

EINLADUNG

Zeit: 02. Dezember 2008, 16:00 Uhr

Ort: Raum 5056, Ahornstr. 55

Referent: Dipl. Informatiker Thomas Deselaers

Titel: **Image Retrieval, Object Recognition, and Discriminative Models**

Abstract:

In this thesis, we present approaches to image retrieval, object recognition, and discriminative models.

For image retrieval, we evaluate a large variety of different descriptors and answer the questions how descriptors can be combined and which descriptor should be chosen according to which criterion. We suggest a set of local descriptors that have been used successfully for object recognition and combine these with textual information and several other descriptors. Additionally, we present methods to optimally fuse visual and textual data for retrieval.

For object recognition, we propose different models and investigate and analyse their relationships and their individual advantages and disadvantages. In particular, we try to avoid heuristics in the creation of the models and incorporate all available knowledge cues. We extend the bag-of-visual words approach into several directions in order to overcome its limitations. In total, we present eight different models for object recognition including a nearest neighbour-based model, two variants of bag-of-visual-words models, and a model based on geometric matching incorporating spatial relationships. We also present a model based on Gaussian mixtures which abandons vector quantisation, can be trained discriminatively, and can incorporate spatial relationships. This model is then rewritten and extended toward log-linear mixtures and support vector machines. We also present a random-forest-based approach that fuses appearance, shape, and depth cues for human computer interaction.

Regarding discriminative models, we delve deeper into some aspects of image retrieval and object recognition. We propose a novel model for optical character recognition. We extend log-linear models to incorporate hidden variables, thus allowing for modelling image deformations and multi-modal data. Furthermore, we investigate the relationship between certain support vector machines and Gaussian mixtures in order to achieve a joint model that fuses their advantages. All approaches proposed in this work were evaluated on standard benchmarks. For image retrieval, we experimentally evaluated the performance of a large variety of descriptors, how they perform on different tasks, and how they can be combined to achieve different results. We participated in several Image-CLEF evaluations and obtained excellent results using content-based image retrieval techniques. In particular, we achieved the best result using visual retrieval in the Image-CLEF 2007 medical retrieval task using our discriminatively trained feature combination. The object recognition approaches were evaluated on the Caltech and PASCAL tasks and it could be shown that Gaussian mixtures and related approaches incorporating spatial information and avoiding vector quantisation outperform all other approaches. The methods proposed in the chapter on discriminative models were evaluated on the standard USPS and MNIST tasks and our deformation-aware log-linear model achieves very competitive results while using an order of magnitude fewer parameters than competing approaches.

The talk will mainly focus on the parts about object recognition and discriminative models.

Es laden ein: Die Dozenten der Informatik