Fully Automatic Model Creation for Object Localization utilizing the Generalized Hough Transform

Heike Ruppertshofen[1,2,3], Cristian Lorenz[2], Peter Beyerlein[4], Zein Salah[3], Georg Rose[3], Hauke Schramm[1]

[1]Institut für Angewandte Informatik, Fachhochschule Kiel
[2]Sector Medical Imaging Systems, Philips Research Europe - Hamburg
[3]Institut für Elektronik, Signalverarbeitung und Kommunikationstechnik, Otto-von-Guericke-Universität Magdeburg
[4]Fachbereich Ingenieurwesen, Technische Fachhochschule Wildau
heike.ruppertshofen@fh-kiel.de

Abstract. An approach for automatic object localization in medical images utilizing an extended version of the generalized Hough transform (GHT) is presented. In our approach the shape model in the GHT is equipped with specific model point weights, which are used in the voting process. The weights are adjusted in a discriminative training procedure, which aims at a minimal localization error of the GHT. Furthermore, these weights yield information about the importance of points, such that non-discriminative points can be excluded from the model. Thereby, the size of the model is decreased, thus reducing processing time. The algorithm is presented in 2D but should be easily extendable for 3D images.

1 Introduction

The task of object localization in medical images is of high interest in many applications in the area of computer-aided diagnosis, e.g., for a fully-automatic segmentation of organs or bones.

For model-based segmentation an initial positioning of the model in the image is required. This initialization is often carried out manually or specialized solutions are developed for the particular problem at hand using, e.g., gray-value thresholding, morphological operators or anatomical knowledge [1]. More advanced, automatic approaches are given by an atlas-based registration [2] or an initial global search of the object conducted, e.g., with evolutionary algorithms [3], template matching [4] or the generalized Hough transform (GHT) [5].

In this work, we will focus on the GHT, which, in general, is a robust method and therefore well suited even for the localization of partially occluded or noisy objects. However, one drawback is the high computational demand given by the vast global search, as it is the case for all aforementioned automatic procedures. To reduce its computational complexity two measures have been undertaken. On the one hand, we will not search for scaled or rotated occurrences of our object;

but instead, will try to include this variation into our model. On the other hand, a weight will be assigned to each model point relative to its importance such that nonrelevant points can be removed from the model; thus reducing model size and runtime. Model point weights are determined through a log-linear combination of model point knowledge and a discriminative training [6] with respect to a minimal localization error.

2 Materials and Methods

2.1 Generalized Hough Transform

The GHT [7] is a model-based method for object localization utilizing a voting process to identify the likeliest position of the object of interest. To this end, the image space is mapped to a transformation parameter space, known as Hough space, which depicts possible locations of the object.

The shape model of the object is represented by vectors from each model point to a given reference point, usually the center of the model. Furthermore, directional information obtained from the gradient direction is added to each point. This information is stored in a look-up-table, the so called r-table, where model points with similar gradient directions are grouped together.

To perform the localization, an edge image is created from the original image. For each edge point e_i the model points m_j with similar gradient direction are retrieved from the r-table. Through the relation $c = e_i - m_j$ the position c of a possible reference point is determined and the vote in the corresponding cell in the Hough space is increased by the weight of the model point m_j.

At the end of the voting process the Hough space is searched and the cell with the highest vote is returned as the assumed position of the object.

The GHT can be extended to find rotated or scaled occurrences of the object of interest as well. However, due to a vast rise in complexity these extensions will not be considered here.

The model used in our extended GHT is built by extracting a number of edge points from a predefined region of interest around the reference point, which avoids the manual generation of a shape model. Yet, the model may contain points, which are not relevant for object localization. To reduce its size, the importance of model points is determined using a discriminative training algorithm as described below and points with low importance are eliminated thereafter.

2.2 Discriminative Training of Model Point Weights

Let us regard the model points of our initial model as individual sources of knowledge via their contribution to the Hough space. The aim of the training approach is to determine the importance of points which is achieved by an adequate combination of knowledge sources and a minimal localization error criterion.

For the derivation of the training approach, we have to take a probabilistic view of the GHT. Instead of searching for the maximal number of votes in the

Hough space, we define a posterior probability

$$p(c_i|X) = \frac{N_i}{N}, \tag{1}$$

with c_i being a cell i in the Hough space, X a set of features extracted from a given image and N_i, N being the number of votes in cell i and in the complete Hough space. By identifying the cell with the highest posterior probability the localization task is now realized with a Bayesian classifier $\hat{c} = \arg\max_{c_i} p(c_i|X)$.

The posterior probability of the Hough cells can be further divided into contributions from the individual model points:

$$p_j(c_i|X) = \frac{N_{i,j}}{N_j}, \tag{2}$$

where $N_{i,j}$ and N_j are the number of votes cast by model point j. A recombination of the model point posteriors can be achieved through a log-linear modeling following the maximum entropy principle [8]:

$$p_\Lambda(c_i|X) = \frac{\exp\left(\sum_j \lambda_j \cdot \log p_j(c_i|X)\right)}{\sum_k \exp\left(\sum_j \lambda_j \cdot \log p_j(c_k|X)\right)}. \tag{3}$$

The coefficients $\Lambda = \{\lambda_j\}_j$ are the model point weights, which need to be estimated. To this end an error function E is defined, which accumulates the localization error on the different images X_n:

$$E(\Lambda) = \sum_n \sum_i \varepsilon(\tilde{c}_n, c_i) \cdot \frac{p_\Lambda(c_i|X_n)^\eta}{\sum_k p_\Lambda(c_k|X_n)^\eta}. \tag{4}$$

Here, the Euclidian distance of the correct cell $\tilde{c}_n$ to a cell in the Hough space c_i is chosen as error measure ε weighted by an indicator function. The exponent η in the indicator function controls the influence of the alternative hypotheses c_k on the error measure. By applying a gradient descent scheme to (4) the optimal weights λ_j can be determined with respect to a minimal localization error [6].

2.3 Material and Design of Experiments

The algorithm was tested on a set of 30 thorax radiographs, which were acquired in a follow-up study of 11 patients. The images have an isotropic resolution of 0.185 mm with varying image sizes. To evade noise in the training process, the images were downsampled twice with a Gaussian filter to a spacing of 0.74 mm.

The given task is to localize the collar bone in all images. For this purpose the intersection of collar bone and lung wall was annotated to obtain a ground-truth for training as well as evaluation of the algorithm.

The dataset was divided into a training and a test dataset, each consisting of 15 images. The model used for the localization task was created from a set of 8 training images.

Table 1. The table shows a comparison of the initial and the final model regarding the number of model points, the average runtime per image in minutes and the localization error on the training and test data in mm displayed as "mean ± std (max)".

	model size	runtime	error on training data	error on test data
initial model	2684	1.2	2.2 ± 2.1 (6.9)	3.8 ± 2.6 (10.2)
final model	920	0.7	1.9 ± 2.1 (8.2)	3.0 ± 2.1 (6.7)

Using this model the GHT is performed on the training images with equal weights for all points, followed by the computation of the optimal model point weights as described above. Based on these weights a subset of points necessary for object localization is determined as final model. Subsequently, the weighted final model is tested on the complete dataset.

Experiments were run on a desktop pc with 2.4 GHz. The GHT is implemented in Matlab and not yet optimized for speed.

3 Results

The initial model extracted from part of the training images and the final model resulting after the discriminative training are shown in Figure 1. It is clearly visible that the computed point weights allow a strong reduction of the model size. Only about 35% of the initial points were kept, reducing the processing time by almost 50% as can be seen in Table 1.

By further examining the model it becomes obvious that not only the edge of the collar bone is important for its detection; but also the lung wall and surrounding bones are of high interest. The inclusion of these structures helps to distinguish the collar bone from the ribs, which show a similar edge response.

Since the initial model was extracted from the training data a very good localization result was achieved on these images, which even grew slightly better for the final model. As one could expect the error on the unknown test data is slightly higher, but should be sufficient for the initialization of a segmentation.

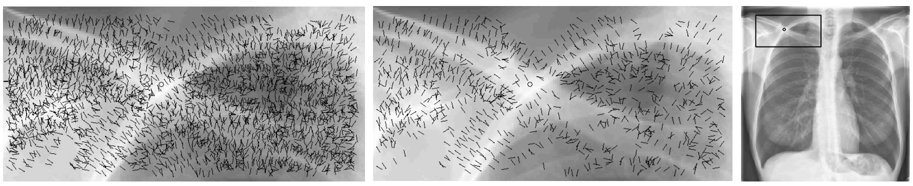

Fig. 1. Comparison of the initial model (left) extracted from the images and the final model (middle). Shown are the point locations with corresponding gradient directions. For illustration an annotated example image is displayed on the right.

4 Discussion

The combination of the GHT with the discriminative training algorithm to determine adequate model point weights shows promising results. The introduction of point weights into the model brings about several advantages. First, no initial shape model needs to be extracted from the data via automatic or manual segmentation, but a number of edge points from the region of interest are sufficient for the GHT algorithm. Second, large models can be exploited during training, since the number of model points can be diminished afterwards keeping only discriminative points. In addition, neighboring structures, which may be helpful in the localization task, can easily be included in the model by extracting a larger ROI around the object of interest. Furthermore, by extracting the model directly from the dataset, the variation of the object of interest in the underlying images is captured. Thus no rotation or scaling of the model was necessary in the process of the GHT. This implies a considerable reduction of runtime and memory need. Runtime was further decreased through the reduction of model size, while the localization error slightly improved to about 3.0 mm on average.

The algorithm is expected to be applicable to detect arbitrary objects, which show a clear response in the edge image. Experiments will be conducted to reveal further possible features to be able to detect objects, which are not characterized by their edge image. Future work will also include an iterative approach, where random points are gradually included in the model and evaluated by DMC to further improve localization accuracy.

Acknowledgment We thank Dr. Cornelia Schaefer-Prokop, AMC Amsterdam for the abundance of radiographs.

References

1. Heimann T, van Ginneken B, Styner M, et al. Comparison and Evaluation of Methods for Liver Segmentation from CT Datasets. IEEE Trans Med Imaging. 2009;28(8):1251–1265.
2. Seghers D, Slagmolen P, Lambelin Y, et al. Landmark based liver segmentation using local shape and local intensity models. In: Proc MICCAI Workshop on 3D Segmentation in the Clinic: A Grand Challenge; 2007. p. 135–142.
3. Heimann T, Münziger S, Meinzer HP, et al. A Shape-Guided Deformable Model with Evolutionary Algorithm Initialization for 3D Soft Tissue Segmentation. In: Proc IPMI; 2007. p. 1–12.
4. Barthel D, von Berg J. Robust automatic lung field segmentation on digital chest radiographs. Int J CARS. 2009;4(Suppl 1):326–327.
5. Schramm H, Ecabert O, Peters J, et al. Towards Fully Automatic Object Detection and Segmentation. In: SPIE; 2006. p. 614402.
6. Beyerlein P. Discriminative Model Combination. In: Proc ICASSP; 1998. p. 481–484.
7. Ballard DH. Generalizing the Hough Transform to Detect Arbitrary Shapes. Pattern Recognit. 1981;13(2):111–122.
8. Jaynes ET. Information Theory and Statistical Mechanics. Phys Rev. 1957;106(4):620–630.