Detection of Vertebrae in CT Slices by Bunch Graph Matching

B. Roeschies and S. Winter

Institut für Neuroinformatik, Ruhr-Universität Bochum, 44780 Bochum, Germany

Abstract— **We present an approach to the detection of anatomical landmarks in CT data sets. Our first application is detection of vertebrae and landmarks on vertebrae in 2D slices. The core technique is a feature driven matching of a graph created from one or more vertebrae with unseen data. This is based on an approach that has been successfully applied to face recognition in the past.**

Keywords— **Anatomical Landmark Detection, Bunch Graph Matching, Wavelet Features, Object Recognition.**

I. Introduction

The automatic detection of anatomical landmarks in medical data is a challenging task. Many computer assisted procedures in modern medicine use landmarks e.g. for intraoperative registration in navigated surgery. Anatomical landmarks can also serve as starting points in image segmentation for automatic diagnosis. Different methods for anatomical landmark detection have been introduced in the past [2,4].

We use a model based approach to detect anatomical landmarks in medical image data. The models consist of several landmarks, that are detected in parallel. As a first application for this technique we detect vertebrae in 2D CT slices. In this case, landmarks of interest may be the spinous process, pedicles, or the articular processes.

Our approach is based on a face recognition system [3]. The main principle for feature extraction is a gabor wavelet transformation. The use of this transformation has a biological background, because the responses gained by filtering an image using a gabor transformation resemble a good model for neural activity patterns in the primary visual cortex [1].

As gabor responses mainly contain local texture information, they alone are not sufficient to detect objects within images. Thus, a topological graph structure is used as a model to represent the shape of an object – in our case a vertebra – and gabor responses of model images are attached to the nodes of the graph, which are usually placed at specific points of interest. This means, that an object is represented by a set of points that carry texture information and have specific local relations to each other given by the graph structure.

The model graph is then matched onto new images searching for a position that yields the best fit with respect to the combined similarity of all graph points. During this step, global deformations, like dilatation or rotation of the graph, take place as well as local movement of single points. The latter process has to incorporate a compromise between local point similarity and deformation strength of the whole graph.

II. Material and methods

A. 2D-Wavelet Transformation

The *wavelet transformation* W of a two-dimensional signal $I(\mathbf{x})$ is defined as its scalar product with a wavelet family ψ:

$$\mathrm{W}_{\psi} I(\mathbf{x}) = \left\langle \psi_{\mathbf{x}_0,s,\varphi}(\mathbf{x}), I(\mathbf{x}) \right\rangle \tag{1}$$

The members of the wavelet family are given by scale s, orientation φ, and translation $\mathbf{x}_0$ with respect to a *mother wavelet* Ψ:

$$\psi_{\mathbf{x}_0,s,\varphi}(\mathbf{x}) = \frac{1}{s^2}\Psi(\mathbf{x})\left[\frac{1}{s}\mathbf{Q}(\varphi)(\mathbf{x}-\mathbf{x}_0)\right]. \tag{2}$$

The rotation is applied using the rotation matrix

$$\mathbf{Q}(\varphi) = \begin{bmatrix} \cos\varphi & \sin\varphi \\ -\sin\varphi & \cos\varphi \end{bmatrix}. \tag{3}$$

An image can be transformed using a wavelet transformation by considering it as a two-dimensional signal.

B. Gabor Wavelets

A special type of wavelet transformation can be created by using a *gabor function* as the mother wavelet. In general, gabor functions are windowed exponential oscillators. To generate feature sets, we use a gabor function with a gaussian window function

$$\Psi(\mathbf{x}) = \frac{1}{\sigma^2} e^{-\frac{1}{2\sigma^2}\|\mathbf{x}\|^2}\left(e^{i\mathbf{x}^T\mathbf{e}_1} - e^{-\frac{\sigma^2}{2}}\right), \tag{4}$$

where the width of the gaussian is determined by σ and $\mathbf{e}_1$ =

[1 0]. The gaussian window function induces the local boundedness of the gabor responses by causing the filter function to fall off with displacement from the center. In our experiments we use $\sigma = 2\pi$.

C. Discretization

As a gabor wavelet transformation is a scalar product with an infinite amount of wavelets, the transformation is discretized. It is shown that only a limited amount of scales and orientations is needed to prevent a significant loss of information [1]. To calculate features, we transform images using $M = 5$ scales and $L = 8$ orientations according to the scheme

$$s = s_{\min} s_0^m, \quad m = 0,...,M-1 \tag{5}$$

$$\varphi = \frac{\pi l}{L}, \qquad l = 0,...,L-1 \tag{6}$$

with $s_{\min} = 0.3$ and $s_0 = 1.392$.

D. Feature Sets and Similarity

The procedure described above leads to a total of 40 complex valued images, i.e. each pixel of the original image receives a *feature set* of 40 complex values, which are represented by their absolute value and phase. The similarity S of two feature sets with absolute value vectors **a** and **b** and phase vectors **α** and **β** is measured by

$$\mathrm{S}(\mathbf{a},\boldsymbol{\alpha},\mathbf{b},\boldsymbol{\beta}) = \frac{\sum_i a_i b_i \cos(\alpha_i - \beta_i)}{\sqrt{\sum_i a_i^2 \sum_i b_i^2}}, \tag{7}$$

which leads to similarity values between 0 and 1.

E. Graph Matching

The gabor wavelet transformation extracts local texture information from the image and allows to calculate similarities of single points from different images. However, these features alone do not allow for a precise detection, because a lot of points have very similar local texture properties. Therefore, points are not matched one-by-one, but according to an underlying topographic structure represented as an undirected graph.

First, a graph consisting of 20-50 nodes is placed by hand on a training image of a vertebra. After that, the matching process is carried out by moving the graph over an image of a different vertebra and calculating the total similarity of the graph as the mean of single point similarities at the graph node positions. In the end, the graph will be placed at the position with the best total similarity. During the moving process the graph may also be transformed by dilatation or rotation to accommodate the respective vertebra variations. In our experiments we created a raster with 4 pixels distance between raster points and moved the center of the graph to all raster positions. At each position 20 scales from a scaling factor of 0.8 up to a factor of 2.2 with respect to the model graph were checked.

The best graph position after this procedure is considered the starting position for the last step, where every single node may be moved separately to find an even better position. In this step, similarities are calculated differently, taking displacement with respect to the starting position into account. Similarities are calculated according to

$$\mathrm{S}' = \mathrm{S}(\mathbf{a},\boldsymbol{\alpha},\mathbf{b},\boldsymbol{\beta}) \cdot e^{-\eta d^2}, \tag{8}$$

where d resembles the euclidian distance of the respective point from the starting position and η is a parameter to define the penalty strength of a large displacement. We used $\eta = 0.001$ for our experiments, which causes similarity to be halved at a displacement of about 26 pixels.

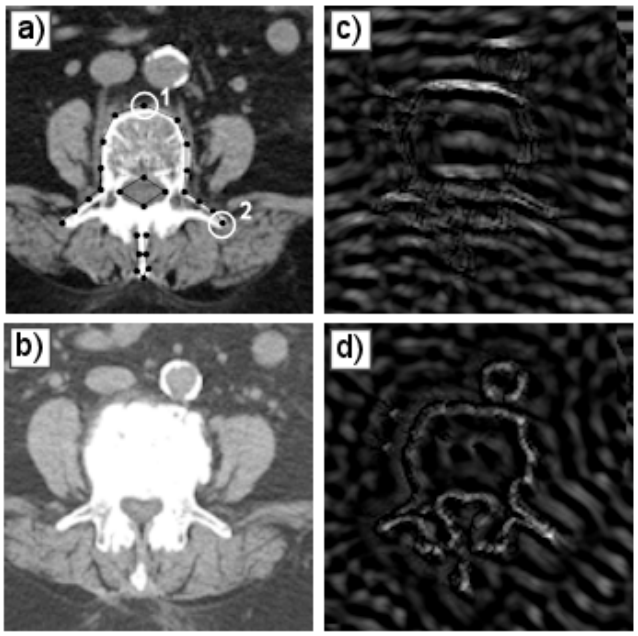

Fig. 1 Local similarity distribution for two points (1 and 2). a) Graph on the training image with the two points marked by circles; b) Target image; c) Local similarity image for point 1; d) Local similarity image for point 2.

Fig. 1 gives an impression of the local similarity distribution on a test image. In Fig. 1 a) one of the slices

that contributed to the model is shown with the respective model graph. Two points (1 and 2) are marked by circles. For these points, similarities were calculated with respect to a test image that is shown in Fig. 1 b). In Fig. 1 c) and d) local similarities for points 1 and 2 are displayed, where a high intensity corresponds to a high similarity.

The images Fig. 1 c) and d) demonstrate that a lot of locations show a high similarity, so that it is necessary to combine this information with topological information. Especially in c), a large area with high similarity values along the vertebral body is present.

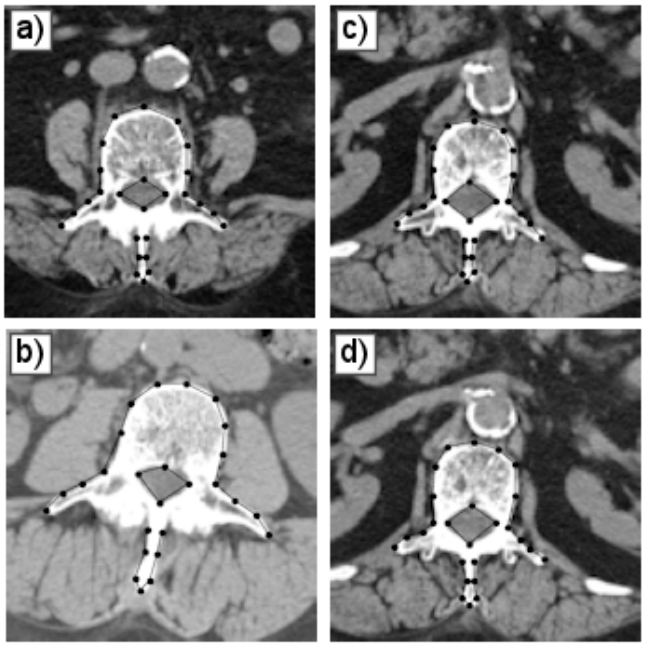

Fig. 2 The 4 vertebra slices used to create the bunch graph that serves as the model for the matching process

To further improve performance, several training graphs are combined to form a *bunch graph* which is depicted in Fig. 2. This is reasonable, because a specific point of a vertebra may have different local texture properties than the respective point of another vertebra. To create a bunch graph, several images were hand-labeled using the same number of graph nodes with distinct indices. Each specific node was placed on corresponding points on each of the training images. After that, every node received the feature sets of all underlying points from the training images. The node positions were taken from one arbitrary vertebra of the bunch graph.

During the matching process single point similarities of a specific node were calculated by taking the best similarity among all feature sets attached to that node.

F. Data and Preprocessing

The graph matching was carried out on a collection of CT slices from 20 different patients of about 50 different vertebrae. As we used several slices of the same vertebra, the total number of images used was about 150. At first, all CT slices were sampled down to a resolution of 128×128 pixels. Then, a window level transformation was applied to the raw CT data with window and level both set to 1000. The results were saved as 8 bit gray level images. In a second step we set all gray values below a threshold of 216 to zero. This value resembles the center of the rightmost minimum in the combined gray level histogram of our test data set. Applying the threshold accommodates the fact that vertebrae generate relatively high intensities in CT data and provides a low-level segmentation.

For our experiments we created a bunch graph consisting of 26 nodes from 4 different vertebrae shown in Fig. 1. Graph nodes were connected by edges that may be used in the future to further constrain the deformation of the graph.

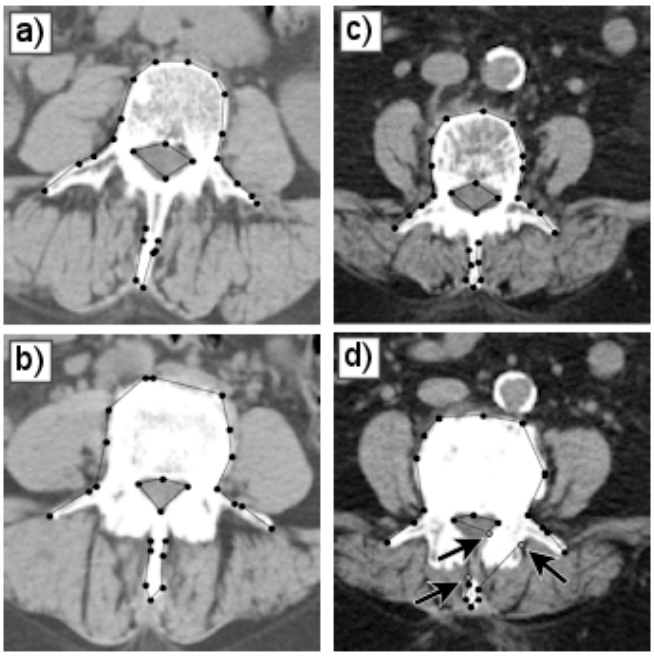

Fig. 3 Some results after the graph matching process. d) shows a graph with three points labeled as missing. The first is located at the posterior part of the vertebral arch and two more at both sides of the spinous process in the model graph (marked by arrows).

III. Results

In all our image slices the detection of the vertebrae was successful. Figure 3 a)-c) show three examples, where all the graph points match to reasonable positions. To

accommodate for the fact that some points that are present in the model graph may be missing in the target image, we defined a similarity threshold of S' = 0.5. Points that received similarity values below this threshold after the graph matching process were considered as missing in the target image. Fig. 3 shows an example where some of the defined landmarks of the model were missing. The points were placed at wrong positions.

The lower right picture illustrates the detection of missing points. For example, the point located at the posterior part of the vertebral arch in the model graph is missing in this picture. A part of the spinous process is missing as well and thus the two points that are located at the sides of the process within the missing part were also labeled as missing points.

IV. Discussion

We were able to detect vertebrae with good accuracy based on visual evaluation. The single point similarities could even be used to determine missing points in the images. These results are pretty satisfying considering the small model size and the fact that some of the test vertebrae had a strongly different shape compared to the ones contained in the bunch graph.

The images in Fig. 3 show that concise points of the graph – like the tip of the spinous process and the pedicles – could be detected with good accuracy, while points that are located on a longer edge and are pretty close were sometimes mistaken for each other. This can be seen in Fig. 3 b), where two points on the front of the vertebral body were moved almost on top of each other.

Images of the same vertebra from a different layer most of the times showed an almost perfect match and very high local similarity values. These values could possibly be used to detect, if two slices are images of the same vertebra from a different layer.

The results indicate that a bunch graph matching using gabor wavelet features is a valid approach to the detection of anatomical structures and landmarks. Additionally, valid statements about missing points can be made, which may lead to the detection of pathological deformations as a further application field for this technique.

V. Conclusions

We have implemented a technique that is well-researched within the application field of face recognition for the detection of vertebrae in CT data. This approach yielded good results in our experiments and motivates further research for anatomical structure detection.

For the future we intend to apply the approach on different types of data, e.g. MRT images, and different anatomical structures. Additionally, we want to extend the approach to work on volume data using three-dimensional wavelet features. The results may further be improved by incorporating a higher number of model images in the bunch graph and using a more than 26 nodes. Also, a more elaborated similarity measure incorporating the compromise between graph deformation and local point similarity may be implemented, possibly making use of edges connecting the graph nodes.

Acknowledgment

This work was an activity of the OrthoMIT consortium. It was supported by the Bundesministerium für Bildung und Forschung (Az. 01EQ0424).

References

1. Jones JP and Palmer LA (1987) An evaluation of the two-dimensional gabor filter model of simple receptive fields in cat striate cortex. J Neurophysiol, 58(6):1233-1258
2. Liu J et al. (2008) A Model-based, Semi-Global Segmentation Approach for Automatic 3-D Point Landmark Localization in Neuroimages. IEEE Transactions on Medical Imaging, 27(8):1034-1044
3. Wiskott L et al. (1999) Face recognition by elastic bunch graph matching. In Jain LC et al., editor, Intelligent biometric techniques in fingerprints and face recognition, pp 355-396
4. Wörz S and Rohr K (2004) Localization of anatomical point landmarks in 3D medical images by fitting 3D parametric intensity models. Medical Image Analysis, 10(1):41-58

- Author: Benjamin Roeschies
- Institute: Institut für Neuroinformatik
- Street: Universitätsstr. 150
- City: 44801 Bochum
- Country: Germany
- Email: Benjamin.Roeschies@neuroinformatik.rub.de