Hierarchical Parsing and Semantic Navigation of Full Body CT Data

S. Seifert[a], A. Barbu[b], K. Zhou[c], D. Liu[c], J. Feulner[a], M. Huber[a], M. Suehling[d], A. Cavallaro[e] and D. Comaniciu[c]

[a]Integrated Data Systems, Siemens Corporate Technology, Erlangen, Germany, [b]Department of Statistics, Florida State University, Tallahassee, FL, USA, [c]Integrated Data Systems, Siemens Corporate Research, Princeton, NJ, USA, [d] Computed Tomography, Siemens Healthcare, Erlangen, Germany, [e]University Hospital, Erlangen, Germany

ABSTRACT

Whole body CT scanning is a common diagnosis technique for discovering early signs of metastasis or for differential diagnosis. Automatic parsing and segmentation of multiple organs and semantic navigation inside the body can help the clinician in efficiently obtaining an accurate diagnosis. However, dealing with the large amount of data of a full body scan is challenging and techniques are needed for the fast detection and segmentation of organs, e.g., heart, liver, kidneys, bladder, prostate, and spleen, and body landmarks, e.g., bronchial bifurcation, coccyx tip, sternum, lung tips. The problem becomes even more challenging when partial body scans are used, where not all organs are present. We propose a new approach to this problem, in which a network of 1D and 3D landmarks is trained to quickly parse the 3D CT data and estimate what organs and landmarks are present as well as their most probable locations and boundaries. Using this approach, the segmentation of six organs as well as the detection of 19 body landmarks can be obtained in about 20 seconds with state of the art accuracy. To illustrate its effectiveness, the new approach has been validated on 80 CT full or partial body scans with very good results.

Keywords: organ segmentation, probabilistic atlas, full body navigation

1. INTRODUCTION

Whole body scanning is a common practice for diagnosis of systemic diseases such as cancer and for differential diagnosis. But the large amount of data makes it difficult for the clinician to work efficiently. Thus, automatic tools are strongly needed to support the clinician in reading and navigation. Recently, robust solutions to the heart[1] and liver[2] segmentation have been proposed. But they are organ specific and costly to transfer to new organs. In this paper we present a new framework for hierarchically parsing whole body CT images and efficiently segmenting multiple organs, taking contextual information into account. Our framework enables the user to segment six organs and to detect 19 landmarks very fast and robustly in about 20 seconds. New anatomy is easy to incorporate since the framework can be trained and handles the segmentation of organs and the detection of landmarks in a unified way. The detected landmarks and segmented organs facilitate the semantic navigation inside the body.

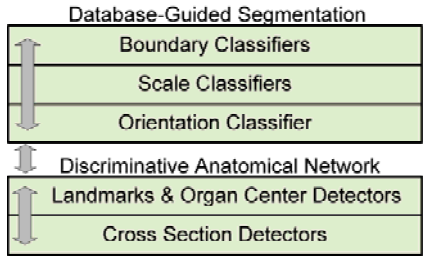

Figure 1. System Diagram.

Most approaches to segmentation of multiple structures rely on probabilistic atlases[3] that are registered non-rigidly to the current scan. Such approaches have two disadvantages. First, a non-rigid registration cannot handle topological changes in the organ boundaries. Such changes occur when an organ has moved from its usual place or is greatly enlarged due to disease. Second, registration techniques are usually computationally intensive.

In this paper, we present a different approach that is based on the Marginal Space Learning[1] idea recently introduced in medical imaging. In Marginal Space Learning, the organs are detected using a sequence of learned classifiers, starting with classifiers with a few parameters (e.g., organ position without orientation and scale) and ending with a classifier that models all the desired organ parameters (e.g., position, orientation and scale). Each learned classifier is a Probabilistic Boosting Tree[4] with 2D[5] or 3D[6] Haar features and trained by AdaBoost.[7]

2. DISCRIMINATIVE ANATOMICAL NETWORK

The body landmarks serve three purposes. First, some landmarks have clinical importance (e.g. aortic bend) and help the clinician find valuable information about the patient. Second, some landmarks are very salient and can be robustly detected, thus help improve the system's robustness through the anatomical network. Third, the 1D landmarks / slices help speed up the detection process by serving as reference points for where to search for the other landmarks. Integrating the 1D landmarks result in a speedup of about two orders of magnitude in detecting all the landmarks and organ centers.

2.1 Slice Detectors

The CT and MR full body scans have a natural alignment of the z-axis with the body since the scans are acquired as axial cross-sections (slices). Browsing these slices, one finds that some are very salient and can be easily detected.

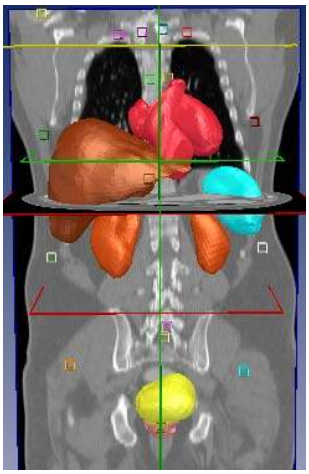

Figure 2. Whole body scan with organs and 19 body landmarks automatically obtained by our method.

Based on this idea, we aligned 44 volumes at 4mm resolution along the z-axis by the z-location of the lung tips, aortic arch, liver center, right kidney center and coccyx. Then we generated 100 slices from each volume by linear interpolation between the above mentioned alignment positions. Thus for example, slice 10 corresponds in all 44 volumes to the aortic arch (Figure 4, top-right). Finally, we trained 100 slice detectors using PBT[4] and 2D Haar features,[5] with the k-th detector using slice k from all volumes as positive examples and all other slices at a distance at least 10 from slice k as negatives.

An evaluation of the 100 detectors using 5-fold cross-validation revealed that three slices that are the most salient: slice 4 - the beginning of the lungs, slice 44 - the beginning of the liver and slice 93 - in the lumbar region. Examples of these slices are showed in Figure 3. In each slice, the bounding box around the body (shown in Figure 3) is found by simple thresholding and pixel-counting operations. We retrained these three slices, denoted by z_0, z_1, z_2, with all the available data from 110 manually annotated volumes. The learned slice detector probabilities are denoted by $p(Z_i = z_i|I)$. An example of detected slices can be seen in Figure 2. The three slice detectors are connected into an anatomical network which ensures that the relative position of the slices is correct. Basically this network is a Markov Random Field with the energy:

$$E(z_0, z_1, z_2) = \begin{cases} \infty & \text{if } z_1 < z_2 + D_0 \text{ or } z_0 < z_1 + D_0 \text{ or if } z_1 = -1 \text{ and } z_0, z_2 \geq 0 \\ -\sum_{i=0}^{2} \log p(Z_i = z_i|I) + (z_0 - z_1 - \mu_{01})^2/\sigma_{01}^2 + (z_1 - z_2 - \mu_{12})^2/\sigma_{12}^2 & \text{otherwise.} \end{cases}$$

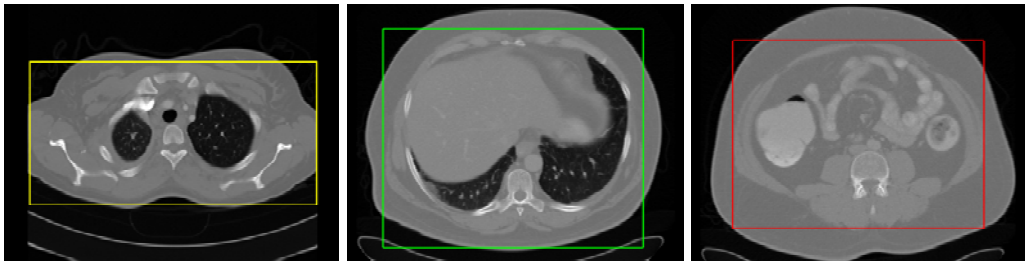

Figure 3. The three cross-sections (slices) that are used as reference for the landmark detectors.

For inference, slices z_0 and z_1 are restricted to have two values: -1 (when the slice does not exist) and the location that maximizes the probability $p(Z_i = z_i|I)$. Then the last slice location z_2 is found to minimize the energy $E(z_0, z_1, z_2)$. Using this inference algorithm, the Discriminative Anatomical Network is an Active Random Field.[8] Training the MRF parameters is an optimization procedure based on simulated annealing to minimize the misclassification error. In 10000 steps, the parameters were obtained with a misclassification error of 0.91%.

2.2 Landmark Detectors

The landmarks and organ centers are 3D points and treated the same way in our system. Below we simply call them landmarks. The landmark detectors are based on PBT and 3D Haar features.[6] Each landmark is connected to one of the three detected slices. If the slice is reported as present, the detector is started with a search range relative to the slice location. The search ranges are typically 300 times smaller than the full body, hence a speedup of two orders of magnitude is obtained.

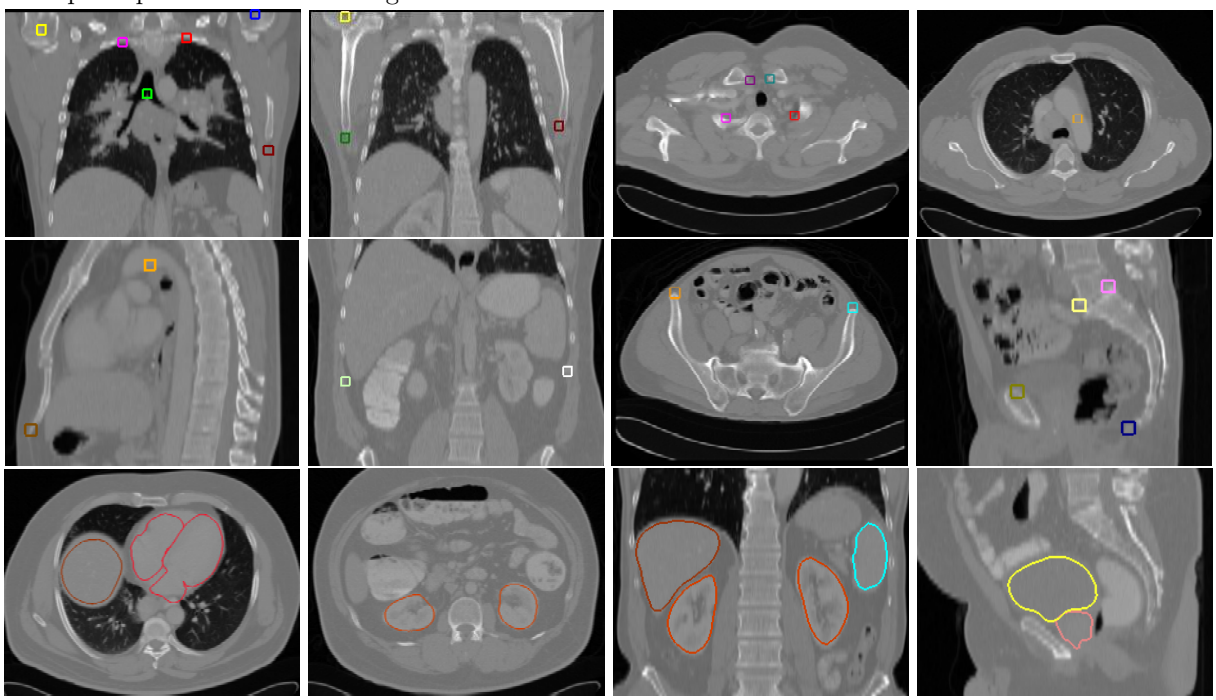

Figure 4. Detected landmarks and segmented organs: heart(red), liver(brown), kidneys(orange), spleen(light blue), bladder(yellow) and prostate (pink) in our system.

The following 19 landmarks have been trained in our system: left and right lung tips, left and right humerus heads, bronchial bifurcation, left and right shoulder blade tips, inner left and right clavicle tips, sternum tip bottom, aortic arch, left and right endpoints of rib 11, bottom front and back of the L5 vertebra, coccyx, pubic symphysis top and the left and right front corners of the hip bone. The following ten organ centers have also been trained: four heart chambers, liver, kidneys, spleen, prostate and bladder.

To obtain a fast and robust system, the landmarks (including organ centers) are connected in a graph (network). Information regarding the location of each landmark is propagated across the edges of the graph, which not only speeds up detection, but also increases detection accuracy. This is because the edges encode geometric relationships such as "to the right of", "close to", etc., and thus constrain the search problem into a smaller domain.

Denote the landmark detectors as $\phi_i(x_i|V)$, where random variables x_i, $i = 1, ..., N$, denote the unknown position, scale, and orientation parameters of the N landmarks, depending on the input volume V. The landmark detectors are incorporated into the DAN through a Markov Random Field as follows:

$$P(x_1, ..., x_N) = \frac{1}{Z} \prod_i \phi_i(x_i|V) \prod_{(i,j)\in E} \psi_{ij}(x_i, x_j)$$

where E denotes the set of edges in the network between landmarks, and $\psi_{ij}(x_i, x_j)$ expresses the pairwise geometric relationship between landmarks. Based on the Belief Propagation algorithm, the marginal distribution of x_i is $P(x_i) \propto \phi_i(x_i|y_i) \prod_{j\in E(i)} m_{ji}(x_i)$, where $E(i)$ contains the neighbors of landmark i in the network, and $m_{ji}(.)$ is computed iteratively according to the schedule:

$$m_{ij}(x_j) = \sum_{x_i} \Big(\phi_i(x_i|y_i) \psi_{ij}(x_i, x_j) \prod_{k\in E(i)\setminus j} m_{ki}(x_i) \Big)$$

2.3 Organ Segmentation

The organ segmenters contain a trained object detector together with a shape refinement module. The object detector, based on Marginal Space Learning, finds the position, orientation and scale of the organ. For robustness and performance reasons, the detectors take contextual information into account coming from the anatomical network. This information is used as a priori knowledge to constrain the marginal space for subsequent organ segmentations. The shape refinement module consist of a trained boundary model and a PCA shape model. The multi-chamber heart segmenter and the liver segmenter have been described in previous work.[1,2] The other segmenters, similar in concept to the liver segmenter, use a hierarchical refinement and have own customizations specific to each organ (configuration paths, thresholds, pyramid levels for detection and segmentation, etc.).

3. RESULTS AND CONCLUSION

Evaluation of the system is done at three levels, namely slice, landmark and organs.

At the slice level, an evaluation on 588 volumes revealed that in 586 volumes the correct slices were reported, obtaining an accuracy of 99.7%.

At the landmark level, an evaluation on 80 volumes revealed that all landmarks were detected in all volumes, with no false detections. The error histogram is shown in Figure 5, with an average of 7mm and 95% of errors less than 18.5mm. Landmark and slice detection together take about 400ms on a regular PC.

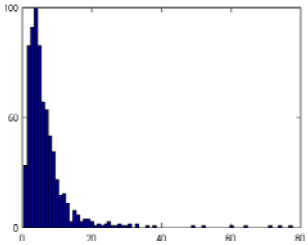

Figure 5. Histogram of landmark errors in mm.

At the organ level, the precision of the organ segmentation was evaluated with cross-validation using a mesh-to-mesh error metric (see Table 6). The heart and liver detectors and segmenters are well tested and already in use in practical applications. The training of the kidney, spleen, bladder and prostate is ongoing. Consistent with our expectations, the mesh error decreases as the number of annotations increases. In terms of detection speed, the approach using the anatomical network, yielded a speedup of about 50%, compared with one using detectors individually. Future work involves the inclusion of other organs, landmarks, bones and vessels.

organs	volumes	mesh error [mm]	detect. time [s/vol]	segm. time [s/vol]
liver	174	1.26	0.67	16.0
heart	457	0.84	1.07	3.55
kidneys	97	1.65	0.67	0.20
spleen	69	2.20	1.06	0.20
bladder	56	2.40	1.03	0.57
prostate	40	2.40	2.00	0.59

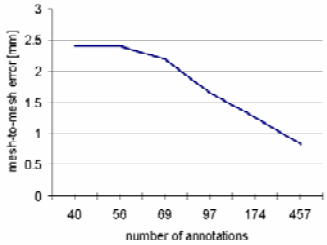

Figure 6. Evaluation results of the precise organ segmentation.

REFERENCES

[1] Y. Zheng, A. Barbu, B. Georgescu, M. Scheuering, and D. Comaniciu, "Fast Automatic Heart Chamber Segmentation from 3D CT Data Using Marginal Space Learning and Steerable Features," *ICCV* , 2007.

[2] H. Ling, K. Zhou, Y. Zheng, B. Georgescu, M. Suehling, and D. Comaniciu, "Hierarchical, Learning-based Automatic Liver Segmentation," *CVPR* , 2008.

[3] P. Thompson, D. MacDonald, M. Mega, C. Holmes, A. Evans, and A. Toga, "Detection and Mapping of Abnormal Brain Structure with a Probabilistic Atlas of Cortical Surfaces.," *JCAT* .

[4] Z. Tu, "Probabilistic Boosting-Tree: Learning Discriminative Models for Classification, Recognition, and Clustering," *ICCV* **3**(5), 2005.

[5] M. Oren, C. Papageorgiou, P. Sinha, E. Osuna, and T. Poggio, "Pedestrian detection using wavelet templates," *Proc. Computer Vision and Pattern Recognition* **97**, pp. 193–199, 1997.

[6] Z. Tu, X. Zhou, A. Barbu, L. Bogoni, and D. Comaniciu, "Probabilistic 3D Polyp Detection in CT Images: The Role of Sample Alignment," *CVPR* , pp. 1544–1551, 2006.

[7] R. Schapire and Y. Singer, "Improved Boosting Algorithms Using Confidence-rated Predictions," *Machine Learning* **37**(3), pp. 297–336, 1999.

[8] A. Barbu, "Training an Active Random Field for Real-Time Image Denoising," *Technical Report* , 2008.