

Multi-Level Approach for the Discriminative Generalized Hough Transform

H. Ruppertshofen^{1,2}, D. Künne¹, C. Lorenz³, S. Schmidt^{4,2}, P. Beyerlein⁴, Z. Salah², G. Rose², H. Schramm¹

¹ University of Applied Sciences Kiel, Institute of Applied Computer Science, Kiel, Germany

² Otto-von-Guericke University, Institute of Electronics, Signal Processing and Communication Technology, Magdeburg, Germany

³ Philips Research Laboratories, Hamburg, Germany

⁴ University of Applied Sciences Wildau, Department of Engineering, Wildau, Germany

Contact: heike.ruppertshofen@fh-kiel.de

Abstract:

The Discriminative Generalized Hough Transform (DGHT) is a method for object localization, which combines the standard Generalized Hough Transform (GHT) with a discriminative training technique. In this setup the aim of the discriminative training is to equip the models used in the GHT with individual model point weights such that the localization error in the GHT becomes minimal. In this paper we introduce an extension of the DGHT using a multi-level approach to improve localization accuracy and to reduce processing time. The approach searches for the target object on multiple resolution levels and combines this information for better and faster results. The advantage of the approach is demonstrated on whole-body MR images, which are intended for PET attenuation correction.

Keywords: Object Localization, Generalized Hough Transform, Machine Learning, Discriminative Training, Multi-level Approach

1 Problem

In the area of radiation therapy planning or computer-assisted interventions and diagnosis, object localization is a prerequisite for many applications in medical image processing, e.g., segmentation algorithms, where an initial position of the segmentation model in the image needs to be known. In many applications this step is still performed manually or specialized solutions are developed for each individual problem [1]. In order to obtain fully automatic processing chains, automatic and general algorithms need to be developed, which can localize objects of arbitrary shape.

Recently, we have proposed a general algorithm for object localization, called Discriminative Generalized Hough Transform (DGHT) [2], which combines the Generalized Hough Transform (GHT) [3] with a discriminative training algorithm. The procedure runs fully automatic and is independent of the target object, which is searched for in the image. The only restriction for the target is that it should be well defined by its shape. For the medical localization tasks considered by us, our algorithm achieves high quality results, mainly tested on 2D images [2]. Until now the GHT has been applied only rarely to 3D images due to its computational complexity [4]. Nevertheless, we have shown in [5] that the application of our algorithm to 3D images becomes practicable due to sparse models and the restriction of transformation parameters, i.e. estimating only translation.

However, when the region of the target object exhibits only low contrast or when many similar objects are visible in the image, the localization is hampered. Furthermore, if the images are large and high localization accuracy is needed, the feasibility of the procedure is questioned due to long processing times. To solve these issues, we introduce an extension of the DGHT using a multi-level approach for more robust and faster results. The target object is first searched for on a low resolution image and the result of this search is used as input for the next higher resolution level, meanwhile reducing the size of the search region.

The extension of the algorithm is tested on whole-body MR images, which were acquired for a PET attenuation correction [5]. Since patients often have to undergo several subsequent MR examinations, fast acquisitions and low specific absorption rates (SAR) are of higher interest for this task than a detailed mapping of the anatomy. The resulting image resolution, resembling the resolution of PET, is rather low, which complicates the image processing due to the lack of anatomical detail available, but also renders it very interesting for the proposed algorithm.

2 Methods and Material

The GHT [3] is a standard method for object localization, which employs a point model to represent and search for a target object in an image. The model is thereby moved across the edge image corresponding to the original image and the co-incidences of model and edge points are counted in a voting process and accumulated in the Hough space. The Hough cell, which obtained the highest vote, is assumed to represent the true target location.

In the DGHT the models are furthermore equipped with individual model point weights. These weights are trained with a discriminative training algorithm [6], based on the information available in the Hough space, i.e. which point has voted in which Hough cell, with the aim to obtain a low localization error in the GHT.

In order to obtain a meaningful model and to be able to capture the variability contained in a training dataset, the model for the GHT is generated directly from the image data by taking the edge points from a given volume of interest (VOI) around the target point from a number of images and is refined in an iterative approach. The procedure starts on a small set of training images, on which preliminary model point weights are trained. The current model is then evaluated on a larger development dataset. Images where the model performs poorly are added to the training dataset, further model points are created from these images, and another iteration is performed until the error on all development images is below a certain threshold or no further improvement is achieved. For more detailed information on the iterative training technique and the DGHT, we refer the reader to [2].

In this paper, we introduce the combination of the DGHT with a multi-level approach. To this end, a Gaussian pyramid of the image is created and the localization is performed on each level. To speed up the procedure the localization is first executed on the lowest resolution level, where only little detail is visible and the localization is fast, due to the small image size. For the next higher resolution level it is assumed that the previous localization result is near the target point such that the search can be constrained to a smaller region. In the following experiments an extract with half the side lengths of the previously considered image extract is cut out around the localized point, such that the number of pixels remains almost constant, while the resolution of the image increases. Thus more and more detail is taken into account on each level while zooming into the target object. The idea of the approach is illustrated in Fig. 1.

For each level of the pyramid an individual model is created using edge points from a VOI as stated above. While the VOI needs to be given for the standard approach, here it is chosen to be centered at the target point with a side length of 75% of the current image extract. Only part of the extract is used for model generation in order to reduce model size and to prevent the algorithm from learning the exact field of views, which might be different on test images.

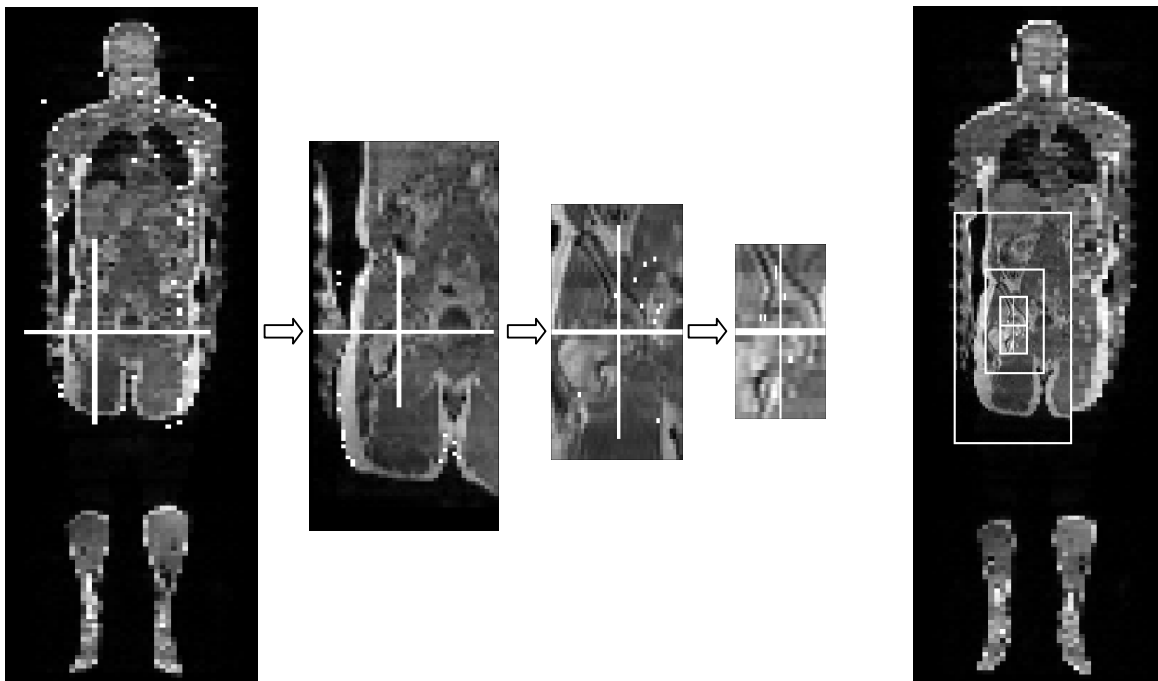


Fig. 1: Illustration of the steps of the multi-level approach. The localization procedure starts on the low resolution image on the left. In the subsequent steps (left to right), the procedure zooms into the target by performing the localization on regions with decreasing size and increasing resolution around the previously localized point (white cross hairs). The right image shows an overlay of the different image extracts used for the localization.

The method is tested on 22 whole-body MR images, which were acquired on a Philips Achieva 3T X-Series MRI system using a whole-body protocol suitable for attenuation correction. As was said earlier, the images are not intended for diagnostic purposes but for the attenuation correction of PET images; therefore a sequence with fast acquisition is applied, which results in images with the rather low resolution of approximately 1.875 mm in plane and a slice thickness of 6 mm. Example images are displayed in Fig. 2.

The given task for these images is to localize the femur for a subsequent segmentation. To this end the center of the femoral head of the right leg was chosen as target point, which is marked in the left images in Fig. 2.

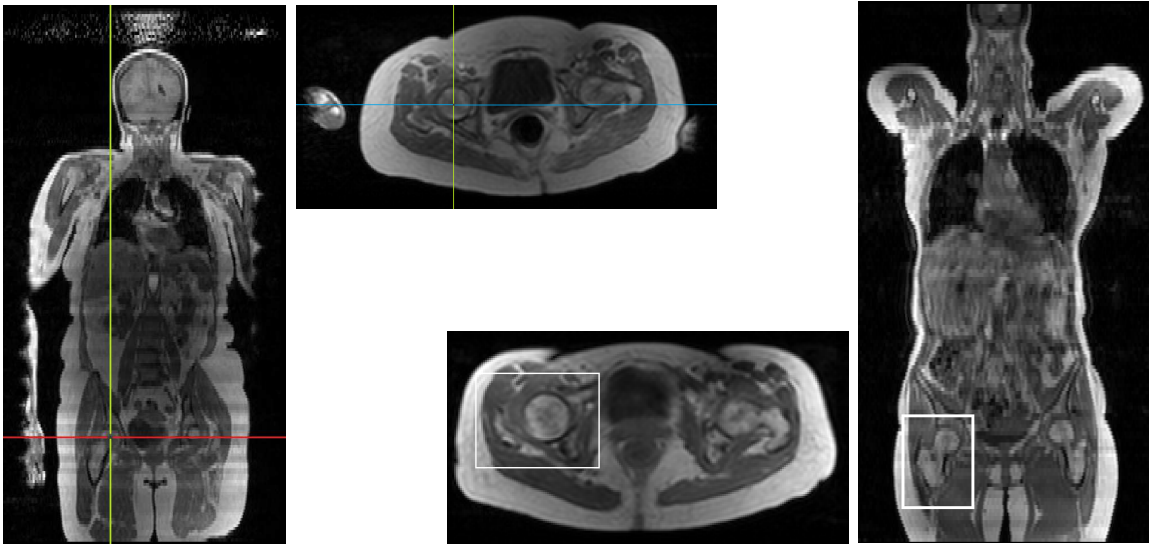


Fig. 2: Coronal and transversal view of two example images. The images on the left additionally show the target point in the center of the femoral head, while the right images display the VOI, which was used for model generation in the standard approach.

From the dataset 10 images are chosen randomly as development dataset, while the remaining images are used for testing purposes. The image pyramid is created with 4 levels and since the resolution in z-direction is much larger than in-slice, the image is downsampled only in x- and y-direction for the first levels of the pyramid to obtain a rather isotropic resolution.

To be able to compare the new multi-level approach with the former approach, a second experiment is performed using the same parameter setting. Yet, instead of utilizing an image pyramid, only one resolution level is employed. To reduce processing time and memory need of the GHT, the image is downsampled once in-slice. For the model creation a VOI around the femoral head is defined, which can be seen in Fig. 2 (right).

3 Results

The results of the two experiments are stated in Table 1. For the training images the results differ only slightly with a good mean localization error of 3.1 mm and 2.2 mm for the standard and multi-level approach, which is not surprising since the models were trained on these images. However, regarding the unknown test data, the standard approach was substantially outperformed by the multi-level approach. The latter achieves a much better localization error with a mean distance of 3.8 mm of the localized and annotated point, while the former obtains only 6.7 mm and even fails on one of the test images.

	Training		Test		Proc. time
	Local. rate	Mean error	Local. rate	Mean error	
Standard approach	100 %	3.1 mm	91.6 %	6.7 mm	28 s
Multi-level approach	100 %	2.2 mm	100 %	3.8 mm	3 s

Table 1: Comparison of the two localization approaches. The table states the localization rate and mean error of the correct localizations on the training and unknown test data and the localization time for the two approaches.

The advantage of the multi-level approach becomes even more obvious when considering the processing time. Due to the much smaller image extracts (only 1-3 % of the pixels of the original image) on which the localization is performed on each resolution level, the multi-level approach takes only about 10 % of the processing time of the standard approach.

4 Discussion

The multi-level approach has proven to be significantly better and faster than the standard approach for the given task. The main reason for the smaller localization error is the higher resolution used in the multi-level approach. While the standard approach, due to computational reasons, performs the localization only on the second resolution level of the image pyramid, the multi-level approach employs the original resolution in its final stage. The obtained localization error of 3.8 mm is exceptional, considering the low resolution of the images, especially the slice distance of 6 mm.

Another advantage of the newly proposed approach, which zooms into the target object, is that it takes the neighborhood of the target into account on a larger scale and thereby facilitates to localize objects with low contrast, high variability or which can be easily confused with further objects visible in the image. One of the test images has a smaller field of view compared to the rest of the dataset. This image covers the body only from the head to the upper part of the femur, still showing the femoral head but not the remainder of the femoral bone. The standard approach, which relies on the whole bone being visible, fails to deal with this occlusion, while the multi-level approach, which orients itself on a larger scale, is not affected by the limited field of view.

In the presented example the algorithm achieved to localize the target with the necessary accuracy on all resolution levels. However, if necessary, it would be conceivable to keep several candidate points on each level, which could be discarded later on higher resolution levels, when identified as false-positives.

Besides facilitating the object localization and making it more robust, the multi-level approach has another large advantage, which lies in the shorter processing times. Since the image extracts, which are used for the localization, are much smaller than the original image, only a fraction of the run time is needed, depending on the size of the image and the number of zoom levels used. In the presented example, a reduction of processing time of 10 % was achieved. With the runtime of 3 s the application of the algorithm in 3D becomes really feasible. Furthermore, the procedure is not yet optimized for speed, so that a further reduction of processing time is to be expected.

In future work, the usage of the demonstrated localization procedure in combination with the segmentation for the attenuation correction will be examined. Since only little anatomical detail is visible in the image, a precise positioning of the segmentation models is needed, which we are confident to fulfill with the presented approach.

5 Acknowledgments

The authors would like to thank the Department of Radiology, Mt. Sinai School of Medicine, New York, the Department of Imaging Sciences and Medical Informatics, Geneva University Hospital and the Business Unit NM/CT, Philips Healthcare for providing the data used in this study. This work is partly funded by the Innovation Foundation Schleswig-Holstein under the grant 2008-40 H.

6 References

- [1] T. Heimann, B. van Ginneken, M. Styner et al., Comparison and Evaluation of Methods for Liver Segmentation from CT Datasets, *IEEE Transactions on Medical Imaging* 28(8), 2009
- [2] H. Ruppertshofen, C. Lorenz, S. Schmidt, P. Beyerlein, Z. Salah, G. Rose, H. Schramm, Discriminative Generalized Hough Transform for Localization of Joints in the lower extremities, *Computer Science – Research & Development* 26, Springer, 2011
- [3] D. H. Ballard, Generalizing the Hough Transform to Detect Arbitrary Shapes, *Pattern Recognition* 13(2), 1981
- [4] H. Schramm, O. Ecabert, J. Peters, V. Philomin, J. Weese, Towards Fully Automatic Object Detection and Segmentation, *Proceedings of SPIE medical imaging*, 2006
- [4] H. Ruppertshofen, C. Lorenz, S. Schmidt, P. Beyerlein, Z. Salah, G. Rose, H. Schramm, Lokalisierung der Leber mittels einer Diskriminativen Generalisierten Hough Transformation, *Proceedings of CURAC*, 2010
- [5] Z. Hu, N. Ojha, S. Renisch, et al., MR-based Attenuation Correction for a Whole-body Sequential PET/MR System, *Proceedings of IEEE Nuclear Science Symposium*, 2009
- [6] P. Beyerlein, Discriminative Model Combination, *Proceedings of IEEE International Conference on Acoustics, Speech and Signal Processing*, 1998