a method has been developed that effectively compensates partial volume effects based on the image and a given segmentation result, independently of the actual segmentation algorithm, making volumetric measurements more accurate and reproducible.

Overall, the methods proposed in this thesis make volumetric tumor size measurement applicable to a wider range of tumor lesions as well as more reliable and less time-consuming, compared to manual segmentation or slice-wise 2D editing. Therefore, the discussed algorithms have the potential to increase the acceptance and relevance of volumetric measurements in clinical practice.

References