# Model-based Segmentation of Anatomical Structures in MR Images of the Head and Neck Area

Marcelo G. Bacher[1,2], Vladimir Pekar[1] and Michael R. Kaus[1]

[1]Philips Research Laboratories, Sector Technical Systems,
Röntgenstrasse 24-26, 22335 Hamburg
[2]University of Applied Sciences Mannheim, Institute of Digital Signal Processing,
Windeckstrasse 110, 68163 Mannheim
Email: {marcelo.bacher, vladimir.pekar, michael.kaus}@philips.com

**Abstract.** Contouring of the target and risk anatomy is one of the most time consuming procedures in radiotherapy planning (RTP). The main imaging modality used in RTP is the computer tomography (CT), where the application of automated segmentation methods in certain treatment areas, such as head and neck, is difficult due to insufficient soft tissue contrast. Magnetic resonance imaging (MRI) generates images with better soft tissue contrast, and it is expected that MRI will be more extensively used in RTP. Owing to the image formation principles, the feature variability of MR data is much higher compared to CT. In this paper, we present an approach that combines a model- based segmentation method with the pattern classification framework to segment organs in MR images of the head and neck area. A validation study demonstrates that the proposed approach is feasible for the organs tested.

## 1 Introduction

The determination of the treatment parameters in radiation therapy planning (RTP) requires the contouring of the patient's anatomy. This task is usually carried out manually, which is tedious, does not guarantee reproducible results and is currently one of the most time-consuming procedures in RTP. Therefore, the development of fast and robust automated segmentation tools is crucial for improving the RTP process. However, this task is difficult, in particular due to high organ variability, noise in the images and in special due to insufficient soft tissue contrast in computer tomography (CT) images, which is the main imaging modality in RTP. The goal of developing 3D model-based segmentation methods was to compensate for the insufficient image information by using prior knowledge. Among the model-based segmentation approaches which are of special interest for RTP applications one can mention sectored snakes [1], deformable m-reps [2], and shape constrained deformable models [3]. Although being generic, these methods were specifically tuned to use CT data features. Nevertheless, their application is difficult for certain treatment areas, such as

head and neck, owing to soft tissue structures which are poorly visible in CT. Compared with CT, magnetic resonance imaging (MRI) provides a much better soft tissue resolution, and it is expected that MRI will be more extensively used in RTP. However, due to the image formation characteristics (e.g. protocol), the anatomical structures in MRI present a substantial spatial variability of their features. The proposed approach combines, analogously to the method presented in [2], the advantages of using prior shape knowledge and the pattern classification framework to model the spatial feature variability along the boundaries of the organs. A validation study for the risk organs in the head and neck area demonstrates the feasibility of the proposed method to segment complex anatomical structures in 3D MR datasets.

## 2 Method

In our approach, deformable organ models are represented by triangulated surface meshes. The models include prior knowledge by integrating statistical point distribution models (PDM) which consist of a mean organ shape with a set of variation modes obtained by principal component analysis of a training set [3]. The proposed approach is based on a learning process that requires reference information in form of manually segmented anatomical structures, referred as ground truth masks. The method consists of four steps: (i) preprocessing of data, (ii) feature extraction, (iii) clustering process, and (iv) deformable model adaptation. The preprocessing step is needed to remove the background noise. This is achieved by computing the grey value histogram of the MR volume, where the background noise is represented as a Rayleigh distribution in the low intensity range of the histogram. This noise distribution is removed by automatically selecting a gray value threshold, which is computed by the optimization method presented in [4]. Next, the deformable model corresponding to the organ of interest is adapted towards the surface of the ground truth superposed to the corresponding training image as described in [3]. The deformable model adaptation is carried out by searching the strong gradient between the mask and the rest of the image. After the deformable model adaptation, the original MR image corresponding to the ground truth is filtered with a low pass Gaussian filter, and the feature extraction is carried out by acquiring feature profiles $\mathbf{p}_i$ passing perpendicular to the deformable model triangles on all images to be used for learning:

$$\mathbf{p}_i = F(\hat{\mathbf{x}}_i + j\mathbf{n}_i\delta),\ j = -l, \ldots, l, \tag{1}$$

where $\mathbf{n}_i$ is the unit normal vector at the $i$-th triangle, $\hat{\mathbf{x}}_i$ is the triangle barycenter and $\delta$ is the step size in mm, and $2l+1$ is the profile length. Each acquired feature profile is normalized independently to have the zero mean value and a unit variance. The investigated feature functions $F(\mathbf{x})$ were the intensity value, gradient projection and Gabor filter bank response, where three frequencies were empirically selected to be $\sqrt{2}$, $2\sqrt{2}$ and $4\sqrt{2}$, and the deviation was $\sigma = 0.6/f$ to avoid redundancies in the responses. After the acquisition of the feature profiles, they are clustered to model the surface of anatomical structures; this means,

the resulting clusters should categorize areas of the anatomical structure on the basis of image features. Due to its simplicity and computational efficiency, the k-means clustering algorithm was chosen. A normalized cross-correlation measure between the profile and the mean profile is used to achieve shift invariant matching between the deformable model boundary and the organ boundary during the profile classification process. In order to account for possible segmentation inaccuracies after the triangular mesh adaptation to the ground truth masks, the similarity measure is computed within a confidence interval $\gamma$.

The determination of the initial number of clusters is carried out as follows. The clustering method is repeated several times for a different number of clusters. Two dispersion functions are computed afterwards: $JWSS_C$ and $JBSS_C$. The first function measures the data dispersion within a cluster:

$$JWSS_C = \sum_{k=1}^{C} \sum_{\mathbf{p}_i \in k} [1 - \max r(\mathbf{p}_i, \bar{\mathbf{p}}_k)]^2, \tag{2}$$

where $C$ is the number of clusters, $\mathbf{p}_i$ is a feature profile belonging to the cluster $k$, $\bar{\mathbf{p}}_k$ is the mean profile corresponding to the cluster $k$, $r(.,.)$ is the cross-correlation function, and $\max r(.,.)$ is the maximum cross-correlation value within the confidence interval. The second function measures the dispersion of the mean profiles:

$$JBSS_C = \frac{\sum_{t=1}^{T} \left\{ \sum_{i=1}^{V} [1 - \max r(\mathbf{p}_i, \bar{\mathbf{p}})]^2 \right\} - JWSS_C}{\sum_{t=1}^{T} \left\{ \sum_{i=1}^{V} [1 - \max r(\mathbf{p}_i, \bar{\mathbf{p}})]^2 \right\}}, \tag{3}$$

where $\bar{\mathbf{p}}$ is the mean profile computed over all acquired feature profiles, $T$ is the number of training images, and $V$ is the number of triangles. The initial number of clusters is a trade-off between the dispersion within and between the clusters. After the cluster means are computed, each triangle of the deformable model is assigned to a cluster according to the similarity between the feature profile and the cluster mean.

The computed cluster means model the spatial feature variability along the anatomical structure. Therefore, it is expected that this model represents the investigated object in all training images. In order to quantify how well the model represents the anatomical structure using a particular feature function, the feature profiles are separated w.r.t. the images they stemmed and their labels are ordered in a vector in a way that each index corresponds to a triangle in the mesh. After that, the percentage of the label coincidences in the vectors is computed. The feature function is accepted to be used for segmentation if this measure lies above a certain threshold. The "goodness" of the cluster means is assessed by computing within the confidence interval the maximum cross-correlation value between the cluster means and the corresponding feature profiles in each training image. If the average maximum cross- correlation value is below a certain threshold, the number of clusters can be changed according to the dispersion functions. As a result of clustering, homogeneous sectors on the deformable model corresponding to specific feature profile means are specified.

**Table 1.** Mean and maximum distance error between the deformable model and the manual reference segmentation. All values are given in millimeters.

| Organ | Image | Feature function | Profile length | Number of clusters | Initialization mean / max | 30 Iterations mean / max |
|---|---|---|---|---|---|---|
| Eye | MR 1 | Gray value | 13 | 4 | 7.12 / 11.59 | 0.75 / 3.26 |
| Eye | MR 2 | Gray value | 13 | 4 | 7.13 / 12.28 | 1.12 / 3.26 |
| Eye | MR 3 | Gray value | 13 | 4 | 7.44 / 11.44 | 1.16 / 3.35 |
| Brain stem | MR 1 | Gabor response | 8 | 3 | 4.71 / 9.68 | 1.00 / 5.70 |
| Brain stem | MR 2 | Gabor response | 8 | 3 | 6.07 / 13.40 | 1.25 / 5.06 |
| Brain stem | MR 3 | Gabor response | 8 | 3 | 6.63 / 20.10 | 2.03 / 10.36 |
| Brain cord | MR 1 | Gradient projection | 13 | 4 | 3.27 / 6.82 | 0.53 / 1.82 |
| Brain cord | MR 2 | Gradient projection | 13 | 4 | 3.38 / 8.53 | 1.40 / 7.50 |
| Brain cord | MR 3 | Gradient projection | 13 | 4 | 4.26 / 11.17 | 1.36 / 8.44 |
| Cerebellum | MR 1 | Gray value (*) | 20 | 3 | 6.79 / 18.31 | 2.22 / 8.95 |
| Cerebellum | MR 2 | Gray value (*) | 20 | 3 | 4.71 / 14.06 | 1.34 / 6.32 |
| Cerebellum | MR 3 | Gray value (*) | 20 | 3 | 4.54 / 13.06 | 1.70 / 6.77 |
| Brain | MR 1 | Gabor response | 32 | 3 | 9.48 / 20.00 | 1.90 / 6.56 |
| Brain | MR 2 | Gabor response | 32 | 3 | 9.06 / 26.88 | 2.26 / 8.95 |
| Brain | MR 3 | Gabor response | 32 | 3 | 8.56 / 22.58 | 2.21 / 9.69 |

(*) Due to the high spatial variability of the cerebellum and to the spatial variability of the surrounding anatomical structures, an alternative method for collecting the feature profiles was used that only considered the feature values inside the cerebellum.

During deformable model adaptation, at each triangle of the deformable model a feature profile is acquired and then compared with the corresponding cluster mean by computing the cross-correlation value. If the maximum cross-correlation value within the confidence interval is equal to or higher than the value obtained during the training phase, an attraction point is delivered for reconfiguring the deformable model as described in [3].

## 3 Results

The proposed segmentation method was tested on three 3D MR datasets taken from patients with brain cancer, where the eyes, brain stem, brain cord, cerebellum and brain were segmented. The image dimensions were 256x256x124, where the geometrical voxel dimensions were 1.07 mm x 1.07 mm x 0.67 mm. For acquiring the training data as well as for segmentation, a constant confidence interval $\gamma$ of 1mm was selected. The validation was carried out as follows: (i) a 3D MR image that was not used for learning (leave-one-out test) was loaded, (ii) the background noise was removed, (iii) the corresponding deformable model was loaded and manually positioned close to the anatomical structure to be segmented, and (iv) 30 iterations of the deformable model adaptation algorithm were applied. The validation results of fully automated segmentation after 30 iterations of the algorithm are presented in Table 1.

In all investigated organs, three initial clusters were selected to represent the number of different feature regions as a trade-off between the dispersion func-

tions and taking into account that a small set of training images was available. For the eye and brain cord modeling, the number of clusters could be increased, which indicated that a more specific model could be obtained. During modeling, several feature profile lengths were tested, where the criteria for choosing a particular profile length were: (i) a sufficient amount of information about the object of interest taking into account its shape variability and the neighboring structures was collected, and (ii) the acquired feature profiles that had opposite spatial positions in the deformable model did not overlap with each other. Each segmentation process including the image preprocessing step and deformable model adaptation (30 iterations) took less than 45 seconds on a 2.4 GHz Dell workstation. The average segmentation error was in the order of 1-2 mm for the organs tested. In the clinical use, the remaining problematic areas can be corrected by using, e.g. manual manipulation tools [5].

## 4 Conclusions and Future Work

The presented approach uses the principles of statistical pattern classification to model the spatial feature variability of anatomical structures and combines them with shape constrained deformable modes to address the segmentation of MR data. A validation study has been carried out to segment the eyes, brain stem, brain cord, cerebellum and brain. Despite the lack of a representative set of reference images, the achieved results show that the proposed approach is able to quickly and accurately (with 1-2 mm mean error) segment complex anatomical structures in MR images of the head and can, therefore, be potentially applied in RTP based on MR data.

In future work, we plan to address the validity of the proposed algorithm from the statistical perspective by using more data, and also to demonstrate its feasibility in other treatment areas. A more extensive study of different feature functions can be done to investigate the possibility of obtaining a better model of the spatial feature variability.

## References

1. Fenster SD, Kender JR. Sectored snakes: Evaluating learned-energy segmentations. In: Procs Int Conf Computer Vision (ICCV'98); 1998. p. 420–426.
2. Stough J, Pizer S, Chaney E, Rao M. Clustering on image boundary regions for deformable model segmentation. In: Procs Int Symp Biomedical Imaging (ISBI'04); 2004. p. 436–439.
3. Pekar V, McNutt TR, Kaus MR. Automated model-based organ delineation for radiotherapy planning in prostatic region. Int J Radiation Oncology Biol Phys 2004;60(3):973–980.
4. Brummer ME, Mersereau RM, Eisner RL, Lewine RRJ. Automatic detection of brain contours in MRI data sets. IEEE Trans Medical Imaging 1993;12(2):153–166.
5. Timinger H, Pekar V, von Berg J, Dietmayer K, Kaus MR. Integration of interactive corrections to model-based segmentation algorithms. In: Procs BVM; 2003. p. 171–175.