Published in final edited form as:

Acad Radiol. 2013 November ; 20(11): 1413–1421. doi:10.1016/j.acra.2013.08.007.

Computer-aided Assessment of Regional Abdominal Fat with Food Residue Removal in CT

Sokratis Makrogiannis, PhD, Giorgio Caturegli, Christos Davatzikos, PhD, and Luigi Ferrucci, MD, PhD

Longitudinal Studies Section, National Institute on Aging, National Institutes of Health, 3001 South Hanover St., 5th Fl. Baltimore, MD 21225 (S.M., G.C., L.F.); and Section of Biomedical Image Analysis, Department of Radiology, University of Pennsylvania, Philadelphia, PA 19104 (C.D.); Department of Mathematical Sciences, Delaware State University, Dover, DE 19901 (S.M.).

Abstract

Rationale and Objectives—Separate quantification of abdominal subcutaneous and visceral fat regions is essential to understand the role of regional adiposity as risk factor in epidemiological studies. Fat quantification is often based on computed tomography (CT) because fat density is distinct from other tissue densities in the abdomen. However, the presence of intestinal food residues with densities similar to fat may reduce fat quantification accuracy. We introduce an abdominal fat quantification method in CT with interest in food residue removal.

Materials and Methods—Total fat was identified in the feature space of Hounsfield units and divided into subcutaneous and visceral components using model-based segmentation. Regions of food residues were identified and removed from visceral fat using a machine learning method integrating intensity, texture, and spatial information. Cost-weighting and bagging techniques were investigated to address class imbalance.

Results—We validated our automated food residue removal technique against semimanual quantifications. Our feature selection experiments indicated that joint intensity and texture features produce the highest classification accuracy at 95%. We explored generalization capability using k -fold cross-validation and receiver operating characteristic (ROC) analysis with variable k . Losses in accuracy and area under ROC curve between maximum and minimum k were limited to 0.1% and 0.3%. We validated tissue segmentation against reference semimanual delineations. The Dice similarity scores were as high as 93.1 for subcutaneous fat and 85.6 for visceral fat.

Conclusions—Computer-aided regional abdominal fat quantification is a reliable computational tool for large-scale epidemiological studies. Our proposed intestinal food residue reduction scheme is an original contribution of this work. Validation experiments indicate very good accuracy and generalization capability.

Keywords

Body composition assessment; false positive reduction

Abdominal obesity is strongly related to metabolic disorders, hypertension, and cardiovascular disease (1,2). Separate quantification of visceral and subcutaneous fat in the abdomen is important because of their different metabolic consequences and because

visceral fat more than subcutaneous fat has been associated with numerous age-related problems, such as insulin resistance, chronic inflammation, and cardiac diastolic dysfunction. Computed tomography (CT) is a widely adopted technique for assessing abdominal fat. Measurements of subcutaneous and visceral fat areas are typically computed from CT images using manual or semimanual analyses operated by clinical specialists. Even though these techniques usually provide reliable results, they are time-consuming and affected by intra- and inter-operator variability. Therefore, a regional fat quantification approach is highly desirable to increase the analysis throughput and reproducibility, especially in large epidemiological studies (3–6).

Earlier abdomen fat quantification approaches (7,8) required user intervention to define a region of interest and detected the fat voxels using density thresholding techniques in the domain of Hounsfield units (HUs). More recent techniques make use of joint intensity and texture information in a fuzzy affinity framework to identify the fat class (9), employ shape and appearance models for quantification of regional fat and muscle compartments (10), or use radial analysis to delineate the boundary between subcutaneous and visceral fat (11).

Despite the high level of sophistication and accuracy that has been achieved by the state-of-the-art methods, the presence of food residues in the intestines that can greatly reduce precision and even introduce a bias in the measure has not been explicitly addressed in the literature. The pixels of the food particles have HUs that greatly overlap with those of the visceral fat voxels and are usually identified as fat when simple intensity criteria are used. The quantification error induced by food residues can introduce a relevant bias in studies of aging because the permanence of food residues is likely to occur in larger quantities in older than in younger individuals.

Herein we present a method for automated quantification of abdominal fat in CT that concentrates on reducing false positives in the estimation of visceral fat resulting from food residues. The fat identification and separation stage uses intensity-based clustering and a gradient vector flow (GVF) active contour model. In several instances, the detected visceral fat includes regions of food residues. To exclude these areas from the quantification of visceral fat, we applied a machine learning scheme that employs texture, local shape features, and a support vector machine (SVM) classifier to identify and remove the food residues within the intestines. The food residue removal stage is applied as a final step in our methodology. In the experiment section, we report performance evaluation of our method against semimanually segmented reference masks of fat regions and draw conclusions on the applicability of this approach in a large-scale study.

MATERIALS AND METHODS

Our methodology can be divided into three main processes: identification of global fat voxels, separation of visceral from subcutaneous fat, and removal of food residues. The main steps of these processes are depicted in Figure 1 and are described in the following section.

Participants and CT Imaging

In our experiments, we used 168 CT abdomen scans of participants of the Baltimore Longitudinal Study of Aging (BLSA), an ongoing epidemiological study that focuses on aging and age-related traits. We selected participants over a wide range of body mass index and visceral and subcutaneous fat areas. We applied our scheme to single-slice CT images acquired at the umbilical level (i.e., at the L4–L5 vertebrae on the lumbar column using a Siemens Somatom Sensation 10 CT scanner [Siemens, Malvern, PA]). All scans were performed with a 500×500 mm field of view, 512×512 image matrix, and 10-mm slice

thickness. The considerable slice thickness was employed to achieve a more extensive superior/inferior coverage at the umbilical level and reduce the variability typical of a thin slice.

Subcutaneous and Visceral Fat Segmentation

We first removed the bed and calibration phantom from the trunk by applying foreground/background thresholding and selecting the largest connected component as the trunk. Then, we applied Fuzzy C-Means clustering (12) to the pixel intensities of the trunk and divided the pixels into four groups corresponding to air, muscle, fat, and bone tissues. Fuzzy C-Means produces vectors of cluster membership values in the range $[0,1]$ for all input samples. Each sample was subsequently assigned to the cluster of maximum membership. We mapped the labeled samples back to the image plane and identified the fat group based on the average cluster intensity.

Our next task was to separate the subcutaneous from visceral region of fat using a GVF snakes-based approach (13). The main difficulty was that the contrast between the subcutaneous region and visceral cavity is much lower than the contrast between the subcutaneous region and air thus affecting the GVF computation. To overcome this limitation, we computed the GVF on the fat membership map produced by Fuzzy C-Means that enhances the contrast between fat and other tissues. After snake convergence, we labeled the interior fat pixels as visceral and the exterior fat pixels as subcutaneous.

Removal of Food Residues

This fat segmentation technique is expected to produce acceptable measurements of subcutaneous and visceral fat. However, a frequent source of error in the quantification of visceral fat is the presence of food residues in the intestines during the scan. Therefore, correction of this error is of critical importance. An example of manual food residue removal is displayed in Figure 2.

Feature Analysis

In our preliminary experiments, we observed that the food residues may take on variable shapes and sizes in the intestines and subsequently in the visceral cavity. Therefore we expect shape and size to be of limited separation capability. In addition, the HUs of food residues are greatly overlapping with the HUs of the visceral fat as seen in Figure 3, indicating that the use of original intensities only, would yield high classification error. Therefore we decided to use texture-based and local shape features along with intensity features for separation of food residues from visceral fat.

We used Gabor filters to extract textured image features in multiple resolutions. The wavelet transform has been widely used for multiscale representation and analysis of signals in medical and biological applications. The main difference from traditional Fourier techniques is that wavelets can localize the information in the time-frequency plane (14). More specifically, the wavelet transforms provide a time-frequency decomposition of a signal over a range of characteristic frequencies that separates individual signal components more effectively than conventional Fourier analysis techniques. The Gabor filters have the advantage of optimizing the joint two-dimensional uncertainty in time and frequency and can be considered as edge and line bar detectors that can be tuned for different orientations and scales (15,16). Gabor functions form a complete but nonorthogonal basis. The two-dimensional Gabor function is the product of an elliptical Gaussian and a complex sinusoid and is given by:

$$G(x, y) = \frac{1}{2\pi\sigma_x\sigma_y} e \left[-\frac{1}{2} \left(\frac{x^2}{\sigma_x^2} + \frac{y^2}{\sigma_y^2} \right) + j2\pi Wx \right] \quad (1)$$

In this equation, σ_x and σ_y are the standard deviations of the elliptical Gaussian along x and y orientations and W is the sinusoid frequency. The filter dictionary can be produced by dilations and rotations of the mother Gabor wavelet (17). We used a filter bank of four scales and six orientations that is illustrated in Figure 4. Details on derivation of two-dimensional Gabor filters are given elsewhere (18).

The Laws' texture energy filters (18) were designed to measure the amount of variation within a sliding window over the image plane. The basic operators were one-dimensional arrays that can be transposed and multiplied among each other to generate two-dimensional kernels. We computed 24 features based on pair-wise combinations of the one-dimensional kernel bank of $\{L5, E5, S5, R5, W5\}$ used for averaging, and edge, spot, ripple, and wave detection, respectively. We use the average $L5L5$ for contrast normalization and then we omit it from the texture representation producing a texture vector with 24 components for each pixel.

We generated multiscale representations of the original CT images according to scale-space theory. Starting from the original formulation of scale-space theory (19), Witkin formalized the scale-space representation of a signal as its convolution with Gaussian filters of increasing widths (20). Koendering later generalized the idea and pointed out that the scale-space is produced by the solution of heat diffusion equation (21). Several applications have been proposed since then, including multiscale segmentation (22). Here we generate linear scale-space using Gaussian kernels and a nonlinear domain by anisotropic diffusion filtering. The nonlinear space smooths the interior of local pixel neighborhoods and preserves edge localization. We generate five scales in each domain resulting in 10 additional features.

In our feature domain, we included the spatial coordinates of each sample on the image plane. To reduce variability resulting from the pose or size of each subject, we decided to map the spatial coordinates to a canonical space. We applied principal axes registration (23,24) to the abdominal mask—produced in the segmentation stage—to estimate a linear transformation model with 6° of freedom corresponding to translation, rotation, and scaling. We first compute the translation using the trunk's centroid. The Principal Component Analysis (PCA) eigenvector orientations determine the rotation parameters and the eigenvalues determine the scaling factors. We then mapped the spatial coordinates of our samples to the canonical space using the computed transforms.

We also added the original CT intensities, median filtered and gradient magnitude of original CT image features to our texture, scale-space and spatial coordinates features, forming a feature space of dimensionality $D = 63$. An example of an original CT image and the computed texture features is depicted in Figure 5.

Classification Using SVMs

Our goal was to train the classifier to distinguish the food residues from visceral fat based on the computed features and to integrate these results in the last step of our fat quantification methodology to reduce the false positives produced by the fat segmentation stage. In the classification step, we used a two-class support vector machine classifier (25) with a radial kernel function.

SVM is a statistical machine learning technique that can be used for regression or classification. It can be applied to linear or nonlinear problems by means of a kernel function that maps the data samples to a high-dimensional feature space in which the problem becomes linearly separable (25). In the binary classification problem, we typically have a training data set comprising N data samples $\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_N$, with targets $t_1, t_2, \dots, t_N$ such that $t_n \in \{-1, 1\}$. Our objective is to define a decision function $y(x)$, which in the linear form is $y(\mathbf{x}) = \mathbf{w}^T \phi(\mathbf{x}) + b$, and to be able to classify new data points $\mathbf{x}$ according to the sign of $y(\mathbf{x})$. This is considered to be a constrained optimization problem and can be solved by use of Lagrangian multipliers a_n . The function to be optimized becomes:

$$L(\mathbf{w}, b, \mathbf{a}) = \frac{1}{2} \mathbf{w}^2 - \sum_{n=1}^N a_n \{t_n (\mathbf{w}^T \phi(\mathbf{x}_n) + b) - 1\} \quad (2)$$

where $\mathbf{a} = \{a_1, \dots, a_N\}$. For nonlinear decision functions, the solution is to define a kernel function $K(\mathbf{x}_i, \mathbf{x}_j) = \phi(\mathbf{x}_i) \cdot \phi(\mathbf{x}_j)$. Advantages of this technique are good generalization capability, robustness in high dimensions, and convexity of the objective function. Because of these advantages, the literature of SVM theory and applications is very broad. More detailed and formal descriptions of SVMs can be found elsewhere (25–27).

Modeling for Imbalanced Classification

As mentioned previously, because visceral fat is more prevalent than food residues at an approximate rate of 11.4:1, our dataset is imbalanced. Although SVM is known to be able to handle moderately imbalanced data, we tested two different approaches for classification. The first one was to use cost-weighting to each class and the other was based on resampling. In cost-weighting, we followed the method reported by Chang et al (28) that produced encouraging results previous work (29). We also implemented bootstrap aggregation (25) (ie, bagging) in the training stage as an alternate strategy for addressing imbalanced classes. For each subject, we randomly undersampled the majority class (ie, visceral fat) to match the sample size of the minority class (ie, food residues), and trained our classifier over the undersampled train dataset. We repeated this process n times ($n = 5$) thus producing n different models. At the testing stage, we predicted the class of each pixel by each model and applied classifier decision voting to produce the final classification.

RESULTS

To illustrate the significance of the food residue removal process, we computed the percentages of food residues over the visceral fat area and the visceral fat areas from reference measurements. A clinical specialist produced reference masks of subcutaneous fat, visceral fat, and food residues using a semimanual workflow. This step produced $n = 516,340$ pixel samples extracted from 168 CT images. We then plotted the food residue percentages versus the visceral fat areas as depicted in Figure 6. From this figure, we conclude that there is no significant correlation between food residue areas and visceral fat areas; therefore, we cannot predict the food residue areas from visceral fat measurements. We also note that although the average percentage of food residues is relatively small (ie, about 3.6%), it has a considerable variance (minimum = 0% and maximum = 81%) that may bias visceral fat quantification. As a result, this source of error needs to be addressed.

In our next experiment, we trained and tested the classification scheme using different feature selection/extraction techniques. Our classification targets were the reference masks of visceral fat and food residues produced by the semimanual workflow. We compared the classification performance of the complete feature domain described in the previous section with the use of linear or nonlinear scale-space, Gabor-based, or Laws-based features only.

We also tested a feature selection methodology that selects the top ranking features according to the Fisher's distance measure, the hierarchical dimensionality reduction (HDR) feature extraction method that iteratively computes the correlation matrix between features and merges the most correlated feature pair until the lower correlation threshold is reached (25), and PCA-based dimensionality reduction. To evaluate these techniques, we report the classification rates using six-fold cross-validation (Table 1). More specifically, we report values of true-positive rate (TPR), true-negative rate (TNR), classification accuracy (ACC), and area under the receiver operating characteristic (ROC) curve (AUC). The original feature space produces the most accurate classification indicating that all employed features contribute the separation between visceral fat and food residues.

We then used k -fold cross-validation with $k \in \{2, 4, 6\}$ in the complete feature space ($D = 63$) and computed the classification rates. In Table 2, we observe that the classifier achieves very encouraging rates of accuracy. Although the sensitivity is very high, the lower specificity rates can be expected from the very diverse texture profiles of food residues and their considerable intensity overlapping with the visceral fat. We also note that as we use fewer samples for training—for example, 50% of the data for training and 50% for testing in twofold cross-validation—the loss in specificity (TNR) becomes more apparent. Still, the classification rates remain high, indicating solid generalizability. We make similar observations about generalization performance in Figure 7 that displays ROC curves for k -fold cross-validation and Table 2 that reports AUC measurements for the ROC experiments. The AUC is slightly reduced for smaller values of k .

Furthermore, we tested the two different approaches for classification of imbalanced data, namely the cost-weighting and classifier-bagging technique. The results in Table 2 were produced by cost-weighting in the training stage, whereas the results of resampling and bagging approach are displayed in Table 3. The cost-weighting approach produces a bit higher classification accuracy, whereas bagging has higher specificity. The AUC measurements are approximately equal.

Another experiment was to assess the segmentation accuracy of the subcutaneous fat, the visceral cavity—that includes all tissues besides the subcutaneous region and the visceral fat after food residue removal. We compared the segmentations produced by our approach to semimanual delineations produced by a human operator over our dataset of 168 images. We first estimated the error between areas computed by the tested and reference workflows. The measures of agreement were the area error relative to the manual segmentation and the Pearson correlation coefficient between the two groups of measurements. Because the area errors do not account for the overlap between the segmented regions, we also computed the Dice similarity coefficient (30) to assess the segmentation accuracy. The results in Table 4 show that the automated segmentation algorithm achieves a high level of accuracy.

DISCUSSION

We have introduced a method for assessment of regional abdominal fat from CT images aiming for application to large scale datasets of epidemiological studies. The focal point of this article is the food residue identification and removal scheme that was developed to address the occurrence of visceral fat false positives when using conventional thresholding/clustering methods of the literature. Having formulated this task as a pattern classification problem, we computed and extracted intensity- and texture-based features, and subsequently evaluated the classification accuracy with respect to different feature domains.

Our experiments indicated that the combined use of intensity, textural, and spatial domain information produces better separation between the food residues and visceral fat. We note

that the sole use of intensity-based information cannot separate the two classes efficiently—as expected by the large overlap in HU space. The introduction of spatial coordinates in canonical space produced by principal axes registration clearly improves the classification performance, but still the classification error is more than 10%. We added texture representation features produced by Gabor and Law's kernels and achieved a considerable increase in sensitivity. The Gabor features and Law's features provide equivalent separability but their joint use improves the classification performance. The inclusion of scale-space features further improves the specificity of classification. The Gabor features offer a multiscale and multiorientation texture representation, whereas the Law's features measure the local texture energy. The scale-space analysis was included to describe the morphological intensity changes in multiple scales. The morphology of food residues is expected to be dependent on the specific scale of analysis, whereas the visceral fat pixels are expected to have more uniform intensity distribution in their neighborhood on the image plane.

In the HDR-based feature selection process, several Law's and some scale-space features were found more correlated and merged first. The HDR method is characterized by increased specificity and slightly decreased sensitivity. The PCA dimensionality reduction also improved specificity and decreased sensitivity relatively to the complete feature space. The Fisher-based feature selection method shows increased separation for the Law's and Gabor features and the anteroposterior spatial coordinates. The overall classification produced by Fisher-based feature selection is less accurate than the complete feature space, HDR, and PCA methods because it lacks sensitivity. Overall, these results indicate that the complete feature space is the most suitable solution because our clinically relevant variable is the visceral fat that corresponds to the majority class.

In the second set of experiments, we evaluated the generalization performance of food identification and observed a limited loss of specificity because we used fewer samples for training. This indicates that the food detection generalizes in a predicted manner and can be used for the complete database of the BLSA. These results were supported by a clinical scientist who visually inspected individual segmentation and quantification results over an extensive set of 1849 BLSA CT images and observed very limited visceral fat misclassifications.

Another interesting consideration is the selection of an appropriate model to handle imbalanced data. We tested cost-weighting and classifier-bagging approaches and report the results in Tables 2 and 3. In these results, we observe that the cost-weighting method performs better in terms of sensitivity and classification accuracy, whereas bagging significantly improves specificity and produces similar but more stable AUC values. The decision on which approach to use is mainly driven by the application. When we are interested in sensitivity, the cost-weighting method may be more appropriate. On the other hand, when specificity is more important, then classifier-bagging is more suitable.

In addition, we conducted segmentation validation experiments to assess the separation of subcutaneous fat region from the visceral cavity. The results in Table 4 show a high level of agreement between the automated and semimanual segmentation methods. The subcutaneous fat and visceral cavity segmentations produce lower relative errors and higher Dice scores than the visceral fat segmentation. This is expected because visceral fat segmentation is susceptible to errors mainly caused by the partial volume effect and food residue false positives. Still, the segmentation accuracy is quite encouraging and applicable to clinical studies.

The main advantage of our technique is that it significantly reduces a type of error that was not addressed in previous published work to the best of our knowledge. This approach also is automated and therefore time-efficient and reproducible in nature. In addition, we note that the automated separation of visceral fat from subcutaneous fat is largely insensitive to anatomical differences and that the main parameter modifications are related to the thickness of subcutaneous region. On the other hand, this method was developed using cross-sectional study frame of thinking. Therefore an extension of this technique would be to use prior segmentations of each participant as baselines for subsequent time points in longitudinal analyses. Another limitation is related to the data that are two-dimensional; therefore, the estimated regional fat may be subject to variability of anatomical site selection during acquisition.

In conclusion, we presented an automated abdominal fat quantification scheme with removal of false positives that is applied to CT imaging data. The method can be divided into the stages of global fat identification, segmentation of subcutaneous fat, and removal of food residues. The key contribution of this work is the development of a statistical learning scheme aiming for detection and removal of food residues from the visceral fat region. Our method was found to meet the accuracy requirements when validated against a reference standard over a set of 168 CT images selected from the BLSA. The reduction of food residues improves the quantification accuracy of visceral fat that is expected to be critical in resolving scientific questions related to clinical outcomes.

Acknowledgments

This research was supported entirely by the Intramural Research Program of the National Institutes of Health, National Institute on Aging. We also thank Dr. Ramona Ramachandran for producing reference masks of visceral fat and food residue regions.

REFERENCES

1. Kissebah AH, Vydelingum N, Murray R, et al. Relation of body fat distribution to metabolic complications of obesity. *J Clin Endocrinol Metabol.* 1982; 54:254–260.
2. Seidell JC, Hautvast JG, Deurenberg P. Overweight: fat distribution and health risks. Epidemiological observations. a review. *Infusionstherapie.* 1989; 16:276–281. [PubMed: 2696728]
3. Beydoun MA, Lhotsky A, Wang Y, et al. Association of adiposity status and changes in early to mid-adulthood with incidence of Alzheimer's disease. *Am J Epidemiol.* 2008; 168:1179–1189. [PubMed: 18835864]
4. Ding J, Kritchevsky SB, Hsu F-C, et al. Association between non-subcutaneous adiposity and calcified coronary plaque: a substudy of the multi-ethnic study of atherosclerosis. *Am J Clin Nutr.* 2008; 88:645–650. [PubMed: 18779279]
5. Lee K, Lee S, Kim Y, et al. Waist circumference, dual-energy X-ray absorptiometrically measured abdominal adiposity, and computed tomographically derived intra-abdominal fat area on detecting metabolic risk factors in obese women. *Nutrition.* 2008; 24:625–631. [PubMed: 18485667]
6. Kanaya AM, Lindquist K, Harris TB, et al. Total and regional adiposity and cognitive change in older adults: the health, aging and body composition (ABC) study. *Arch Neurol.* 2009; 66:329–335. [PubMed: 19273751]
7. Jensen MD, Kanaley JA, Reed JE, et al. Measurement of abdominal and visceral fat with computed tomography and dual-energy x-ray absorptiometry. *Am J Clin Nutr.* 1995; 61:274–278. [PubMed: 7840063]
8. Yoshizumi T, Nakamura T, Yamane M, et al. Abdominal fat: standardized technique for measurement at CT. *Radiology.* 1999; 211:283–286. [PubMed: 10189485]
9. Bandekar A, Naghavi M, Kakadiaris I. Performance evaluation of abdominal fat burden quantification in CT. *Conf Proc IEEE Eng Med Biol Soc.* 2005; 3:3280–3283. [PubMed: 17282946]

10. Chung H, Cobzas D, Birdsell L, et al. Automated segmentation of muscle and adipose tissue on CT images for human body composition analysis. *Proc SPIE*. 2009;7261. 72610K.
11. Zhao B, Colville J, Kalaigian J, et al. Automated quantification of body fat distribution on volumetric computed tomography. *J Comput Assist Tomogr*. 2006; 30:777–783. [PubMed: 16954927]
12. Bezdek, JC. Pattern recognition with fuzzy objective function algorithms. Norwell, MA: Kluwer Academic Publishers; 1981.
13. Xu C, Prince JL. Snakes, shapes, and gradient vector flow. *IEEE Trans Image Proc*. 1998; 7:359–369.
14. Unser M, Aldroubi A. A review of wavelets in biomedical applications. *Proc IEEE*. 1996; 84:626–638.
15. Daugman J. Complete discrete 2-D Gabor transforms by neural networks for image analysis and compression. *IEEE Trans Acoustics. Speech Signal Proc*. 1988; 36:1169–1179.
16. Manjunath B, Ma W. Texture features for browsing and retrieval of image data. *IEEE Trans Pattern Anal Machine Intell*. 1996; 18:837–842.
17. Lee TS. Image representation using 2D Gabor wavelets. *IEEE Trans Pattern Anal Machine Intell*. 1996; 18:959–971.
18. Shapiro, LG.; Stockman, GC. Computer vision. Prentice Hall: 2001.
19. Iijima T. Basic theory on normalization of a pattern (in case of typical one-dimensional pattern). *Bull Electrical Lab*. 1962; 26:368–388. [in Japanese].
20. Witkin A. Scale-space filtering. *Int Jt Conf Artificial Intell*. 1983:1019–1022.
21. Koenderink JJ. The structure of images. *Biol Cybernetics*. 1984; 50:363–370.
22. Makrogiannis S, Vanhamel I, Fotopoulos S, et al. A watershed-based multiscale segmentation method for color images using automatic scale selection. *SPIE J Electronic Imaging*. 2005; 14
23. Alpert NM, Bradshaw JF, Kennedy D, et al. The principal axes transformation-a method for image registration. *J Nucl Med*. 1990; 31:1717–1722. [PubMed: 2213197]
24. Arata LK, Dhawan AP, Broderick JP, et al. Three-dimensional anatomical model-based segmentation of MR brain images through Principal Axes Registration. *IEEE Trans Biomed Eng*. 1995; 42:1069–1078. [PubMed: 7498910]
25. Duda, RO.; Hart, PE.; Stork, DG. Pattern classification. Wiley-Interscience Publication; 2000.
26. Burges CJC. A tutorial on support vector machines for pattern recognition. *Data Min Knowl Discovery*. 1998; 2:121–167.
27. Cortes C, Vapnik V. Support-vector networks. *Mach Learn*. 1995; 20:273–297.
28. Chang CC, Lin CJ. LIBSVM: a library for support vector machines 2006. Software available at: <http://www.csie.ntu.edu.tw/~cjlin/libsvm>.
29. Tang Y, Zhang YQ, Chawla NV, et al. SVMs modeling for highly imbalanced classification. *IEEE Trans Syst Man Cybernetics, Part B: Cybernetics*. 2009; 39:281–288.
30. Dice LR. Measures of the amount of ecologic association between species. *Ecology*. 1945; 26:297–302.

Figure 1.

Flow diagram of our abdomen fat quantification algorithm (*top*). It can be divided into three main processes: identification of global fat voxels, separation of visceral from subcutaneous fat, and removal of food residues from the visceral region using a statistical learning scheme. Original computed tomography image, fat membership map produced by Fuzzy C-Means, separation of visceral cavity from subcutaneous region after snake convergence, and generation of final tissue label map (*bottom*). (Color version of figure is available online).

Figure 2. Manual labeling and removal of food residues. The visceral fat region includes food residues (false positives) that are removed and stored for the classifier training stage. (Color version of figure is available online).

Figure 3.

Histogram of Hounsfield units of manually labeled visceral fat and food particle pixels computed from the reference masks. The significant overlap between the two classes complicates the identification of food residues directly from pixel intensities. (Color version of figure is available online).

Figure 4.
The real components of the Gabor filter bank used for representing the texture signatures of food residues in our classification scheme. (Color version of figure is available online).

Figure 5.

Example of an original computed tomography image from our database and the computed Gabor and Laws' texture energy features. We note here that the features are computed over the visceral cavity after exclusion of the subcutaneous region. (Color version of figure is available online).

Figure 6.

Scatter plot of the food residues percentage against the visceral fat area after supervised regional fat quantification. We observe that food residues and visceral fat areas are not linearly correlated and that the estimated food residue areas are characterized by high variation. (Color version of figure is available online).

Figure 7.

Received operating characteristic of k -fold cross-validation for $k \in \{2, 4, 6\}$ using all computed features ($D = 63$). We observe only a slight decline in classification performance when we use fewer samples for training (ie, for smaller values of k). This observation is consistent with results reported in Table 2. (Color version of figure is available online).

TABLE 1

Comparison of Different Feature Computation and Selection Schemes

Feature Domain	D	TPR (%)	TNR (%)	ACC (%)	AUC
I, S, G, L, LS, NLS	63	97.6 ± 0.3	65.8 ± 3.8	95.0 ± 0.5	0.954
I, S, G, L, NLS	58	97.5 ± 0.3	65.5 ± 3.9	95.0 ± 0.5	0.953
I, S, G, L, LS	58	97.5 ± 0.3	65.4 ± 3.9	94.9 ± 0.7	0.953
I, S, G, L	53	97.5 ± 0.4	63.9 ± 4.4	94.7 ± 0.6	0.950
I, S, G	29	94.2 ± 1.0	73.3 ± 2.8	92.6 ± 1.0	0.941
I, S, L	29	93.5 ± 1.0	80.0 ± 3.4	92.4 ± 1.0	0.946
I, S	5	85.0 ± 1.8	84.9 ± 4.4	85.0 ± 1.5	0.923
I	3	69.5 ± 2.9	81.7 ± 2.4	70.5 ± 2.5	0.846
HDR-based	44	97.1 ± 0.8	68.5 ± 9.3	94.8 ± 0.6	0.953
Fisher-based	30	93.1 ± 1.2	80.1 ± 3.9	92.0 ± 1.1	0.943
PCA	25	96.7 ± 0.5	68.4 ± 4.2	94.4 ± 0.6	0.949

ACC, classification accuracy; AUC, area under the receiver operating characteristic curve; Fisher, Fisher's distance ranking; G, Gabor features; HDR, hierarchical dimensionality reduction; I, intensity features; L, Law's features; LS, linear scale-space; NLS, nonlinear scale-space; PCA, principal component analysis; S, spatial coordinates; TNR, true-negative rate; TPR, true-positive rate.

TABLE 2

Classification Rates and AUC of k -fold Cross-Validation Experiment for $k \in \{2, 4, 6\}$ Using All Computed Features ($D = 63$)

k -fold	TPR (%)	TNR (%)	ACC (%)	AUC
2	97.8 ± 0.2	61.3 ± 0.6	94.9 ± 0.5	0.951
4	97.6 ± 0.9	64.7 ± 0.3	94.9 ± 0.7	0.952
6	97.6 ± 0.3	65.8 ± 3.8	95.0 ± 0.5	0.954

ACC, classification accuracy; AUC, area under the receiver operating characteristic curve; TNR, true-negative rate; TPR, true-positive rate.

TABLE 3

Classification Rates and (AUC) of k -fold Cross-Validation with Repeated Undersampling and Bagging for $k \in \{2, 4, 6\}$ Using All Computed Features ($D = 63$)

k -fold	TPR (%)	TNR (%)	ACC (%)	AUC
2	95.7 ± 0.5	74.6 ± 0.7	94.0 ± 0.7	0.951
4	95.7 ± 0.6	75.9 ± 2.4	94.1 ± 0.8	0.953
6	95.8 ± 0.8	76.4 ± 6.8	94.2 ± 0.6	0.953

ACC, classification accuracy; AUC, area under the receiver operating characteristic curve; TNR, true-negative rate; TPR, true-positive rate.

TABLE 4

Segmentation Validation for Subcutaneous Fat, Visceral Cavity, and Visceral Fat

	Subcutaneous Region	Visceral Cavity	Visceral Fat
Pearson correlation coefficient (ρ^2)	0.983	0.993	0.956
Relative area error (%)	5.4 ± 4.5	1.7 ± 1.1	8.2 ± 8.8
Dice coefficient (%)	93.1 ± 2.9	98.4 ± 5.5	85.6 ± 5.7