IAM@ImageCLEFphoto 2009: Experiments on Maximising Diversity using Image Features

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

This paper describes the diversity-enabled retrieval system constructed at Southampton for the ImageCLEFphoto 2009 task. The task aimed to promote diversity in image search. The task incorporated two separate query types. The part 1 topics were described as a main topic (i.e. 'David Beckham'), together with a set of clusters or sub-topics (i.e. 'Manchester United', 'Real Madrid', etc.). The part 1 topics also included detailed information about what might be expected in the results of a search for each of the clusters. The part 2 topics provided a single topic with no context.

Our retrieval system used Terrier as the underlying textual indexing and retrieval platform, and combined it with a technique for re-ranking the results by maximising the visual dissimilarity of retrieved images. Southampton's baseline system used standard text retrieval techniques for the part 2 topics. The baseline handling of the part 1 topics augmented the standard text search with multiple sub-queries (one per cluster) followed by a merge phase in order to build a complete ranking for the topic. On top of the baseline system we developed a re-ranking procedure for the results lists that leveraged visual features extracted from the images and attempted to re-order the list such that the first images in the list were highly visually dissimilar. We hypothesised that the use of the visual re-ranking would increase the diversity of the images at the top of the result set of each query.

The results showed that our visual re-ranking method does indeed work at increasing the diversity in the top results. However, at the same time it causes a slight drop in precision. The text-based approach designed for handling the 'part 1 topics' of the task was shown to perform very well even without visual features.