Attribute Feedback

Hanwang Zhang, Zheng-Jun Zha, Jingwen Bian, Yue Gao, Huanbo Luan and Tat-Seng Chua
School of Computing, National University of Singapore
Singapore
{hanwang, zhazj, bian_jingwen, gaoy, luanhb, chuats}@comp.nus.edu.sg

ABSTRACT

This demonstration presents a new interactive Content Based Image Retrieval (CBIR) system, termed Attribute Feedback (AF). Unlike traditional relevance feedback purely founded on low-level features, AF system shapes user's search intents more precisely and quickly by collecting feedbacks on intermediate-level semantic attribute. At each interaction iteration, the AF system first determines the most informative binary attributes for feedbacks and then augments the binary attribute feedbacks by a new type of attributes, "affinity attributes", each of which is learnt offline to describe the distance/similarity between user's envisioned image(s) and a retrieved image with respect to the corresponding affinity attribute. Based on the feedbacks on binary and affinity attributes, the images in corpus are further re-ranked towards better fitting user's search intents. The experimental results on two real-world image datasets have demonstrated the superiority of the AF system over other state-of-the-art relevance feedback based CBIR approaches.

Categories and Subject Descriptors

H.3.3 [**Information Search and Retrieval**]: Retrieval Model

General Terms

Algorithms, Experimentation, Performance

Keywords

Image Search, Attribute Feedback, Relevance Feedback

1. INTRODUCTION

Interactive Content Based Image Retrieval (CBIR) has attracted significant attentions in both academia and industry recently. The most popular interactive search scheme is Relevance Feedback (RF), which collects users' feedbacks on the search results (i.e., images) and use the feedbacks to refine search system. Typically, users are asked to label the top results as "relevant" or "irrelevant". Through iterative user feedback and system refinement, RF attempts to capture users' search intents and improve search results.

Although RF has shown encouraging potential in CBIR, it usually suffers from the following problems. First, traditional RF relies on the search system to infer users' search intents from the low-level features of the "relevant" and/or "irrelevant" images, and hence is usually ineffective in narrowing down the search to users' target. Second, the search results in the first few rounds of feedbacks are usually unsatisfactory. The top results usually contain rare or even no relevant image. With no/rare relevant images, most RF approaches become ineffective or even no longer applicable. In this demonstration, we presents a novel interactive CBIR system, named Attribute Feedback (AF). Here, attributes could be shape, color, texture, material, or parts of objects, such as "*round*", "*red*", "*metal*", and "*wheel*" etc. [1]. The AF system allows users to provide binary feedbacks on attributes to state which attributes are in their search intent or not. These binary feedbacks compose a semantic description of users' search intent, such as "*has ear and leg, has no hand*". However, the binary feedbacks on some attributes, which are shared by many categories, are not discriminative enough. For example, the feedback "*has ear*" is not discrim-

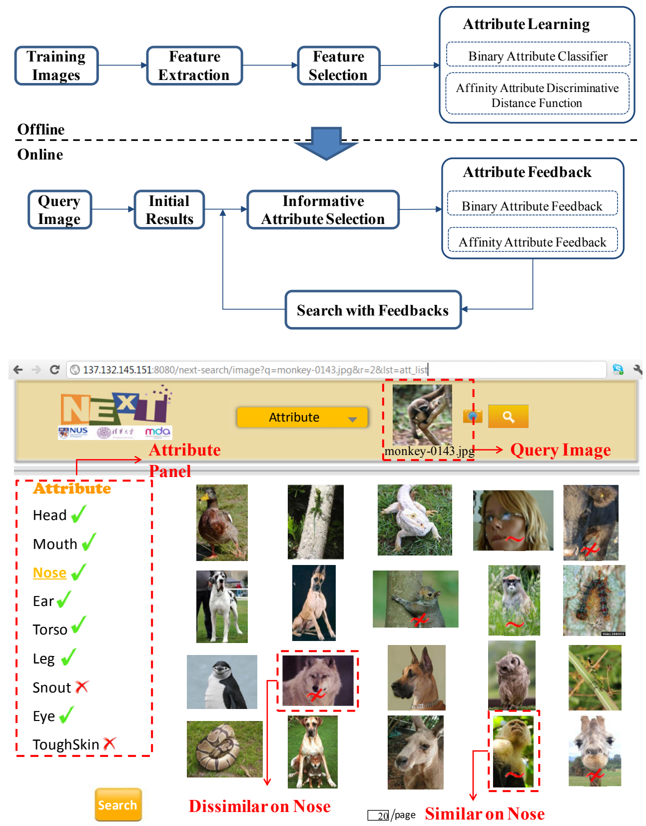


Figure 1: The flowchart of the proposed AF (above) and the user interface (below).

inative enough to distinguish users' target (e.g. *cat*) from other animals, since "*ear*" is shared by many animal categories. Hence, we propose a new type of attributes, termed "affinity attributes", which refer to the attributes that are shared by many categories and have large variance in appearances. For each affinity attribute of interest, the AF system allows users to give affinity judgments on the images containing this attribute to indicate which images are (dis-)similar to their envisioned target images with respect to this attribute. Based on users' binary and affinity attribute feedbacks, the AF system can effectively narrow down the search to users' target with less interaction effort. Moreover, even the top search results contain no relevant images, some of them might be partially relevant to users' envisioned images with respect to certain attribute(s). By accumulating feedbacks on such attribute(s), the AF system can push the search close to users' target gradually and thus can overcome the no/rare relevance sample problem.

2. SYSTEM OVERVIEW

The flowchart and the user interface of the AF system is illustrated in Fig. 1. In the offline part, we learn a set of binary classifiers, each of which predicts the presence of an attribute in images. We also learn a set of discriminative distance functions for the affinity attributes. Each function computes the distance between two images with respect to a particular affinity attribute. In the online part, the AF system interactively collects users' binary and affinity feedbacks, and refines the search results. At each iteration, a set of informative attributes are selected for feedbacks. Users are allowed to give positive/negative feedbacks on the presented attributes. For each affinity attribute of interest, users can further provide affinity feedbacks on the images containing this attribute to indicate whether each of them is similar to the target with respect to the attribute. Based on these binary and affinity feedbacks, a search model is then executed to refine the search results.

3. SYSTEM IMPLEMENTATION

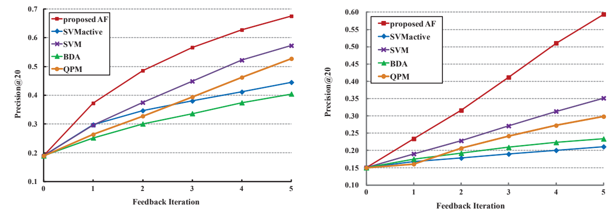
The AF system consists of three components: attribute learning, informative attribute selection and search with feedbacks.

3.1 Attribute Learning

We extract four types of visual features, including color, texture, shape, and scale-invariant feature transform (SIFT) [1]. All these features are represented in a bag of words style. Feature selection is then performed to select discriminative features for each attribute by ℓ_1-norm logistic regression. Finally, we learn a set of SVM classifiers for the binary attributes. For each affinity attribute a_i, we learn a distance function $D_{a_i}(x_j, x_k)$ to compute the distance between two images x_j and x_k with respect to a_i. $D_{a_i}(x_j, x_k)$ is defined as the Mahalanobis distance $\sqrt{(x_j - x_k)^T \mathbf{M}(x_j - x_k)}$.

3.2 Informative Attribute Selection

At each online iteration, a list of most informative attributes are selected for feedbacks with an expectation that users' feedbacks on these attributes can improve the search performance effectively. An attribute is considered to be informative if it frequently (rarely) appears in current search results but is unlikely (likely) to be user interest. Here, we argue that the informativeness of an attribute is correlated to the search results at current iteration as well as users' feedbacks at previous iterations. We thus develop and maintain a set of Bayesian models, each of which infers the probability of an attribute being informative given the current search results and previous feedbacks. The attributes with high posterior probabilities are then selected and presented for user feedback.



(a) Pascal-Yahoo! (320 queries) (b) Web Image (700 queries)

Figure 2: Retrieval Performance (top 20) of the AF and four baseline methods at five feedback iterations on (a) Pascal-Yahoo! and (b) Web Image.

3.3 Search with Feedbacks

Based on users' binary and affinity feedbacks, the AF system updates the search results as follows. The relevance of each image in database is first updated based on binary feedbacks as an aggregation of its probability of containing the positive attributes and that of not containing the negative attribute. The affinity feedbacks are then incorporated to refine the relevance. Considering an affinity attribute, the relevance of an image is reinforced by its similarity to the images with "similar" affinity feedback on the affinity attribute, and also penalized by its similarity to the images with "dissimilar" feedback. Finally, the AF system presents the images with a descending order to their relevance scores.

3.4 Experimental Results

We evaluated the proposed AF system on two real-world image corpus: (a) the Pascal-Yahoo! dataset [1] with 15,339 images and 64 attributes; and (b) an image corpus collected from the Web containing 76,303 images and 67 attributes. We compared the AF system to four state-of-the-art relevance feedback approaches [3][2][5][4]. For the sake of fair comparison, all the approaches were constrained to collect 20 feedbacks on the top 20 search results at each iteration. As shown in Fig. 2, the proposed AF system outperforms the state-of-the-art methods significantly.

4. CONCLUSION

We have demonstrated a new interactive content-based image retrieval system, termed Attribute Feedback (AF). This system allows users to provide feedbacks on semantic attributes. By collecting attribute feedbacks, the system can shape users' search intents precisely and quickly. It thus can produce search results meeting users' intent with less interaction effort.

5. REFERENCES

[1] A. Farhadi and et al. Describing objects by their attributes. In *CVPR*, 2009.

[2] G. D. Guo and et al. Learning similarity measure for natural image retrieval with relevance feedback. *TNN*, 2002.

[3] Y. Rui and et al. Content-based image retrieval with relevance feedback in mars. In *ICIP*, 1997.

[4] S. Tong and E. Chang. Support vector machine active learning for image retrieval. In *ACM MM*, 2001.

[5] X. S. Zhou and T. S. Huang. Small sample learning during multimedia retrieval using biasmap. In *CVPR*, 2001.