THE DISCRIMINATIVE GENERALIZED HOUGH TRANSFORM FOR FULLY AUTOMATIC OBJECT DETECTION AND LOCALIZATION

Gabriel E.[a], Hahmann F.[a], Meyer C.[a,b], Schramm H.[a,b], Koch R.[b]

[a]Institute of Applied Informatics, University of Applied Sciences Kiel, Germany;
[b]Institute of Computer Science in the Technical Faculty of Christian-Albrechts-University Kiel, Germany

eric.gabriel@fh-kiel.de

Abstract

The Discriminative Generalized Hough Transform (DGHT) is a general and robust automatic object detection and localization technique. We outline the basic algorithm and demonstrate its performance in an eye localization task on the FERET database, nearly achieving state-of-the-art results without any task-specific optimizations. In the next project phase, we aim to extend the DGHT framework towards multi-object detection in a real-life scenario, and illustrate a basic concept.

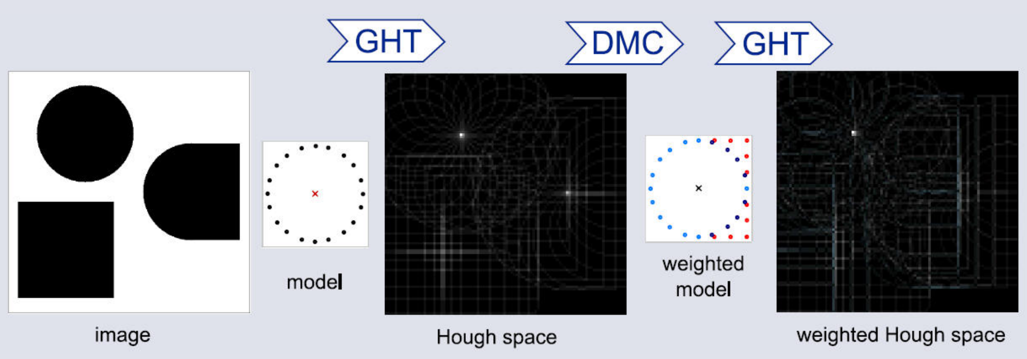
1. Goals and Motivation

- DGHT is a powerful and general object detection and localization technique
  - Model-based
  - Can handle occlusion, clutter and noise
- Goal of the project:
  - Extension of the DGHT framework towards multi-object detection / localization scenarios
- First step:
  - Quantitative evaluation of DGHT performance on an eye localization task
  - Comparison to state-of-the-art approaches

2. Basic Idea of the DGHT

The basic system for model training consists of two main modules:

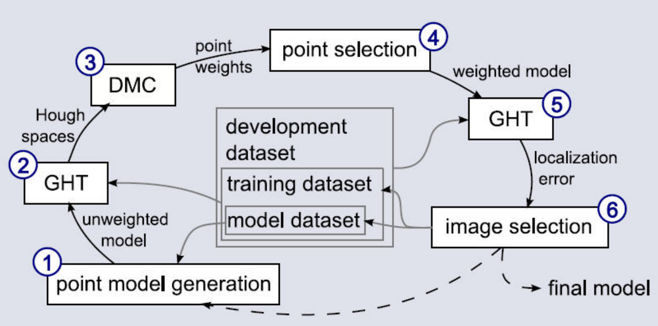
1. The Generalized Hough Transform (GHT) [1] for object localization
2. A discriminative training technique (DMC) [2] to train individual model point weights, which are utilized in the GHT



Model point weights: ● = 1, ● = 0.5, ● = -1

The usage of model point weights results in more discriminative models and more robust results. Negative weights allow to include confusable structures in the model.

3. Overview of DGHT Training Procedure



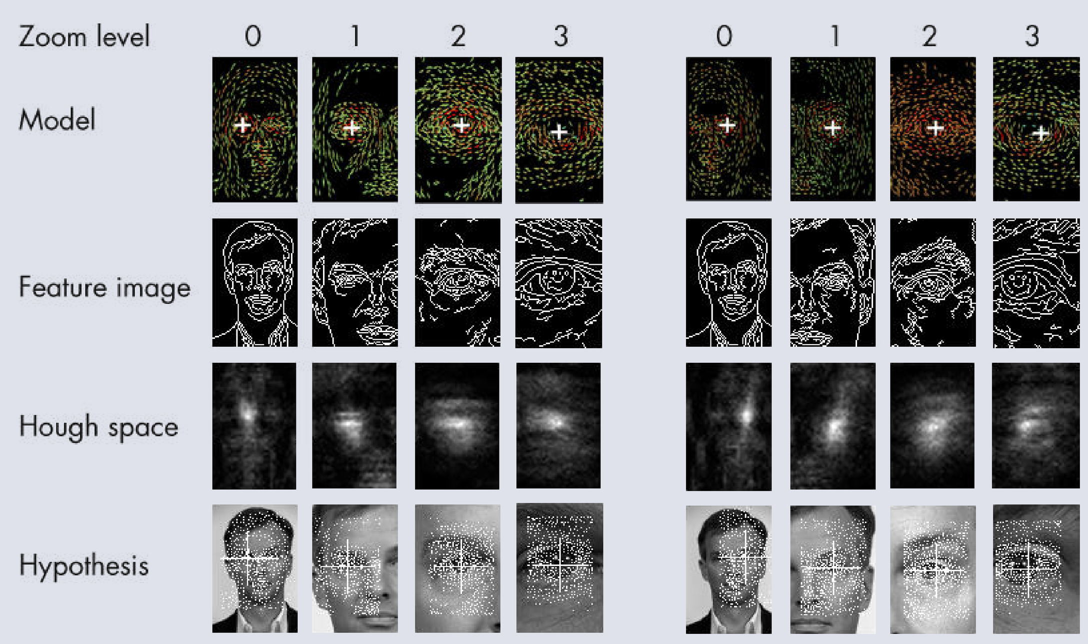
1. Generate an initial model from images in model dataset
2. Perform the GHT on training dataset with unweighted model
3. Use Hough space information to compute model point weights
4. Remove points with low absolute weight
5. Perform GHT on training dataset with weighted model
6. Add images with high localization error to model datasets.

⇒ If all training images yield satisfying results or no new images can be included into the model, return final model.

4. DGHT on Eye Localization Task

- Database: FERET[3], 2409 images
- Setup: 594 training images, 1815 test images (subjects separated)
- Database variability:
  - Eyes opened / (partly) closed, glasses / no glasses
  - Background, illumination
  - Age, gender, ethnic groups
- Sample images:

- Training of two separate models (right / left eye) based on edge features in four zoom levels (subsampling) [4]
- Illustration of DGHT models and application to test image:



(1) right eye (2) left eye

5. Experimental Evaluation

- Precision of eye detection given by relative error measure proposed by Jesorsky[5]:

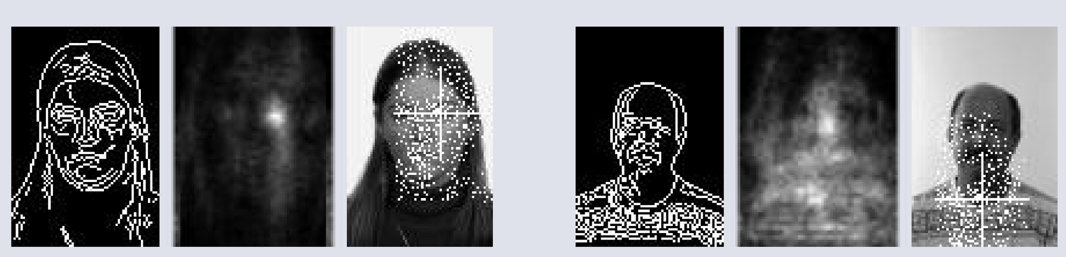
$$d_{eye} = \frac{\max(||C_l - \tilde{C}_l||, ||C_r - \tilde{C}_r||)}{||C_l - C_r||}$$

C_l, C_r: true eye positions
$\tilde{C}_l, \tilde{C}_r$: estimated eye positions

Method*	$d_{eye} \leq 0.1$	$d_{eye} \leq 0.15$	$d_{eye} \leq 0.25$
Hassaballah[6]	83.0 %	96.7 %	97.6 %
Qian[7]	91.9 %	96.9 %	97.7 %
Kroon[8]	97.6 %	—	99.6 %
DGHT	96.2 %	97.5 %	98.3 %

- $d_{eye} \leq 0.1$ = within the iris
- $d_{eye} \leq 0.25$ = within half of an eye width

- Other approaches are specificly designed for eye detection
  - First step: face detection
  - Further processing of the proposed eye regions within the detected face
- DGHT result examples:



(a) Good localization (b) Bad localization

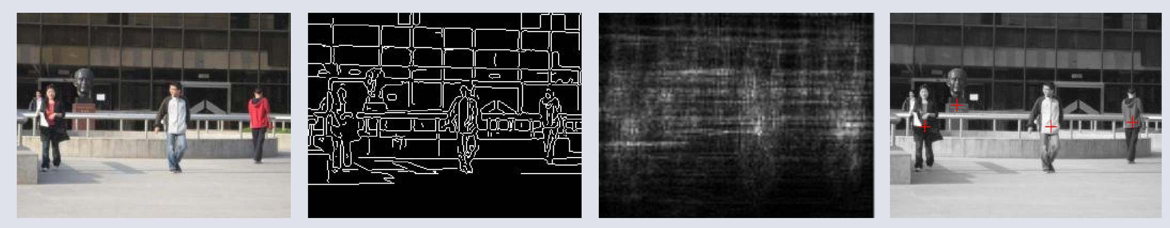
* The methods differ in the definition of the test to some extent.

6. Discussion

- DGHT performs comparably to state-of-the-art approaches in eye localization and other tasks [9, 10]
  - In general without task-specific optimizations
- Limitations of current DGHT algorithm:
  - No general multi-object localization (only with prior knowledge)
  - No general multi-class localization (only with classifying DGHT)
  - Limited treatment of variability in size and pose
  - Training and test conditions have to be closely matched

7. Extension: Multi-Object Detection

- Extension of DGHT framework towards multi-object localization
- Example: Person localization (e.g. IAIRCarPed database[11])



- Concept: DGHT provides list of interest points (N positions in Hough space with highest number of votes) for additional classification:
  - Pattern classification techniques applied to bounding box around interest point
  - Modeling geometric relationships (graphical models) e.g. distance between left and right eye
- Issues:
  - Handling of variability
  - Integrating other than edge features (in DGHT / classification)
  - Using Hough space in classification / graphical model

References

[1] Ballard, D. H., "Generalizing the hough transform to detect arbitrary shapes," *Pattern Recognit.* **13**(2), 111–122 (1981).

[2] Beyerlein, P., "Discriminative model combination," *Proc. IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, 481–484 (1998).

[3] Phillips, P.J., et al., "The FERET database and evaluation procedure for face-recognition algorithms", *Image Vision Comput.* **16**(5), 295–306 (1998).

[4] Ruppertshofen, H., et al., "Multi-level approach for the discriminative generalized Hough transform", *CURAC* **10** (2011).

[5] Jesorsky, O., et al., "Robust face detection using the hausdorff distance", *Audio-and video-based biometric person authentication*, 90–95 (2001).

[6] Hassaballah, M., et al., "An Automatic Eye Detection Method for Gray Intensity Facial Images", *International Journal of Computer Science Issues (IJCSI)* **8**(4) (2011).

[7] Qian, Z., & Xu, D., "Automatic eye detection using intensity filtering and K-means clustering", *Pattern Recognition Letters* **31**(12), 1633–1640 (2010).

[8] Kroon, B., et al. "Eye localization in low and standard definition content with application to face matching", *Computer Vision and Image Understanding* **113**(8), 921–933 (2009).

[9] Hahmann, F., et al., "Bone Age Assessment Using the Classifying Generalized Hough Transform", *Pattern Recognition*, 313–322 (2013).

[10] Ruppertshofen, H., et al., "Discriminative generalized hough transform for localization of joints in the lower extremities", *Comp. Sci.- Research and Dev.* **26**(1-2), 97–105 (2011).

[11] Yang, W., et al., "IAIR-CarPed: A Psychophysically Annotated Dataset with Fine-Grained and Layered Semantic Labels for Object Recognition", *Pattern Recognition Letters* **33**(2), 218–226 (2012).