

Segmentation and Evaluation of Adipose Tissue from Whole Body MRI Scans

Yinpeng Jin¹, Celina Z. Imielińska², Andrew F. Laine¹, Jayaram Udupa³, Wei Shen⁴,
and Steven B. Heymsfield⁴

¹Department of Biomedical Engineering, Columbia University

²College of Physicians and Surgeons, Office of Scholarly Resources
Department of Medical Informatics and Department of Computer Science
Columbia University

³Medical Image Processing Group, Department of Radiology, University of Pennsylvania

⁴Obesity Research Center, St. Luke's-Roosevelt Hospital and Institute of Human Nutrition,
Columbia University College of Physicians and Surgeons

Abstract. Accurate quantification of total body and the distribution of regional adipose tissue using manual segmentation is a challenging problem due to the high variation between manual delineations. Manual segmentation also requires highly trained experts with knowledge of anatomy. We present a hybrid segmentation method that provides robust delineation results for adipose tissue from whole body MRI scans. A formal evaluation of accuracy of the segmentation method is performed. This semi-automatic segmentation algorithm reduces significantly the time required for quantification of adipose tissue, and the accuracy measurements show that the results are close to the ground truth obtained from manual segmentations.

1 Introduction

Adipose tissue quantification plays a central role in studying obesity in children and adults. The statistics show that 60% of American adults are either overweight or obese, and the rates are skyrocketing in children and adolescents [1]. Obesity is strongly linked to morbidity and mortality rate [2, 3]. The recent report of a striking prevalence of metabolic syndrome is directly related to excess visceral adiposity [4]. There is great interest in studying regional adipose tissue. At present, CT scans provide the best imaging for *visceral adipose tissue* (VAT), but radiation exposure is prohibitive, especially in children and young women. Therefore multiple CT scans are not admissible. The availability of MRI technology for whole body scans made it an attractive imaging modality for quantification of adipose tissue. At present, most investigators quantify VAT from a single CT or MRI slice.

The current proprietary software tools for segmentation of adipose tissue from whole body MRI scans is very laborious and time consuming and requires highly trained experts/technicians [5, 6]. The methods rely heavily on manual delineation and/or low level semi-manual segmentation, (e.g. thresholding and histogram-based region growing) that are generally successful with high-contrast images [7]. Since

most MRI scanners are not free of magnetic field inhomogeneities and chemical shift artifacts, segmentation of MRI images requires substantial operator involvement. This approach is very time consuming, since it takes 5-6 hours per subject to separate subcutaneous adipose and visceral adipose tissue, skeletal muscle and lungs from the remaining organs. Regional adipose tissue has a high coefficient of variation (CV), a study shows that the CV of visceral and adipose tissue is 9.4-17.6% and 2.1-4.9 %, respectively [8]. No sufficient data on the evaluation of intermuscular adipose tissue quantification is published. There are problems related to image quality and artifacts: non-uniform RF coil response, Figure.1(a,c,d); poor contrast, Figure.1(b); chemical shift, Figure.1(c); that pose challenge for automated, accurate and efficient image segmentation [9]. Finally, while image segmentation remains a very challenging problem, evaluation of segmentation methods generally lacks a consensus within the medical image processing community.

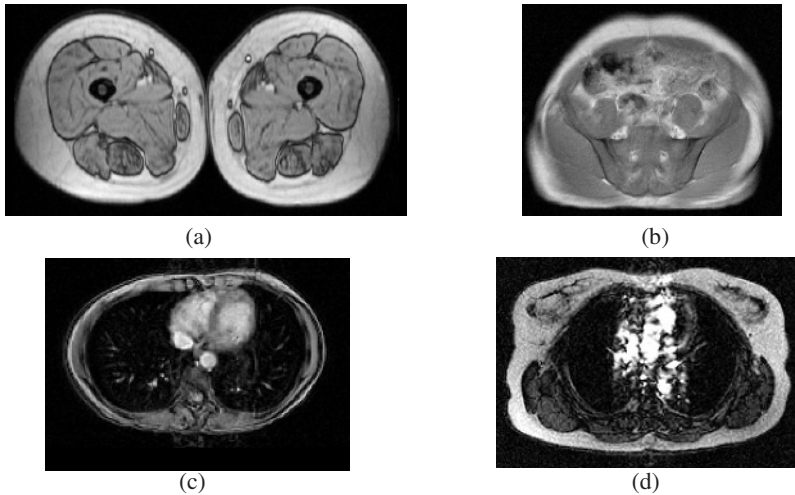


Fig. 1. (a) Non-uniform RF coil response and black boundary artifacts. (b) Poor contrast and non-uniform RF coil response. (c) Chemical Shift artifacts. (d) Blood flow artifacts, non-uniform RF coil response and poor signal-to-noise ratio.

2 Methods and Materials

We use a hybrid segmentation method for processing both radiological and the Visible Human data [10, 11]. We follow the evaluation framework, as reported in [12], to assess accuracy of the segmentation of adipose tissue. There is no single segmentation method that can yield acceptable results for every application domain (i.e. an Application, a Body region and an imaging Protocol $\langle A, B, P \rangle$). Therefore, we can only evaluate performance of a segmentation method in the context of a specific $\langle A, B, P \rangle$ [12].

2.1 Hybrid Image Segmentation Methodology

Hybrid segmentation integrates boundary-based and region-based algorithms that amplify the strength and reduce the weakness of both approaches, and can yield high precision, accuracy and efficiency [10, 11, 13, 14]. We have built and tested, as a part of a large effort to provide an open source Segmentation and Registration Toolkit (ITK, <http://www.itk.org>) funded by the National Library of Medicine, a hybrid segmentation method [11] that combines fuzzy connectedness segmentation [15], Voronoi Diagram classification [16] and a deformable model based smoothing algorithms [18].

2.1.1 Fuzzy Connectedness Segmentation

Fuzzy connectedness was introduced by Udupa [15] and has been successfully used for segmentation of multi-channel images in several applications. This method uses the fact that medical images are inherently inhomogeneous. Object segmentation is achieved by defining a group of pixels that show a certain level of global hanging togetherness (fuzzy connectedness). We define *affinity* between two *elements* in an image (e.g. pixels, voxels) via a degree of *adjacency* and the *similarity* of their intensity values. A *global* fuzzy relation, called fuzzy connectedness, is defined in the image by assigning to every pair of elements a strength measurement of *global* hanging togetherness. The strength of a path connecting two elements is defined as the “weakest” link, e.g. the lowest *affinity* value along the path. The strength of fuzzy connectedness between two pixels is defined as the strongest path among all paths connecting them. A fuzzy scene map representing fuzzy connectedness value between each pixel in the image and a seed pixel is computed using dynamic programming [15]. A user-defined threshold applied to the fuzzy scene map results in a segmented object with the selected level of fuzzy connectedness (See Figure 2).

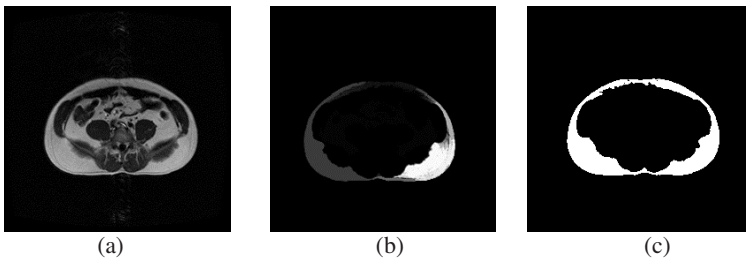


Fig. 2. Fuzzy Connectedness Segmentation. (a) Input image. (b) Fuzzy Scene Map. (c) Result: segmented object with fuzzy connectedness level value (a user-defined threshold) of 0.025.

2.1.2 Voronoi Diagram Based Classification and Segmentation

This algorithm, described in detail in [16], is based on repeatedly dividing an image into regions using Voronoi diagram [17] and classifying the Voronoi regions based on an homogeneity classifier for the segmented organ/tissue. The region-based homogeneity operator determines the performance of the segmentation. The simplest homogeneity test can be achieved by measuring the mean and standard deviation of the target object [16]. More accurate object description and homogeneity operator has

been tested to provide more accurate segmentation [10]. In Figure 3, we show an example of segmentation of visceral adipose tissue in a single slice of color Visible Human data.

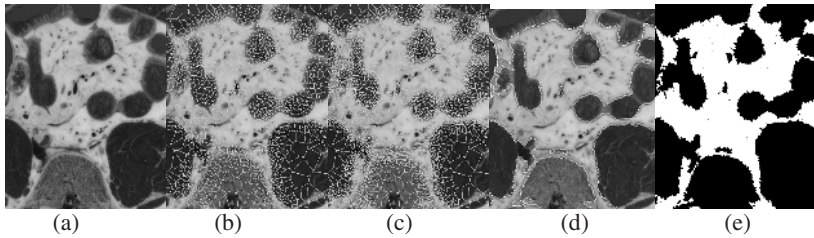


Fig. 3. Voronoi Diagram classification. (a) Input image. (b)(c): Voronoi regions after 2 and 8 iterations, respectively. (d) The final boundary. (e) The final segmented region: a binary object.

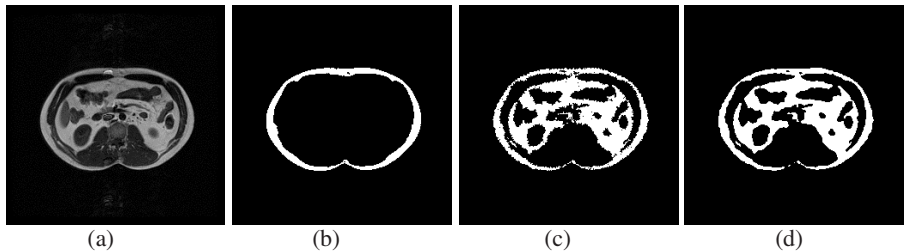


Fig. 4. Hybrid segmentation. (a) Input image (MRI T1). (b) Fuzzy connectedness segmentation. (c) Voronoi diagram segmentation using (b) as a *prior*. (d) A volume-preserved smoothing on (c).

2.1.3 Hybrid Segmentation: Integration of Fuzzy Connectedness and Voronoi Diagram Classification Methods

In our hybrid segmentation, we integrate the fuzzy connectedness algorithm with the Voronoi Diagram classification. We aim to improve the robustness and performance of the segmentation, and to reduce the need for user interactions. First, a human operator manually picks few pixels inside the object, and small regions, e.g. 5 by 5 pixels centered at the pixels are collected to compute an estimation of the mean and variance of the pixel intensity for adipose tissue. Such a procedure needs to be done once for one whole body MRI scan. With the estimated mean and variance, we invoke the fuzzy connectedness algorithm and generate a fairly reasonable segmentation for a sample connected adipose tissue in the image, a *prior*. From the segmented sample tissue, a more accurate homogeneity operator is derived to classify regions in the Voronoi diagram segmentation [11]. We generate Voronoi diagram from randomly distributed seed points over the image. The Voronoi regions are classified as interior, exterior and boundary regions. The boundary regions are subdivided by adding seed points to their edges and re-compute the Voronoi. Then, the Voronoi regions are classified again. We iterate the algorithm until the boundary regions converge to the final segmentation and each boundary region reaches an area less than a selected threshold

[11]. The Voronoi diagram classification yields a boundary with a “noisy” appearance. Finally, a level-set based volume-preserved smoothing algorithm [18] is applied to smooth final result. In Figure 4 we show an example of segmentation of adipose tissue of a single MRI T1 weighted image.

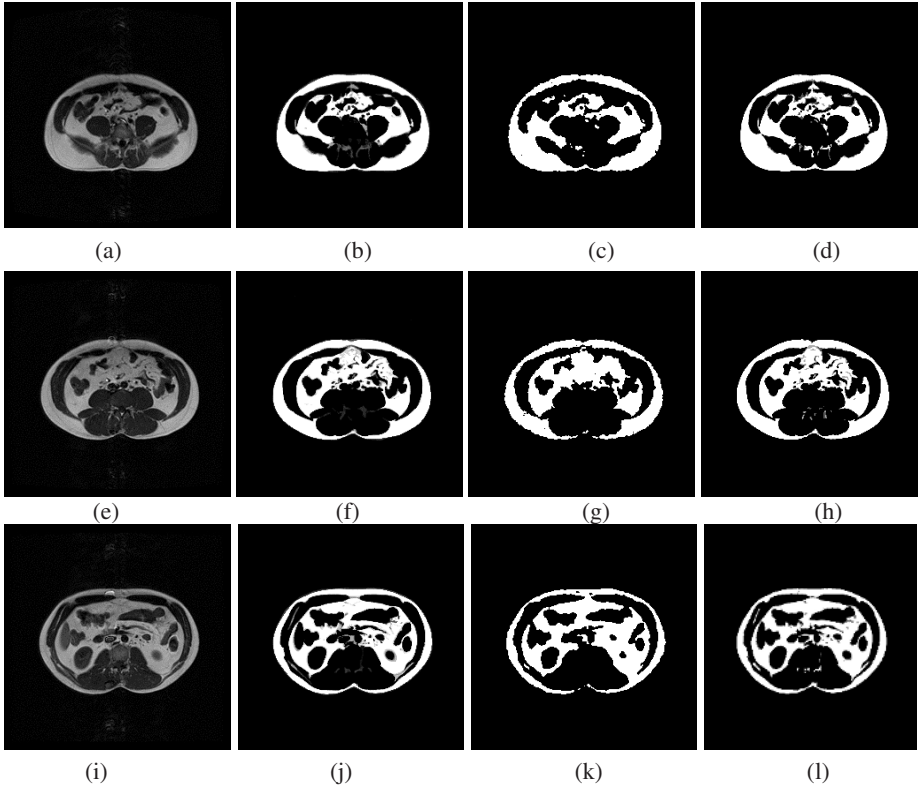


Fig. 5. Results depicting hybrid and manual segmentation: (a)(e)(i) input images (MRI T1); (b)(f)(j): ground truth images; (c)(g)(k): hybrid segmentation; (d)(h)(l): manual segmentation.

2.2 Evaluation of Segmentation Algorithms

We have recently developed a comprehensive segmentation evaluation methodology [12] in a joint effort between UPenn and Columbia based on our previous experience with related applications in medicine and segmentation methods. Any method of evaluation of segmentation should specify the *application domain* under consideration that is determined by three entities: (A) *an application* (e.g. quantification of adipose tissue), (B) *a body part* (e.g. visceral adipose tissue), and *imaging protocol* (MRI T1 weighted image). A segmentation method in an application domain $\langle A, B, P \rangle$ should be evaluated in terms of three factors: *Precision* which represents repeatability of segmentation taking into account all subjective actions required in producing the

result; *Accuracy*, which denotes the degree to which the segmentation agrees with the ground truth; *Efficiency*, which describes the practical viability of the segmentation method. In this paper, we limit evaluation of our segmentation method to evaluation of *delineation accuracy* (where delineation is the low-level process of determining the precise spatial extent of the object in the scene), due to fact that we are using data from another clinical study. The following measures are defined to characterize the delineation accuracy of a segmentation method: *False Negative Volume Fraction* (FNVF, fraction of tissue that was missed), *False Positive Volume Fraction* (FPVF, the amount of tissue falsely identified) and *True Positive Volume Fraction* (TPVF, fraction of the total amount of tissue in the ground truth with which the delineation obtained by the method overlaps), for mathematical details, see [12]. Those measurements require carefully generated ground truth [12].

2.3 Data Description and the Generation of Ground Truth

In our study, we use the data acquired at the Obesity Research Center, St. Luke's Roosevelt Hospital in New York, under different clinical study of hand-segmentations of hundreds of whole body MRI T1 weighted scans. From the multiple, human subject datasets, two abdomens (*reference* datasets) were selected and segmented repeatedly by all experts involved in the study. Each such reference dataset consists of six slices, with slice thickness of 10mm and inter-slice separation of 40 mm. The reference datasets were segmented (semi)-manually by six experts under a rigorous protocol. Each set was segmented three times, by each expert, with three month intervals between segmentations. All the manual segmentations were aided with an interactive segmentation tool utilizing simple histogram-based thresholding operator and paint-brush. We found the resulting manual segmentation to be suitable for defining a surrogate for true delineation (ground truth). For our study, we used 16 manual delineations (binary masks) of the two abdomens (2 segmentations out of 18 were disqualified due to mislabeling problem), and the ground truth was established by simple averaging of the corresponding binary images into a fuzzy object (with pixel value between 0 and 1). Then, we segmented the data with our semi-automatic hybrid method and compared the results to the ground truth. We also selected three manual segmentations, averaged them to provide another segmentation result, that we call manual segmentation. Then, the three factors (FNVP, FPVF, TPVF) for measuring the segmentation accuracy were computed for both segmentations, and for individual slices in the data (the inter-slice distance in the dataset was too large to treat as a contiguous volume).

3 Experimental Results

The results of hybrid and manual segmentations for one of the reference datasets, are presented in Figure 5. It shows three slices from the reference dataset (a)(e)(i), and corresponding slices of: ground truth (b)(f)(j), hybrid segmentation (c)(g)(k), and manual segmentation (d)(h)(l).

Table 1. Accuracy measurements of the hybrid segmentation shown in Figure 5 (c), (g) and (k).

Ground Truth	Area difference (%)	FNVP (%)	FPVF (%)	TPVF(%)
Fig. (b)	1.3	8.2	9.4	91.8
Fig. (f)	0.2	7.7	7.9	92.3
Fig. (j)	3.9	8.1	11.9	91.9

Table 2. Accuracy measurements of the manual segmentation shown in Figure 5 (d), (h) and (l).

Ground Truth	Area difference (%)	FNVP (%)	FPVF (%)	TPVF(%)
Fig. (b)	2.4	5.3	7.4	94.7
Fig. (f)	2.0	6.3	5.9	93.7
Fig. (j)	4.1	4.1	8.2	95.9

Table 1 and Table 2 give the accuracy measurements for the hybrid and manual segmentations compared with the ground truth, respectively. Since the ground truth was built upon the manual delineations, these “self” consistency measurements in Table 2 should be the best accuracy measurement one can expect from any segmentation results. As we can see in Table 1 and Table 2, the simple measurement of area difference does not provide accurate evaluation in terms of overall performance. The three factors: FNVP, FPVF, TPVF give better evaluation of the accuracy measurement. Our hybrid segmentation provides about 8~9% accuracy, while the best accuracy result for the manual segmentation is 5~7%. To assess efficiency of segmentation, an experienced human operator delineates adipose tissue of a 6-slice MRI T1 dataset in about 15~20 minutes. While using our hybrid segmentation, only a few mouse-clicks are needed, and adipose tissue in a 2D image is delineated in real time (less than 1 second).

4 Discussion

We have built and tested a hybrid segmentation tailored to delineate adipose tissue from whole body MRI T1 scans. This approach requires significantly less human involvement than any existing conventional interactive tools. We have evaluated delineation accuracy using three factors: FNVP, FPVF, TPVF. Under the framework, we’ve demonstrated that the results from hybrid segmentation are close to those obtained from manual delineation. The ground truth generated for this study is not ideal, therefore it is recommended to acquire, in the future, new data that will allow a complete evaluation under all three factors: precision, accuracy and efficiency. The same evaluation framework can also be used for testing the variability (CV index) of the hand segmentations, among controlled group of human operators.

References

1. K. Flegal, M. Carroll, R. Kuczmarski, and C. Johnson, "Overweight and obesity in the United States: prevalence and trends," *Int J Obesity*, vol. 22, pp. 38–47, 1998.
2. "NIH Clinical Guidelines on the Identification: Evaluation and Treatment of Overweight and Obesity in Adults—the Evidence Report.," *Obesity Res*, vol. 6, pp. 51S–209S, 1998.
3. F. X. PiSunyer, "Medical Hazards of Obesity.," *Annals of Internal Medicine*, vol. 119, pp. 644–660, 1993.
4. Y. W. Park, S. K. Zhu, L. Palaniappan, S. Heshka, M. R. Carnethon, and S. B. Heymsfield, "The metabolic syndrome: prevalence and associated risk factors.," *In press, Archives and Internal Medicine*.
5. R. Ross, L. Leger, R. Guardo, J. D. Guise, and B. G. Pike, "Adipose tissue volume measured by magnetic resonance imaging and computerized tomography in rats.," *J App. Physiol*, vol. 70, pp. 2164–2172, 1991.
6. R. Ross, "Magnetic resonance imaging provides new insights into the characterization of adipose and lean tissue distribution.," *Can. J. Physiol. Pharmacol.*, vol. 74, pp. 778–785, 1996.
7. J. Rogowska, K. Batchelder, G. S. Gazelle, E. F. Halpern, W. Connor, and G. L. Wolf, "Evaluation of selected two-dimensional segmentation techniques for computed tomography quantitation of lymph nodes.," *Invest Radiol.*, vol. 31, pp. 138–45, 1996.
8. J. Elbers, G. Haumann, H. Asscheman, J. Seidell, and L. R. Gooren, "reproducibility of fat area measurements in young, non-obese subjects by computerized analysis of magnetic resonance images.," *Int J Obes Relat Metab Disord.*, vol. 21, pp. 1121–9, 1997.
9. G. Z. Yang, S. Myerson, F. Chabat, D. J. Pennell, and D. N. Firmin, "Automatic MRI adipose tissue mapping using overlapping mosaics.," *MAGMA.*, vol. 4, pp. 39–44, 2002.
10. E. Angelini, C. Imielinska, Y. Jin, and A. Laine, "Improving statistics for hybrid segmentation of high-resolution multichannel images," *SPIE Annual meeting on Medical Imaging*, vol. 4684:1, pp. 401–411, 2002.
11. C. Imielinska, D. Metaxas, J. Udupa, Y. Jin, and T. Chen, "Hybrid Segmentation Methods of Anatomical Data," *MICCAI, Utrecht, Netherlands*, vol. 2208, pp. 1048–1057, 2001.
12. J. Udupa, V. LaBlanc, H. Schmidt, C. Imielinska, P. Saha, G. Grevera, Y. Zhuge, P. Mollholt, Y. Jin, and L. Currie, "A Methodology for Evaluating Image Segmentation Algorithm," *SPIE Conference on Medical Imaging, San Diego CA.*, vol. 4684, pp. 266–277, 2002.
13. T. N. Jones and D. N. Metaxas, "Automated 3D segmentation using deformable models and fuzzy affinity," *15th International conference, Information Processing in Medical Imaging, Vermont, USA*, pp. 113–126, 1997.
14. T. N. Jones and D. N. Metaxas, "Image Segmentation based on the integration of pixel affinity and deformable models," *Computer Vision and Pattern Recognition, Sant Barbara, CA*, 1998.
15. J. K. Udupa and S. Samarasekera, "Fuzzy connectedness and object definition: Theory, algorithms and applications in image segmentation," *Graphical Models and Image Processing*, vol. 58, pp. 246–261, 1996.
16. C. Imielinska, M. S. Downes, and W. Yuan, "Semi-automated color segmentation of anatomical tissue," *Computerized Medical Imaging and Graphics*, vol. 24, pp. 173–180, 2000.
17. F. P. Preparata and M. I. Shamos, *Computational Geometry*. New York: Springer, 1985.
18. J. Sethian, *Level set methods and fast marching methods*, 1999.