Relevance Feedback for Content-Based Image Retrieval Using Bayesian Network

Jing Xin[1] and Jesse S. Jin[1,2]

[1]School of Computer Science and Engineering, University of New South Wales, NSW 2052, Australia
[2]School of Information Technologies, University of Sydney, NSW 2006, Australia

{jingx, jesse}@cse.unsw.edu.au

Abstract

Relevance feedback is a powerful query modification technique in the field of content-based image retrieval. The key issue in relevance feedback is how to effectively utilize the feedback information to improve the retrieval performance. This paper presents a relevance feedback scheme using Bayesian network model for feedback information adoption. Relevant images during previous iterations are reasonably incorporated into the current iteration and the chosen relevant images can better capture user's information need.

Keywords: relevance feedback, content-based image retrieval, relevant image adoption, Bayesian network.

1 Introduction

The rapid growth of internet and multimedia information has resulted in the development of multimedia information retrieval techniques, especially the content-based image retrieval (CBIR). CBIR systems extract visual features from the images automatically. Similarities between two images are measured in terms of the differences between the corresponding features.

To take into account the subjectivity of human perception and bridge the gap between the high-level concepts and the low-level features, relevance feedback has been proposed to enhance the retrieval performance. During the process of image retrieval, the user specifies the relevance of the retrieved objects, and the system will then refine the query results by learning from this information. A variety of relevance feedback techniques have been proposed. The main algorithms include feature re-weighting and query point movement (Rui, Huang, Ortega and Mehrotra 1998, Ishikawa, Subramanya and Faloutsos 1998), Bayesian target search (Cox, Miller, Minka, Papathomas and Yianilos 2000), support vector machine active learning (Hong, Tian and Huang 2000), and two class learning (MacArthur, Brodley and Shyu 2000).

The key issue in relevance feedback is how to effectively utilize the information provided from user feedback to increase the retrieval accuracy. Most current CBIR systems use the information given in one pass at each iteration and the refined query is treated just like a starting query. Since the feedback information from the previous iterations and the next iteration does not directly connect, the major problem of their approach is that when the next iteration of retrieval is started the information from previous iterations is almost lost. Alternatively we may always use all the relevant images accumulated form the very beginning to fully utilize the information. However, it is still not an ideal resolution. With the progress of the retrieval process, there might exist different levels of relevance among all the objects that have been marked "relevant" before. Some of them are still highly relevant while others might become neutral or even irrelevant. In another word, the user inconsistency might happen at each iteration of relevance judgments which will impair the retrieval results.

The contribution of this paper is to propose a scheme for optimization in utilizing the user feedback information. We explore Bayesian network as a relevant image adoption model to select a number of good points composing the positive feedback information. It is based on the evaluation of the belief values of the relevant image nodes in the network such that these belief values can be used as probabilistic measure of usability of the relevant images. By this approach, objects deemed relevant during previous iterations are reasonably incorporated and the chosen relevant images can better capture user's information need than former methods.

The remainder of this paper is organized as follows: Section 2 presents an overview of Bayesian network and the proposed model for relevant image adoption. Section 3 describes the process of inference propagation and our main algorithm. Conclusions are given in Section 4 with a discussion of future research works.

2 A Bayesian Framework for Relevance Feedback

2.1 Bayesian Belief Network

A Bayesian belief network is a graphical representation of a set of random variables and their dependencies. It provides an effective knowledge representation which imitates knowledge structures used by the human mind. It has been widely used in textual information retrieval for a long time (Baeza-Yates and Ribeiro-Neto 1999).

In a Bayesian network, vertices in the graph represent the random variables $X = \{x_1, x_2, \ldots, x_n\}$ and direct arcs correspond to conditional dependencies between variables. The parents of a node x_i can be denoted by $Pa(x_i)$. Each variable is independent of its non-descendants given its immediate parents in the network. The interactions between nodes are represented by the conditional probability $P(x_i \mid Pa(x_i))$. Other components that are contained in a Bayesian network include: (a) Prior beliefs: the initial beliefs of nodes in the network; (b) Evidences: Observations that are inputs to the network; (c) Posterior probabilities: the final computed beliefs after the evidences have been propagated through the network (Luo and Savakis 2001). By applying the conditional probabilities between nodes and the indecencies between a node and its non-descendants, the joint probability distribution over the complete set of variables can be expressed as:

$$P(X) = P(x_1, x_2, \ldots, x_n) = \prod_{i=1}^{n} P(x_i \mid Pa(x_i)) \qquad (1)$$

Therefore, a complicated joint probability distribution can be reduced to a set of conditional probability which are easier to characterize. More details of Bayesian network can be found in (Pearl 1988, Heckerman 1995, Jensen 1998).

We found that for our relevant image adoption problem, Bayesian network offers distinct advantages. Due to the subjectivity of human perception and the mechanism of relevance feedback, user's feedback judgement is not performed through exact match. User's preference is always reflected by the probabilities. Bayesian network models the probabilistic relationships among the objects, which make it suitable to handles situation relating to probability distributions over variables. Furthermore, the architecture representation of Bayesian network is highly adaptive and easy to build. Bayesian network provides a good framework for integrating the feature representations and it offers easy maintenance when adding new features. The flexibility of feature incorporation is preferable in our application.

2.2 Bayesian Network Relevance Feedback Model

To apply Bayesian Network to our application for relevant image adoption, we need to first formulate the problem in terms of creating a set of variables representing the distinct elements. Then settle the dependency relationships between the variables, that is, creating the links from the parent nodes to child nodes. After that, the numeric probabilities for each variable and link need to be assessed.

In (Wilson, Srinivasin and Indrawan 2000), they proposed a Bayesian network image retrieval engine searching the whole image database to find the retrieval results. In our work, we build a similar network, but only to process the relevant images. Our network aggregates all the features compared with their separate networks for each feature in the feature set. Moreover, our Bayesian network model is a dynamic model which can be adjusted with different relevant images specified by the user. We have also developed a different mechanism for the updating of conditional probabilities associated with the direct links in the network for our purpose. The diagram of the proposed Bayesian network is shown in Fig 1. Generally, the network consists of three layers: query layer, feature index layer and relevant image layer. The root node is the query layer representing the query example image given by the user. The intermediate layer is feature index layer which can be further divided into two levels. The first level contains low-level feature representations, such as color, texture and shape. The second level is composed by the components of the feature vectors. The relevant image layer consists of the individual relevant images specified by the user. The network is different with various query examples and different relevant images. New specified relevant images get added to the relevant image layer as we move from one iteration of feedback to the next. All nodes in the network are associated with a belief value and all the direct connections between the nodes are associated with a link weight, which is represented by a conditional probability.

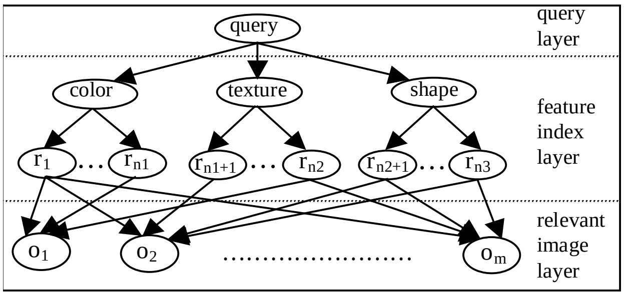

Figure 1: Bayesian network for relevant image adoption

The link weight from the query node to the feature index node specified by $p(f_i \mid q)$ reflects the emphasis of that feature representation in user's information need. The two levels of feature index layer are joined by weighted links $p(r_j \mid f_i)$ representing the different contributions of the component to that feature vector. The link weights $p(f_i \mid q)$ and $p(r_j \mid f_i)$ can be calculated by inter-weight updating and intra-weight updating algorithms proposed in (Rui, Huang, Ortega and Mehrotra 1998).

The weight of a link from the feature vector node to the relevant image node represented by $p(o_k \mid r_j)$ is hard to be obtained directly. However, it can be calculated using Bayes' rule:

$$p(o_k \mid r_j) = \frac{p(r_j \mid o_k)\, p(o_k)}{p(r_j)} \qquad (2)$$

Here $p(r_j)$ and $p(o_k)$ are prior belief values of the component r_j and the relevant image o_k. All components of feature vectors and all relevant images are assumed to have equal prior beliefs. Therefore, $p(r_j)$ is equivalent to $1/n_3$ and $p(o_k)$ is equivalent to $1/m$. The weight $p(r_j \mid o_k)$ represents the importance of component r_j in the relevant image o_k. We propose a distance-based approach to calculate this weight. Intuitively, if the distance between the component r_j of the relevant image and that of the query image is small, it means that the component r_j is important for the relevant image o_k. On the other hand, if the distance is large, then r_j is not an ideal component. Based on this analysis, the inverse of the distance can be a good estimation of weight $P(r_j \mid o_k)$. Thus $P(r_j \mid o_k)$ is calculated as follows:

$$p(r_j \mid o_k) = \frac{1}{d_{jk}} \tag{3}$$

where d_{jk} is the distance of relevant image o_k and the query example on component r_j. These weights are then normalized to make the sum of the weights equal to 1:

$$p(r_j \mid o_k) = \frac{p(r_j \mid o_k)}{\sum_j p(r_j \mid o_k)} \tag{4}$$

After the weights of links and the prior beliefs of all the nodes in the network have been assigned, we can initialize the query node and perform the inference propagation throughout the network to update the belief values of all the relevant image nodes. Then theses belief values can be used as the probabilistic ranking of the relevant images. The details of the inference process and how to select relevant images will be discussed in the next section.

3 Network Evaluation for Relevant Image Adoption

The approach for relevant image adoption in the Bayesian network model is based on the idea that, by learning user's feedback information, the system obtains new evidence about the current relevance level for all the relevant images to the query. At each stage of iteration, the belief values of all the relevant image nodes in the network are re-evaluated according to the result of query and user's relevance feedback.

The query example image can be treated as evidence, which will be instantiated with $p(q) = 1$ when the network is evaluated. The belief values across the network are updated given the initial evidence. According to the topology of our Bayesian network, the inference process propagates from the query layer through the feature index layer to relevant image layer. Currently, there are many efficient methods for exact inference and approximate inference in Bayesian network. Although exact inference in general for an arbitrary Bayesian network is NP-hard (Cooper 1990), it is still efficient for some classes of Bayesian networks. Considering that the size of our network is not large, we employ exact inference method. One of the most influential methods for exact inference is tree-clustering that transforms the network into a so-called junction tree (Jensen, Lauritzen & Olesen, 1990). The junction tree basically clusters the variables in such a way that all loops in the network are removed and the clusters are as small as possible. After converting the network to a tree, the message passing scheme (Pearl 1988) can be run on this tree to update the beliefs of each node in the network given the observation of evidence.

After the inference propagation, the belief values of all the relevant image nodes represented by $\{b(o_k) = P(o_k \mid q), k = 1,2,\ldots,m\}$ are assessed and used as the probabilistic ranking of the relevant images. The candidate relevant image set then can be obtained as follows:

$$S = \{o_1, o_2, \ldots, o_l\}, \quad b(o_k) > \tau \tag{5}$$

where τ is a constant threshold which can be estimated by the training process. Using these relevant images as our positive feedback information, we can employ an arbitrary relevance feedback algorithm to obtain the next iteration of retrieval objects.

After the inference process is performed, all the weights of links are updated for the next iteration of feedback. Also the network beliefs at time t simply become the prior beliefs for the iteration t+1. This process repeats at each stage of feedback until the retrieval process is over and the user is satisfied with the retrieval results. The algorithm for our relevant image adoption is described as follows:

1. After the initial feedback, all the relevant images are used to build the Bayesian network and the link weights are calculated.
2. Perform the inference propagation to update the belief values of all the nodes in the network.
3. Select the relevant images whose belief is above the threshold as the positive feedback information.
4. Update the link weights using the chosen relevant images.
5. The updated belief values are set as new prior beliefs for the next iteration.
6. Start a new iteration of retrieval using the updated positive feedback information.
7. **IF** the user continues to give feedback judgement
8. New relevant objects get added to the network.
9. Go back to step 2 with the updated weights and prior beliefs.
10. **Else** stop the retrieval process and wait for the new query.

4 Conclusion and Future Directions

In this paper, we have focused on the problem of improving the effectiveness in utilizing the relevant images given by user's feedback. We propose an approach using Bayesian network as the relevant image adoption model to find the ideal objects. We believe that Bayesian network is a powerful tool that is applicable in the point of view of relevance feedback in image retrieval. Currently the performance of the system is under test and the effectiveness of our approach is being evaluated. The precision rate as a function of the number of feedback iterations will be used as the performance measure.

The ongoing extensions to the work involve the partition of the relevant images. Besides the relevant image adoption, the clustering of the chosen relevant images can be done to further clarify the user's information need. One or more clusters which the chosen relevant images belong to will be set as positive feedback information. In addition, the possible modification of the network structure includes the semantic level feature index being incorporated into the network and mapped to the low-level features. The utilization of negative feedback will also be investigated.

5 References

Baeza-Yates, R. and Ribeiro-Neto, B. (1999): *Modern information* retrieval. New York, Addison Wesley.

Cooper, G. (1990): Computational complexity of probabilistic inference using Bayesian belief networks. *Artificial Intelligence* 42:393-405.

Cox, I. J., Miller, L., Minka, P., Papathomas, V. and Yianilos P. (2000): The Bayesian image retrieval system, PicHunter: Theory, Implementation and Psychophysical Experiments. *IEEE Transaction on Image Processing* 9(1):20-37.

Heckerman, D. (1995): A tutorial on learning with Bayesian Networks. Technical Report MSR-TR-95-06. Microsoft Research, Redmond, Washingtons.

Hong, P., Tian, Q. and Huang, T.S. (2000): Incorporate Support Vector Machines to Content-Based Image Retrieval with Relevance Feedback, *IEEE International Conference on Image Processing*. Vancouver, Canada.

Ishikawa, Y., Subramanya, R. and Faloutsos, C. (1998): MindReader: Query databases through multiple examples". *Proceeding of the 24th VLDB Conference*, New York.

Jensen, F.V. (1996): *An introduction to Bayesian Networks*. New York, Springer Verlag.

Jensen, F.V., Lauritzen, S.L. and Olesen, K.G. (1990): Bayesian updating in causal probabilistic networks by local computations. *Computational Statistics Quarterly* 4:269-282.

Luo, J. and Savakis, A. (2001): Indoor vs outdoor classification of consumer photographs using low-level and semantic features. *Proceeding of International conference on Image Processing.* Thessaloniki, Greece.

MacArthur, S. D., Brodley, C. E. and Shyu, C. (2000): Relevance feedback decision trees in content-based image retrieval. *IEEE Workshop on CBAIVL*, South Carolina.

Pearl, J. (1988): *Probabilistic reasoning in intelligent systems: networks of plausible inference*. San Mateo CA, Morgan Kaufmann.

Rui, Y., Huang, T.S., Ortega, M. and Mehrotra, S. (1998): Relevance feedback: a power tool in interactive content-based image retrieval. *IEEE Transaction on Circuits and Systems for Video Technology* 8(5):644-655.

Wilson, C., Srinivasan, B. and Indrawan, M. (2000): A general inference network based architecture for multimedia information retrieval. *IEEE International Conference on Multimedia.* New York.