Proceedings of the 28th IEEE
EMBS Annual International Conference
New York City, USA, Aug 30-Sept 3, 2006

Unsupervised Segmentation of Cardiac PET Transmission Images for Automatic Heart Volume Extraction

Anu Juslin and Jussi Tohka

Abstract— **In this study, we propose an automatic method to extract the heart volume from the cardiac positron emission tomography (PET) transmission images. The method combines the automatic 3D segmentation of the transmission image using Markov Random Fields (MRFs) to surface extraction using deformable models. Deformable models were automatically initialized using the MRFs segmentation result. The extraction of the heart region is needed e.g. in independent component analysis (ICA). The volume of the heart can be used to mask the emission image corresponding to the transmission image, so that only the cardiac region is used for the analysis. The masking restricts the number of independent components and reduces the computation time. In addition, the MRF segmentation result could be used for attenuation correction. The method was tested with 25 patient images. The MRF segmentation results were of good quality in all cases and we were able to extract the heart volume from all the images.**

I. INTRODUCTION

Positron emission tomography (PET) is a medical imaging technique providing quantitative information about the physiological properties of living tissue. Because in PET emission images only the physiological properties of the tissues are seen, it is difficult to identify the exact position of some anatomical structures, like the myocardium in the thorax region. A transmission measurement using an external radioactive source is commonly performed to compensate for the effect of the attenuation of annihilation photons in the subjects's body. In comparison to the emission image, soft tissue and lungs are more identifiable in the transmission image, see Fig. 2. Because of this the transmission images have been used, in addition to attenuation correction, for image alignment and for motion detection and correction [1].

The aim of this study is to automatically extract the heart volume as it is seen in the transmission image to find the heart region for further analysis of the emission image. The extraction of the heart region is needed e.g. in independent component analysis (ICA) and factor analysis (FA) of the PET emission images. ICA and FA can be used to automatically extract the time activity curves (TACs) to calculate the perfusion of the myocardium [2], [3]. The volume of the heart is used to mask the emission image in such way that only the cardiac region is used for the analysis. This way it is easier to define the number of components to be separated from the emission image that is necessarily for ICA and FA [4]. In addition, the masking reduces the computation time. The completely automatic heart volume extraction from transmission images is challenging, because the heart tissue has the same attenuation factor than other soft tissues like the muscles. However, the position heart surface can be deducted based on position of the lungs. The lungs feature an attenuation factor different from the body, and therefore, the lungs can be identified in transmission images.

Our heart volume extraction method first segments the transmission image using Markov Random Fields (MRFs) [7]. Secondly, the heart volume is extracted based on the shape and position of lungs using the deformable models with dual surface minimization (DM-DSM) algorithm [6]. The search of the heart volume is automatically initialized using the segmentation result calculated in the first phase. The method was tested with 25 patient images.

This study was supported by TEKES Drug 2000 technology program, Graduate School of Tampere University, and the Academy of Finland, project No. 108517, No. 104824 and No. 213462 (Finnish Centre of Excellence program (2006-2011)).

A. Juslin and J. Tohka are with Institute of Signal Processing, Tampere University of Technology, FIN-33101 Tampere, Finland anu.juslin@tut.fi

II. METHODS AND MATERIALS

As already mentioned, the heart cannot be directly segmented from a transmission image, because we can identify only the soft tissue and the lungs based on the intensity values. However, anatomically the heart is located between the lungs, whose position can be defined by segmentation. We utilized this prior knowledge and first segmented the transmission image to 4 tissue classes; background (BG), lungs (LU), soft tissue (ST) and region between thorax and background including the bed. The last class was required because of the partial volume effect (PVE). In the second phase, the segmentation result was used to define the initialization for the deformable models based surface extraction. Before the segmentation, the transmission images were preprocessed using $3 \times 3 \times 3$ median filter to reduce noise in the images.

A. Segmentation of the transmission image

We assumed to have 4 tissue classes in our data, where one of the class is the partial volume (PV) class between the background and the thorax. To take into account the spatial correlations between the neighboring voxel labels in the segmentation, we used Markov random fields (MRFs). MRFs are statistically based models that use the fact that the probability of a voxel belonging to some tissue class depends also on its neighborhood.

The transmission image to be segmented is denoted by $X = \{x_i : i = 1 \ldots, N\}$, where N is the number of voxels in image. All the intensities in the image X are drawn from one of the K classes (in our study the $K = 4$). A context

image consisting of the labels of each voxel is denoted by $C = \{c_i \in \{1, 2, 3, 4\} : i = 1, \ldots, N\}$.

The intensities of the pure voxels (BG, LU, ST) were assumed to follow the normal density

$$p(x_i|\mu_k, \sigma_k^2) = [(2\pi)\sigma_k^2)]^{-1/2} \exp[-\frac{(x-\mu_k)^2}{2\sigma_k^2}] \quad (1)$$

where μ_k is the mean and σ_k^2 is the variance of the class k. The PV class voxels contains two types of tissues (ST, BG). The probability density function (pdf) of the PV class is constructed by marginalizing the densities over all possible values of partial volume coefficient (PVC) w, which defines the fraction of the background tissue in a mixed voxel. For the PV class, the marginal density is [5]

$$p(x_i|c_i = \{BG/ST\}) = \int_0^1 g(x_i; \mu(w), \Sigma(w))dw, \quad (2)$$

where

$$\mu(w) = w\mu_{BG} + (1-w)\mu_{ST}; \Sigma(w) = w^2\Sigma_{BG} + (1-w)^2\Sigma_{ST}, \quad (3)$$

for the PVC $w \in [0, 1]$. To estimate the parameters (μ_k, σ_k^2) for the class-conditional pdfs the K-means algorithm was used to cluster the intensities in the transmission image to 3 classes. The parameters (μ_k, σ_k^2) were estimated based on this clustering. The context image can be then estimated as

$$C^* = \arg\max_C P(C|X) = \arg\max_C P(C) \prod_{i=1}^{N} p(x_i|c_i), \quad (4)$$

where the prior term $P(C)$ is modeled by an MRF. Here we use a Potts model

$$P(C) \propto \exp(\beta \sum_{i=1}^{N} \sum_{j \in \mathcal{N}_i} \frac{a_{ij}}{d(i,j)}), \quad (5)$$

where β is 0.5, $\mathcal{N}_i$ is the 26 -neighborhood around voxel i, $d(i, j)$ is the distance between centers of voxels i and j. The coefficient a_{ij} is based on labels c_i and c_j: if $c_i = c_j$ then $a_{ij} = 1$; if either c_i or c_j is LU and the other is either BG or BG/ST then $a_{ij} = -100$, otherwise $a_{ij} = -1$. These coefficient values mean that in practice we do not allow the lung voxels to be neighbors of the background or the BG/ST class voxels.

The Iterated Conditional Modes (ICM) algorithm [7] was used to solve (4) locally. The standard ICM algorithm initialized by maximum-likelihood estimation was not applicable here, because the lungs and the region between the soft tissues and background feature highly similar intensity values. However, the ST class behaves differently from other classes and is reasonably well extracted with the K-means clustering. We can base our initialization on this: the voxels outside the body that were labeled as lungs in maximum-likelihood initialization were relabeled as BG/ST PV class in our initialization and the voxels inside the body labeled as BG/ST were relabeled as lungs. We used the flood filling algorithm to automatically decide which voxels were outside and inside the body.

B. Extraction of the heart volume

The heart volume was extracted using the DM-DSM (Deformable Models with Dual Surface Minimization) surface extraction algorithm [6], [8]. The surface extraction is reformulated as an energy minimization problem. The energy $E(S; I)$ of the surface S given an image I depends on the image data and the properties of the surface itself. It is a weighted sum of the internal energy penalizing surfaces that are not smooth and the external energy that couples surfaces with the image data. The total energy of the surface S is

$$E(S; I) = \lambda E_{int}(S) + (1 - \lambda)E_{ext}(S; I), \quad (6)$$

where $E_{int}(S)$ is the internal energy, $E_{ext}(S; I)$ is the external energy, and $\lambda \in [0, 1]$ is the regularization parameter controlling the tradeoff between external and internal energies. We set $\lambda = 0.3$.

The external energy was obtained using 3D Sobel gradient operator [9]. We improved the external energy component masking out the thorax surface and the bed using the segmentation result. The removal of the thorax surface from the gradient image made it easier to initialize the algorithm, because between the thorax and background there is a strong edge. This way the thorax surface could not interfere with the heart surface extraction.

We used an ellipsoid as the initialization for DM-DSM. Both the center point and the size of the ellipsoid were computed based on the segmented lungs from the transmission image (Fig. 1). The anatomy of the lungs was used as a prior information based on the fact that the heart is located within mediastinum and bordered laterally by the lungs. We explain the determination of the center point, the determination of the semi-axes was similar in spirit and it is not explained here. The z-coordinate of the center point was defined as the z-coordinate of the mass center point of the lungs. The x- and y-axis coordinates were calculated using the lung contours on the already defined z-plane. We computed the distance between the left and right lung, when y-coordinate was varied (above the calculated mass center point in the y-axis.). The y-coordinate, which gave the maximum distance was selected as the y-coordinate of the center point. The x-coordinate was defined to be in the center of the lungs.

C. Transmission images

The test data consisted of 25 transmission images acquired with a GE Advanced scanner (General Electric Medical Systems, Milwaukee, Wisconsin, USA) from healthy female volunteers. The images were reconstructed using the MRP method [11]. In the reconstructed 3D image (size of 128*128*35) each voxel had the volume of 4.3x4.3x4.25 cubic millimeters.

III. Results

The MRF based segmentation was excellent in all cases in visual inspection (Figs. 2, 3, 4). This despite the fact that the test data set consisted of subjects of very different sizes, as it can be seen from the Figs 2-4. These three examples of 25

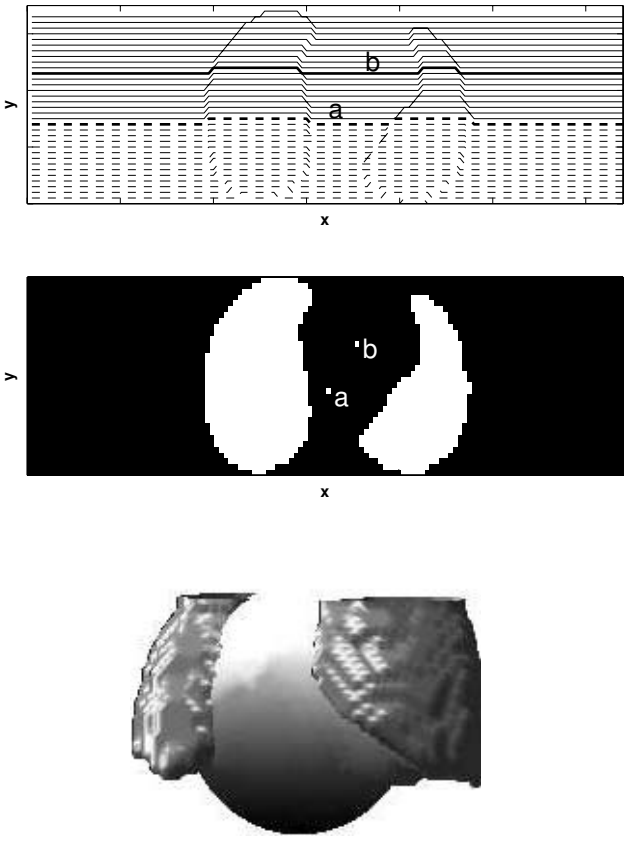

Fig. 1. The center point (b) for the initial ellipsoid was calculated above the mass center point of the lungs (a) using the lung countours (top and middle rows), the 3D rendering of the initial ellipsoid (white) and lung surfaces (gray)(bottom row).

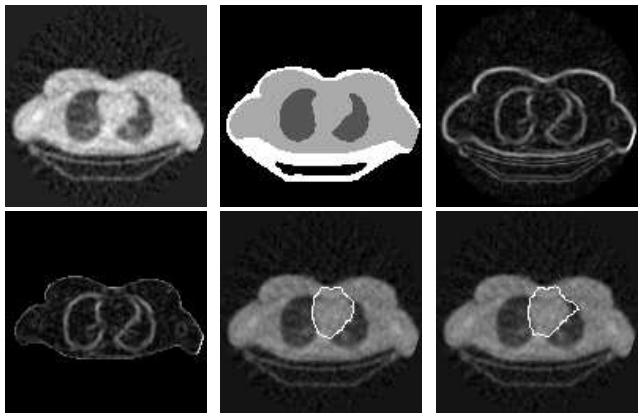

Fig. 2. The original transmission image (left), segmentation result of the transmission image (middle) and gradient image before modification (right) in the top row. In the bottom row the energy image after the removal of the thorax surface(left) and the extracted heart surface with automatic (middle) and manual (right) initializations. For all images the same transaxial slice is shown. The Jaccard coefficient between volumes was 0.86.

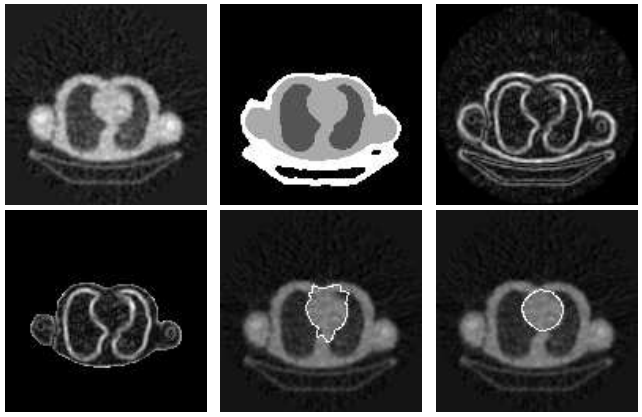

Fig. 3. An example of heart volume extraction result (cf. Fig 2). The Jaccard coefficient between volumes was 0.65.

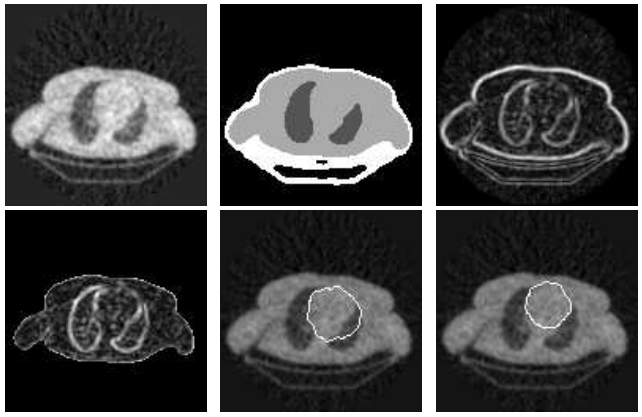

Fig. 4. An example of heart volume extraction result (cf. Fig 2). The Jaccard coefficient between volumes was 0.56.

cases were selected, because they depict the big variability among the images.

We performed quantitative evaluation of the automatically extracted heart volumes against the heart volumes extracted with DM-DSM, but with a manual initialization. We compared the similarity of the extracted volumes using Jaccard coefficient [10], which ranges from 0 for volumes that have no common voxels to 1 for volumes that are identical. The mean value for the coefficient between the extracted volumes was 0.70 with standard deviation 0.09.

When analyzing the visual result of the surface extraction we can see that the surface was sometimes stuck on the outer surface of the lungs (Figs. (2- 4)). This happened both with the manual (Fig. 2) and automatic (Fig. 4) initializations. The problem could be eliminated using the knowledge of the exact position of the lung surfaces and by removing the lateral parts of the lung surfaces in a similar manner than we did for the thorax surface. However, this is not a problem for masking the emission image for ICA as far as the extracted volume covers the whole heart region.

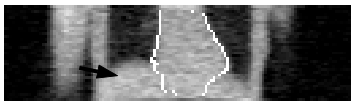

Fig. 5. An example of cross section of the extracted heart volume overlaid on coronal image slice. The arrow points to the liver.

IV. Discussion

In this study, we developed an automatic 3D segmentation method for extracting the heart volumes from PET transmission images. In the method the transmission image was first segmented to 4 tissue classes; background, lungs, soft tissue, and region between thorax and background with bed using Markov Random Fields (MRFs). The segmentation result was then used to automatically define the initialization for a deformable surface, which was used for the extraction of the heart volume. We tested the algorithm with 25 patient images and the method was capable to automatically extract the heart volume in all cases. The extracted heart volumes were compared to volumes extracted with manual initialization and the volumes were found to be similar. The advantage of using automatic method for initialization is that the extraction result is always the same, whereas, with manual initialization, the results can vary along with the initialization. We defined the manual initialization for some images second time after two weeks (both definitions by A.J.) and compared the extracted volumes. The average Jaccard coefficient was approximately 0.8, which is not much higher than the average Jaccard coefficient between automatic and manual initializations. Furthermore, it was time consuming to define manually the initialization from the 3D transmission images. Note that ground truth of the real heart volumes are not available in our subject data and we cannot quantitatively say which initialization produced the best result.

The heart volume is needed in independent component analysis (ICA) and factor analysis (FA) to restrict the image area in emission image [2], [3]. The masking reduces the computation time and simplifies the selection of the number of components to be separated. For example in oxygen-15-labeled water images the liver tissue is problematic for ICA, because its functional behaving is similar to the myocardium. Using the anatomic heart volume mask most of the liver can be eliminated from the analysis. The problem of using transmission images to define the mask is that there is no visible border between myocardium and liver, and their intensity values are similar in the transmission image. For this reason, part of the liver was included in the masks (Fig. 5). Additionally, the heart volume included some other parts of the mediastinum than just heart tissue. This is not a problem for the component analysis, because there occur heart and respiratory movement during the emission acquisition, which need to be taken into account.

The segmentation of transmission images have been previously used to calculate more accurate attenuation factors [12], [13], [14]. In comparison to previously proposed segmentation algorithms, the segmentation algorithm developed in this study is unsupervised and it does not require predefined threshold values. Using the MRF for the segmentation the local characteristics of the 3D structures were also taken into account. Our algorithm did not have problems with larger subjects, as it was reported using 1D Hidden Markov Models and the Viterbi algorithm [14]. In addition, we included explicit modeling of the partial volume effect (PVE) between the background and soft tissue (body), which could provide even more accurate estimation of the attenuation factors for the thorax region.

V. Acknowledgments

The authors would like to thank Turku PET centre for providing the data.

References

[1] D. L. Bailey, Transmission Scanning in Emission Tomography, *Eur J Nucl Med*, vol 25, no. 7, 1998, pp 774–787.

[2] J.S. Lee, D.S. Lee, J.Y. Ahn, G.J. Cheon, S-K. Kim, J.S. Yeo, et al., Blind Separation of Cardiac Components and Extraction of Input Function from $H_2^{15}O$ Dynamic Myocardial PET Using Independent Component Analysis, *J Nucl Med*, vol 42, no. 6, 2001, pp 938–943.

[3] H-M. Wu, C. K. Hoh, Y. Choi, H. R. Schelbert, R. A. Hawkins, M. E. Phelps, and S-C Huang, Factor Analyses for Extraction of Blood Time-Activity Curves in Dynamic FDG-PET Studies, *J Nucl Med*, vol 36, no. 9, 1995, pp 1714–1721.

[4] A. Hyvärinen, J. Karhunen, E. Oja, *Independent Component Analysis*, John Wiley Sons, Inc; 2001.

[5] J. Tohka, A. Zijdenbos, and A. Evans, Fast and Robust Parameter Estimation for Statistical Partial Volume Models in Brain MRI, *NeuroImage*, vol 23, 2004, pp 84–97.

[6] J. Tohka, J. Mykkänen, Deformable Mesh for Automated Surface Extraction from Noisy images, *Int. J. Image and Graphics*, vol 4, no. 3, 2004, pp 405–432.

[7] J. Besag, On the Statistical Analysis of Dirty Pictures, *J. R. Stat. Soc., Ser B*, vol 48, no 3, 1986, 259–302.

[8] A. Juslin, A. Reilhac, M. Magadán-Méndez, E. Albán, J. Tohka, and U. Ruotsalainen, Assessment of Separation of Functional Components with ICA from Dynamic Cardiac Perfusion PET Phantom Images for Volume Extraction with Deformable Surface Models, *In Proc. of Third International Conference on Functional Imaging and Modelling of the Heart (FIMH'05), Lecture Notes in Computer Science*, vol. 3504, 2005, pp 338–347.

[9] S. Zucker, and R. Hummel, A Three-Dimensional Edge operator, *Trans. Patt. Ana. Mach. Intell*, vol 3, no. 3, 1981, pp 324–331.

[10] D.A. Jackson, K.M. Sommers, H.H Harvey, Similarity coefficients: Measures of Co-occurence and Association or Simplymeasures of Occurrence, *The American Naturalist*, vol 133, no. 3, 1989, pp 436–453.

[11] S. Alenius, U. Ruotsalainen, and J. Astola, Using Local Median as the Location of the Prior Distribution in Iterative Emission Tomography Image reconstruction, *IEEE Trans Nucl Sci*, vol 45, no 6(2), 1998, pp 3097–3107.

[12] M. Xu, N. A. Mullani, K. L. Gould, and W. L. Anderson, A Segmented Attenuation Correction for PET, *J Nucl Med*, vol 32, 1991, pp 161–165.

[13] H. Zaidi, M. Diaz-Gomez, A. Boudraa, and D. O. Slosman, Fuzzy Clustering-Based Segmented Attenuation Correction in Whole-Body PET Imaging, *Phys. Med. Biol.*, vol 47, 2002, pp 1143–1160.

[14] J.M.M. Anderson, R. Srinivasan, B.A. Mair, and J. Votaw, Hidden Markov Model Based Attenuation Correction for Positron Emission Tomography, *IEEE Trans. Nucl. Sci*, vol. 49, 2002, pp 2103–2111.