Study of higher order image descriptors
Project description

Malte Stær Nissen
tgq958@alumni.ku.dk

Benjamin Michael Braithwaite
cpg608@alumni.ku.dk

Supervisor: Kim Steenstrup Pedersen, kimstp@diku.dk
Co-supervisor: Sune Darkner, darkner@diku.dk

February 6, 2014

1 Problem statement

Features/interest points are a central concept in computer vision. These points mark areas of an image containing a significant amount of information about the local structure. A visual descriptor is a representation of the local structure of a feature. Such a descriptor can be modelled in numerous ways depending on the application and desired properties. In this thesis we will study state of the art visual descriptors and develop, implement, test, and discuss our own descriptors based on this study.

2 Motivation

Computer scientists have for many years tried to make computers perceive the world as humans do. This is done by performing analysis of real world images. The analysis is often based on extraction of features in the images. These features can then be be used for applications such as: object detection of pedestrians [FMR08], content based image retrieval [SWS+00], feature matching for solving the image correspondence problem [DAP11], etc. The state of the art methods are however not perfect and hence we will try improve upon the accuracy of the methods.

3 Proposed solution

We will try to design and implement a variant of the HOG descriptor utilizing higher order differential information. We will select a few of the following

applications for testing our solution: Texture recognition, pedestrian detection, general object detection using deformable parts models, image correspondence, medical image registration, and content based image retrieval.

At first we will try to experiment with a HOG descriptor extended to various orders of differential information and various parameters for the number of cells per block, number of pixels per cell and number of channels per cell histogram. Furthermore we will experiment with the combination of HOG and the shape index [KvD92] as well as different ways of processing the higher order information.

4 Learning goals

The following learning goals will be the success criteria for our project.

- Perform a literature study of visual descriptors and their applications.
- Develop our own visual descriptor based on state of the art work within the field.
- Select and implement solutions of some of the mentioned applications of visual descriptors.
- Perform a parameter study of our solutions.
- Perform an empirical evaluation of our solutions on selected problems compared to state of the art solutions.
- Discuss the results of the empirical studies.

5 Risk assessment

The focus of the thesis is to gain insight into the field of descriptors and develop our own non-trivial descriptor on the basis hereof. This makes the risk of failure low since no requirements for the accuracy of the final product are stated. There is however a risk of developing a solution which performs worse than state of the art methods, which will however still have given us significant insight into the field in question.

6 Time schedule **(preliminary)**

- **Literature study: Deadline:** Monday the 24th of February.
- **HOG and shape index:** Basic descriptor implementation and preliminary report section about HOG extended with shape index.
  Deadline: Wednesday the 5th of March.

- **Image correspondence implementation:** Feature matching implementation and test with descriptor on DTU robot dataset. Chosen to get a simple initial estimate of the viability of our descriptor choices. Report section about the subject written.
  Deadline: Wednesday the 19th of March.
- **Unknown workload:** Further work will be determined continously.
- **Final experiments/results: Deadline:** Sunday the 3rd of August
- **Final report: Deadline:** Sunday the 17th of August

References

[DAP11] Anders Lindbjerg Dahl, Henrik Aanæs, and Kim Steenstrup Pedersen. Finding the best feature detector-descriptor combination. In *3D Imaging, Modeling, Processing, Visualization and Transmission (3DIMPVT), 2011 International Conference on*, pages 318–325. IEEE, 2011.

[FMR08] Pedro Felzenszwalb, David McAllester, and Deva Ramanan. A discriminatively trained, multiscale, deformable part model. In *Computer Vision and Pattern Recognition, 2008. CVPR 2008. IEEE Conference on*, pages 1–8. IEEE, 2008.

[KvD92] Jan J Koenderink and Andrea J van Doorn. Surface shape and curvature scales. *Image and vision computing*, 10(8):557–564, 1992.

[SWS+00] Arnold W. M. Smeulders, Marcel Worring, Simone Santini, Amarnath Gupta, and Ramesh Jain. Content-based image retrieval at the end of the early years. *IEEE Trans. Pattern Anal. Mach. Intell.*, 22(12):1349–1380, December 2000.