

A Concept for the Application of a Hierarchical Image Subdivision to the Segmentation Editing Problem

Frank Heckel^{1,2} and Stefan Braunewell¹

¹ Fraunhofer Institute for Medical Image Computing MEVIS,
Universitaetsallee 29, 28359 Bremen, Germany
{frank.heckel, stefan.braunewell}@mevis.fraunhofer.de

² Innovation Center Computer Assisted Surgery,
University of Leipzig, Semmelweisstraße 14, 04103 Leipzig, Germany

Abstract. Efficient 3D segmentation editing tools are important components in the segmentation process for cases where automatic segmentation algorithms fail to provide sufficient results. In this paper, we propose a novel generic concept for 3D manual correction of a given segmentation mask. It is based on a hierarchical subdivision of the image generated by the Interactive Watershed Transform. The user adds missing parts to the given segmentation or removes parts from it by coarsely defining foreground and background regions via markers. Our method is independent of the algorithm by which the initial segmentation has been generated, without the requirement of mapping its result to the concepts and data structures of the algorithm used in the editing step. This allows modifications to be kept local. In addition, only a few assumptions on the object of interest are made. We have successfully applied our algorithm to bone segmentation in CT angiography.

Keywords: segmentation editing, interactive segmentation, interactive watershed transform, bone removal, CTA

1 Introduction

Segmentation is one of the essential problems in medical imaging. In order to reduce the effort for the user, fully-automatic segmentation algorithms for specific tasks are desirable and a lot of research has been done in this field during the past decades. However, such algorithms could fail to provide an acceptable segmentation result in pathological cases or due to low contrast and a low signal-to-noise ratio for example. On the other hand, a lot of interactive segmentation algorithms exist that could be used as an alternative in such cases, which requires the user to start a new segmentation from scratch, though.

Given an unacceptable automatic segmentation result, most parts of the so called segmentation mask are often correct. Consequently, locally and efficiently correcting the *errors* in such segmentation masks would be the preferable solution. This is known as the segmentation editing problem. Segmentation editing

tools are important components in the segmentation process, particularly for applications used in clinical routine. Efficient and intuitive editing is a challenging problem, however. The user expects to be able to generate a sufficient segmentation for all cases with such tools, while only few interactions are accepted in routine. Also note that the segmentation problems that should be solved by such tools are typically difficult, as the preceding algorithm already failed. Editing is typically done in several steps, where each step is defined by one user interaction that results in a new (temporary) segmentation mask. Editing tools can be either part of the automatic segmentation algorithm, i.e., the automatic algorithm is able to consider additional information given by the user (constraints), or dedicated algorithms that are independent of the automatic segmentation method. Dedicated tools have the advantage that they can be combined with any other (semi-)automatic segmentation algorithm. However, their segmentation results could be inconsistent with the initial algorithm.

In this paper, we propose a simple and generally applicable method for combining an interactive segmentation from a hierarchical subdivision, such as the Interactive Watershed Transform (IWT) for example [7], with an arbitrary initial segmentation, providing a dedicated segmentation editing tool. User interaction is done by placing markers on incorrectly segmented foreground and background regions. From these markers, regions to be removed from the given segmentation or added to it are computed and the segmentation result is immediately displayed. Our algorithm guarantees to only change the given segmentation in regions that have been explicitly modified by the user. We discuss the proposed algorithm in the context of bone segmentation in computed tomography angiography (CTA). In order to get a clear view on contrast enhanced vessels in 3D visualizations, bone structures need to be removed from the image (cf. Fig. 3). This is particularly challenging as contrast enhanced vessels and bones look very similar in CTA and they could even touch [8]. As a consequence, automatic algorithms often leak into vessels or miss bone structures. From a computational point of view, bone removal requires a segmentation of the whole CTA image, making fast algorithms necessary.

2 Related Work

Segmentation editing is a relatively new area of research. A comprehensive overview on segmentation editing approaches in medical imaging has been given by Heckel et al. [10]. Dedicated editing approaches include morphological approaches [13, 16, 18], deformable models [3], shortest-path approaches [10], object reconstruction methods [12, 10, 14], machine learning approaches [9] and graph-based algorithms [4]. Interaction concepts cover paint brushes [18] contours [16, 12, 10, 14, 4, 9] as well as surface bulging and dragging [13, 3]. A dedicated editing method based on the Random Walker algorithm, where the user corrects errors of the segmentation by drawing markers (or seeds), has been proposed by Grady and Funka-Lea [5]. The authors formulate the editing problem as the minimization of an energy functional that explicitly considers the given initial

segmentation as its global optimum. A similar approach that uses the Image Foresting Transform has been suggested by Miranda et al. [17].

Approaches to bone segmentation in the context of bone removal include thresholding [2] and atlas-based methods [19] for example. The IWT combined with an automated placement of markers has also been successfully used by Hahn et al. [8]. Note that this approach includes manual correction possibilities by definition. In order to distinguish between bones and contrast enhanced vessel, dual-energy CTA has shown to be a useful imaging technique [21, 20]. Besides bone segmentation, the IWT has also proven to work well for skull stripping in MRI [6] and lung lobe segmentation in CT [15].

3 Methods

The basis for the proposed segmentation editing algorithm is a subdivision of the 3D image \mathcal{I} into homogeneous 3D *regions*. Based on this subdivision, a general marker-based interactive segmentation approach can be defined, which is summarized in the Sec. 3.1. In order to provide a general and efficient editing tool that requires only a few interactions, our goal was the utilization of this interactive segmentation method to any given initial segmentation mask. This segmentation editing algorithm is described in Sec. 3.2.

3.1 Hierarchical Image Subdivision

The subdivision algorithm first assembles voxels with similar properties into (coherent) regions. In the simplest case, only the intensity of each voxel is considered. Additional information such as edges, texture or the shape of the resulting region could be used as criteria as well. A well-known example for such a subdivision is the watershed transform. It uses both intensity and gradient information and refers to the regions as *basins*. In a second step, regions are iteratively merged into larger regions giving a tree structure that we refer to as *hierarchical subdivision* $\mathcal{H}(\mathcal{I})$. The IWT does this by incrementally flooding the watershed “landscape”. If a watershed between adjacent basins is found during this process, these basins are merged [7].

Based on a subdivision of the image, an object of interest can be segmented by identifying (labeling) all regions belonging to it. We refer to the object of interest as *foreground*, while the remaining regions define the *background*. The labeling can be done automatically based on knowledge about the object or interactively by the user, e.g., by setting foreground and background markers on specific regions. From the information of these manual markers, additional regions could also be labeled automatically. The specific *level* l at which the segmentation is performed influences the accuracy of the segmentation result, while it also makes the segmentation algorithm more efficient in terms of the required computation time and the time required by the user for the segmentation, respectively, because less data needs to be processed. The IWT calls this level *preflooding height*. Note that foreground and background markers might contradict, i.e., they could be

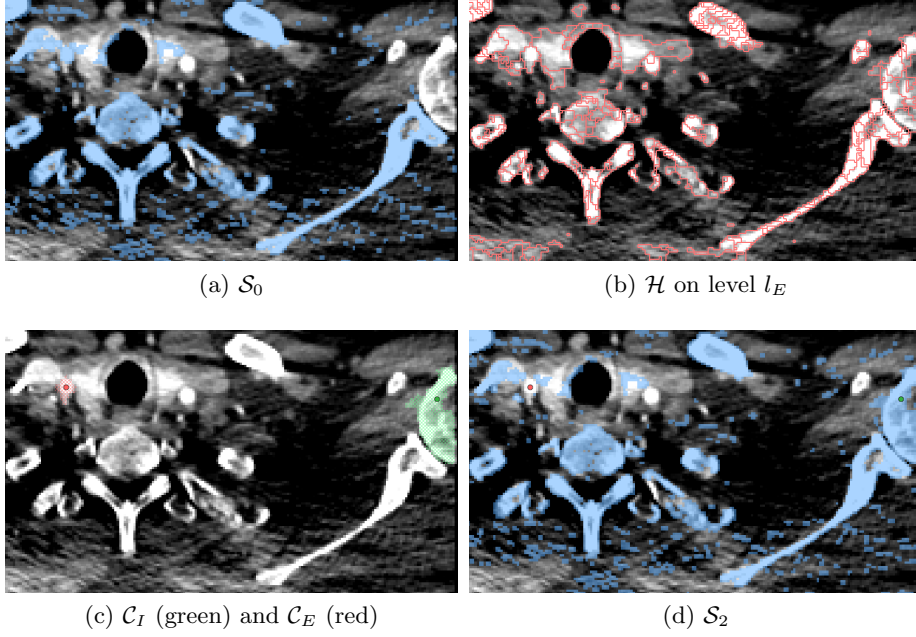


Fig. 1: Editing example showing the initial segmentation (a), the subdivision (b), the added and removed regions (c) and the result after two editing steps (d). Note that only bone-like voxels ≥ 175 HU are considered.

located in the same region with respect to the chosen level in \mathcal{H} . In this case, the affected regions have to be split according to the underlying hierarchy until all contradictions are resolved or the lowest level has been reached. Summing this up, a segmentation mask \mathcal{S} is given by the following equation, where $s(\cdot)$ denotes the segmentation algorithm as described above:

$$\mathcal{S} = s(\mathcal{H}(\mathcal{I}), \mathcal{M}_{\text{foreground}}, \mathcal{M}_{\text{background}}, l). \quad (1)$$

3.2 Segmentation Editing

Let \mathcal{S}_0 be the initial, erroneous 3D segmentation result. In contrast to existing methods [5, 17], we do not map \mathcal{S}_0 to concepts and data structures of the algorithm utilized for the editing step (e.g., the IWT). Mapping \mathcal{S}_0 to \mathcal{H} would require a proper pre-labeling of \mathcal{H} such that \mathcal{S}_0 is recovered. This is difficult in the general case, however, as \mathcal{S}_0 might have been created by a method (e.g., a model- or atlas-based algorithm) that results in completely different borders than what is inherent to \mathcal{H} . Hence, simply replacing \mathcal{S}_0 with the mapped result could cause significant changes in regions that the user did not intend to change. A solution could be to introduce artificial borders that are considered on all levels of \mathcal{H} . However, the most generic approach is to combine the result of

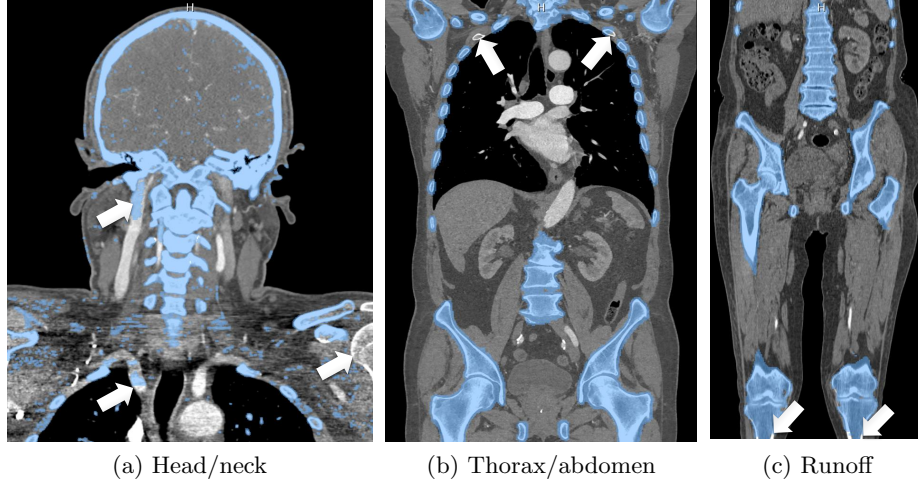


Fig. 2: Bone segmentation examples used for evaluation. The initial segmentation mask is overlaid in blue. The arrows indicate some prominent errors.

the interactive step with \mathcal{S}_0 a posteriori. This guarantees that \mathcal{S}_0 is only changed in regions where the user explicitly modified it, while other regions are kept.

Given a segmentation mask \mathcal{S} , two editing tasks can be distinguished: adding missing object voxels (*include*) and removing incorrectly included background voxels (*exclude*). Errors are iteratively resolved by correctly marking them as foreground or background, giving a set of include markers \mathcal{M}_I and a set of exclude markers \mathcal{M}_E . Based on the quintuple $(\mathcal{H}(\mathcal{I}), \mathcal{M}_I, \mathcal{M}_E, l_I, l_E)$, two *correction masks* are computed: an include mask \mathcal{C}_I on level l_I of the subdivision hierarchy and an exclude mask \mathcal{C}_E on level l_E . These masks are given by:

$$\begin{aligned}\mathcal{C}_I &= s(\mathcal{H}(\mathcal{I}), \mathcal{M}_I, \mathcal{M}_E, l_I) \\ \mathcal{C}_E &= s(\mathcal{H}(\mathcal{I}), \mathcal{M}_E, \mathcal{M}_I, l_E),\end{aligned}\tag{2}$$

After each user interaction (i.e., adding, removing or moving a marker), a new segmentation result \mathcal{S}_i is computed as

$$\mathcal{S}_i = (\mathcal{S}_0 \cup \mathcal{C}_I) \setminus \mathcal{C}_E,\tag{3}$$

where i is the number of editing steps (cf. Fig. 1). l_E and l_I are application specific. They could be set manually by the user or automatically for the specific image and segmentation problem by a machine-learning approach for example.

4 Results and Discussion

We have applied our editing concept on automatically generated bone segmentations in three different CTA images shown in Fig. 2. Evaluation was done

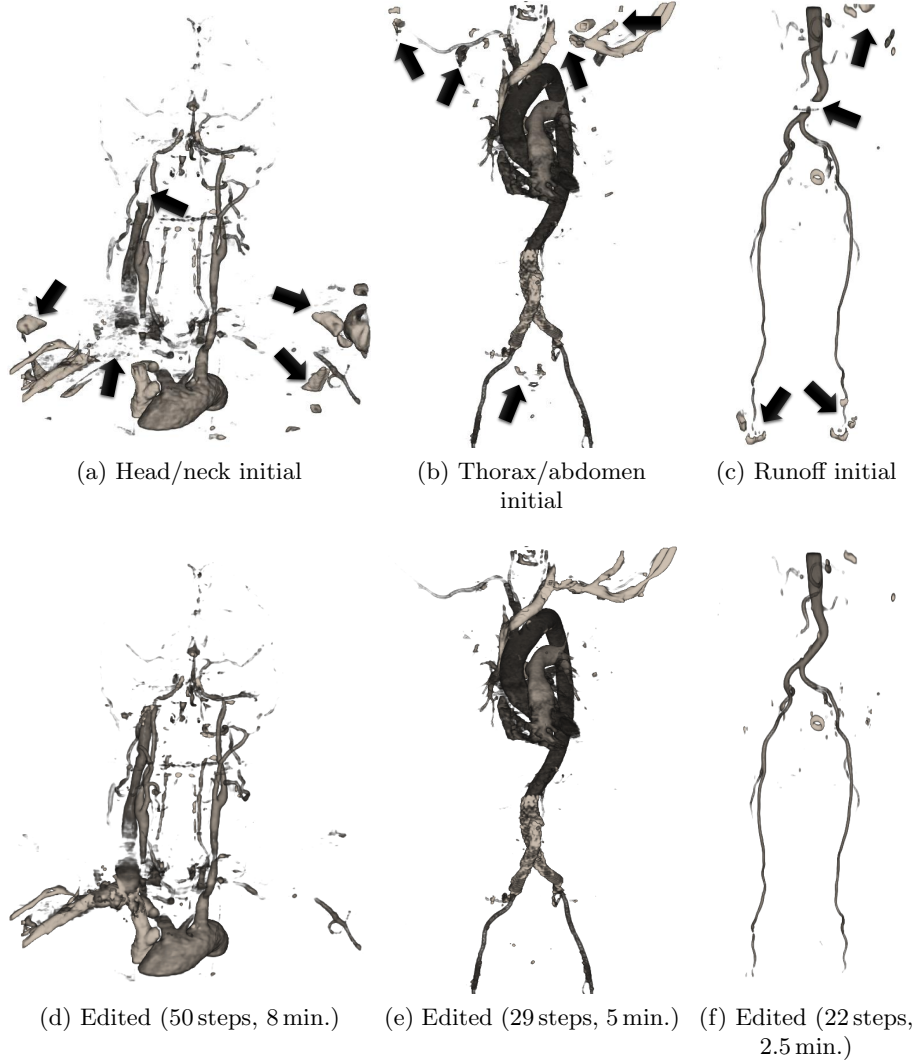


Fig. 3: 3D renderings of the evaluation data with removed bones before and after editing of the automatically generated bone segmentations. The arrows indicate the most prominent errors.

qualitatively based on 3D visualizations of the bones and the vessels before and after editing (cf. Fig. 3). Reference segmentations were not available. The computationally expensive generation of \mathcal{H} was performed in a pre-processing step using the IWT. As described in Sec. 1, bone segmentation in CTA is used for getting a clear view on the vessels. Thus, a coarse segmentation is often sufficient. For including bone structures, a relatively high level l_I can be chosen,

Table 1: Computation times.

Dataset	Size in voxels	Pre-processing	Computation time per editing step
Head/neck	$256 \times 256 \times 363$	5.9 s	1.1 ± 0.06 s
Thorax/abdomen	$256 \times 256 \times 652$	13.6 s	4.4 ± 0.9 s
Runoff	$256 \times 256 \times 189$	4.6 s	0.4 ± 0.05 s

while removing small vessels requires a rather low level l_E . Therefore, l_I was set to a preflooding height of 600 HU and l_E to a preflooding height of 80 HU. In addition, both \mathcal{C}_I and \mathcal{C}_E were dilated by a $3 \times 3 \times 3$ kernel in order to consider partial volume effects as well. Note that these settings can also be interactively adapted during editing by the user in order to find the best configuration, which was not done during evaluation, though. Because bones and enhanced vessels have a relatively large density in CT, all computation were restricted to voxels ≥ 175 HU in order to reduce both memory consumption and computation times (cf. Fig. 1b). Editing was performed until all obvious errors were fixed or no better result could be achieved. Fig. 3 shows 3D renderings of the CTA images with removed bones, both before and after editing. Tab. 1 gives an overview on the computation times. All measurements were performed on a laptop with an Intel Core i7-2620M (2.7 GHz) and 8 GB RAM running Windows 7 64 Bit.

Our evaluations have successfully proven the suitability of the proposed segmentation editing concept for the challenging task of bone segmentation in CTA. The parameters of the algorithm have been chosen very conservative, such that large bones can be added with a few markers and small vessels can be removed. Excluding large vessels, like in the head/neck example, takes several markers, though. Less interactions can be expected with optimized parameters and subdivision strategies. Even though the implementation is currently not optimized for low computations times and memory usage, reaction times of ≤ 1 s are already possible for images of up to 20 million voxels on mid-range hardware. In addition, the computation time is almost independent of the number of markers. The proposed algorithm does not depend on intermediate states of the editing algorithm (cf. Eq. 3). Therefore, the order in which the markers are given doesn't influence the editing result and undo operations as well as deleting and moving markers in any order is seamlessly possible. Each modification only affects a single region in \mathcal{H} on a specific level, i.e., modifications are guaranteed to be local with respect to \mathcal{H} . However, this is not necessarily true for the segmentation problem from a user's point of view, as a single 3D region could contain several parts of the object of interest that might look separated on a specific 2D slice. The main drawback of our method is the limited accuracy defined by the lowest level in \mathcal{H} . Therefore, correcting errors where the subdivision did not find a proper border between fore- and background as well as voxel-accurate modifications are not possible. Also, the dilation of \mathcal{C}_E turned out to be problematic in some cases, in which it removed correctly segmented voxels that cannot be re-added.

5 Conclusion and Future Work

We have proposed a novel generic concept for a dedicated segmentation editing algorithm. Our approach can be combined with any automatic segmentation algorithm, without the requirement of mapping its result to the concepts and data structures of the algorithm used in the editing step. This allows modifications to be kept local. It also provides deterministic results given the same markers, independently of their order. Despite the homogeneity of regions belonging to the object and the background and their separability by \mathcal{H} , no assumptions on the segmentation task are made. First qualitative evaluations in the context of the bone removal problem have shown that the suggest method allows an efficient correction of bone segmentation errors. Using the IWT as basis for the hierarchical subdivision, computation times are interactive (≤ 1 s) for 3D images of up to $2 \cdot 10^7$ voxels.

Our algorithm is currently in a proof-of-concept state. Comprehensive, quantitative evaluations are necessary to prove its benefit for real-world applications. It also needs to be investigated, how well the proposed concept works in other applications, such as tumor, organ or vessel segmentation for example, which should be possible as long as foreground and background can be separated by \mathcal{H} . The watershed algorithm has shown to generate proper results for our application. In recent years, however, a representation called superpixels has emerged as a convenient primitive in computer vision [1]. Combining this representation with the concept of object-based image analysis (OBIA) [11] for the merging step could serve as a more powerful and flexible alternative in the future. The OBIA framework would allow an improved hierarchical subdivision as additional information such as texture and region shape could be considered. Finally, a local definition of the level parameter and an intelligent editing algorithm that dynamically adapts its parameters for a specific editing step would be favorable in order to make the editing more efficient and accurate.

Acknowledgement

We thank Markus Harz from Fraunhofer MEVIS for discussions concerning the bone removal problem and the interactive watershed transform. This work was supported by Siemens AG, Healthcare Sector, Imaging & Therapy Division, Computed Tomography & Radiation Oncology, Forchheim, Germany.

References

1. Achanta, R., et al.: SLIC superpixels compared to state-of-the-art superpixel methods. *IEEE T. Pattern. Anal.* 34(11), 2274–2282 (2012)
2. Alyassin, A.M., Avinash, G.B.: Semiautomatic bone removal technique from CT angiography data. In: *Proc. SPIE Med. Imaging*. vol. 4322, pp. 1273–1283 (2001)
3. Beichel, R., et al.: Liver segmentation in CT data: A segmentation refinement approach. In: *Proc. 3D Segmentation in The Clinic: A Grand Challenge, MICCAI Workshop*. pp. 235–245 (2007)

4. El-Zehiry, N., et al.: A splice-guided data driven interactive editing. In: Proc. ISBI. pp. 1086–1089 (2013)
5. Grady, L., Funka-Lea, G.: An energy minimization approach to the data driven editing of presegmented images/volumes. In: Proc. MICCAI. vol. 2, pp. 888–895 (2006)
6. Hahn, H.K., Peitgen, H.O.: The skull stripping problem in MRI solved by a single 3D watershed transform. In: Proc. MICCAI. vol. 1935, pp. 134–143 (2000)
7. Hahn, H.K., Peitgen, H.O.: IWT - interactive watershed transform: A hierarchical method for efficient interactive and automated segmentation of multidimensional gray-scale images. In: Proc. SPIE Med. Imaging. vol. 5032, pp. 643–653 (2003)
8. Hahn, H.K., et al.: HWT – hybrid watershed transform: Optimal combination of hierarchical interactive and automated image segmentation. In: Proc. SPIE Med. Imaging. vol. 6512, pp. 65120Z–1–11 (2007)
9. Harrison, A.P., et al.: IntellEditS: Intelligent learning-based editor of segmentations. In: Proc. MICCAI. vol. 8151, pp. 235–242 (2013)
10. Heckel, F., et al.: Sketch-based editing tools for tumour segmentation in 3D medical images. *Comput. Graph. Forum* 32(8), 144–157 (2013)
11. Homeyer, A., et al.: A generic concept for object-based image analysis. In: Proc. VISIGRAPP. pp. 530–533 (2010)
12. Ijiri, T., Yokota, H.: Contour-based interface for refining volume segmentation. In: Proc. Pacific Graphics. vol. 29, pp. 2153–2160 (2010)
13. Kang, Y., et al.: Interactive 3D editing tools for image segmentation. *Med. Image Anal.* 8(1), 35–46 (2004)
14. Karimov, A., et al.: Vivisection: Skeleton-based volume editing. In: Proc. EuroVis. vol. 32, pp. 461–470 (2013)
15. Kuhnigk, J.M., et al.: Lung lobe segmentation by anatomy-guided 3D watershed transform. In: Proc. SPIE Med. Imaging. vol. 5032, pp. 1482–1490 (2003)
16. Maleike, D., et al.: Interactive segmentation framework of the medical imaging interaction toolkit. *Comput. Methods Programs Biomed.* 96(1), 72–83 (2009)
17. Miranda, P.A.V., et al.: How to complete any segmentation process interactively via image foresting transform. In: Proc. SIBGRAPI. pp. 309–316 (2010)
18. Silva, S., et al.: A 3D tool for left ventricle segmentation editing. In: Proc. ICIAR. vol. 6112, pp. 79–88 (2010)
19. Straka, M., et al.: Bone segmentation in CT-angiography data using a probabilistic atlas. Tech. Rep. TR-186-2-04-01, Institute of Computer Graphics and Algorithms, Vienna University of Technology (2004)
20. Thomas, C., et al.: Automatic bone and plaque removal using dual energy CT for head and neck angiography: Feasibility and initial performance evaluation. *Eur. J. Radiol.* 76(1), 61–67 (2010)
21. Watanabe, Y., et al.: Dual-energy direct bone removal CT angiography for evaluation of intracranial aneurysm or stenosis: comparison with conventional digital subtraction angiography. *Eur. Radiol.* 19(4), 1019–1024 (2009)