A NEW METHOD FOR ROBUST ORGAN POSITIONING IN CT IMAGES

Torbjørn Vik *Daniel Bystrov* *Nicole Schadewaldt* *Heinrich Schulz* *Jochen Peters*

Philips Research Hamburg

ABSTRACT

Robust initialization is essential for any successful segmentation process of medical images. For CT images, initialization is challenging because quality, appearance, content, and field-of-view of the images are highly variable. Furthermore, high execution speed is desirable, whereas the user tolerance to errors is low in clinical applications. We present a new method for efficient and robust positioning of organs in CT images. It is based on a novel probabilistic atlas that, given a tissue type, describes the probability density of the random variable *spatial location*. Random sampling is then employed to select the most informative points for matching. We present results on pelvic and abdominal images acquired for radiotherapy planning.

Index Terms— atlas-image registration, probabilistic registration, CT, radiotherapy

1. INTRODUCTION

Our aim is the automated organ-at-risk segmentation on CT images used for radiotherapy planning. The segmentation of a given image is obtained after applying several algorithms in a sequence. The first step is the initialization step, where we pre-position multiple organ surface meshes. A good initialization assures that the initial solution is in the capture range of the subsequent active shape algorithms. Any error in this positioning reduces the overall segmentation performance. For a commercial system, we therefore require a very robust algorithm that is capable of dealing with images from different institutions showing high variability in quality, appearance (e.g. contrast), content, and the field-of-view (FOV).

Several methods for initialization exist. One, often used for brain MR image segmentation, is to use low-parametric atlas–image registration (intensity-based using sum-of-squared differences, normalized mutual information, or class-label-based using the Kullback-Leibler divergence [1]). Direct application of such approaches do not seem well suited for soft-tissue organs and need integration over the whole image domain, which is inefficient. In our method, we choose a different similarity measure that allows us to develop an efficient optimization scheme. Another line of possible approaches are based on discriminative learning (boosting and Haar wavelets) [2, 3, 4]. These methods are very efficient, but require large amounts of training data with ground truth segmentation. In [5], the authors proposes a probabilistic model for the abdomen region, but it depends on the segmentation and identification of spinal vertebras. Finally, template matching based on the image intensity gradient and implemented using the generalized Hough transform [6] has also been shown to be a very effective and efficient method when the dimension of the accumulator remains small (e.g. 3-dimensional for translational positioning of a single organ). However, this is not the case in our aplication where we position organ groups such as bladder, prostate, rectum, left and right femoral heads simultaneously.

We present a new, registration-based approach. The main novelty of this approach is that we model the probability density of a spatial location given a tissue class. This allows us to develop an inference method that is attention-driven, robust and efficient. We describe the method in Section 2 and show experimental results in Section 3.

2. METHOD



(a) Image $u(x)$ (b) $P(t = \text{bone}|u(x))$ (c) $q(x|t = \text{bone})$

Fig. 1. General idea: An image (a) is decomposed into tissue class probabilities based on gray values. In (b), we show the bone tissue probability of (a) together with samples of typical bone positions (red dots). These samples are sufficient to find the anatomic structure in a tissue atlas (c).

2.1. General idea

We start by giving an intuitive explanation before proceeding to a more formalized description of our matching procedure. Given an observed image as in Fig. 1(a), the task is to position one or more organ models (surface meshes) such as bladder,

prostate, etc. Given an atlas (a reference space), in which the relative positions and sizes of the models are coded, this task can be solved by determining a diffeomorphic, spatial transformation (e.g. similarity transformation: translation, rotation and isotropic scaling) between the atlas and the given image.

The key idea behind our approach is that only very little information from the image is necessary in order to determine such a transformation. Consider Fig. 1(b), which shows a certain set of points. If these points are representative random samples of positions of bony structure in the pelvis area, we are able to give a good estimate of the transformation to the atlas Fig. 1(c).

Many approaches are based on an averaged gray value atlas, which is again based on a Gaussian model of the probability density function (pdf), say $q(u|x)$. The main novelty of our approach is to consider $q(x|u)$, the pdf of a spatial location in atlas space, x, given a gray value u. This distribution can be considerably simplified if we make use of a hidden variable $t =$ *tissue type* in order to decouple spatial information x, and intensity u,

$$q(x|u) = \sum_t q(x|t,u)q(t|u) \approx \sum_t q(x|t)q(t|u). \quad (1)$$

This approximation is based on the assumption that for the spatial probability density the intensity can be neglected when the tissue type is known, i.e. $q(x|t,u) \approx q(x|t)$. Examples of tissue type definitions are shown in Fig. 2.

2.2. Matching

For the derivation of the matching criterion, let us define the following:

- $u \in R$ is image gray value.
- $f_\theta : R^3 \to R^3$ is a spatial transformation parametrized by $\theta \in \Theta \subset R^D$. For the similarity transform, $D = 7$ (translation, rotation, and istropic scaling).
- $P, Q : T \to [0,1]$ designates a probability of tissue type, $t \in T$. P relates to the image, Q to the atlas.
- $p, q : x \in \Omega \subset R^3 \to [0,\infty)$ is the pdf of the continuous variable x. p relates to the image, q to the atlas.

Next, we consider tissue types. Since the Hounsfield units in CT images are calibrated, we consider the probability $P(t|u)$ of a tissue type $t \in T$ (e.g. $T = \{$air, bone, fat and water$\}$) given a gray value u. For a given image the gray value at a fixed position x is known, so that we can use $P(t|u)$ to determine $P(t|x)$, compare Fig. 1(b). From Bayes formula, we have that

$$p(x|t) = \frac{P(t|x)p(x)}{P(t)} \propto P(t|x), \quad (2)$$

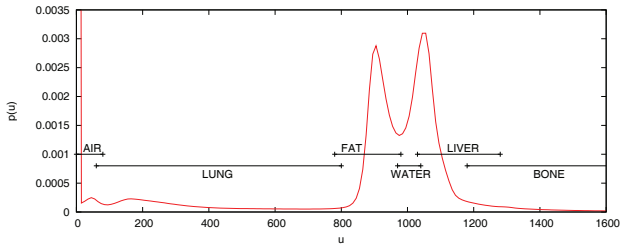


Fig. 2. Histogram and tissue type definition. Shown is a typical histogram (normalized so that the integral is 1) of CT gray values (+1000) in abdominal scans as well as the ranges (pairwise overlapping) defining the tissue types.

if samples (x,t) are drawn with a spatially uniform distribution $p(x)$. This is the case when the image is uniformily sampled. We see that there is a strong dependency between spatial location x, and the tissue type t, see Fig. 1 and Fig. 3.

We propose to determine the unknown transformation, f_θ, by comparing the pdf of the image, $p(x|t)$, with the pdf of the transformed atlas, $q(x|t,\theta)$:

$$-\sum_t \int_{x\in\Omega} p(x|t) \log q(x|t,\theta) \quad (3)$$

The rationale behind this choice is to deform the atlas q such that tissue positions in the image p are well predicted. This approach has the advantage to be robust to varying image information such as varying FOV, and varying relative tissue content: due to the singularity of the $\log$ function, (3) has the following property: if $p(x|t) > 0$ then θ must be choosen such that $q(x|t,\theta) > 0$ as well. Therefore it is important that the atlas covers a FOV larger than any image. Intuitively, (3) tries to *explain* all tissue observations in the image with the atlas. Note that it is essential to have more than one tissue class, $t > 1$, to avoid a degenerate transformation to be optimal.

When both pdfs p and q are normalized over the same domain, (3) is identical to cross-entropy (and related to Kullback-Leibler divergence) [7]. This is only the case when the same anatomy is covered in the image and the atlas, e.g., whole body or same field-of-view (FOV).

Finally, in order to implement the matching, we make an efficient approximation to (3) using random sampling. The approximation is based on Monte Carlo sampling and results from rewriting the terms of (3) as expectations

$$E_{p(x|t)}\{\log q(x|t,\theta)\} \approx \sum_i \log q(x_i|t,\theta), \quad (4)$$

where each x_i is drawn from $p(x|t)$. Before minimizing (3), we draw a set of samples x_i from $p(x|t)$ and then optimize the transformation parameters θ so that the transformed points have high probabilities in the atlas space. To draw spatial samples, we draw uniform samples on $[0,1]$ and map to spatial location, x, using the inverse of the associated cumulative

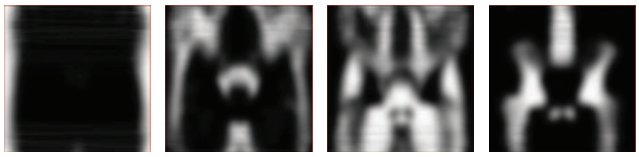
density function: $\int_{-\infty}^{\tilde{x}} p(x|t)$. Global optimization is done with the particle swarm algorithm [8].

To summarize the approach: We rely on intelligent information reduction before solving a minimization problem. The information reduction is based on (1) reducing continuous gray values u to discrete tissue types t, and (2) selecting only a representative number of samples typical of the spatial distribution of a tissue type.

2.3. Learning

Learning in our approach consists of determining the atlases $q(x|t)$, $t \in T$ as well as the tissue type probabilites $P(t|u)$ from a set of aligned training images. To determine the tissue type probabilites, $P(t|u)$, we start by creating a histogram of typical gray values in the region of interest of the training images. Then, we define a set of pair-wise overlapping ranges based on analysis of the histogram and experiments, see Fig. 2. In the overlap regions we define a linear transition probability so that $\sum_t P(t|u) = 1$.

We then proceed to warping all training images into a reference space, e.g., an average space or the space of one chosen reference image. In this space, we calculate the probabilities of each tissue type for each voxel and average the statistics voxel-wise across all training images to obtain $Q(t|x)$ which can be renormalized to obtain $q(x|t)$ as in (2). Example atlases can be seen in Fig. 3. Due to averaging, the resulting tissue-type atlas is smooth, which is desirable for optimization, and the FOV is as large as possible.



air fat water bone

Fig. 3. Tissue atlases for pelvic region created by registering a training set of N = 60 images, u_i, with a similarity transformation and subsequent smoothing and averaging of the tissue probabilities $P(t|u_i(x))$.

3. EXPERIMENTS

The application for our approach is to initialize a segmentation pipeline for automatic delineation of normal structures in planning CT images for radiotherapy. The challenge hereby is the large variability in FOV, clinical protocol and image quality (different institutions and scanner types). The most appropriate validation of our method with respect to this application is to assess whether the initialization is within the capture range of the next step of the segmentation pipeline. To avoid a complete description of the pipeline, however, we limit the experiments section to give quantitative results on ground truth landmarks. As transformation model, we use the similarity transform.

3.1. Pelvis

For pelvic matching we use 500 samples from each of the tissue types (min, max) HU + 1000: air (0, 800), fat (780, 980), water (970, 1100), bone (1080, 2500). For validation, we have defined seven landmark positions in 89 images: left- and right femoral head, end of sacrum, pubis, left- and right most anterior position of the iliac crest, and the most anterior point between L5/S1. These landmarks were used to calculate a ground truth transformation for each image. During training, we used the ground truth transformation to calculate mean landmarks in the atlas space. For evaluation, these mean landmarks were then mapped back to the image space of the testing images with (1) the ground truth transformation and (2) the transformation estimated with our criterion. The difference between (1) and (2) was then analyzed using three-fold cross-validation (three data groups, cyclically using two for training and one for testing). Tab. 1 shows the mean error over all datasets, both euclidean error d_μ and error with standard deviation ($\mu \pm \sigma$) in x (left-right), y (anterior-posterior) and z (head-feet) direction.

The average error is low: below 7mm for each landmark with standard deviations in each direction up to 8mm. No outliers were observed. Most of the error is due to a systematic under-estimation of the scaling parameter. This places femurs (x-error) and iliac crests (x- and y-error) too close to the pelvic center. However, the central area is matched very well, as seen on the L5/S1 landmark, the y-error for the femoral heads, the z-error for iliac crests, and the x-error for the sacrum and pubis landmarks. The reason for this under-estimated scaling is that the atlas has higher probabilities in the center than in the peripheral parts. The ground truth transformation brings the central region better into registration than the peripheral regions. The result is sufficient to successfully initialize our segmentation pipeline for all organs and cases and runs in approximately 5 seconds (Intel Xeon 2.67 GHz RAM 6 cores).

3.2. Liver

For liver matching we use 2000 samples from each of the tissue types (min, max) HU + 1000: air (0, 80), lung (60, 800), fat (780, 980), water (970, 1040), liver (1030, 1280). The runtime was similar to the pelvis case. We used 105 abdominal datasets and did cross-validation with three groups again. It is more difficult to find anatomical landmarks relevant for the liver position. Instead, we use ground truth liver segmentations and compare the liver centers after transforming the ground truth liver of the testing data sets with (1) the ground truth transformation and (2) the estimated transformation. The average error is below 2mm (last line of Tab. 1).

Landmark	d_μ	$\mu_x \pm \sigma_x$	$\mu_y \pm \sigma_y$	$\mu_z \pm \sigma_z$
pubis	3.7	-0.1 ± 2.3	-1.8 ± 7.1	3.2 ± 4.7
l. femur	4.0	3.5 ± 4.6	-0.9 ± 5.9	1.7 ± 4.2
r. femur	3.9	-3.5 ± 4.6	-1.1 ± 5.9	1.4 ± 4.7
l. iliac crest	6.4	4.8 ± 5.9	-4.2 ± 6.2	-0.0 ± 7.8
r. iliac crest	6.2	-4.5 ± 5.8	-4.2 ± 6.7	-0.4 ± 8.0
sacrum end	2.9	-0.1 ± 2.4	2.7 ± 6.7	1.1 ± 5.6
L5/S1	2.8	0.1 ± 2.0	-2.0 ± 6.0	-1.9 ± 6.9
liver center	1.8	-1.6 ± 5.6	-0.8 ± 6.2	0.3 ± 21.1

Table 1. Landmark errors: distance in mm between mean landmarks mapped with (1) the ground truth transformation, and (2) the estimated similarity transformation. There were no outliers. The scaling is slightly under-estimated, the central region is well matched. See text for details.

The largest part of the variance with 21mm is in z-direction. This is caused by the large anatomical variability in this direction. The z-scaling of the liver cannot always be compensated for without violating other anatomical constraints. Example results of liver positioning on datasets of variable FOV and quality are shown in Fig. 4.

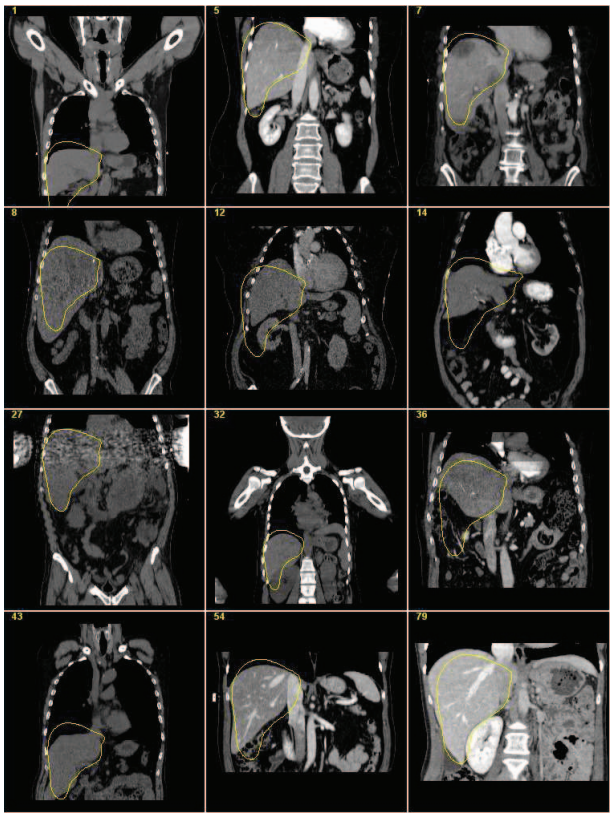


Fig. 4. Example liver results: a mean surface model of the liver positioned with similarity transform in different images of varying FOV, content and contrast from different scanners.

4. CONCLUSIONS

We present a novel method for low-parametric matching of an image with an atlas. We use this method for positioning atlas surface models in the image. The key novelty of our method is that we model the spatial position x as a random variable instead of the usual approach where local gray values are modeled. We derive an efficient and very robust method in the sense that the matching is insensitive to varying FOV and image content. It is possible to use the probabilistic atlas for higher dimensional transformation spaces. Future work should address an optimal learning method for the choice of atlases (tissue class ranges, most informative atlases). The model can be extended from scalar gray values to more general appearance features.

5. REFERENCES

[1] E. D'Agostino, F. Maes, D. Vandermeulen, and P. Suetens, "An information theoretic approach for non-rigid image registration using voxel class probabilities," *Medical Image Analysis*, vol. 10, no. 3, pp. 413–431, 2006.

[2] S. Seifert, A. Barbu, S.K. Zhou, D. Liu, J. Feulner, M. Huber, M. Suehling, A. Cavallaro, and D. Comaniciu, "Hierarchical parsing and semantic navigation of full body CT data," *Proc. SPIE*, p. 725902, 2009.

[3] 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," in *Int'l Conf. Computer Vision*, 2007.

[4] A. Criminisi, J. Shotton, D. Robertson, and E. Konukoglu, "Regression Forests for Efficient Anatomy Detection and Localization in CT Studies," in *MICCAI Workshop MCV, LNCS 6533*, Springer-Verlag Berlin Heidelberg, 2010, pp. 106–117.

[5] Jianhua Yao and Ronald M. Summers, "Statistical location model for abdominal organ localization," in *MICCAI (2), LNCS 5762*, 2009, pp. 9–17.

[6] A. Saalbach, I. Wächter-Stehle, R. Kneser, S. Mollus, J. Peters, and Jürgen Weese, "Optimizing GHT-Based Heart Localization in an Automatic Segmentation Chain," in *MICCAI (3)*, 2011, pp. 463–470.

[7] R. M. Gray, *Entropy and Information Theory*, Springer-Verlag, New York, 1991, available online.

[8] R. C. Eberhart and J. Kennedy, "Particle Swarm Optimization," in *IEEE Int'l Conf. Neural Networks*, 1995, vol. 4, pp. 1942–1948.