

Self-Assessed Contrast-Maximizing Adaptive Region Growing

Carlos S. Mendoza¹ *, Begoña Acha¹, Carmen Serrano¹, and Tomás Gómez-Cía²

¹ Universidad de Sevilla
Av. de los Descubrimientos s/n
41092 Sevilla, Spain

² Grupo Unidad Clínica de Cirugía Plástica y Grandes Quemados
Hospitales Universitarios Virgen del Rocío de Sevilla
Avda Manuel Siurot, s/n
41013 Sevilla, Spain

{csanchez1, bacha, cserrano}@us.es, tgomezcc@gmail.com

Abstract. In the context of an experimental virtual-reality surgical planning software platform, we propose a fully self-assessed adaptive region growing segmentation algorithm. Our method successfully delineates main tissues relevant to head and neck reconstructive surgery, such as skin, fat, muscle/organs, and bone. We rely on a standardized and self-assessed region-based approach to deal with a great variety of imaging conditions with minimal user intervention, as only a single-seed selection stage is required. The detection of the optimal parameters is managed internally using a measure of the varying contrast of the growing regions. Validation based on synthetic images, as well as truly-delineated real CT volumes, is provided for the reader’s evaluation.

Key words: CT, segmentation, region-growing, seed, muscle, bone, fat, surgical planning, virtual reality

1 Introduction

One of the most promising applications of medical image computerized visualization is virtual reality surgical planning. Traditional surgical planning uses volumetric information stored in a stack of intensity-based images, usually from computerized tomography (CT) scanners. Based on a number of these image slices, surgeons build their own mental 3D model of the relevant tissues. This task is difficult, even for experienced surgeons. As a consequence, they can miss important information or draw incorrect conclusions due to anatomical variability, either of which can lead to suboptimal treatment strategy decisions [1].

* This work was supported by "Fundación Reina Mercedes" from "Hospital Universitario Virgen del Rocío" (Sevilla), and "Consejería de Salud de la Junta de Andalucía". Carlos S. Mendoza was supported by a doctoral scholarship financed by Universidad de Sevilla.

Using volumetric renderings of anatomical structures, and the appropriate virtual tools for basic surgical operations (like tissue excision and repositioning), the complexity of many plastic surgery interventions can be addressed prior to the actual physical procedure.

A main bottleneck for these computer environments is the delineation of the tissues involved, to such an extent that automated approaches become mandatory. Automatic segmentation is a fundamental problem exhaustively addressed in the literature. Any inaccuracies in the process can distort the simulated measures and surgical operations. In this paper we propose a novel segmentation strategy in the context of the development of a virtual surgical planning environment.

The environment under study was initially conceived for the simulation of head and neck reconstructive surgery, providing virtual tools for tissue excision and repositioning, tissue quantification and stereolithographic prototyping. In such a framework the need for proper delineation of diverse tissues like skin, fat, muscles, organs and bone becomes crucial. Although formerly relying on very simple segmentation methods like thresholding and simple region growing [2–5], only further developments in automatic segmentation approaches, deprived of user parameters, can remain useful for the clinical practitioners and surgeons. According to our on-field inquiries, only a seed-selection stage seems to be tolerated in such a non-technical environment. There is a strong restriction in the operating human time, so parameter selection becomes unaffordable.

In order to pay back the cost and burden of this virtual surgical planning platform development, a wide range of situations should be covered with the proposed technique. Most available physical resources, like imaging devices, should be compatible with the method, even ensuring backwards compatibility (for images acquired in the past). As a consequence neither resolution, contrast nor SNR specific standards can be expected. Further, no imaging protocol can be presumed, as related to patient positioning in the scanner, presence of radioactive contrast, body segment of the patient to be imaged and so on.

Few authors have referred to the issue of classifying a concrete set of tissues using a common method. In their work, Zhou et al. [6] developed a technique for skin, fat, muscle/organs and bone segmentation. Their approach consisted mainly of threshold selection except for bony tissue, for which they made use of a self-assessed adaptive region growing algorithm. Their threshold selection method, based on hierarchical discriminant analysis made assumptions on the histogram that turned out to be unaffordable in our less predictable context. Their strategy for bony tissue, that had been earlier proposed for bronchus segmentation by Law and Heng [7], computed the optimal adaptive threshold by detecting sudden increases in the segmented volume. Its main weakness is the need for an empirical range in this increase for distinguishing routinary growth from undesired leaking. This range would hardly be established in our more general problem. Apart from manual trial-and-error adaptive threshold selection [8], other self-assessed adaptive region growing strategies, outside our context of application, have been proposed in the past. In their work [9], Hojjatoleslami and

Kittler proposed a method based on finding the global maxima for two different contrast measures which they computed iteratively as intensity-decreasing pixels were added to the segmented region. The success of the assessment was founded on the assumption that maximal contrast occurred on region boundaries, which is a reformulation of approaches assuming that the variation of the gray values within regions is smaller than across regions, an inherent assumption in all region growing techniques [10]. Unfortunately, the exhaustivity of their per-pixel approach entailed very low computational efficiency. Revol-Muller et al. [11] used morphological measures to assess the multiplier of the adaptive range for region growing. Instead of computing their assessment function for every pixel addition to the region, they sampled the function for an evenly-spaced set of values.

In our method we propose an assessment function based on the evolving contrast of the region-growing sequence. This strategy allows for segmentation of images without a bimodal histogram requirement as opposed to the assessment measure proposed by Revol-Muller et al. [11]. To make this approach computationally feasible in 3D we produce only evenly-spaced samples of this function along the values of the multiplier for the adaptive ranges that extend around the iteratively estimated mean. We guarantee the sufficiency of the sampling resolution by setting it to a small percentage of the continuously updated standard deviation of the grown region. Globally, our goal is to provide fast automatic segmentation based only on a seed selection step. The results should be comparable to those obtained by manually-tuned region-growing approaches, in a very wide variety of imaging conditions.

2 Method

2.1 Tissue Model

Since our goal is providing a mechanism for segmentation of skin, fat, muscle/organs and bone tissues with minimal user intervention, we have established a model for these tissues that takes into account their intensity distributions in CT images.

First of all, we consider here only tissue segments that exhibit an inherent density and thus an average intensity in the image domain. We model then our object of interest as a connected region whose pixel intensities are sampled from a Gaussian distribution with unknown mean and standard deviation. We may presume that our tissues of interest are surrounded by other tissues derived from other, sometimes adjacent, intensity distributions, like other authors have stated [12]. Although common in the literature, this assumption for the intensities is rarely met in practice, in the sense that the intensity distributions of tissues are only approximately Gaussian. To deal with this inconvenience, and also with partial overlap of distributions between tissues, we propose the use of an assessment function that is to be evaluated along a sequence of region growing stages (region growing sequence).

2.2 Segmentation Algorithm

Normalization and Denoising. Since our method was conceived for images from a wide range of scanners and acquisition protocols, we decided to develop a normalizing stage that could account for such variability. As we will introduce later on, for the self-assessed region growing stage of the algorithm we require the input intensity dynamic range to be normalized with respect to some parameter estimates of the objective intensity distribution.

In the following equations N is a cubic neighborhood of radius R around the seed, \mathbf{x} is a voxel position, $f(\mathbf{x})$ is the intensity for voxel at \mathbf{x} , \bar{f}_N is the mean intensity estimate in N and $|N|$ is the cardinality of N . Moreover, σ_{f_N} is the estimated standard deviation for intensities in N , K is a constant parameter, and $f(\mathbf{x})$, $f'(\mathbf{x})$ are the input and output intensities for the non-linear mapping described below.

$$\bar{f}_N = \frac{1}{|N|} \sum_{\mathbf{x}_k \in N} f(\mathbf{x}_k) \quad , \quad (1)$$

$$\sigma_{f_N} = \sqrt{\frac{1}{|N|} \sum_{\mathbf{x}_k \in N} (f(\mathbf{x}_k) - \bar{f}_N)^2} \quad , \quad (2)$$

$$f'(\mathbf{x}) = \left(1 + \exp \left(-\frac{f(\mathbf{x}) - \bar{f}_N}{\left(\frac{K\sigma_{f_N}}{3}\right)} \right) \right)^{-1} \quad . \quad (3)$$

In a first step we proceed by maximum-likelihood (ML) estimation of the mean and standard deviation as in (1-2), and then perform a non-linear normalization using a sigmoidal transfer function centered on the estimated mean as in (3). The width of the sigmoidal window extends $K\sigma$ around the center \bar{f}_N of the mapping. For $K = 3$ the width of the window would be enough to map 99.7% of the samples, of a Gaussian distribution with similar mean and standard deviation. Greater values of K ensures robust mapping for the estimated distribution. The sigmoidal mapping has been chosen because of its smoothness, and its ability to focus the output dynamic range on a given input intensity range of interest.

Finally, we perform non-linear denoising using an in-slice bidimensional median filter with kernel radius Γ . Other denoising schemes would be valid, always keeping in mind that edge preservation is crucial in our approach. We chose to use bidimensional median filtering because it is as fast as using a smoothing kernel, and preserves edges better than most linear filtering techniques.

Self-Assessed Region Growing. Departing from a normalized version of the image under study, whose intensities lie in the range $[0, 1]$, and a manually provided seed, we perform the self-assessed contrast-maximizing algorithm, for which a generic iteration is described in the following steps:

1. Update multiplier $k_i = k_0 + i\Delta k$
2. Compute, in last iteration grown region R_{i-1} , ML estimates for the mean (available from last iteration) and standard deviation $(\bar{f}'_{R_{i-1}}, \sigma_{f'_{R_{i-1}}})$

3. For every candidate voxel $\bar{\mathbf{x}}_{c_{i-1}}$ being 26-connected to R_{i-1} , $\mathbf{x}_{c_{i-1}} \in R_i$ if

$$f'(\mathbf{x}_{c_{i-1}}) \in \left[\bar{f}'_{R_{i-1}} \pm k_i \sigma_{f'_{R_{i-1}}} \right] \quad (4)$$

4. Compute the assessment function $O_i(\bar{f}'_{R_i}, \bar{f}'_{P_i})$ using the intensity average \bar{f}'_{R_i} in R_i and the intensity average \bar{f}'_{P_i} in the external perimeter P_i of R_i according to (1) and the following eqs.:

$$P_i = \{x_{c_i}\} \cap R_i^C, \quad (5)$$

$$O_i(\bar{f}'_{R_i}, \bar{f}'_{P_i}) = \left| \frac{\bar{f}'_{P_i} - \bar{f}'_{R_i}}{\bar{f}'_{P_i} + \bar{f}'_{R_i}} \right| \quad (6)$$

5. If O_{i-1} was a local maximum, when compared to O_{i-2} and O_i (only when $i \geq 2$), then the algorithm stops and the output is R_{i-1} . Otherwise another iteration takes place

Of all aforementioned parameters only k_0 and Δk are critical for the performance of the algorithm. k_0 affects computational efficiency requiring a greater number of iterations before a local maximum of $O(\bar{f}'_{R_i}, \bar{f}'_{P_i})$ is found. Therefore it should be set to the largest possible value that guarantees that only a small percentage of tissue intensity samples are included in the first iteration, in order not to miss the first local maximum. In what concerns Δk , the choice must guarantee that the assessment function is being sampled adequately in order to detect its local variations. Since the estimates for the mean and standard deviation are continually updated as the region grows, the estimates become increasingly close to the theoretical values. We argue that setting Δk below one tenth of 3 (which is the theoretical value multiplying the standard deviation of a Gaussian distribution for 99.7% of its samples to be included in a range of that width around the mean) is enough for the segmentation process to be able not to miss the available local maxima of the assessment function. This postulate is supported by our experimental results.

3 Results

We have implemented our algorithm using open source medical image processing libraries, more precisely the Insight Toolkit for algorithm development, and the command line executable module infrastructure provided by 3DSlicer for fast prototyping, calibration, evaluation, and manual segmentation on real images for further validation [13]. The algorithm that we will validate, and that was finally implemented in the virtual reality platform that motivated its development, uses the following parameter values: $R = 2$, $K = 12$, $\Gamma = 1$, $k_0 = 1$ and $\Delta k = 0.1$.

For validating our algorithm we have proposed several experiments, based on synthetic as well as real images. In their inspiring work, Udupa et al. [14] proposed a methodology for the validation of medical volume segmentation algorithms. For a segmentation algorithm to be proven useful it has to demonstrate its accuracy (quality of results), precision (parameter independence) and efficiency (human and computational times).

Accuracy is evaluated in terms of False Positive and False Negative Volume Fraction ($FPVF$ and $FNVF$). $FPVF$ compares the number of voxels assigned to the object which actually belong to the background, with the number of voxels which compose the background. $FNVF$ quantifies the number of voxels assigned to the background, which actually belong to the object, as compared to the number of voxels in the object. To produce such a comparison a true delineation of the object must be available. This ground-truth can be accomplished either by manual human-expert segmentation; or creating synthetic images, corrupting them with simulations of typical acquisition artifacts, and then performing segmentation on the result.

$$FPVF = \frac{|\text{Segmented Region} - \text{True Region}|}{|\text{Entire Volume} - \text{True Region}|}, \quad (7)$$

$$FNVF = \frac{|\text{True Region} - \text{Segmented Region}|}{|\text{True Region}|}, \quad (8)$$

where $|\cdot|$ indicates again cardinality and subtracting means performing a set intersection with the complement of the subtrahend.

Precision is evaluated according to the so-called precision quotient (PR). Precision can be computed in terms of inter/intra-operator variability and inter-scanner variability. Since our method relies on only a seed selection procedure, and is supposed to work for a variety of acquisition devices, we decided to compute precision for inter-seed variability. This variability is accounted for by comparing the resulting segmentation from differently placed seeds.

$$PR = \frac{|\text{Segmentation 1} \cap \text{Segmentation 2}|}{|\text{Segmentation 1} \cup \text{Segmentation 2}|} \quad (9)$$

Several seeds can be used so that (9) is computed for all possible combinations of outputs, and then averaged to obtain a more representative measure.

Efficiency relates to the segmentation performance time, human as well as computational time. According to Udupa et al. [14], it is extremely hard to provide an efficiency measure that proves useful when comparing different methods. One possible way of summing up all efficiency factors is related to the economic cost of usage for an algorithm. Human time is much more expensive than computer time. In our method, usage is designed to be extremely simple, and reduces to placing one pin point on the desired object. For that reason we have computed here only computational times for the execution of our algorithm.

3.1 First Experiment. Constant-valued spheres

To validate the claim that our contrast-based assessment function accurately detects homogeneous-intensity regions, we have created synthetic volume images and corrupted them with some typical CT acquisition artifacts. We have created a 3D volume composed by two ideal tissues, with constant intensity. The first tissue is shaped as a ball centered in the image domain, with radius $r = 20$ voxels.

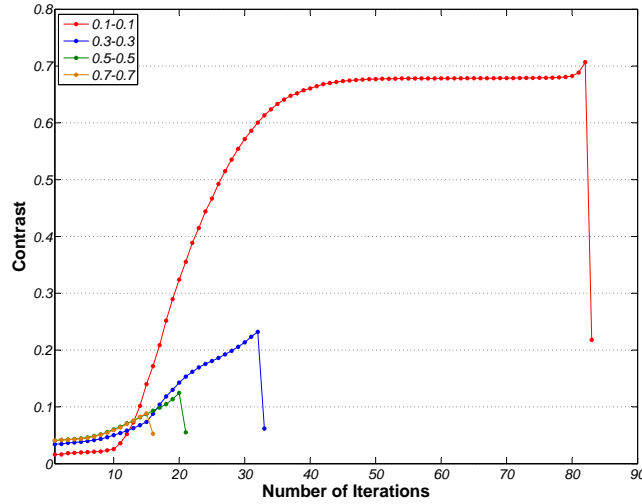


Fig. 1. Contrast evolution for some constant-valued spheres

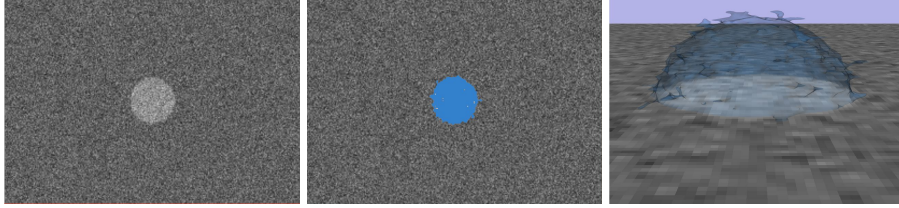


Fig. 2. Synthetic volume slice, segmentation and reconstruction

The second tissue is the background which extends to a final image domain with size $512 \times 512 \times 100$ voxels. We argue that this simple phantom is enough to test the effect of artifacts on boundary detection as long as connectivity is preserved. Our method depends only on connectivity and intensity shifts, so the precise shape of the tissue is irrelevant as long as it is connected.

For simulating CT acquisition, we proceed by blurring the image (accounting for partial volume effect) and adding Gaussian noise. Blurring was performed using a Gaussian smoothing kernel with width σ_b . Zero-mean Gaussian noise is generated with standard deviations σ_n . We produce segmentations for all combinations for $\sigma_b \in [0.1, 0.9]$ in increments of 0.2, and $\sigma_n \in [0.1L, 0.7L]$ (with L the absolute intensity difference between the two tissues) in increments of 0.2.

From all the segmentations we compute $FPVF$ and $FNVF$. The average value for these measures is provided in Table 1. For a representative case ($\sigma_b = 0.5$, $\sigma_n = 0.5$), we have computed PR for all possible combinations of three

different seeds, selected randomly inside the region of interest. The average PR is presented in Table 1. For efficiency evaluation, computational time t_c has been computed and averaged across all images. Notice in Table 1 how the obtained

Table 1. Computed measures for constant-valued spheres

$FPVF$	$2.46 \cdot 10^{-4} \pm 1.06 \cdot 10^{-4}$
$FNVF$	0.0205 ± 0.0313
PR	0.925 ± 0.0413
t_c (s)	72.6 ± 40.8

values for $FPVF$ and $FNVF$ are close to zero, indicating very high segmentation fidelity. PR is close to 1, indicating weak dependence on seed placement. Times show great variance due to the iterative nature of the algorithm.

For illustrative purposes we have included in Fig. 1 a graphical representation of the evolution of the assessment function as the algorithm iterates. Notice how the fall after the peak decreases as noise and blurring increase.

Also, in Fig. 2 we can see a slice of the generated volume for the case 0.5-0.5 and also a surface reconstruction of the segmented region. Notice the good results facing noise with standard deviation as high as half the intensity difference between the two tissues.

3.2 Second Experiment. Continuous-valued spheres

For this experiment we produced a similar synthetic image, only now the intensity inside the ball varies from 0 to 1 proportionally to the Euclidean distance from the center of the voxel to the center of the ball. The intensity value for the background was set to 1. In this scheme, no clear boundary is available, because we wanted to prove that our method does not require abrupt intensity changes for boundaries to be detected. We corrupted the image only with Gaussian noise of standard deviation $\sigma_n = 0.1, 0.3$ and 0.5 . We computed the same accuracy and efficiency measures, as well as precision for the case $\sigma_n = 0.3$. Due to the nature of the values in the regions, seeds must be placed close to the center in order for the condition (which is inherent to any region growing approach) of greater variance across than inside regions to be met [10]. According to Table

Table 2. Computed measures for continuous-valued spheres

$FPVF$	$4.12 \cdot 10^{-4} \pm 4.44 \cdot 10^{-4}$
$FNVF$	0.191 ± 0.110
PR	0.930 ± 0.0176
t_c (s)	257 ± 12.9

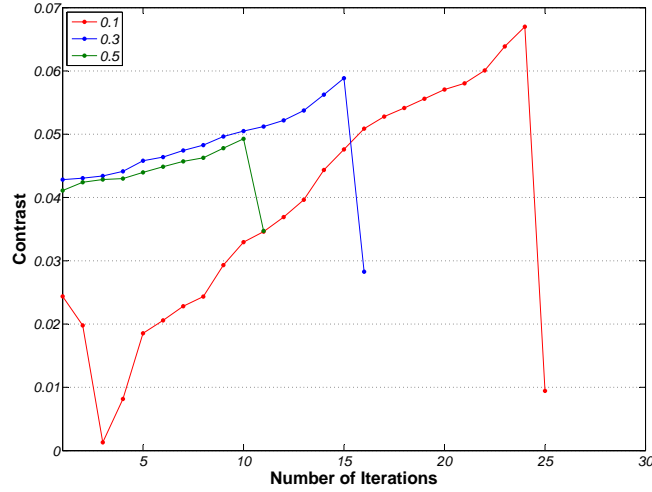


Fig. 3. Contrast evolution for some continuous-valued spheres

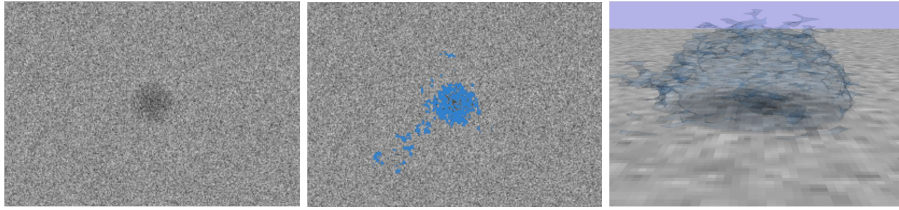


Fig. 4. Synthetic volume slice, segmentation and reconstruction

2, the results for this extremely subtle boundary are still acceptable. $FPVF$ and $FNVF$ stay somewhat close to 0, and PR value is close enough to 1 as to support the claim of low seed location dependence. Time has increased due to the greater variance of the segmented tissue, which forces the algorithm to perform more iterations.

For Fig. 3 we can observe the same effect as in Fig. 1. Notice the increased roughness in these curves as compared to those in Fig. 1. This is due to a greater influence of noise on consecutive iterations of the region growing sequence, due to the non-constant intensity of the tissue.

We present in Fig. 4 again a slice of the generated volume for the case 0.5 and its segmentation reconstruction. The quality of the segmentation is pretty good even for extremely dim boundaries. This proves that our method detects not only intensity shifts, but also intensity evolution shifts.

3.3 Third experiment. Real CT images

For this last experiment we have produced automatic segmentations for 10 real CT images for the tissues described (skin, muscle/organs, fat and bone). The testing set proceeds from all different scanners involved in our clinical setting. As explained above, validation of real CT images requires manual segmentations provided by a clinical expert. In our application context this manual segmentation process can be extremely time-consuming, or even intractable, due to the lack of tissue localization. Just to give an example, manual segmentation of muscular tissue implies manually avoiding all blood vessels and fat traces, for up to 500 slices. For this reason we have computed our accuracy and precision metrics from just 20 slices in 10 cases.

Moreover, the manual accuracy of the segmentation is bounded by human perceptual limitations presenting considerable inter-subject variability, and we (the authors and the clinical practitioners) have observed that in many cases poorer results in accuracy are related to incomplete manual segmentations, rather than incomplete automatic segmentations. Let us say then, that the quality of the segmentation is *at least* as good as the presented results.

For the results presented in Table 3, all segmentations have been used for accuracy and efficiency assessment, and one particular segmentation and several seeds for precision. While *FPVF* stays very low, ensuring self-contained segmented regions, *FNVF* is still reasonably close to 0. Precision is not as good as for the synthetic images, due to the more complex shape of the segmented regions, which produce a slightly greater dependence on seed placement. Computational time ranges between 2 or 3 minutes, which implies a great reduction as compared to previous trial-and-error parameter tuning, according to non-technical users' opinions.

Again illustration for contrast evolution and a rendering of the resulting segmented volume are presented in Figs. 5 and 6. All results were obtained from a single seed selection in each image.

Table 3. Computed measures for real CT images

<i>FPVF</i>	$9.2 \cdot 10^{-3} \pm 1.75 \cdot 10^{-3}$
<i>FNVF</i>	0.151 ± 0.0815
<i>PR</i>	0.733 ± 0.171
t_c (s)	156.0 ± 36.2

4 Concluding Remarks, Limitations and Future Work

Considering the numeric results for the experimental validation we consider our approach successful in solving the particular needs for which it was conceived.

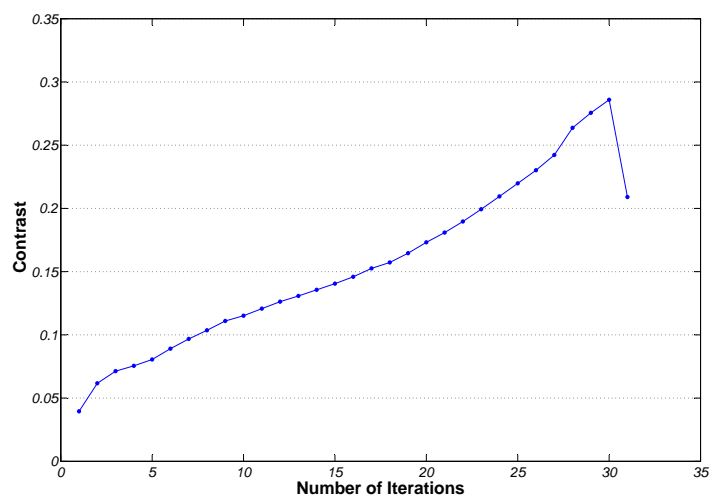


Fig. 5. Contrast evolution for real CT segmentation

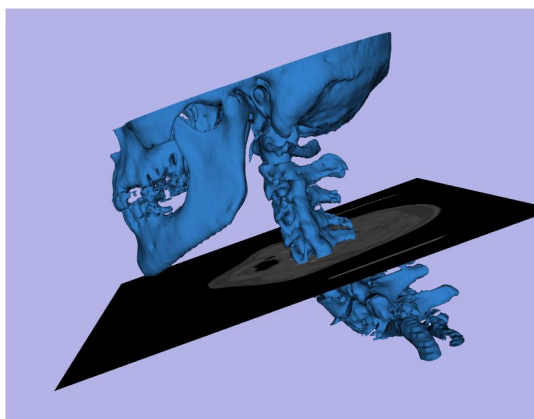


Fig. 6. Real CT volume slice and surface model for performed segmentation of bone tissue

The algorithm has been integrated in the reference platform and prevents from previous trial-and-error segmentation, which was very time-consuming according to its users. Its greatest advantage is thus the absence of tuning parameters and ability to produce nice results for a wide set of acquisition devices.

Its most significant limitation to this region growing approach has to do with the fact that connectivity is sometimes too weak as a requirement for some regions of interest, as they may be connected to other similar intensity regions. For

this reason a future improvement for the technique could be incorporating some morphological limitations to the growth of the region in order to impose some degree of *stiffness*, to avoid flooding towards undesired regions. This improvement could make the algorithm useful in greater variety of situations (individual organs, tumors...).

References

1. Reitinger, B., Bornik, A., Beichel, R., Schmalstieg, D.: Liver surgery planning using virtual reality. *IEEE Comput. Graph. Appl.* **26**(6) (2006) 36–47
2. Zucker, S.W.: Region growing: Childhood and adolescence. *Computer Graphics and Image Processing* **5**(3) (1976) 382–399
3. Sivewright, G.J., Elliott, P.J.: Interactive region and volume growing for segmenting volumes in MR and CT images. *Medical Informatics* **19**(1) (1994) 71–80
4. Sekiguchi, H., Sano, K., Yokoyama, T.: Interactive 3-dimensional segmentation method based on region growing method. *Systems and Computers in Japan* **25**(1) (1994) 88–97
5. Suárez, C., Acha, B., Serrano, C., Parra, C., Gómez, T.: VirSSPA- a virtual reality tool for surgical planning workflow. *International Journal of Computer Assisted Radiology and Surgery* **4**(2) (2009) 133–139
6. Zhou, X., Kamiya, N., Kara, T., Fujita, H., Yokoyama, R., Kiryu, T., Hoshi, H.: Automated recognition of human structure from torso CT images. In: *Proceedings of the Fourth IASTED International Conference on Visualization, Imaging, and Image Processing*. (2004) 584–589
7. Law, T.Y., Heng, P.A.: Automated extraction of bronchus from 3D CT images of lung based on genetic algorithm and 3D region growing. In: *Proceedings of SPIE - The International Society for Optical Engineering*. Volume 3979. (2000) 1/–
8. Adams, R., Bischof, L.: Seeded region growing. *IEEE Trans. Pattern Anal. Mach. Intell.* **16**(6) (1994) 641–647
9. Hojjatoleslami, S.A., Kittler, J.: Region growing: A new approach. *IEEE Trans. Image Process.* **7**(7) (1998) 1079–1084
10. Haralick, R.M., Shapiro, L.G.: Image segmentation techniques. *Computer Vision, Graphics, & Image Processing* **29**(1) (1985) 100–132
11. Revol-Muller, C., Peyrin, F., Carrillon, Y., Odet, C.: Automated 3D region growing algorithm based on an assessment function. *Pattern Recognition Letters* **23**(1-3) (2002) 137–150
12. Jian, W., Feng, Y., Ma, J.L., Sun, X.P., Jing, X., Cui, Z.M.: The segmentation and visualization of human organs based on adaptive region growing method. In: *Proceedings - 8th IEEE International Conference on Computer and Information Technology Workshops, CIT Workshops 2008*. (2008) 439–443
13. Pieper, S., Lorensen, B., Schroeder, W., Kikinis, R.: The NA-MIC Kit: ITK, VTK, pipelines, grids and 3D Slicer as an open platform for the medical image computing community. In: *2006 3rd IEEE International Symposium on Biomedical Imaging: From Nano to Macro - Proceedings*. Volume 2006. (2006) 698–701
14. Udupa, J.K.: Multiple sclerosis lesion quantification using fuzzy-connectedness principles. *IEEE Trans. Med. Imag.* **16**(5) (1997) 598–609