CORRELATED PROBABILISTIC LABEL PROPAGATION FOR REGION-BASED IMAGE RETRIEVAL

Fei Li, Qionghai Dai, Wenli Xu, Guihua Er

Department of Automation, Tsinghua University, Beijing 100084, China

ABSTRACT

Label propagation and manifold ranking have been successfully adopted in content-based image retrieval (CBIR) in recent years. However, while the global low-level features are widely utilized in current systems, region-based features have received little attention. In this paper, a novel transductive framework based on correlated probabilistic label propagation is proposed for region-based images retrieval (RBIR), which can be characterized by three key properties: (1) Unified feature matching (UFM) is chosen to measure the similarity between two segmented images. (2) To represent the segmented images in a uniform feature space, a generative model is adopted and the probabilistic labels of each image can be obtained. (3) In the retrieval process, multiple probabilistic labels of training samples are propagated simultaneously on the weighted graph, and the correlation among different labels are explored. Experimental results on 10000 images show that our algorithm can greatly improve the retrieval performance of the RBIR system.

Index Terms— Image databases, region-based image retrieval, manifold ranking, relevance feedback

1. INTRODUCTION

With the rapid increase of the volume of digital image collections, content-based image retrieval (CBIR) has been an active research area in recent years. At the beginning, much work has been focused on looking for effective low-level representation of images, and varieties of similarity measures have been proposed. But the retrieval performance is far from satisfactory because of the gap between high-level semantic concepts and low-level visual features [1]. To narrow down the gap, relevance feedback [2] has been introduced to involve the user in the retrieval process and its significant effectiveness has been proved by lots of researchers.

Recently, many machine learning methods have been applied to relevance feedback. The supervised methods consider CBIR as a classification problem. Their goal is to train a classifier with the labeled samples, which can ultimately separate all the images into two classes: one is relevant to the query while the other not. Lots of work has been dedicated to the construction of effective classifier, however, due to the lack of training samples, the performance of the obtained classifier may be unstable. Therefore, semi-supervised learning, which does not only utilize the labeled examples, but also the unlabeled ones, has attracted more and more attention. Manifold ranking algorithm [3], which can well explore the relationship among all the samples, has been successfully integrated into CBIR [4]. Although global features are successfully adopted for retrieval [4], it is believed that the performance will be highly improved if we introduce manifold ranking algorithm to the context of region-based image retrieval (RBIR), for region-based low-level features accord with human perception better.

In this paper, we present a novel transductive framework based on correlated probabilistic label propagation for RBIR. In order to exploit the information in region-based features, on the one hand, unified feature matching (UFM) [5] is adopted to measure the similarity between two segmented images and the weighted graph is constructed for label propagation; on the other hand, based on a generative model, the segmented images are represented in a uniform feature space and the probabilistic labels of each image can be calculated. Considering the correlation among labels, the algorithm of correlated label propagation [6] is introduced to overcome the disadvantage of propagating multiple labels independently on the weighted graph. Then we calculate the similarity between the images in the database and the user's query, according to which the retrieval results and the label set for relevance feedback can be obtained.

The rest of the paper is organized as follows: Section 2 gives a brief description of manifold-ranking based image retrieval. Then in Section 3, we describe our algorithm in detail. Our experimental results are presented in Section 4, which is followed by some conclusions in Section 5.

2. MANIFOLD-RANKING BASED IMAGE RETRIEVAL

In [4], a learning framework named manifold-ranking based image retrieval (MRBIR) is proposed, and the whole process can be summarized as follows.

Let $X = \{\mathbf{x}_1, \mathbf{x}_2, \cdots, \mathbf{x}_n\}$ be the set of images in the database, f be a ranking function which assigns $\mathbf{x}_i$ a ranking score f_i, $y = [y_1, y_2, \cdots, y_n]^T$ be a binary vector where y_i

indicates whether $\mathbf{x}_i$ is a query image or not, $d : X \times X \rightarrow R$ be the distance measure between two images.

In the framework of MRBIR, the K nearest neighbors of each image are found out at first. By connecting the two neighboring images $\mathbf{x}_i$ and $\mathbf{x}_j$ with a edge weighted by $W_{ij} = \exp[-d^2(\mathbf{x}_i, \mathbf{x}_j)/2\sigma^2]$, a graph, which takes each image as a vertex, is formed to explore the relationship of all the images in the feature space. Note that W_{ii} is set to be 0 to avoid self-reinforcement.

To achieve convergence, W is symmetrically normalized by $S = D^{-1/2}WD^{-1/2}$, where D is the diagonal matrix with (i, i)-element equal to the sum of the i-th row of W.

After constructing the weighted graph, all the images in the database spread their ranking scores to their neighbors via the graph. This is done by iterating $f(t + 1) = \alpha Sf(t) + (1 - \alpha)y$. The spread process does not stop until it reaches a global stable state f^*, then the images with largest ranking scores are selected as the retrieval results.

In the process of relevance feedback, considering the asymmetry between relevant and irrelevant images, they should be treated differently. Three different schemes are proposed in MRBIR, and the scheme which simply reduces the contribution of the negative ranking scores performs slightly better.

Furthermore, three active learning methods are incorporated into MRBIR with difference principles when generating label set in each round of relevance feedback. More details can be found in [4].

3. CORRELATED PROBABILISTIC LABEL PROPAGATION IN RBIR

3.1. Image segmentation and representation

To segment an image, the system first partitions the image into non-overlapping blocks, and low-level features, such as color and texture, are extracted in each block. In order to make a tradeoff between the effectiveness of features and the computational complexity, the block size is set to be 16×16 pixels. In image segmentation, JSEG algorithm [7] is adopted for its flexibility of adjusting the number of regions. After segmentation, an image $\mathbf{x}_k$ can be represented by a set of regions $\{R_1(\mathbf{x}_k), R_2(\mathbf{x}_k), \cdots, R_{N_k}(\mathbf{x}_k)\}$. The feature of the region $R_i(\mathbf{x}_k)$ is calculated as the mean feature vector of all the block-based features in it. As in [8], each region $R_i(\mathbf{x}_k)$ corresponds to a saliency membership $v[R_i(\mathbf{x}_k)]$, which is directly proportional to the area of the region $R_i(\mathbf{x}_k)$, and inversely proportional to the average distance between each member block of $R_i(\mathbf{x}_k)$ and the image center. The saliency membership of all the regions of image $\mathbf{x}_k$ is normalized so that $\sum_{i=1}^{N_k} v[R_i(\mathbf{x}_k)] = 1$.

3.2. Unified feature matching

In order to measure the similarity between two segmented images, a novel fuzzy logic approach named unified feature matching (UFM) is proposed [5]. In UFM, each region of the segmented image $R_i(\mathbf{x}_k)$ is characterized by a fuzzy set $\tilde{\mathcal{R}}_i(\mathbf{x}_k)$. Under the assumption that the fuzzy membership function is Cauchy function, the fuzzy similarity measure for two fuzzy sets $\mathcal{S}\left[\tilde{\mathcal{R}}_i(\mathbf{x}_m), \tilde{\mathcal{R}}_j(\mathbf{x}_n)\right]$ can be calculate efficiently. Then we can calculate the similarity between two images $\mathbf{x}_m$ and $\mathbf{x}_n$ as

$$\mathcal{S}(\mathbf{x}_m,\mathbf{x}_n) = \frac{1}{2}\sum_{i=1}^{N_m} v[R_i(\mathbf{x}_m)]\; l_i^{\mathbf{x}_n} + \frac{1}{2}\sum_{j=1}^{N_n} v[R_j(\mathbf{x}_n)]\; l_j^{\mathbf{x}_m} \quad (1)$$

where

$$l_i^{\mathbf{x}_n} = \max_{1\le j\le N_n} \mathcal{S}\left[\tilde{\mathcal{R}}_i(\mathbf{x}_m), \tilde{\mathcal{R}}_j(\mathbf{x}_n)\right] \quad (2)$$

$$l_j^{\mathbf{x}_m} = \max_{1\le i\le N_m} \mathcal{S}\left[\tilde{\mathcal{R}}_j(\mathbf{x}_n), \tilde{\mathcal{R}}_i(\mathbf{x}_m)\right] \quad (3)$$

As shown in [5], UFM measure greatly reduces the influence of inaccurate segmentation and provides very good performance, so it is chosen in our system to combine with the manifold ranking algorithm. Note that UFM represents the similarity between images, while manifold ranking algorithm needs a distance measure. We use the reciprocal of UFM as the distance measure d for simplicity in our system.

3.3. Calculation of probabilistic labels

As in [8], we represent the segmented images in a uniform feature space based on a generative model. After extracting all the region-based features of the images in the database, the feature dimensionality is reduced by principle component analysis (PCA). Assume that each region-based features R is generated by a Gaussian mixture model (GMM) composed of M components,

$$p(R) = \sum_{n=1}^{M} \alpha_n\, p\left(R|\mu_n, \Sigma_n\right) \quad (4)$$

where $p\left(R|\mu_n, \Sigma_n\right) \sim N(\mu_n, \Sigma_n)$ is a normal distribution. We can use Expectation Maximization (EM) algorithm to calculate the parameters of the model.

Consider each Gaussian component as a label, we can calculate the probabilistic labels of each region as $L(R) = [L_1(R), L_2(R), \cdots, L_M(R)]^T$,where $L_j(R)$ is given by

$$L_j(R) = \frac{\alpha_j\, p\left(R|\mu_j, \Sigma_j\right)}{\sum_{n=1}^{M} \alpha_n\, p\left(R|\mu_n, \Sigma_n\right)} \quad (5)$$

Then the probabilistic labels of image $\mathbf{x}_k$, $L(\mathbf{x}_k) = [L_1(\mathbf{x}_k), L_2(\mathbf{x}_k), \cdots, L_M(\mathbf{x}_k)]^T$, can be calculated as the weighted sum of the probabilistic labels of its regions:

$$L_j(\mathbf{x}_k) = \sum_{i=1}^{N_k} v[R_i(\mathbf{x}_k)]\, L_j[R_i(\mathbf{x}_k)] \quad (6)$$

3.4. Correlated probabilistic label propagation

After constructing the weighted graph by UFM measure and calculating the probabilistic labels of each image, we can propagate the probabilistic labels via the graph. The most direct method is to spread all the probabilistic labels of the query image $\mathbf{x}_q$ independently. Then each image $\mathbf{x}_k$ will correspond to a vector, instead of only a number representing ranking score as in [4]. By sorting the L_2 distance between $\mathbf{x}_k$ and $\mathbf{x}_q$, we can get the retrieval results.

However, in GMM, some of the Gaussian components may overlap each other in a certain degree. Therefore, some of the probabilistic labels may be correlated. Obviously, the above method does not take the correlation into account.

An algorithm of correlated label propagation is proposed in [6], it formulates the framework as a linear programming problem with an exponential number of constraints, and utilizes the properties of submodular functions to solve the problem efficiently. The algorithm can be adopted and further modified in our system for correlated probabilistic label propagation, which is detailed as follows.

Let S be the symmetrically normalized weight matrix, Y be a $n \times M$ matrix, where n is the total number of the images in the database, and M is the total number of the components in GMM. The element of Y is defined as: $Y_{ij} = L_j(\mathbf{x}_i)$, if $\mathbf{x}_i$ is a query; $Y_{ij} = 0$, otherwise. Denote the label frequency vector as $p = [p_1, p_2, \cdots, p_M]^T$, where $p_i = \sum_{k=1}^{n} Y_{ki}$. Then a new matrix Y' is obtained by sorting the columns of Y according to p_i in ascending order. Define matrix $\widehat{Y}$ as

$$\widehat{Y}_{ij} = \Omega\left(\sum_{k=1}^{j} Y'_{ik}\right) \tag{7}$$

where $\Omega(\cdot)$ is the label kernel function [6].

Let $\widehat{F}(0) = \widehat{Y}$, and iterate $\widehat{F}(t+1) = \alpha S\widehat{F}(t) + (1-\alpha)\widehat{Y}$ until the process converges. The limit of the sequence $\{\widehat{F}(t)\}$ is defined as $\widehat{F}^*$, and matrix $\overline{F}^*$ can be obtained by

$$\overline{F}^*_{ij} = \begin{cases} \widehat{F}^*_{ij}, & j = 1; \\ \widehat{F}^*_{ij} - \widehat{F}^*_{i,j-1}, & j > 1. \end{cases} \tag{8}$$

then each image $\mathbf{x}_k$ in the database will correspond to a vector $\overline{F}^*_k = [\overline{F}^*_{k1}, \overline{F}^*_{k2}, \cdots, \overline{F}^*_{kM}]^T$. We calculate the L_2 distance between image $\mathbf{x}_k$ and the query image $\mathbf{x}_q$ as $\|\overline{F}^*_k - \overline{F}^*_q\|$, then the images with smallest distances will be selected as the retrieval results.

3.5. Relevance feedback and active learning methods

In the process of relevance feedback, the matrix Y is defined as: $Y_{ij} = L_j(\mathbf{x}_i)$, if $\mathbf{x}_i$ is a positive sample; $Y_{ij} = -L_j(\mathbf{x}_i)$, if $\mathbf{x}_i$ is a negative sample; $Y_{ij} = 0$, otherwise. Taking the asymmetry between relevant and irrelevant images into account, we reduce the contribution of negative samples by a parameter γ ($0<\gamma<1$) . This method is similar to that used in [4], which has been proved effective.

Let the positive and negative sample sets be $\mathcal{P}$ and $\mathcal{N}$, their size be $|\mathcal{P}|$ and $|\mathcal{N}|$, respectively. The distance between image $\mathbf{x}_k$ and the user's query can be defined as

$$\mathcal{D}(\mathbf{x}_k, \mathcal{P}, \mathcal{N}) = \frac{1}{|\mathcal{P}|}\sum_{\mathbf{x}_i \in \mathcal{P}} \|\overline{F}^*_k - \overline{F}^*_i\| - \frac{1}{|\mathcal{N}|}\sum_{\mathbf{x}_j \in \mathcal{N}} \|\overline{F}^*_k - \overline{F}^*_j\| \tag{9}$$

the return set consists of the images with smallest distances.

To effectively form the label set, active learning method is adopted in our system. As in MRBIR, the most relevant unlabeled images is chosen for user to label, which is the best active learning scheme in [4].

3.6. Implementation issues

In our system, the weighted graph is constructed by connecting only neighboring images, so the matrix S is sparse. However, there may be hundreds of mixture components in GMM, therefore, the matrix $\widehat{F}$ is large, and the computational load of the iteration process is heavy. To address this problem, a candidate set consisting of the positive samples, the negative samples, and the K nearest neighbors of each positive samples is constructed. Then the probabilistic labels of training samples are propagated via the sub-graph corresponding to the candidate set, the return set and the label set are also selected from the candidate set. As the candidate set may exclude the false relevant images, the algorithm can also improve the retrieval results, which will be demonstrated in the experiments.

If the query image is not in the database, as the method in [4], we can connect it with its K nearest neighbors, add one row and one column to W, and perform the other operations similarly with the enlarged matrix W.

4. EXPERIMENTAL RESULTS

The proposed algorithm is evaluated on the database of 10000 real-world images from Corel gallery. All the images belong to 100 semantic categories and 100 images in each category. The region-based features adopted in the experiments are color moments in LUV color space, color histogram in HSV color space, coarseness vector and directionality. In the experiments, the performance measurement used is the top-*k* precision P_k, which is the percentage of the relevant images in the top-*k* returned images. In order to make a reasonable and fair comparison, P_k is averaged by 1000 query sessions, in which the query images are selected randomly from the whole database and kept the same in different algorithms. 4 rounds of relevance feedback are conducted in each query session. During each round, 10 images are labeled by the user.

The parameters in the experiments are set as follows: the numbers of neighbors when constructing the graph and forming the candidate set are 100 and 50 respectively. The value

of σ is set to be twice of the average distance of all the image pairs. GMM is composed of 100 components. α is fixed at 0.99, consistence with the experiments in [3]. As in [4], the number of iteration step is 50. The parameter γ for reducing the contribution of negative samples is 0.5. And the exponential function $\Omega(x) = 1 - e^{-x}$ is chosen as label kernel function, which provides good performance in [6].

The average precisions of manifold ranking (MR) algorithm using different features and distance measures (global features with L_1 distance, global features with L_2 distance and region-based features with the reciprocal of UFM) are shown in Fig. 1. We can draw a conclusion that the retrieval results of the last experiment are the best, which indicates the power of region-based features and UFM similarity measure.

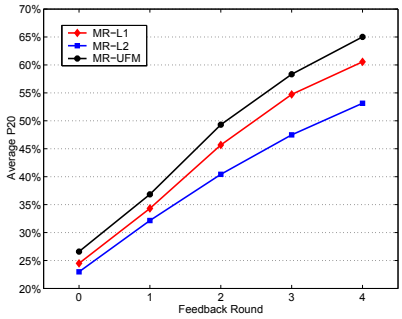

Fig. 1. Comparison of manifold-ranking algorithm using different features and distance measures.

Fig. 2 shows the performance of three algorithms for RBIR: the manifold ranking algorithm, independent probabilistic label propagation (IPLP) and correlated probabilistic label propagation (CPLP). For a fair comparison, all the three algorithms adopt UFM as similarity measure, construct the same weight graph, and introduce the candidate set. The comparison of retrieval performance shows the effectiveness of probabilistic labels, since the latter two algorithms can achieve higher precision than the manifold ranking algorithm. It is also shown that by exploring the correlation among probabilistic labels, the algorithm of correlated probabilistic label propagation can achieve better retrieval results.

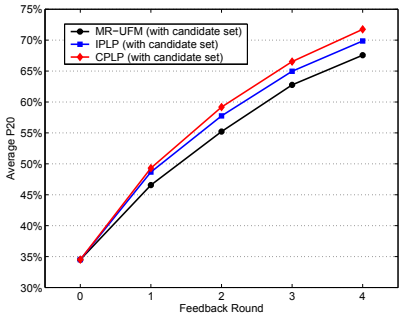

Fig. 2. Comparison of three label propagation algorithms for RBIR.

5. CONCLUSIONS

In this paper, we combine correlated label propagation with manifold ranking algorithm, and present a novel transductive framework for RBIR. In our method, UFM is utilized to measure the similarity between two segmented images effectively, representation based on GMM is introduced to calculate the probabilistic labels of each image. To overcome the disadvantage of propagating multiple labels independently and explore the correlation among labels, the algorithm of correlated label propagation is integrated in our system. Experimental results demonstrate the effectiveness of our proposal.

6. ACKNOWLEDGEMENTS

This work is supported by the Distinguished Young Scholars of NSFC (No.60525111), and the key project of NSFC (No.60432030).

7. REFERENCES

[1] A. Smeulders, M. Worring, S. Santini, A. Gupta, and R. Jain, "Content-based image retrieval at the end of the early years," *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol. 22, no. 12, pp. 1349–1380, 2000.

[2] Y. Rui, T.S. Huang, M. Ortega, and S. Mehrotra, "Relevance feedback: A power tool for interactive content-based image retrieval," *IEEE Trans. Circuits and Systems for Video Technology*, vol. 8, no. 5, pp. 644–655, 1998.

[3] D. Zhou, J. Weston, A. Gretton, O. Bousquet, and B. Schölkopf, "Ranking on data manifolds," in *Advances in Neural Information Processing Systems*, 2003.

[4] J. He, M. Li, H. Zhang, H. Tong, and C. Zhang, "Manifold-ranking based image retrieval," in *Proc. ACM Int. Conf. Multimedia*, 2004, pp. 9–16.

[5] Y.X. Chen and J.Z. Wang, "A region-based fuzzy feature matching approach to content-based image retrieval," *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol. 24, no. 9, pp. 1252–1267, 2002.

[6] F. Kang, R. Jin, and R. Sukthankar, "Correlated label propagation with application to multi-label learning," in *Proc. IEEE Int. Conf. Computer Vision and Pattern Recognition*, 2006, vol. II, pp. 1719–1726.

[7] Y. Deng, B.S. Manjunath, and H. Shin, "Color image segmentation," in *Proc. IEEE Int. Conf. Computer Vision and Pattern Recognition*, 1999, vol. II, pp. 446–451.

[8] W. Jiang, G. Er, Q. Dai, and J. Gu, "Similarity-based online feature selection in content-based image retrieval," *IEEE Trans. Image Processing*, vol. 15, no. 3, pp. 702–712, 2006.