Automated segmentation of muscle and adipose tissue on CT images for human body composition analysis

Howard Chung[1], Dana Cobzas[1], Jessica Lieffers[3], Laura Birdsell[2] and Vickie Baracos[2]

Departments of [1]Computing Science, [2]Oncology, and [3]Agricultural, Food and Nutritional Science, University of Alberta, Canada

ABSTRACT

The ability to compute body composition in cancer patients lends itself to determining the specific clinical outcomes associated with fat and lean tissue stores. For example, a wasting syndrome of advanced disease associates with shortened survival. Moreover, certain tissue compartments represent sites for drug distribution and are likely determinants of chemotherapy efficacy and toxicity. CT images are abundant, but these cannot be fully exploited unless there exist practical and fast approaches for tissue quantification. Here we propose a fully automated method for segmenting muscle, visceral and subcutaneous adipose tissues, taking the approach of shape modeling for the analysis of skeletal muscle. Muscle shape is represented using PCA encoded Free Form Deformations with respect to a mean shape. The shape model is learned from manually segmented images and used in conjunction with a tissue appearance prior. VAT and SAT are segmented based on the final deformed muscle shape. In comparing the automatic and manual methods, coefficients of variation (COV) $(1-2\%)$, were similar to or smaller than inter- and intra-observer COVs reported for manual segmentation.

Keywords: computer-assisted therapy and therapy planning, segmentation, shape

1. DESCRIPTION OF PURPOSE

Body composition, specifically proportions of fat and lean tissues, has many implications for cancer patients. Obesity associates with increased prevalence or recurrence of several cancers.[1,2] A wasting syndrome involving muscle and fat loss frequently accompanies advanced disease and associates with poor outcomes.[3] Moreover, fat and lean tissue compartments represent sites for distribution of lipid- and water-soluble drugs used in cancer treatment, and are likely determinants of chemotherapy efficacy and toxicity.[4,5] Computed tomography (CT) imaging is an accepted gold standard method for human body composition analysis.[6,7] This tool can precisely follow tissue change[7] and single images at specific skeletal landmarks are good correlates of whole body fat and lean tissue mass.[8,9] Cancer patients undergo frequent CT imaging as part of their routine care and we have started to use these images secondarily for body composition assessment.[4,5] However, these images cannot be fully exploited unless there exist practical, fast and precise approaches for tissue quantification.

Conventional manual segmentation of CT images uses defined windows of Hounsfield units (HU, units of radiation attenuation) for each tissue, and is guided by operator knowledge of anatomical structures. Automatic fat segmentation methods have been reported which are relatively straightforward owing to the unique HU ranges of adipose tissues.[10] There are no reports of automatic segmentation of skeletal muscle, despite being highly related to human function and disease outcome. This latter task is particularly challenging owing to the large variability in muscle shape and the overlap in HU between muscle and internal organs.

We propose an automatic method for segmenting muscle, visceral (VAT) and subcutaneous (SAT) adipose tissues from CT images, taking the specific approach of shape modeling for the analysis of muscle. Explicit shape representations based on critical points are dependent on a particular parametrization and cannot easily handle topological changes. One solution is to use an implicit shape representation either as a binary image or a signed distance function.[11,12] This representation is robust to topological changes and not dependent on a particular parametrization. The shape is then parametrized either directly in the space of embedding functions (e.g.[11]) or

Further author information: (Send correspondence to Howard Chung)
Howard Chung: E-mail: hschung@cs.ualberta.ca

as deformation of a lattice defined in the domain of the embedding shape functions.[12] We constructed a series of steps capable of segmentation, which was both accurate and less time consuming than the standard manual methods.

2. METHOD

2.1 Data and preprocessing

In the following section we present methods for segmenting muscle and adipose tissues on CT scans taken during routine cancer patient care. All images are taken at the 3rd lumbar vertebra (L3), a standardized skeletal landmark. The original CT data is thresholded so only points in the predefined HU range for muscle [-29,150] are considered for muscle segmentation. We denote the image after thresholding as I (Figure 1 (a)).

2.2 Shape prior model and deformations

The muscle shape is represented by a deformable binary image $I_s : \Omega \rightarrow \{0, 1\}$ where points inside the shape are 1. Image deformations are parametrized using the Free Form Deformation (FFD) model[12] consisting of B-spline cubic interpolation of regular lattice points. Lattice point deformations are coded with respect to a mean shape estimated from a set of training images. For an efficient representation we further encode control (lattice) point deformations using principal components. In summary, the steps for computing the shape parameters from manually segmented images are: (1) affine alignment and mean shape (I_s) computation; (2) non-rigid alignment with respect to I_s using the FFD model; (3) encode incremental deformations using Principal Component Analysis (PCA).

The *affine alignment* computes parameters $\mathbf{a} = [a_1 \dots a_6]$ that align binary shape images I_i with a reference image I_r (chosen from the same set) by minimizing the energy:

$$E_{affine}(\mathbf{a}) = \int_{\Omega} \left[I_r(\mathbf{x})(\tilde{I_r}(\mathbf{x}) - \tilde{I_i}(\mathbf{x}, \mathbf{a}))^2 + \alpha BI_r(\mathbf{x})\ \tilde{I_i}(\mathbf{x}, \mathbf{a}) \right] d\mathbf{x} + \beta(a_2^2) + \gamma(a_3^2) \tag{1}$$

where $\tilde{I_r}$ denotes the smoothed shape I_r; BI_r is a binary mask of I_r with interior filled; $\tilde{I_i}$ is I_i smoothed and warped by $\mathbf{a}$; α, β, γ are weighting coefficients. The terms of the energy are the usual SSD registration score, a penalty for the shape falling outside the outer boundary of I_r, and terms to prevent excessive rotation of the image. The template shape I_s is computed as the mean of the resulting aligned images.

The *non-rigid deformations* of the training images I_i are encoded as a FFD of a lattice grid $F_{ij}, i = 1 \dots M, j = 1 \dots N$ over an image $I = \{(x, y) | 1 \leq x \leq X, 1 \leq y \leq Y\}$. Dense deformations are achieved using B-spline interpolation based on lattice points. The position of a point $\mathbf{x} = (x, y)$ undergoing the deformation is given by the tensor product:

$$\mathbf{x}' = \sum_{k=0}^{3} \sum_{l=0}^{3} B_k(u(x)) B_l(v(y)) F_{i+k, j+l} \tag{2}$$

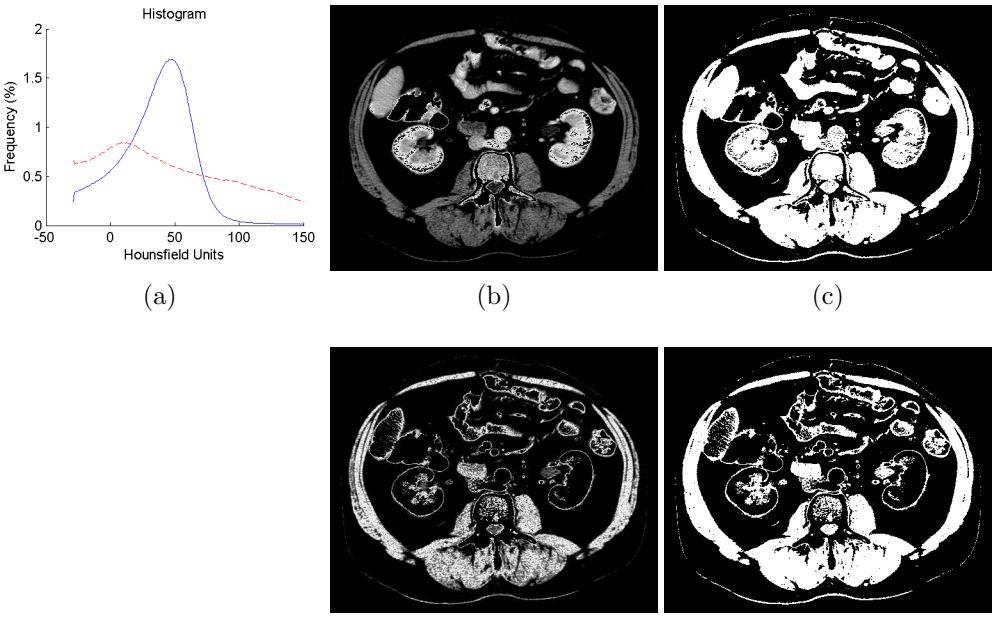
where B_k, B_l are the cubic B-spline basis functions:

$$B_0(w) = \frac{(1-w)^3}{6}, B_1(w) = \frac{3w^3 - 6w^2 + 4}{6}, B_2(w) = \frac{-3w^3 + 3w^2 + 3w + 1}{6}, B_3(w) = \frac{w^3}{6} \tag{3}$$

$$u = \frac{x}{X}(M-1) - \lfloor \frac{x}{X}(M-1) \rfloor, v = \frac{y}{Y}(N-1) - \lfloor \frac{y}{Y}(N-1) \rfloor \tag{4}$$

and $F_{i+k,j+l}$ is the (i+k, j+l) lattice point with:

$$i = \lfloor \frac{x}{X}(M-1) \rfloor - 1, j = \lfloor \frac{y}{Y}(N-1) \rfloor - 1 \tag{5}$$



(a) (b) (c)

(d) (e)

Figure 1. (a) histogram of muscle HU pixels inside muscle shape (solid line) and outside muscle shape (dashed line); (b) grayscale image of muscle HU [-29,150]; (c) thresholded, binary image of (b); (d) grayscale muscle probability image; (e) thresholded, binary image of (d)

The lattice incremental deformations $\theta = (\delta F_{i,j})$ are computed by minimizing the energy:

$$E_{FFD}(\theta) = \int_{\Omega} I_s(\mathbf{x})(\tilde{I}_s(\mathbf{x}) - \tilde{I}_i(\mathbf{x}, \theta))^2 d\mathbf{x} + \mathcal{R}(F((\theta))) \tag{6}$$

$\mathcal{R}(F(\theta))$ is a regularization term that keeps the lattice grid monotonic and plaid:

$$\mathcal{R}(F(\theta)) = \alpha \sum_{i,j} \left| F_{i,j} - \left(\frac{F^x_{i-1,j} + F^x_{i+1,j}}{2}, \frac{F^y_{i,j-1} + F^y_{i,j+1}}{2} \right) \right|^2 + \beta \sum_{i,j} \min(0, F_{i+1,j} \times F_{i,j+1})^2 \tag{7}$$

The first term penalizes grid points from movements too close or too far from it's immediate 4 neighbours, while the second term prevents grid points from crossing over (which would cause the patch normal to flip).

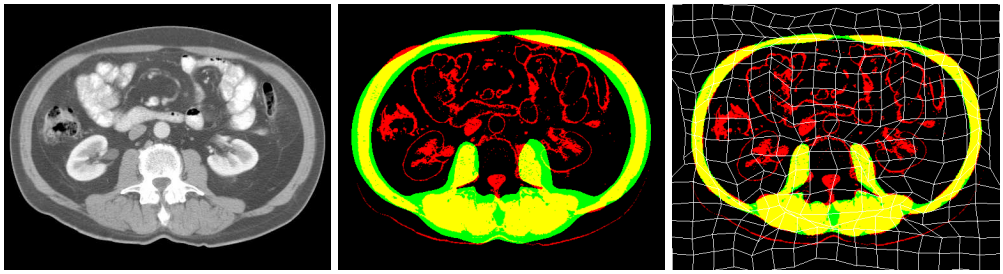
Incremental deformations θ of control points are *encoded using PCA*:

$$\theta = \theta_m + \sum_k \lambda_k \theta_k$$

The final shape parameters are the PCA coefficients $\lambda = (\lambda_k, k = 1 \dots K)$.

2.3 Appearance model

In addition to the shape prior, we compute an appearance prior that uses Parzen histograms of the manually segmented muscle and remaining tissue in the [-29,150] HU to better disambiguate muscle from the organs. We use the histograms to calculate a probability image that shows the likelihood of a pixel belonging to the muscle tissue. The probability image is thresholded to obtain the set of pixels which will be considered for segmentation. Compared to thresholding based on the [-29,150] HU range, this results in an image containing fewer organ pixels while maintaining the relevant muscle (Figure 1).



(a) (b) (c)

Figure 2. (a) initial CT image; (b) affine alignment; (c) muscle segmentation

2.4 Muscle segmentation

To segment the muscle tissue in a new image I, we first threshold the points inside the muscle range and, based on the appearance prior, compute the probability images p_m, p_o (muscle and organ tissue). A thresholded, binary copy of p_m, denoted PI, is used to perform a 6 DOF affine alignment with the mean shape.

For computing the shape deformations λ that give the final muscle tissue segmentation we investigated two approaches. The first approach computes the best shape that balances two energy terms (inside and outside) similar to an active region model.[13,14]

$$E_{FFD}(\lambda) = -\int_{\Omega} I_s(\mathbf{x}) BI(\mathbf{x},\lambda) \log p_m(I(\mathbf{x},\lambda)) d\mathbf{x} - \int_{\Omega} (1 - I_s(\mathbf{x})) BI(\mathbf{x},\lambda) \log p_o(I(\mathbf{x},\lambda)) d\mathbf{x} + \mathcal{R}(F(\lambda)) \quad (8)$$

where BI is the binary mask of I and $\mathcal{R}(F(\lambda))$ is the same regularization as in Equation 7.

In practice a simplified version of the above energy, similar to the one used in Equation 6 gives better results. The energy minimizes the displacement between PI and the shape prior I_s.

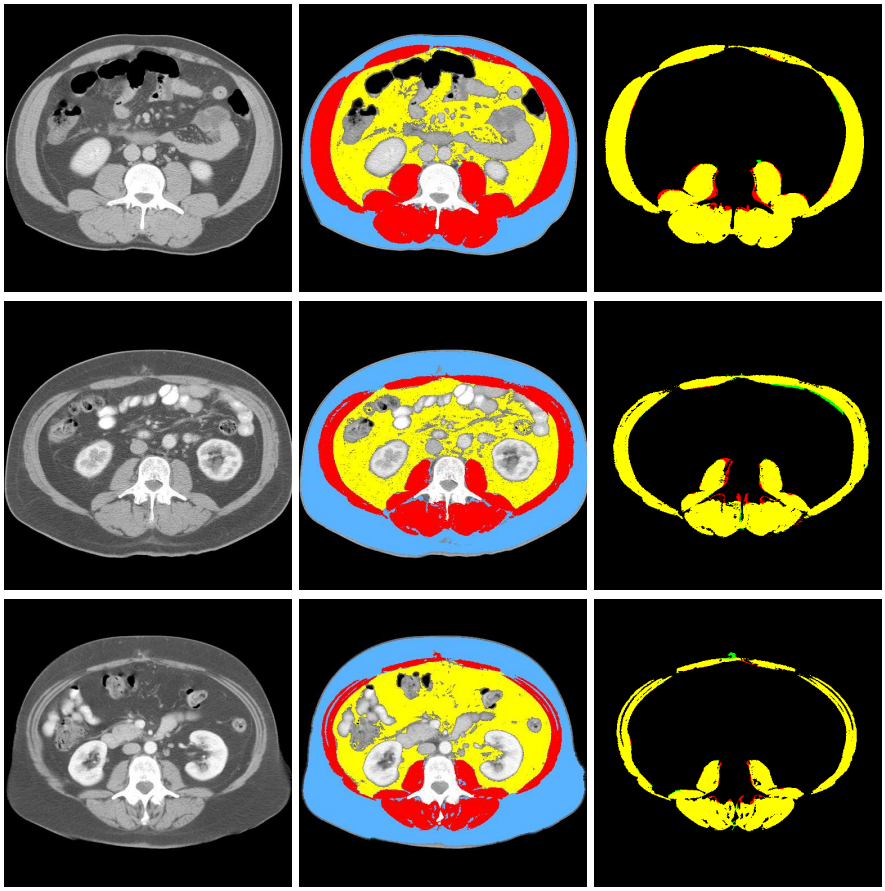
$$E_{FFD}(\lambda) = \int_{\Omega} I_s(\mathbf{x})(\tilde{I}_s(\mathbf{x}) - \tilde{PI}(\mathbf{x},\lambda))^2 d\mathbf{x} + \mathcal{R}(F(\lambda)) \quad (9)$$

To help minimization and performance we use a *multiresolution approach* with 3 resolution levels ($1/16, 1/4$ and full image size). Deformations are first solved using the coarse resolution for both the image and the grid. The solution from the coarse level is interpolated at the next finer level and the procedure is repeated until full resolution is reached. At each subsequent resolution level the best PCA coefficients (in the least square sense) are computed from the interpolated deformations and used to initialize the solution.

Figure 2 illustrates the muscle segmentation steps. The original CT image (Figure 2 (a)) is thresholded based on the probability image and converted to binary, giving PI. The affine parameters for aligning PI to I_s are solved and applied to PI (Figure 2 (c)). Finally, the FFD parameters required to warp I_s to the affine aligned PI are solved and applied to I_s (Figure 2 (d)).

2.5 Adipose tissue segmentation

The final muscle shape is used to segment both SAT (HU [-190, -30]) and VAT (HU [-150, -50]). The region outside the inner shape boundary is considered for SAT segmentation and the region inside the inner boundary for VAT. The muscle, SAT, and VAT regions are expanded to capture pixels near region borders. To avoid labeling AT inside the organs, a mask of the organs is created with holes filled. Large sections of AT enclosed by organ tissue are colored as VAT provided no gas pockets are present within, which likely indicates AT inside the intestines. Final segmentation of each region is performed by taking only pixels in the valid HU range for each tissue. Figure 3 presents complete muscle/fat segmentations for three patients.



(a) (b) (c)

Figure 3. (a) initial CT images; (b) muscle (red), SAT (blue), VAT (yellow); (c) manual muscle segmentation (red), automatic muscle segmentation (green), overlap (yellow)

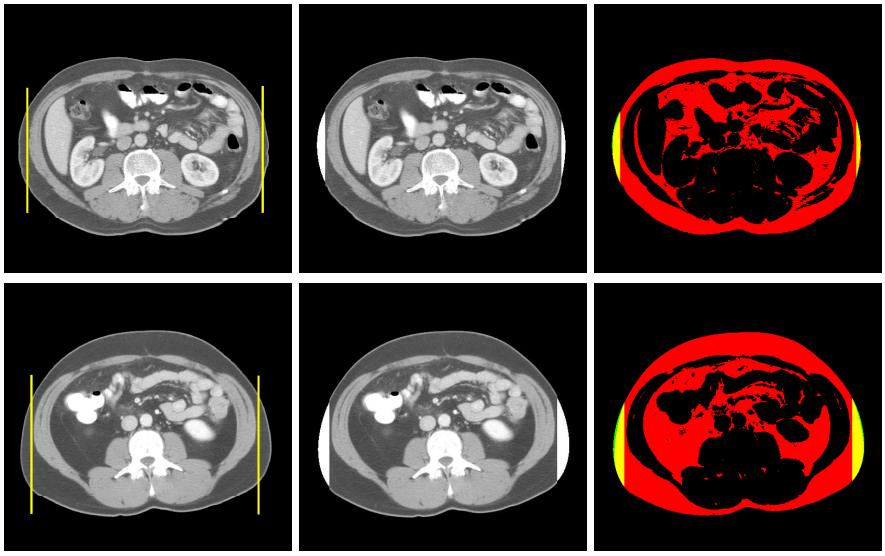
2.6 Missing SAT estimation

In some cases, the SAT in the image is truncated when the persons body exceeds the maximum diameter of the CT scanner. We designed an automatic method to complete SAT tissue which assumes smooth body curvature and calculates a Bezier curve (for each side) based on the two end tangents. The tangents are approximated using finite differences.

3. RESULTS

For evaluating the performance of the segmentation algorithm we compared the manual and automatic segmentation of 20 CT abdominal scans of cancer patients. Two consecutive axial CT image at the level of L3 were selected for each patient. Images were selected to have good contrast with no gross shape abnormalities. Manual segmentation was performed using Slice-O-Matic V4.3 software (Tomovision, Montreal, Canada) which calculates the tissue surface area (cm^2). The results were compared with data generated by the automatic method for the same images (Table 1). Total adipose tissue (TAT) was calculated for both methods as the sum of VAT and SAT. The Jaccard score was used to assess overlap. Results show excellent agreement between manual and automatic segmentations: Jaccard scores of about 95% and Spearman's correlation coeffients of about 0.994.

To determine the performance of the SAT estimation, non-truncated CT images which had been manually segmented were artificially truncated and the missing SAT extrapolated. The difference between the actual and estimated SAT was analyzed. Results are summarized in Table 2 and show reasonable agreement between actual and estimated SAT cut off with Jaccard scores of approximately 84%.



(a) (b) (c)

Figure 4. (a) original CT images with lines showing artificial cut; (b) SAT estimates; (c) manual TAT (red), estimated SAT (green), overlap (yellow)

While the results show good accuracy, there are several sources of error which could be addressed to improve the segmentation. Errors in segmentation come mainly from: 1) fine detail around the spine which cannot be captured by the current system, 2) contact between muscle and internal organs which causes erroneous segmentation of organ as muscle due to overlapping HU, and 3) incorrect segmentation of AT enclosed by organ tissue and inside the intestines. To address these sources of error, one could: 1) increase the resolution of the grid around the spine to achieve finer deformations in that region, 2) analyze the textures of muscle and organ to differentiate between tissues, and 3) identify the organs and segment VAT based on this knowledge.

Inaccuracies in the SAT estimation are mainly due to the assumption that the body will follow a regular, smooth curve (similar to a 'C' shape) beyond the point of truncation. However, not all bodies follow this pattern, resulting in incorrect estimation. Shape learning could be incorporated to improve the system.

Two types of images which our system will fail to segment include: 1) images with low HU contrast and 2) images of abnormal muscle shapes.

Most CT images display good contrast in HU values and their histograms show two clear peaks, one in the muscle range and one in the AT range. Because thresholding is based on probabilities calculated from high contrast data, low contrast images pose a problem. The muscle shape after thresholding a low contrast image is insufficient for proper segmentation. If an appropriate histogram transformation could be determined, taking a low contrast image to high contrast, the image could then be segmented using the existing system.

The majority of patients show a muscle shape which is normal. In images of patients that do not show a normal muscle shape, the segmentation fails due to having been trained using normal muscle shapes. Incorporating abnormal muscle shapes into the training data could result in unnatural deformations when segmenting normal shapes. As a result, images with abnormal muscle shapes would likely need to be analyzed separately.

4. NEW OR BREAKTHROUGH WORK TO BE PRESENTED

The complete automatic segmentation of muscle, VAT and SAT has not been previously reported. This segmentation is key to the discrimination and quantification of important and distinct body composition features. The automated method also offers a new tool to extrapolate SAT areas which have been cut off during scanning. This is a common problem found to occur at a rate of $\sim 20.8\%$ in a population of 850 cancer patients.

Tissue	Area(cm^2)		Raw diff.(cm^2)	Diff.(%)	COV(%)	Jaccard(%)	Spearman's
	M	A					corr.
Muscle	167.5 ± 38.1	164.2 ± 37.6	-3.33 ± 2.49	-2.06 ± 1.71	1.61 ± 1.05	95.2 ± 1.6	.994
VAT	189.3 ± 65.0	191.1 ± 64.9	1.82 ± 4.56	0.57 ± 5.79	1.90 ± 3.66	95.5 ± 6.1	.997
SAT	228.4 ± 81.3	229.0 ± 81.0	0.64 ± 2.11	0.35 ± 0.92	0.52 ± 0.45	97.2 ± 0.8	.998
TAT	417.7 ± 92.0	420.1 ± 94.1	2.46 ± 4.73	0.50 ± 1.05	0.54 ± 0.60	97.8 ± 1.0	1.000

Table 1. Comparison of manual (M) and automated (A) segmentation of CT images. (n=20) all values are reported as mean ±SD. Raw differences are calculated as automated-manual. Correlation coefficients were all significant ($p < 0.001$)

Tissue	Cut Area(cm^2)		Raw diff.(cm^2)	Diff.(%)	Jaccard(%)
	Actual	Estimate			
SAT	15.9 ± 10.3	16.2 ± 10.8	0.30 ± 2.55	1.88 ± 17.80	84.7 ± 7.2

Table 2. Comparison of actual and estimated SAT cut off. All values are reported as mean ± SD. Raw differences are calculated as automated-manual.

5. CONCLUSION

Conventional manual segmentation takes $\sim$ 20-30 minutes per image. For the images that were analyzed, a MATLAB® implementation of the automated version required 138 ± 24 seconds to segment each image running on a 2.4 GHz Intel Core2 Duo T7700 CPU with 3 GB RAM. This analysis is also free of any intra- or inter-observer variability. The coefficients of variation (COV) between automated and manual analyses ($1 - 2\%$) are similar to or smaller than COV between observers using manual segmentation,[8] or between paired consecutive images analyzed by a single observer[5] and as such, equal or exceed manual approaches.

We have taken common image analysis tools and created an automated program which can rapidly and reliably segment muscle, VAT and SAT. Although many body composition studies exist, very few have taken a populationbased approach, due to cost, availability of imaging modalities, and image analysis time. We recently reported manual CT image analysis of muscle in 250 obese cancer patients[5] and demonstrated that muscle loss was prevalent and associated with decreased functional status, shortened survival and increased chemotherapy toxicity. The image analysis for that study represents $\sim$ 250h of work using manual segmentation; automated analysis provides the potential to analyze body composition on large quantities of diagnostic CT images in relation to important cancer-related outcomes.

REFERENCES

[1] Kubo, A. and Corley, D. A., "Body mass index and adenocarcinomas of the esophagus or gastric cardia: a systematic review and meta-analysis," *Cancer Epidemiol. Biomarkers Prev.* **15**, 872–878 (2006).

[2] Luo, W., Morrison, H., de Groh, M., Waters, C., DesMeules, M., Jones-McLean, E., et al., "The burden of adult obesity in canada," *Chronic Dis. Can.* **27**(4), 135–144 (2007).

[3] Dewys, W. D., Begg, C., Lavin, P. T., Band, P. R., Bennett, J. M., Bertino, J. R., et al., "Prognostic effect of weight loss prior to chemotherapy in cancer patients," *Eastern Cooperative Oncology Group. Am. J. Med.* **69**, 491–497 (1980).

[4] Prado, C. M. M., Baracos, V. E., McCargar, L. J., Mourtzakis, M., Mulder, K. E., Reiman, T., et al., "Body composition as an independent determinant of 5-fluorouracil-based chemotherapy toxicity," *Clin. Can. Res.* **13**, 3264–3268 (2007).

[5] Prado, C. M. M., Lieffers, J. R., McCargar, L. J., Reiman, T., Sawyer, M. B., Martin, L., et al., "A population-based study of the prevalence and clinical implications of sarcopenic obesity in patients with solid tumors of the respiratory and gastrointestinal tracts," *Lancet Oncol.* **9**, 629–635 (2008).

[6] Baumgartner, R. N., Koehler, K. M., Gallagher, D., Romero, L., Heymsfield, S. B., Ross, R., et al., "Epidemiology of sarcopenia among the elderly in new mexico," *Am. J. Epidemiol.* **147**, 755–763 (1998).

[7] Ross, R., Dagnone, D., Jones, P. J., Smith, H., Paddags, A., Hudson, R., et al., "Reduction in obesity and related comorbid conditions after diet-induced weight loss or exercise-induced weight loss in men," *Ann. Intern. Med.* **133**, 93–103 (2000).

[8] Mourtzakis, M., Prado, C. M. M., Lieffers, J. R., Reiman, T., McCargar, L. J., and Baracos, V. E., "A practical and precise approach to quantification of body composition in cancer patients using computed tomography images acquired during routine care," *App. Physiol. Nutr. Metab.* **In Press** (2008).

[9] Shen, W., Punyanitya, M., Wang, Z., Gallagher, D., St-Onge, M. P., Albu, J., et al., "Total body skeletal muscle and adipose tissue volumes: estimation from a single abdominal cross-sectional image," *J. Appl. Physiol.* **97**, 2333–2338 (2004).

[10] Kullberg, J., Ahlstrom, H., Johansson, L., and Frimmel, H., "Automated and reproducible segmentation of visceral and subcutaneous adipose tissue from abdominal MRI," *Int. J. Obes.* **31**, 1806–17 (2007).

[11] Rousson, M. and Paragios, N., "Shape priors for level set representations," in [*ECCV (2)*], 78–92 (2002).

[12] Huang, X., Li, Z., and Metaxas, D., "Learning coupled prior shape and appearance models for segmentation," in [*MICCAI*], (2004).

[13] Cremers, D., Rousson, M., and Deriche, R., "A review of statistical approaches to level set segmentation: integrating color, texture, motion and shape," *International Journal of Computer Vision* **72**(2) (2007).

[14] Saddi, K., Chefd'hotel, C., Rousson, M., and Cheriet, F., "Region-based segmentation via non-rigid template matching," in [*MMBIA (in conjunction with ICCV)*], (2007).