iFind—A System for Semantics and Feature Based Image Retrieval over Internet

Hongjiang Zhang, Liu Wenyin, Chunhui Hu
Microsoft Research China
49 Zhichun Road
Beijing 100080, China
{hjzhang, wyliu, i-chhu }@microsoft.com

iFind© (v1.0) is a web based image retrieval system developed at Microsoft Research China. It provides the functionalities of keyword based image search, query by image example, category based image browsing, relevance feedback, and semi-automatic image annotation. The key technology in the system is the integrated semantics and feature based image retrieval and relevance feedback approach, which will be presented in our paper in the ACM Multimedia 2000 Proceedings [1]. When the user provides feedback images, the system can refine the retrieval result based on the user's feedback. In the meantime, the system updates the annotation of feedback images by increasing the linkage to the positive examples' annotation and decreasing the linkage to the negative examples' annotation. The updated annotation can further help to improve image retrieval results of the system in later use.

Traditional image retrieval methods are either based on low-level visual features only, which fail to address the images' semantic content, or based on manually annotated keywords only, which fail to utilize the images' visual properties. By forming a semantics network on top of the keyword association with images in the database, we are able to accurately deduce and utilize images' semantic content for retrieval purposes. Hence, it can further improve the feature-based image retrieval method. In our demo, we show that the combined semantics- and feature-based image retrieval and relevance feedback approach outperform both the keyword-based and the feature-based image retrieval methods. The accuracy and effectiveness of our method is demonstrated with experimental results on real-world image collections.

The main user interface is shown in Figure 1 and typical scenarios are as follows.

If a user wants to use keyword-based search for, for instance, pictures of girls, he/she can type "girl" in the text search box on the up-left corner of the window and hit the return bar or click the button "GO!". The user will see some images displayed in several pages in the browsing page. The user can use "<<prev" and "next>>" to switch among these pages. Occasionally, the user may see many irrelevant images. In this case, the user can confirm some of those relevant images by clicking on the "√" symbol and those irrelevant images by clicking on the "×" symbol below each image, and then click on the button "Feedback" on the bottom of the page. The user is promised to get much better result after a few iterations of such feedbacks.

In the browsing page, the user can use the hyperlink "View" below each image to see the image in full size. The user can also use the hyperlink "Similar" below each image to find those images that are visually similar to this one. The user can also click on the hyperlink "Query by Example" and follow the instructions to submit the query example image from the local machine to find those images that are visually similar to the query image.

The user may want to find the intended image by browsing in a category or browsing randomly all images in the database. In these cases, he can do so by selecting the "Browse by Categories" or "Browse Randomly" hyperlinks, respectively.

The middle two hyperlinks on the menu are for system test. The system can improve its intelligence (the semantics network of the database) on image retrieval in daily practice. If the user wants to save what the system has recently learnt, the user can click the "save" hyperlink. Otherwise, the user may click "load" to discard it.

We have performed a set of solid experimental evaluations of *iFind* performance in term of retrieval accuracy and learning speed. The ground truth dataset is composed of 122 categories, each consisting of 100 images. Therefore, there are totally 12200 images in the database, most from Corel image databases. The experiments show that our system not only outperforms both the keyword-based and the feature-based image retrieval methods in retrieval accuracy, but is also a fast learner: the combination of semantic annotation and feature-based retrieval makes the system learn and memorize from user's feedback and continuously improve the retrieval performance. Figure 2 shows the retrieval performance and the learning curve of the *iFind* system.

The system is implemented on Microsoft Internet Information Server (IIS). The retrieval algorithm is implemented in an ISAPI extension DLL and the user interface is written in DHTML. We use Microsoft Access as our database scheme and ODBC as the database interface.

REFERENCES

[1] Lu Y et al. (2000) A Unified Framework for Semantics and Feature Based Relevance Feedback in Image Retrieval Systems. *To appear in Proc. ACM MM2000*.

Figure 1. A typical interface page of the *iFind* system.

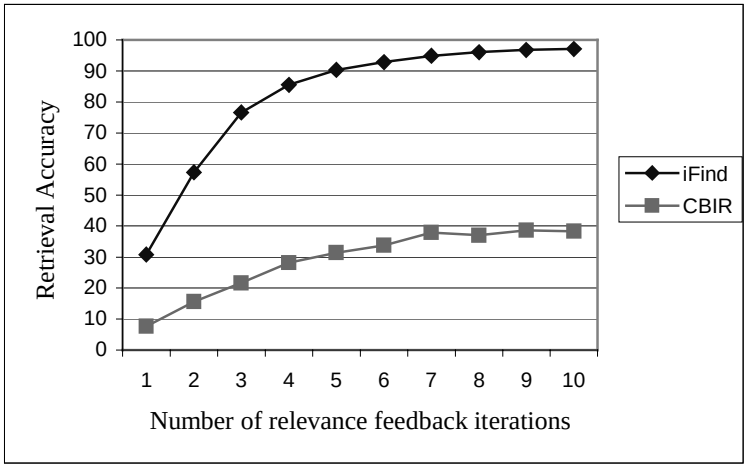

Figure 2: Performance comparison and learning curve.