Photo Retrieval and Photo Annotation Tasks

Jonathon S. Hare, David P. Dupplaw & Paul H. Lewis
Intelligence Agents Multimedia Group
School of Electronics and Computer Science, University of Southampton

Photo Retrieval

Task
The 2009 photo retrieval task aimed to provide further study of the importance of diversity in image search results. Diversity in search results is often helpful to users who drive their search with poorly specified or ambiguous queries. Since the search engine cannot know precisely what the user wants, it can increase the probability of finding relevant images by presenting the complete spectrum of representations within the query. The task required participants to analyze methods for increasing diversity whilst maintaining high relevance.

Research Question
Can we use content-based image features to increase image search result diversity?

Baseline Text-based Indexing and Retrieval

Our baseline retrieval system uses Terrier as the underlying platform. The queries from the two parts of the task were handled differently. Terrier was configured to just use standard TF-IDF weighting for the retrieval. Two term-processing pipelines were tested; one with Porter stemming, and another without.

Image Features

We used a bag-of-visual-terms feature morphology for both the retrieval and annotation tasks. For the retrieval task, we used a multi-scale difference-Gaussian peak detector for region detection, coupled with the SIFT local descriptor and a 3125 term vocabulary. For the annotation task we combined MSER and doG detectors with SIFT and Colour-SIFT features, using 3125 term vocabularies for each detector/feature combination.

Photo Annotation Task

Task
The 2009 photo annotation task aimed to investigate whether the use of a small ontology or hierarchy could be used to improve annotation performance. The task was to annotate 1000 images with 53 concepts from the provided hierarchy. 5000 annotated images were provided for training.

Technique
We used an auto-annotation tool that we had previously developed. The tool uses a matrix factorisation of a multi-lingual (visual-terms and keywords) term-document matrix to build a semantic space. Un-annotated images can be projected into this space (based on their visual-terms), and their placement is such that they occur “near” keywords that describe their content.

Results
The semantic space performs reasonably in terms of EER and AUC, and is also computationally efficient for such a small dataset. Performance with the hierarchical scoring measure is compromised due to the difficulty in generating binary annotation probabilities from the space.

Future Work
Currently we are partners in the FP7 funded FET project LivingKnowledge. LivingKnowledge aims to investigate the use of bias and diversity in future web search. Our work in ImageCLEF is already feeding into the project prototypes. One particular area of current interest is the use of semantic web techniques to automatically generate sub-topics (like those in the part 1 topics illustrated above) from a user provided query. With respect to our visual re-ranking algorithm, the next stage is to improve it so that it incorporates information from the original ranking provided by the text search in order to attempt to reduce the precision loss.