

I A M @ImageCLEF 2009

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



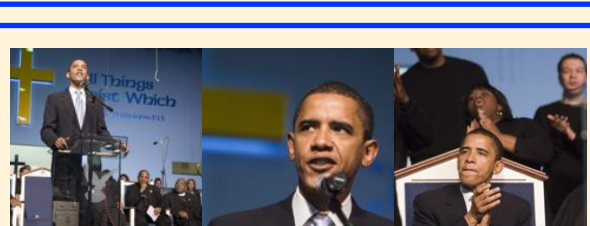
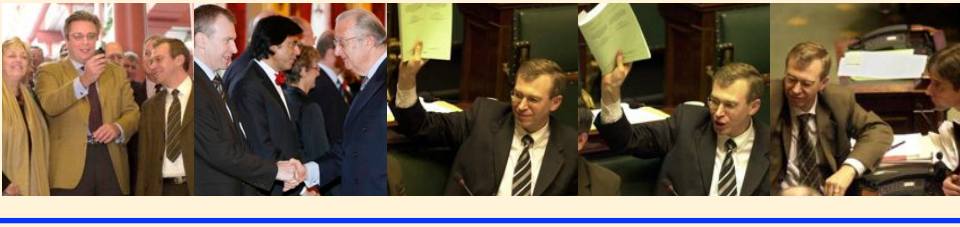
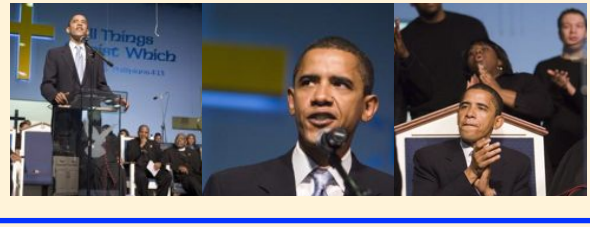
Task
The 2009 **photo retrieval task** aimed to provide further study of the **importance of diversity** in **image search** results. Diversity in search results is often helpful to users who drive their search with **poorly specified** or **ambiguous queries**. Since the search engine cannot know precisely 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 analyse 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.

<small>2007 US District Attorney Johnnie Cochran Jr. (CJ) is surrounded by reporters as he arrives at the Criminal Courts Building in Los Angeles where he is expected to continue his opening statements in the O.J. Simpson double murder trial 26 January. Cochran expressed his concern for the health of prosecutor William Hodgman who was hospitalized late 25 January, complaining of chest pains. 100000 [Louisin-La-Neuve - 10APR95 - Marcel CROCHET (photo) who succeeds Pierre Macq as rector of the LLN-university, pic-Belga-Herwig Vergut. 100000] Peng Ming-min, presidential candidate of the main opposition Democratic Progressive Party (DPP), tells the press 20 March that Taiwan is de facto independent and that it is absurd to say it is a part of China. Peng is running in Taiwan's first-ever direct presidential election 22 March. AEP-DM/TCY</small>		Indexing
Part 1 Topics <top> <num> 1 </num> <title> leterme </title> <clusterTitle> yves leterme </clusterTitle> <clusterDesc> Relevant images contain photographs of Yves Leterme. Images of Leterme with other people are relevant if Leterme is shown in the foreground. Images of Leterme in the background are irrelevant. </clusterDesc> <image> belga28/05980958.jpg </image> <clusterTitle> leterme albert </clusterTitle> <clusterDesc> Relevant images contain photographs of Yves Leterme and King Albert II. Images with only one of them are considered to be irrelevant. </clusterDesc> <image> belga27/05960161.jpg </image> <clusterTitle> leterme albert </clusterTitle> <clusterDesc> Images which contain photographs of Leterme which are not part of the above categories are relevant to this cluster. </clusterDesc> <image> belga32/06229323.jpg </image> </top>	Part 2 Topics <top> <num> 26 </num> <title> obama </title> <image> belga30/06098170.jpg </image> <image> belga28/06019914.jpg </image> <image> belga30/06107499.jpg </image> </top>	Topic Specification
+yves +leterme +leterme +albert +leterme -albert	+obama	Terrier Query (TF-IDF Weighting Model)
		Intermediate results
		Final results

Re-ranking results using visual features

Using visual features extracted from the images, we developed an iterative technique for re-ranking the images based on the idea of maximising the distance between the currently ranked images and the remaining images. This has the effect of forcing highly dissimilar (or diverse) images to the top of the list, and images that are near-duplicates of highly ranked images to the bottom.

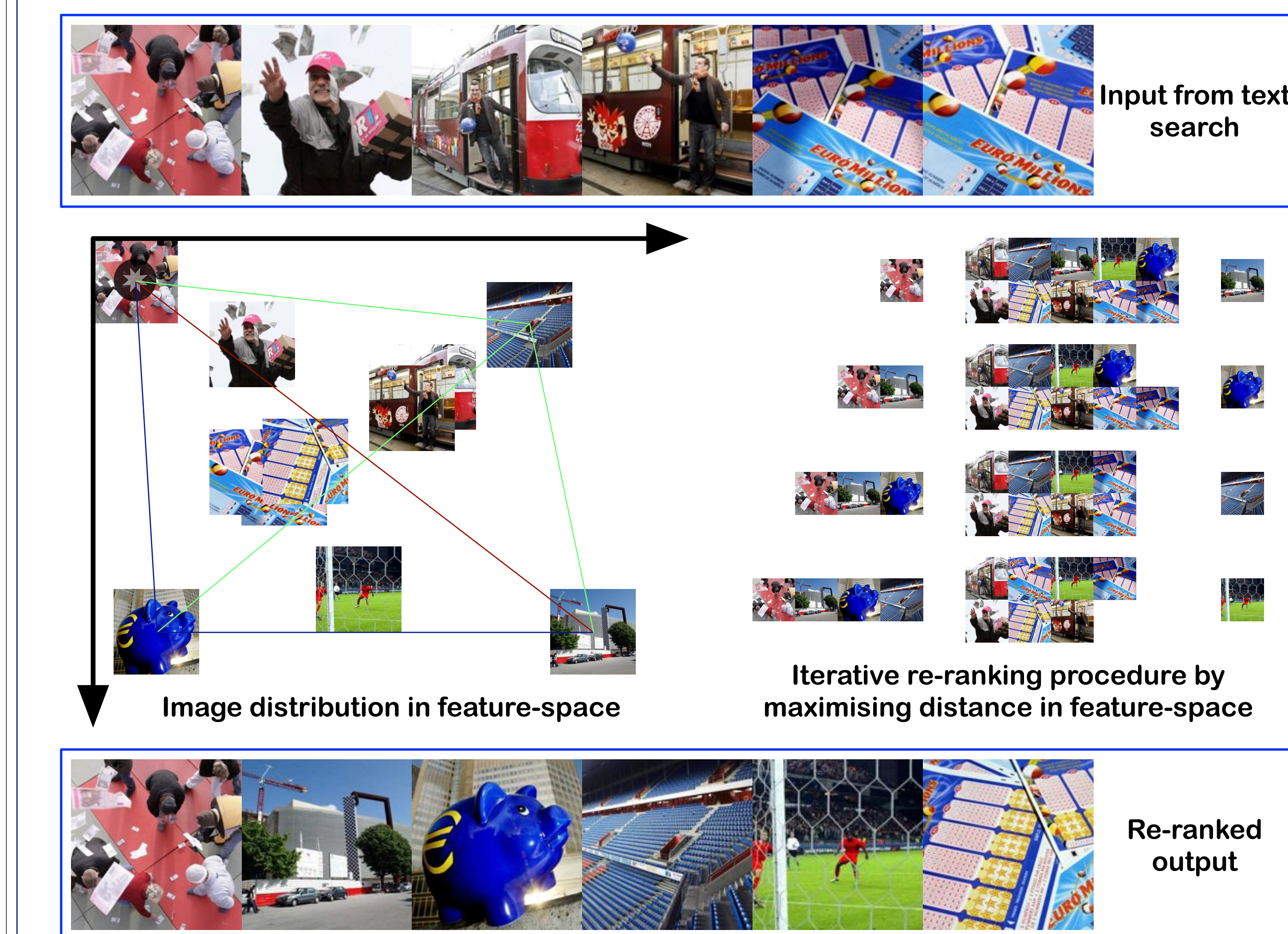


Image Features

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-of-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.

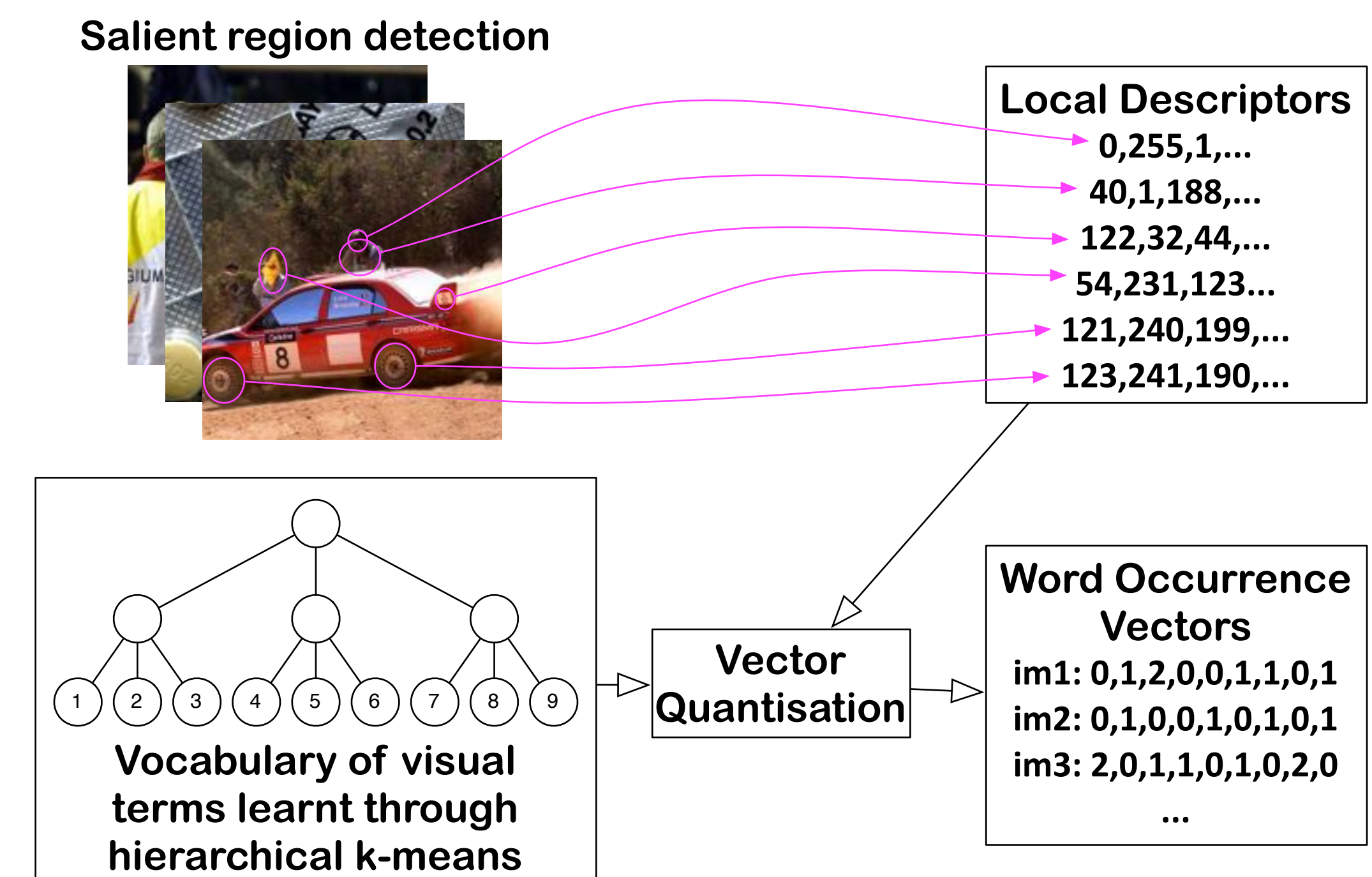
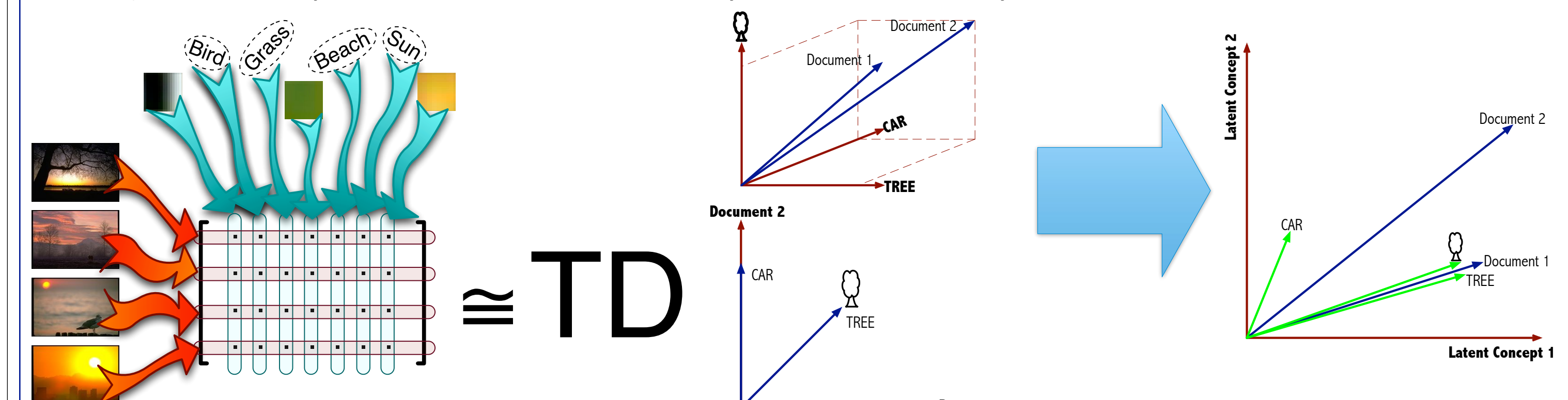


Photo Annotation Task

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 **13000** 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.

Retrieval Results

Four sets of runs were submitted: text only (with stemming), text only (no stemming), text + visual re-ranking (with stemming), and text + visual re-ranking (no stemming).

The results indicated:

- The **text-retrieval** (especially for part 1 topics) was very **effective**.
- The **visual re-ranking** does indeed **boost diversity** (as seen through an increase in cluster recall), however, it also causes a **drop in precision**.
- Disabling stemming** gives a slight **precision increase**; we hypothesise that this is due to the large number of **named entities** (names of people, places, things, etc.) in the queries.

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**.

