Deformable Templates for the Localization of Anatomical Structures in Radiologic Images

Wolfgang Sörgel and Bernd Girod

Telecommunications Laboratory
University of Erlangen-Nuremberg
Cauerstraße 7, 91058 Erlangen, Germany
{wsoergel,girod}@nt.e-technik.uni-erlangen.de

Abstract. This paper describes a method for non-rigid registration of a model template with the corresponding anatomic structure in a radiologic image. The model template is a graph consisting of labeled nodes and edges. Each node is labeled by a feature vector that characterizes the sought structures in its vicinity. The edges connect the nodes in accordance with anatomical topology. The features which describe possible node positions are computed by filtering the image with a set of Gabor filters and comparing the filter responses with those from the model. The final node positions are determined by minimizing a deformation energy term associated with the template. We show results for digitized film orthopantomograms of the jaw which indicate good performance at modest computational effort for such artifact-rich images. The results are used for localization of bone lesions relative to anatomic structures.

Keywords: Deformable Templates, Gabor Filters, Anatomy Localization, Non-rigid Registration

1 Introduction

In the analysis of medical images it is often desirable to detect the position of typical anatomical structures. One way to achieve this is to register the image to a model or atlas image of the depicted structure. In this case, the shape, size and location in the image of the sought structure is roughly known but one needs to apply a non-rigid coordinate transformation for registration. The method proposed in this paper is similar to the one introduced by Amit [1] based on a decomposable graph of landmarks matched to the image by local operators and discrete optimization, which we have combined with feature computation by Gabor filtering as used for face recognition [2, 3]. The approach is based on a deformable template net consisting of nodes placed on anatomical landmarks and edges connecting these nodes. Figure 1 shows an example of a net that has been manually adapted to a schematic drawing of the facial skeleton. The nodes are labeled with features describing how the image should look like in the vicinity of each node. These features are compared with the respective features for each image coordinate point to identify possible locations of each node. The final position of each node is then determined by minimizing a deformation energy

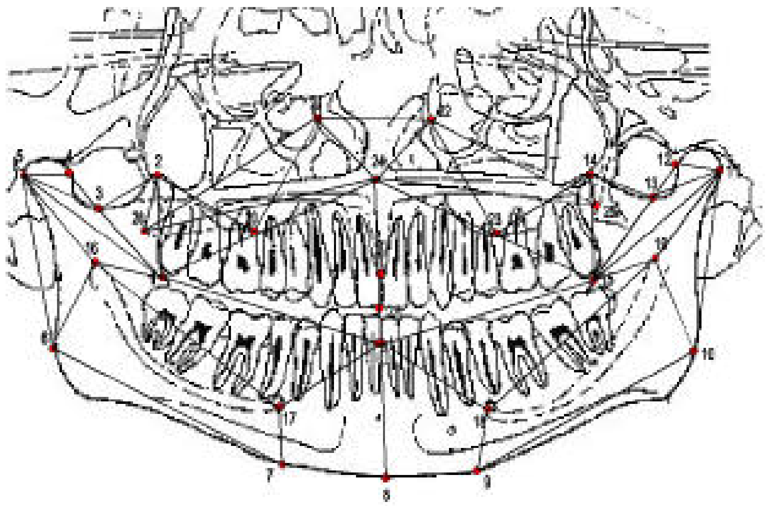


Fig. 1. Model mesh with nodes and edges manually adapted to a schematic drawing of the facial skeleton.

associated with the net. Based on the position of the nodes of the deformed net a non-rigid coordinate transform between model and image can be applied.

2 Image Feature Computation

The localization of different structures in the images is based on Gabor filtering. These filters have found widespread use in texture analysis [4] and various computer vision tasks [2, 3]. They allow detection of localized structures, are sensitive to orientation and spatial frequency while being tolerant against brightness variations and can easily be incorporated into multi-scale approaches. One such complex, zero mean filter kernel is given by

$$\psi(x,y) = \frac{k^2}{4\pi^2} \exp\Big(-\frac{k^2(x^2+y^2)}{8\pi^2}\Big) \cdot \Big(e^{-j(x\cos\phi + y\sin\phi)} - e^{-2\pi^2}\Big). \qquad (1)$$

It has the shape of a plane wave responding best at the spatial frequency or scale $k = 2^{-\frac{\nu+2}{2}}\pi$ and orientation ϕ, restricted by a Gaussian envelope. By linear filtering with filter kernels ψ of different scales and orientations, one obtains a complex feature vector $\mathbf{J}(x,y) = \big[J_1(x,y) \quad J_2(x,y) \ \ldots \ J_n(x,y)\big]^{\mathrm{T}}$ with elements of the form $J_i = a_i \exp(j\varphi_i)$ for each pixel location (x,y) in the image. For the anatomy detection application on panoramic X-ray images we use 4 scales $\nu \in \{0;1;2;3\}$ and 4 orientations $\phi \in \{0^o; 45^o; 90^o; 135^o\}$ resulting in 16-dimensional feature vectors.

Based on these feature vectors we now determine similarity of points in a given image I to those in another model image I'. While the amplitudes a_i are spatially smooth, the phase components ϕ_i vary with approximately the characteristic spatial frequency of the kernel. We use

$$S(\mathbf{J}, \mathbf{J}') = \frac{\sum_i a_i a_i'}{\sqrt{\sum_i a_i^2 \sum_i a_i'^2}} \cdot \left(1 + \frac{\sum_i a_i a_i' \cos(\varphi_i - \varphi_i')}{\sqrt{\sum_i a_i^2 \sum_i a_i'^2}}\right) \tag{2}$$

as a similarity criterion function, resulting in high similarity where amplitude and phase are both similar. Compared to using amplitude only we so get better results. The similarity function is evaluated for each node of the mesh. Using one or more model images with known position of the node we obtain a similarity value from (2) for this node at every pixel within the search area in the image. A number of point positions most similar to the model is chosen as candidate points for possible positions of the node in the deformed net.

3 Deformation of the Mesh

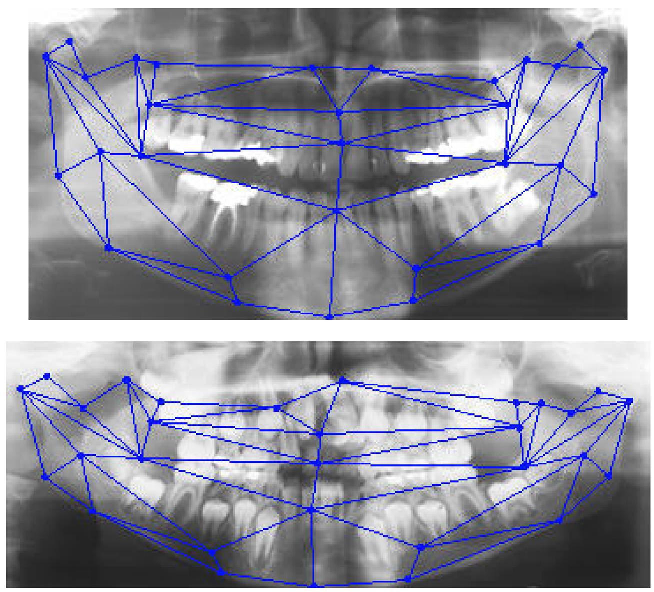
The final position of the nodes in the image is now determined by minimizing a deformation energy associated with the net. This energy $E = E_{\text{distance}} + \lambda E_{\text{angle}}$ is composed of a distance term and an angle term. The distance term E_{distance} models the edges of the net as linear springs, normalized by the model edge length. Distance energy is used for keeping the proportions of the model. The second term E_{angle} is derived from the difference of the angles between the edges at each node from the actual net image to the model net. This term controls local deformations. Since we do not allow intersecting edges we further use a penalty term which increases the energy by a large, fixed amount for each intersection.

In the next step the model net is adapted to the image by minimizing an energy-like objective function. Energy minimization is initialized by placing each node on one of its candidate points. Then a greedy algorithm [5], moving nodes in turn among candidate points, is utilized to find the final position of the nodes by minimizing E. Actual minimization is done in three steps, using only E_{distance} first, then E_{angle} and finally the complete energy. Using 5 candidate points for each node, 4 to 6 iterations over all node points are needed for convergence of each step. Note that feature similarity (2) is not used for energy minimization, only for the selection of the candidate points which are then treated equally. Thus there is no need to trade off mesh deformation against feature similarity.

4 Results and Applications

We apply the proposed algorithm to a collection of 233 digitized orthopantomographs of the jaws, originating from different sources and with variable quality. The structure is well recognized in all cases. Well visible anatomical landmarks, like points at the outer contour of the mandible are correctly found in over 95 % of all images. On average 75 % of the 30 node points we use on these images are placed correctly. Most of the points not found are obscured by imaging artifacts or poor film exposure. Since the topological structure of the net is not affected by misplaced points and the misplacement is typically less than the average distance between node points, this rate is acceptable for our task. It could be further increased by adding another optimization step where the requirement that node points can only be placed on the candidate points determined in section 2 is relaxed. The use of a multi-resolution scheme which can be efficiently implemented within the given framework is another option. Computational effort is modest with an overall execution time of a few minutes on a desktop workstation. Figure 2 shows a typical example of a net adapted to an image of the jaw and an example for an artifact-rich image with some misplaced nodes. Note that the structure is recovered well nevertheless. The deformed template

Fig. 2. Deformable net adapted to an X-ray image of the facial skeleton. Upper image: Typical result. Lower image: Difficult image with some misplaced nodes



is then used to define a non-rigid coordinate transform from the actual image to the model. This transform can either be based on a triangulation of the net and affine transforms of the image points within each triangle or radial basis functions defined in the node points.

We use the proposed template matching method within a system for characterization and diagnosis of bone lesions in the facial skeleton [6]. The goal is to determine the position of a given region within the jaws. For this purpose, the coordinates of the contour of a lesion are transformed to the model coordinates which allows intersecting it with the predetermined anatomic labeling of the regions in the model. As a result we get a fuzzy membership value for each region which enables us to compare the location of different lesions.

5 Conclusions

We have presented a method for non-rigid registration of template models with anatomical structure in radiological images. Based on a labeled mesh we can determine a coordinate transform from the actual image to the model. The Gabor filters used as feature labels in the nodes allow detection of local characteristic structures, sensitive to scale and orientation while being robust against brightness and contrast variations commonly found in X-ray images. We achieve an overall rate of correct node localization of 75 %, with better results for well visible structures. For our application of mapping lesion positions to predetermined regions this performance is sufficient. Improvements by additional fine search, the use of a multi-resolution framework or by the application of heuristics are under investigation. The method is not bound to an image modality or application, extensions to 3D datasets are possible.

6 Acknowledgments

This work was supported in part by a grant from *FRIATEC AG*, Mannheim. We would like to thank the *Klinik und Poliklinik für Mund-,Kiefer-,Gesichtschirurgie* in Erlangen for their competent advise.

References

1. Yali Amit. Graphical shape templates for automatic anatomy detection with applications to MRI brain scans. *IEEE Transactions on Medical Imaging*, 16(1):28 – 40, February 1997.
2. M. Lades, J.C. Vorbrüggen, J. Buhmann, J. Lange, C. von der Malsburg, R. P. Würtz, and W. Konen. Distortion invariant object recognition in the dynamic link architecture. *IEEE Transactions on Computers*, 42(3):300 – 311, 1993.
3. Laurenz Wiskott, Jean-Marc Fellous, Norbert Krüger, and Christoph von der Malsburg. Face recognition and gender determination. In *Proceedings Intern. Workshop on Automatic Face- and Gesture Recognition*, pages 92 – 97, Zürich, 1995.
4. B. S. Manjunath and Y. W. Ma. Texture features for browsing and retrieval of image data. *IEEE Transactions on Pattern Analysis and Machine Intelligence (PAMI)*, 18(8):837 – 841, August 1996.
5. D.J. Williams and M. Shah. A fast algorithm for active contours and curvature estimation. *Computer Vision, Graphics, Image Processing*, 55:14 – 26, 1992.
6. Wolfgang Sörgel, Sabine Girod, Martin Szummer, and Bernd Girod. Computer aided diagnosis of bone tumors in the facial skeleton. In *Bildverarbeitung für die Medizin*, pages 179 – 183. Springer, March 1998.