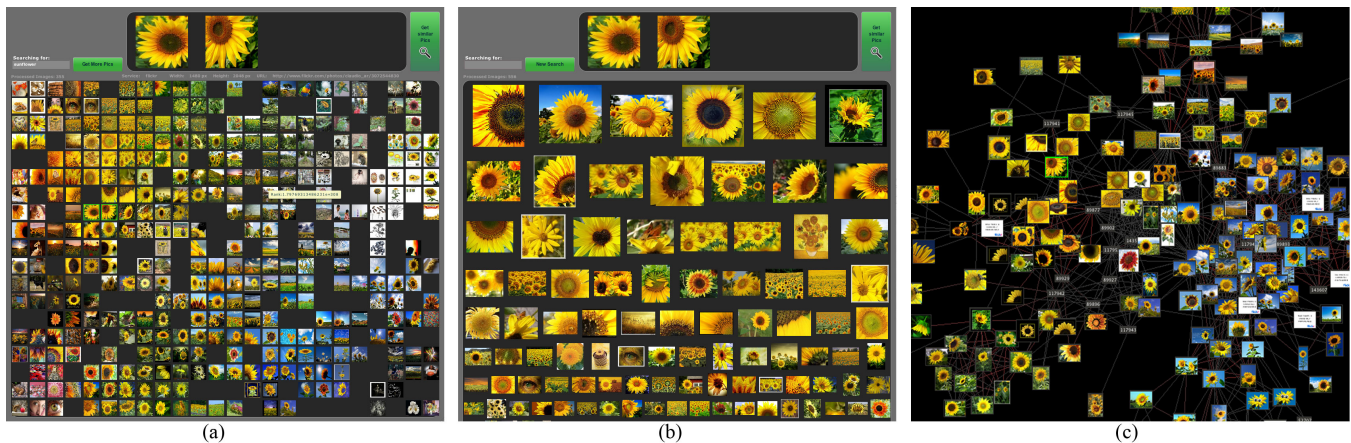# Image Retrieval using Collaborative Filtering and Visual Navigation

Kai Uwe Barthel, Sebastian Müller, David Backstein, Dirk Neumann, Klaus Jung *
HTW, University of Applied Sciences, Berlin

Figure 1: *(a) A visually sorted display of 355 Flickr "sunflower" images. Due to the visually sorted display particular images can easily be found. The user can select candidate images to refine the search. (b) These selected images are used to retrieve visually similar images and to model the semantic relationship between the images. Every time a user selects two or more candidate images a weighted link between these images is updated to build a network of image relationships. (c) This network of semantic image relationships can be navigated visually.*

Keywords: CBIR, collaborative techniques, semantic modeling

1 Introduction

Internet image search systems mostly use words from the context of the web page containing the image as keywords. The performance of these search systems is rather poor, as the search systems neither know the intention of the searching user nor the semantic relationships of these images. Content-based image retrieval (CBIR) systems rely on the assumption that similar images share similar visual features. Despite intense research efforts, the results of CBIR systems have not reached the performance of text based search engines. The main problem of CBIR systems is the semantic gap between the content that can be described with low-level visual features and the description of image content that humans use with high-level semantic concepts. Some image retrieval systems have combined the keyword and the content-based visual search approach. However with this approach many images may be found that semantically do not match. In addition semantically similar images that visually look different cannot be found at all.

Collaborative filtering uses the known preferences of a group of users to make recommendations of the unknown preferences for other users. Recommender systems using these methods have been used very successfully to suggest "similar" music, books or products. Up to now collaborative filtering has not been used for image retrieval. On the one hand this is due to the fact that usually the number of images to be searched is extremely high. On the other hand existing image search systems do not allow to learn relations between different images from the users' interaction.

In this work we propose a new image search system using keywords, low-level visual features, and collaborative filtering techniques to generate a network of semantic inter-image relationships. Unlike other approaches our new system does not try to learn the

degree of confidence between images and associated keywords. We rather propose to model the degree of similarity between images by building up a network of linked images. The weights of the inter-image links are learned from the users' interaction with the search system only.

2 Our Approach

Conventional image search systems typically display 20 - 50 result images at a time. In a first step our image search system retrieves a much larger set of result images. This set will be displayed visually sorted, allowing to inspect up to 1000 images simultaneously. Due to this large number of images it is most likely that the user will find and select several candidate images that are close to his desired search result. In a next step this candidate set is used to retrieve further visually similar images. But even more important is the fact that these candidate images that have been selected together can be used to model the inter-image relationships. Each time two or more images have been selected together their link weights will be increased. By collecting candidate sets from many searches we successively build a semantic network of weighted links of image relationships.

This network of semantic image relationships can be used to improve the image search results by retrieving images that are linked to the query images. Opposed to existing systems our new image search system can retrieve semantically similar images that do not share the same keyword or the same visual appearance. In addition the search results will get better every time a new search is performed.

For images that are linked to a sufficiently large number of other images a visual navigation mode may be enabled. In this mode all connected images are displayed using a force-based graph layout scheme. This visual navigation mode enables the user to navigate through large image collections in a totally new way just by following the semantic links - independent of any language or keywords.

---

*e-mail: barthel@htw-berlin.de