IMAGE RETRIEVAL WITH FEATURE SELECTION AND RELEVANCE FEEDBACK

Yu Sun, Bir Bhanu
Center for Research in Intelligent Systems, University of California, Riverside, CA, 92521, USA.
ysun@vislab.ucr.edu, bhanu@cris.ucr.edu

ABSTRACT

This paper proposes a new content based image retrieval (CBIR) system combined with relevance feedback and the online feature selection procedures. A measure of inconsistency from relevance feedback is explicitly used as a new semantic criterion to guide the feature selection. By integrating the user feedback information, the feature selection is able to bridge the gap between low-level visual features and high-level semantic information, leading to the improved image retrieval accuracy. Experimental results show that the proposed method obtains higher retrieval accuracy than a commonly used approach.

Index Terms — CBIR, Relevance Feedback, Feature Selection

1. INTRODUCTION

Content-based image retrieval (CBIR) [1] has been a significant topic of research in the last decade. In the CBIR context, an image is represented by a set of low-level visual features, which are generally not effective and efficient in representing the image contents, and they also have no direct correlation with high-level semantic information. The gap between high-level information and low-level features is the fundamental difficulty that hinders the improvement of the image retrieval accuracy. Recently, a variety of solutions have been suggested that aim to bridge this semantic gap. Two of the most commonly used methods are online feature selection and user relevance feedback.

The feature selection [2] basically narrows the semantic gap by selecting the feature subset that best represents the query and discards redundant features. Image retrieval uses the selected feature subset to search the image database such that the retrieved images will be closer to a given query.

The relevance feedback [3] narrows the semantic gap by making use of user provided judgments which are the labels (relevant or non-relevant) on the retrieved images for a query. The retrieval performance improves as the user provides more and more feedback information to the CBIR system. Query vector modification (QVM) [4] and feature relevance learning [5] are the two widely used methods to integrate user feedback information into the CBIR system.

Currently, the feature selection and relevance feedback are rarely used together to further narrow the semantic gap. The work in [8] applies feature selection as a form of feature weighting into the query vector modification (QVM) method for relevance feedback. However, it ignores the important classification or mutual information evaluation for feature selection. As a result, the work in [8] does not fully capture the key characteristics required for feature selection. In this paper, a measure of inconsistency from relevance feedback is integrated into feature selection, and combined with the Bayesian classifier to improve CBIR performance.

The feature selection procedure is composed of two steps: searching the combination of feature subsets within a feature space using specified search strategy, and evaluating the performance of the selected subset by a criterion. Existing evaluation criteria are classification performance, mutual information and entropy. In this paper, a new term called the measure of inconsistency from relevance feedback, is combined with the Bayesian classifier to build the overall criterion for feature selection. The combined criterion is able to select the optimal feature subset which leads to improve the image retrieval accuracy and better satisfies the user semantic requirements.

This paper makes the following contributions:

(1) A new term called the measure of inconsistency, from relevance feedback, is combined into feature selection as a new criterion to further improve the image retrieval.

(2) The semantic gap is further narrowed by combining the online feature selection and the user relevance feedback.

The outline of the paper is as follows. Section 2 describes the technical approach of the new CBIR system in detail. Section 3 provides experimental results and analysis. Finally, conclusions are given in Section 4.

2. THECHNICAL APPROACH

The proposed CBIR system that integrates both online feature selection and the user relevance feedback is shown in Figure 1. For a given query, the original features (color, texture and shape) are extracted from the query image, and the K-nearest neighbor (K-NN) algorithm with Euclidean metric searches the image database, and retrieves N top ranked images having features most closed to the query. The session with this query is terminated when the user is satisfied with the retrievals, otherwise, the user provides relevance feedback by labeling the retrievals as relevant (positive feedback) and non-relevant (negative feedback). A measure of inconsistency is computed based on the user feedback and it is given as the input to the feature selection to select the feature subsets which will guide the K-NN search to obtain higher retrieval accuracy in the next CBIR iteration.

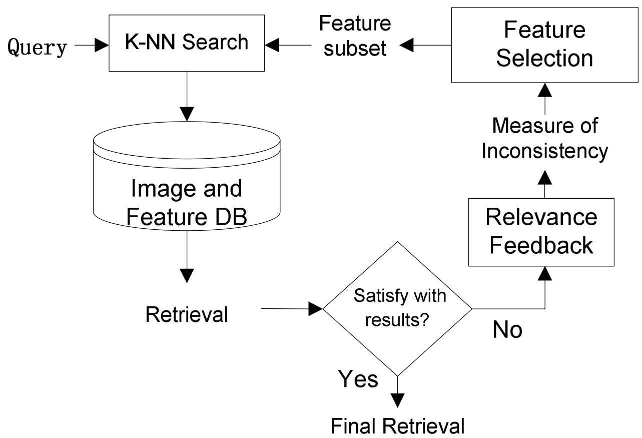

Fig. 1. The overall CBIR system diagram.

2.1. Measure of Inconsistency from Relevance Feedback

For each CBIR iteration, let $\chi = \{x_1, ..., x_N\}$ denotes the N retrieved images. The property $f(x_i)$ of the retrieved image x_i is expressed by its visual feature vector and its relevance feedback label: $f(x_i) = \{f_1(x_i), ..., f_M(x_i), l_i\}$, where $i \subset \{1, ..., N\}$ denotes the ith retrieved image, and M is the dimension of feature, either be full dimension or dimension of the selected subset. The $l_i \subset \{0,1\}$ represents the feedback label of the retrieved image *i*, either be 1 for positive feedback or be 0 for negative feedback. The retrieved images χ are grouped into two clusters, as relevant (positive) and non-relevant (negative), according to user's feedback labels. The mean feature value is then computed for each of the two clusters, as given below.

Mean positive feature vector:

$$Mf_P = \frac{1}{N_P} \sum_{l_k=1} \{f_1(x_k), ..., f_M(x_k)\} \tag{1}$$

Mean negative feature vector:

$$Mf_N = \frac{1}{N_N} \sum_{l_k=0} \{f_1(x_k), ..., f_M(x_k)\} \tag{2}$$

where N_P and N_N is the number of positive and negative feedback images, respectively, satisfying $N_P + N_N = N$. The measure of inconsistency θ_{RF} is computed by the two mean vectors as shown below,

$$\theta_{RF} = arccos\left(\frac{Mf_P \bullet Mf_N}{\sqrt{\|Mf_P\|_2^2 \cdot \|Mf_P\|_2^2}}\right) / 2\pi \tag{3}$$

which is the angle between the mean positive and negative feature vectors, and it is normalized into $\{0,1\}$. $Mf_P \bullet Mf_N$ is the dot product of the mean relevance and non-relevance vectors, and $\|\cdot\|_2^2$ is 2nd order norm operator. The larger the measure of inconsistency the better it is since we need the two mean positive and negative feature vectors to be as separated as possible. The feedback angle is further used as an evaluation criterion to guide the feature subset selection.

2.2. Feature Selection Combined with User Feedback

The feature selection block in Fig 1 starts with the original image features and outputs the optimal feature subset. The realization of the feature selection block is indicated in Fig 2. The block (1) in Fig 2 refers to the feature space search strategy, namely the sequential forward selection [6], in which features are sequentially selected from original features to build the best feature subset, and which dimension of the feature should be selected out into the subset, is uniquely decided by the feature performance evaluation criterion. The evaluation criterion, called the wrapper evaluation, is the most important element in feature selection system. In this paper, the new evaluation criterion is the combination of Bayesian classifier and the measure of inconsistency. The feature dimension having the highest classification results and measure of inconsistency, is selected out and added into the current subset to build the new selected subset. For more detailed realization of this feature selection strategy, please refer to [7]. The wrapper evaluation criterion in this paper is shown below,

$$C_{wrapper} = \partial \cdot C_{Bayesian} + \beta \cdot \theta_{RF} \tag{4}$$

where $C_{Bayesian}$ is the classification result of the Bayesian classifier, and traditionally $C_{wrapper} = C_{Bayesian}$. And θ_{RF} is measure of inconsistency introduced in Section 2.1. The weights are set to $\alpha = \beta = 0.5$. The equation (4) is the improvement of the tradition wrapper evaluation criterion by integrating the user feedback evaluation θ_{RF} . As a result, the best feature subset will be selected by criterion in equation (4), and the selected subset will have both highest classification and feedback inconsistency. Since user feedback is integrated into feature selection, user will provide higher percentage of positive feedback in next feedback iteration, based on images retrieved using the best feature subset selected by equation (4). Measure of inconsistency of a candidate feature subset, with dimension M, is computed by equation (1)-(3). It is worth to note that the positive and negative feedback images are accumulated from all iterations to compute the measure of inconsistency.

The Bayesian classifier is extensively used in the wrapper evaluation criterion. The classifier estimates the label of an image by processing its feature vector, using the maximum a posteriori (MAP) probabilistic approach. From the comparison of the Bayesian classification results with the actual image class labels, an estimate of the correct classification rate (CCR) [7] is obtained as the feature subset evaluation $C_{Bayesian}$. After the search of block (1) in Fig 2, all feature dimensions are ranked according to the results of the evaluation criterion of equation (4). The feature ranking as well as the related performance evaluation are put into block (2) in Fig 2 to select the subset with the highest evaluation, as the final selected subset.

As shown in Fig 1, the feature selection provides more effective feature subset, which is input to the K-NN search for the next retrieval iteration. With the improved feature subset selected by the measure of inconsistency, the K-NN search ranks the images in database that better represents the user feedback information with higher retrieval accuracy.

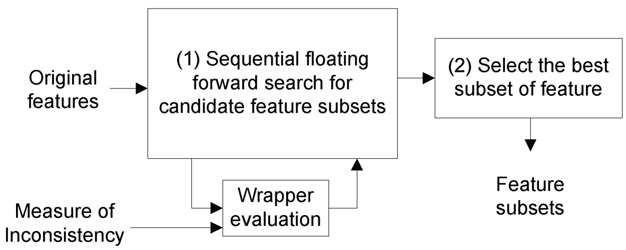

Fig 2. The feature selection diagram with user feedback.

3. EXPERIMENTAL RESULTS

3.1. Datasets

In experiments, we run the CBIR system in two image databases, the first of which is the butterfly image database (http://janzen.sas.upenn.edu/) containing 29 classes and summing to 7600 images, and example images are shown in Fig 3(a). The second database has 210 natural images collected from Google Images, with 5 classes related to semantic concepts as snowy mountains, trees, falls, bridges and sand beaches, respectively, and example images are shown in Fig 3(b). As in Fig 3, the two databases are labeled DB#1 and DB#2, respectively.

We use features covering wide range of image properties. Totally 27 feature dimensions are extracted from the entire image, composed of following 4 sets of feature properties: *(1)* mean and standard deviation of the RGB components of color space, totally 6 dimensions. *(2)* The HSV components of color space, with the same distribution as RGB, so it also has 6 dimensions. *(3)* 8-dimension texture feature derived from the mean and standard deviation of the filtered image by Gabor filters at 4 orientations in steps of 45 degrees. *(4)* The 7-dimension shape feature derived from first 7 central geometric moments of the image. Totally 27 dimensions of features are extracted to build the image and feature database as shown in Fig 1. And they are also used as the original feature sets for feature subset selection. In the experiments, totally 20 images are retrieved in each iteration.

3.2. CBIR Combined with Feature Selection and Relevance Feedback

Fig 3 (a) and (b) provides example images. The two image databases indicate large overlapped properties which hinder the retrieval accuracy. For instance, in Fig 3(a), the images in class #13 & #25 are visually similar. Furthermore, in Fig 3(b), objects of 'snowy mountain' and 'falls' share the similar dominant properties. On the other hand, both class #9 in Fig 3(a) and 'bridge' in Fig 3(b) show intra-class variations. The two databases are challenging cases that have overlapped properties among different classes, as well as high variations within class. By using our method, feature selection will search feature subset that best discriminate among classes and discard overlapped feature properties.

Comparison of feedback precisions: The feedback precision is defined as the percentage of positive feedback in each feedback iteration. In Fig 3(c)&(d), the proposed method is compared to the query vector modification (QVM) scheme, in terms of feedback precision. QVM [4, 8] is one of the most widely used relevance feedback techniques. It modifies the query feature so that it will move closer to the relevant feature points and move away from non-relevant points. The final precision at each feedback iteration is computed by averaging precisions from different query sessions. Our method outperforms QVM in every feedback iteration as in Fig 3(c)&(d). It needs to say that precisions after the 7^{th} iteration are stabilized (almost unchanged).

Comparison of retrieval precisions: The retrieval precision is defined as the percentage of the *true* retrieval in final retrieval results. For systems including the relevance feedback, this precision is the final (optimal) retrieval precision after the last feedback iteration. In table 1, the proposed method is compared to three different methods in terms of the retrieval precision. In order to ensure the comparability, all methods use K-NN search for retrieval and use the same feature data introduced in section 3.1. Additionally, for all the methods, the same queries are tested with the same number of repetitions, and the average precision is computed. Moreover, the *RF_ONLY* method (QVM) in table 1 is the same method used in Fig 3(c) & (d).

Table 1. Comparison of retrieval precisions.

	DB#1	DB#2	Feature Selection	Relevance Feedback
RF_FS: **This paper**	**80.7%**	**71.2%**	YES	YES
RF_ONLY: QVM [8]	78.6%	69.0%	×	YES
FS_ONLY: [7]	76.1%	68.3%	YES	×
RT_ONLY: traditional retrieval	74.9%	66.5%	×	×

In table 1, the *RF_ONLY* method applies QVM in each feedback iteration, but uses the original features without feature selection. The *FS_ONLY* method uses the feature selection [7], which is the same selection scheme as used in our system, to select the best feature subset and input this subset into K-NN search for the retrieval, but no relevance feedback is performed. The *RT_ONLY* method only runs the K-NN search for the retrieval by original feature data, without feature selection and relevance feedback. It is obvious that our method reaches higher precision, than those using only one of the relevance feedback and feature selection schemes.

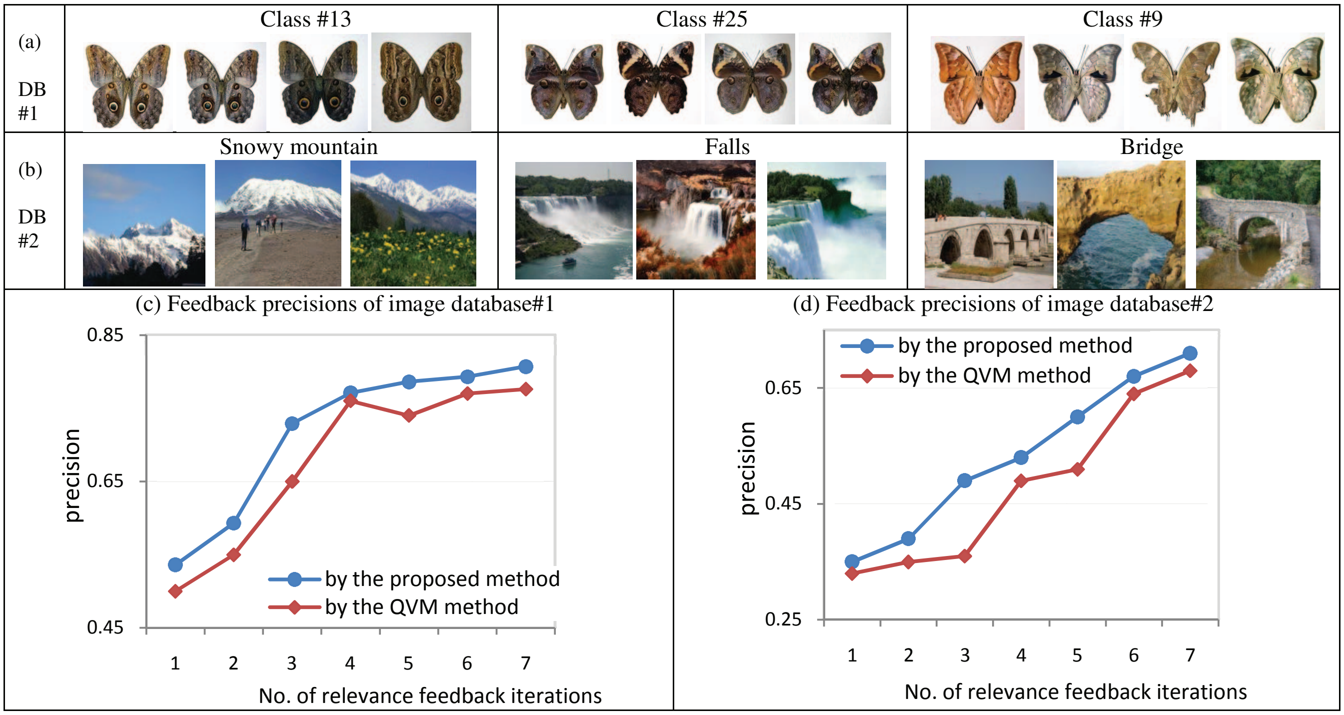

Fig 3. Example images in databases, and performance comparison of the proposed method.

4. CONCLUSIONS

In this paper, we presented a new approach that combines relevance feedback and feature selection to improve the performance of a CBIR system. The approach uses a new criterion called the measure of inconsistency to guide the feature selection, in order to improve the image retrieval by integrating the user relevance feedback information. We performed experiments on different sizes of image databases to indicate the benefits of the proposed method. We showed the improvements in both feedback and retrieval precisions over the other current methods.

Acknowledgement: This research was supported by NSF grant 0641076.

5. REFERENCES

[1] R. Rahmani, S. A. Goldman, H. Zhang, S. R. Cholleti, and J. E. Fritts, "Localized content-based image retrieval," *IEEE PAMI*, vol. 30, no. 11, pp. 1902–1912, Nov. 2008.

[2] H. Liu, and L. Yu, "Toward integrating feature selection algorithms for classification and clustering," *IEEE PAMI*, vol. 17, no. 4, pp. 491-502, April 2005.

[3] M. R. Azimi-Sadjadi, J. Salazar, and S. Srinivasan, "An adaptable image retrieval system with relevance feedback using kernel machines and selective sampling," *IEEE PAMI*, vol. 18, no. 7, pp. 1645-1659, July 2009.

[4] M. R. Widyanto, and T. Maftukhah, "Fuzzy relevance feedback in image retrieval for color feature using query vector modification method," *Journal of Advanced Computational Intelligence and Intelligent Informatics*, Vol.14, No.1, pp. 34-38, Jan. 2010.

[5] B. Ko, J. Peng, and H. Byun, "Region-based image retrieval using probabilistic feature relevance learning," *Journal of Pattern Analysis & Applications*, Vol.4, No.2, pp. 174-184, June 2001.

[6] B. Ko, J. Peng, and H. Byun, "Sequential forward selection approach to the non-unique oligonucleotide probe selection problem," *Proceedings of the Third IAPR International Conference on Pattern Recognition in Bioinformatics*, 2008.

[7] D. Ververidis, and C. Kotropoulos, "Fast and accurate sequential floating forward feature selection with the bayes classifier applied to speech emotion recognition," *Elsevier Signal Processing*, Vol. 88, No. 12, 2008.

[8]C. Tusk, K. Koperski, S. Aksoy, and G. Marchisio, "Automated feature selection through relevance feedback," *Geoscience and Remote Sensing Symposium*, 2003.