

Integrating Relevance Feedback Techniques for Image Retrieval Using Reinforcement Learning

Peng-Yeng Yin, Bir Bhanu, *Fellow, IEEE*, Kuang-Cheng Chang, and Anlei Dong

Abstract—Relevance feedback (RF) is an interactive process which refines the retrievals to a particular query by utilizing the user's feedback on previously retrieved results. Most researchers strive to develop new RF techniques and ignore the advantages of existing ones. In this paper, we propose an image relevance reinforcement learning (IRRL) model for integrating existing RF techniques in a content-based image retrieval system. Various integration schemes are presented and a long-term shared memory is used to exploit the retrieval experience from multiple users. Also, a concept digesting method is proposed to reduce the complexity of storage demand. The experimental results manifest that the integration of multiple RF approaches gives better retrieval performance than using one RF technique alone, and that the sharing of relevance knowledge between multiple query sessions significantly improves the performance. Further, the storage demand is significantly reduced by the concept digesting technique. This shows the scalability of the proposed model with the increasing-size of database.

Index Terms—Content-based image retrieval, long-term learning, reinforcement learning, relevance feedback, short-term learning.

1 INTRODUCTION

TRADITIONAL database systems are not capable of manipulating pictorial data since they manage data in alphanumeric form. Recently, content-based image retrieval (CBIR) systems [1], [2], [3], [4] have been developed that allow the users to submit a query by an image sample, namely, *query_by_example*, such that the database images that are "similar in content" to the query image are retrieved. Other human interaction models such as allowing the user to select multiple regions-of-interest and to specify the relevance of their spatial layout are also proposed [5]. The search for content-based image retrieval can be broadly classified into two classes according to its purpose: *target search* and *category search*. The purpose of target search is to find exactly a specific image, while the purpose of category search is to find a group of similar images. This paper is focused on the latter with the *query_by_example* interface.

Although CBIR systems are queried by image content, a raw query image still needs to be formulated into an abstract form to execute it efficiently. Since end users, in general, do not know the make-up (kinds of images) of the image database and the content representation and search techniques used in the environment (what types of features and indexing methods are employed), it is hard for them to choose an appropriate query at the first trial. Therefore, the

query formulation process is treated as a series of tentative trials until the target images are found.

A relevance feedback (RF) technique is an interactive process which can fulfill the requirements of query formulation. The principal idea behind the RF is as follows: The user initializes a query session by submitting an image. The system then compares the query image to each image in the database and returns t images that are the nearest neighbors to the query. If the user is not satisfied with the retrieved result, he or she can activate an RF process by identifying which retrieved images are relevant and which are nonrelevant. The system then updates the relevance information, such as the reformulated query vector, feature weights, and prior probabilities of relevance, to include as many user-desired images as possible in the next retrieved result. The process is repeated until the user is satisfied or the results cannot be further improved. The general system flow chart of the RF process is depicted in Fig. 1.

The RF techniques provide a way to bridge the gap between the machine subject in terms of low-level features and the human subject that is driven by high-level semantics. Most researchers strive to develop a new RF technique which can attain better retrieval performance than the existing ones. However, they ignore the fact that, for a given image database, an RF technique that brings the best retrievals to a certain class of query images may be inferior to other RF approaches for another class of query images. Even for a specific query, we may need to apply different RF techniques at various feedback iterations to achieve the highest retrieval performance. By integrating existing RF algorithms, this paper demonstrates that a superior performance can be obtained. We develop a new model, named image relevance reinforcement learning (IRRL), that can integrate multiple RF techniques and makes use of their advantages. The IRRL model automatically chooses the best RF approaches at various feedback iterations for a given query. Also, the

• P.-Y. Yin and K.-C. Chang are with the Department of Information Management, National Chi Nan University, 303 University Rd., Puli, Nantou 545, Taiwan. E-mail: {pyyin, s3213518}@ncnu.edu.tw.

• B. Bhanu and A. Dong are with the Center for Research in Intelligent Systems, University of California, Riverside, CA 92521. E-mail: {bhanu, adong}@cris.ucr.edu.

Manuscript received 21 May 2003; revised 2 Dec. 2004; accepted 1 Feb. 2005; published online 11 Aug. 2005.

Recommended for acceptance by C. Schmid.

For information on obtaining reprints of this article, please send e-mail to: tpami@computer.org, and reference IEEECS Log Number TPAMI-0093-0503.

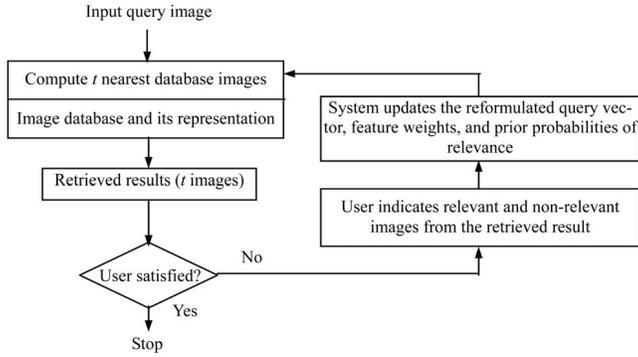


Fig. 1. The general system flow chart of the relevance feedback process.

relevance knowledge created during multiple query sessions is stored in a global memory which is shared during future query sessions and is used to accelerate the relevance learning.

The remainder of this paper is organized as follows: Section 2 describes the related research on relevance feedback and our motivation. Section 3 presents the proposed IRRL model. Section 4 gives the experimental results and provides comparative performances. Finally, Section 5 presents the conclusions of the paper.

Notations. The notations of the variables used in the paper are presented in Table 1.

2 RELATED RESEARCH, MOTIVATION, AND CONTRIBUTIONS

2.1 Relevance Feedback Techniques

Let the query image and a database image be represented by feature vectors $X = (x_1, x_2, \dots, x_d)$ and $Y = (y_1, y_2, \dots, y_d)$, respectively, where d is the number of selected features and x_i and y_i are the values of the i th feature. The system derives the similarity between X and Y by computing the

distance under the given dissimilarity metric. The normalized Euclidean metric

$$Dist(X, Y) = \sqrt{\sum_{i=1}^d (x_i - y_i)^2 / d}$$

is generally used for this purpose. The top t database images that are the nearest neighbors of the query are then returned to the user. If the user is not satisfied with the retrieved result, he or she can activate an iterative RF process until satisfied. In the following sections, the most popular RF techniques are presented.

2.1.1 Query Vector Modification

The query vector modification (QVM) approach repeatedly reformulates the query vector through user's feedback so as to move the query toward relevant images and away from nonrelevant ones. Let a user submit the i th database image as the query and have experienced j RF iterations, and let $X_i^{(j)}$ denote the current query formulation. Also, let the set of relevant images identified at the j th iteration be R , and the set of identified nonrelevant images be N . For the $(j+1)$ th RF iteration, the reformulated query vector [6], [7] is calculated as

$$X_i^{(j+1)} = \alpha X_i^{(j)} + \beta \sum_{Y_k \in R} \frac{Y_k}{|R|} - \gamma \sum_{Y_k \in N} \frac{Y_k}{|N|},$$

where Y_k are images that belong to region R or N , and α , β , and γ are the parameters controlling the relative contribution of each component. The effect of the QVM can be illustrated in Fig. 2. For simplicity, let us assume that only two features, f_1 and f_2 , are used for image similarity matching. All of the database images spreading over the feature space are either relevant (indicated by symbol "+") or nonrelevant (indicated by symbol "-") to the current query $X_i^{(j)}$ according to the user's intention. However, the most "similar" images to the query according to the machine subject (KNN algorithm) in the sense of Euclidean

TABLE 1
The Notations of the Variables in the Paper

Notation	Meaning	Notation	Meaning
t	Number of retrieved images in one display	d	Number of selected features
$X_i^{(j)}$	Query formulation at the j th feedback iteration using the i th image as the initial query	w	Relevance weight of the i th feature
R	Set of relevant images in one display	N	Set of non-relevant images in one display
$p(Y R)$	The conditional probability that an image is Y given that it is relevant to the query	$p(Y N)$	The conditional probability that an image is Y given that it is non-relevant to the query
$p(R Y)$	The <i>a posteriori</i> probability that image Y is relevant to the query	$p(N Y)$	The <i>a posteriori</i> probability that image Y is non-relevant to the query
μ^R	Mean of observed relevant images	μ^N	Mean of observed non-relevant images
σ^R	Standard deviation of observed relevant images	σ^N	Standard deviation of observed non-relevant images
$s_{i,j,k}$	A state of the IRRL model	a_h	An action of the IRRL model
r	An immediate reward of the IRRL model	$\delta(s_{i,j,k}, a_h)$	State transition function
$\tau(s_{i,j,k}, a_h)$	Reward function	π	A strategy for selecting actions
γ	Discounting factor	e	Action entropy threshold
$E(s_{i,j,k})$	Action entropy in state $s_{i,j,k}$	$\Phi(X_i)$	Dominant strategy for the query using the i th image

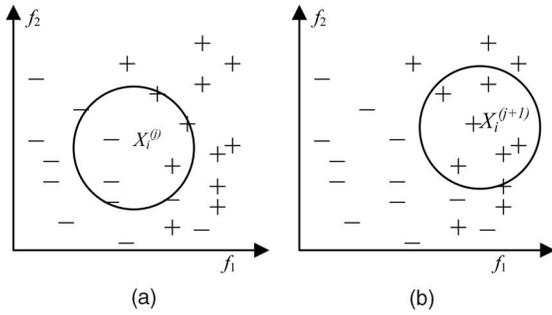


Fig. 2. Illustration of the query vector modification approach. (a) The closest neighbors to the current query are retrieved. (b) A query reformulation can move the query to a region involving more relevant images.

metric are those residing at the interior of the circle centered at $X_i^{(j)}$ (see Fig. 2a). Thus, a simple query reformulation derived from the QVM formula can move the query to a location that is expected to be in a region involving more relevant images (see Fig. 2b).

2.1.2 Feature Relevance Estimation

The feature relevance estimation (FRE) approach assumes, for a given query, some specific features may be more important than others when computing the similarities between images and the query. The relevance of each feature varies in different proportion and should be estimated before the similarity measure is derived. The most natural way of estimating the individual feature relevance is to verify the retrieval ability using each feature alone [8], [9], [10], [11]. The process is as follows: First, all features have equivalent significance of relevance to the query, the set of the top t most similar images can be found by using Euclidean distance. The user is then required to identify the relevant and nonrelevant images from the retrieved t images. To examine the retrieval ability of each feature, all the database images are projected onto the corresponding feature axis, and the new t closest images to the query are computed. Then, the relevance of the feature is evaluated by counting how many of the newly retrieved t images are identified as relevant. The larger the number, the better the retrieval ability of the tested feature and, thus, the feature is more relevant to the query. Finally, the feature relevance is used as a weight incorporated into the dissimilarity metric to express the degree of emphasis on the corresponding feature, viz.,

$$\text{Dist}(X, Y) = \sqrt{\sum_{i=1}^d w_i (x_i - y_i)^2},$$

where w_i is the relevance value or relative weight of the i th feature and $\sum_{i=1}^d w_i = 1$. The process is illustrated in Fig. 3. Again, we assume that only two features are used for image similarity matching and that $X_i^{(j)}$ denotes the current query formulation. If we let $w_1 = w_2$ at the j th RF iteration, the boundary that defines the nearest neighborhood of $X_i^{(j)}$ will form a circle centered at $X_i^{(j)}$ (see Fig. 3a). For the $(j+1)$ th RF iteration, the system learns that feature f_1 is more relevant to the query than feature f_2 by assessing the retrieval ability of each feature, the weightings are adapted so that $w_1 > w_2$. The resulting new boundary of the nearest

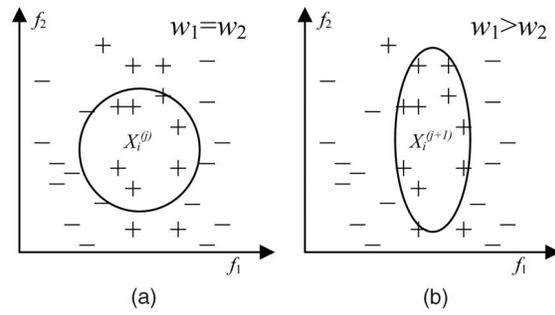


Fig. 3. Illustration of the feature relevance estimation approach. (a) The nearest neighborhood boundary forms a circle when the feature relevance values are equal. (b) The nearest neighborhood boundary forms an ellipse elongated along the direction of the less relevant feature axis to include more relevant images.

neighborhood of $X_i^{(j+1)}$ will include more relevant images and form an ellipse elongated along the direction of the less relevant feature axis. Note that the query formulation is not changed during all RF iterations, i.e., $X_i^{(j+1)} = X_i^j$ for all j .

2.1.3 Bayesian Inference

The Bayesian inference (BI) approaches use a Bayesian framework to estimate the a posteriori probability that a database image is relevant to the query given the prior history of feedbacks [12], [13], [14]. Since the probability distribution over all database images is updated after each feedback iteration, the system is, therefore, able to improve the retrieval performance. Bayesian theory provides a framework to compute the a posteriori probabilities $p(R|Y) = p(Y|R)p(R)/p(Y)$ and

$$p(N|Y) = p(Y|N)p(N)/p(Y).$$

Then, the system can judge whether a database image Y is relevant to the query using the classifier

$$J_\theta(Y) = p(R|Y)/p(N|Y) = p(Y|R)p(R)/p(Y|N)p(N).$$

The conditional probabilities $p(Y|R)$ and $p(Y|N)$ can be approximated by parametric models such as Gaussian kernels using the feature vectors that are identified as relevant or nonrelevant, and $p(R)/p(N)$ can be treated as a small constant since the number of relevant images is much less than that of nonrelevant images. Thus, the retrieval system can output the top t images with the highest values of $J_\theta(Y)$.

2.2 Exploitation of Experience from Multiple Users

The RF techniques discussed in Section 2.1 concentrate on the improvement of retrieval performance for a given query based on the knowledge derived from the retrieval experience and relevance feedback within that query session only. A query session is the period that consists of the submittal of the original query and all the subsequent feedback iterations. So, the relevance process always starts with no assumption about the query formulation, feature weights, and prior probabilities of relevant and nonrelevant images when a query session is initiated. The information generated during this process is erased after the user has terminated the feedback iterations, and it is not used for the next query processing. Nevertheless, there exists agreement, to some extent, among multiple users for judging image relevance

TABLE 2
Shortcomings of RF Techniques and Advantages of Integrating Multiple RF Methods

Main RF Approaches	Disadvantages
<i>QVM</i>	<ul style="list-style-type: none"> • QVM assumes equal weight for each feature, however, not every relevant image is consistently relevant to the query along every feature dimension.
<i>FRE</i>	<ul style="list-style-type: none"> • FRE does not modify the query vector, so the query cannot be moved toward a more desired region of the feature space.
<i>BI</i>	<ul style="list-style-type: none"> • BI is statistics-based and needs to experience more feedback iterations to accurately approximate the probability distribution.
RF Integrations	Advantages
<i>QVM</i> and <i>FRE</i>	<ul style="list-style-type: none"> • The query can be moved to a desired region and each feature is assigned a weight representing the significance of relevance.
<i>QVM</i> and <i>BI</i>	<ul style="list-style-type: none"> • QVM progressively predicts the center of mass of relevant images and if this prediction is correct, it provides good Gaussian kernels of probability estimation for BI.
<i>FRE</i> and <i>BI</i>	<ul style="list-style-type: none"> • FRE usually expedites the convergence of the feature weights through a higher order weighting function such as quadratic or exponential functions, it may provide good predictions for the standard deviations of Gaussians for BI.

and it would benefit future query performances if the retrieval experiences obtained from multiple users can be preserved and carefully exploited. Since real image databases experience retrievals from many users, the metaknowledge derived from such retrieval experiences can be exploited to learn visual concepts and refine them incrementally [15]. Recently, Bhanu and Dong [16] used a semisupervised fuzzy clustering approach in conjunction with prior retrieval experience to partition the database into clusters related to high-level concepts, which can be used for efficient indexing. Yin et al. [17] proposed a systematic framework that extends the ability of traditional relevance feedback techniques to the exploitation of accumulated retrieval experiences by the use of virtual features, which are a form of shared long-term memory storing the semantic concepts learned from multiple users.

2.3 Motivation for Our Approach

There are two motivations inspiring us to develop the proposed approach. First, for a given image database an RF technique that brings the best retrievals to a certain class of query images may be inferior to other RF approaches for another class of query images. Second, the experience from long-term interactions with multiple users could be exploited and the derived knowledge can improve future retrieval performance.

2.3.1 Reinforcement Learning

Reinforcement learning addresses the issue of how an agent can learn a task through a sequence of trial-and-error interactions with its environment [18], [19]. It has been successfully applied to several computer vision problems such as image segmentation, feature extraction, and object recognition [20], [21]. In this paper, we intend to apply reinforcement learning for the optimal integration of multiple RF techniques. The reasons are threefold.

- Each RF technique has its own assumptions and limitations (see Table 2). In particular, the QVM does not have the mechanism that gives various relevance significance to different features, so the shape of the nearest neighborhood of the query cannot be changed. On the other hand, the FRE does not modify the query vector at all iterations, the retrieval

performance cannot be further improved if the initial query is far from most relevant images. The BI is statistics-based and needs experience with more feedback iterations to accumulate statistically sufficient amount of samples to make correct inference. Also, the probability estimation for BI is more complicated than the formulations for QVM and FRE and, thus, requires more computations.

- Different RF techniques applied at consecutive iterations may be complementary to each other (see Table 2). Heesch and Ruger [22] have pointed out that through the integration of QVM and FRE the query center can be updated and each feature can be assigned a weight. As for the BI, if the number of available samples is limited in practice, the selection of these samples plays an important role on the accuracy of probability distribution estimation. Since BI statistically infers the probability distributions based on the observed samples only, it does not make progressive prediction about the unseen samples. QVM progressively predicts the center of mass of relevant images and if this prediction is correct, it provides good Gaussian kernels for probability density estimation for BI. Similarly, FRE usually expedites the convergence of the feature weights through a higher order weighting function such as quadratic or exponential functions [10], it may provide good predictions for the standard deviations of Gaussians for BI. Several previous works [23], [24], [25] have been built on the combination of multiple features and similarity models. However, this paper is focused on the integration for multiple RF techniques.
- It is very common that the collected images of an image database have extremely diverse content, therefore, the distributions of feature vectors for relevant and nonrelevant images can vary significantly from query to query. Some of the cases may be easily modeled by the QVM, while others may be more appropriately modeled by FRE or BI. For a given image database an RF technique that brings the best retrievals to a certain class of query images may be inferior to other RF approaches for another class of query images. Even for a specific query

image, we usually cannot find an RF technique that is the most suitable one to be applied at all feedback iterations. So, the choice of the optimal RF technique is query and iteration-dependent.

Although reinforcement learning improves the next retrieval by applying the optimal RF technique, it does not affect the first retrieval of every query session. With the long-term learning scheme, the reinforcement learning model can provide higher initial retrieval performance.

2.3.2 Long-Term Learning

There are two learning issues when dealing with multiple query sessions. First, it is desirable that a retrieval system conduct an effective relevance learning based on the prior knowledge acquired at previous query sessions. The traditional short-term relevance learning scheme always starts a new query session without any assumption about the query formulations, feature weights, or probability distributions of relevant and nonrelevant images. The original query with equally weighted features is referred to as the first retrievals. In contrast, the long-term relevance learning scheme presented here keeps a global memory for each database image for storing the last query formulation, feature weights, and the a priori probabilities of relevant and nonrelevant images when the database image is previously used as a query. Thus, the relevance learning for a new query session can start from its previous state to expedite the learning process. The long-term relevance learning scheme can greatly improve the retrieval performances, especially for the first retrievals of query sessions.

Second, the task of learning the optimal strategy of selecting multiple RF techniques for a specific query is done by evaluating the retrieval performance obtained using each strategy alone. The accuracy of the evaluation depends upon how many users' retrieval experiences have been considered since there may exist a range of differences between users' subjectivity about relevance. The long-term strategy learning stores the performance estimate of every strategy for each database image based on experiences from multiple users. When the convergence of the learning task has taken place (this status can be detected by a mechanism such as the entropy analysis as will be discussed in Section 3.3), the system can select the optimal RF technique at a specific feedback iteration for a given query to provide the highest possible performance.

2.4 Contributions of the Paper

The original contribution of this paper includes the following aspects.

1. We present two integration schemes, a) combination and b) hybridization, to achieve the maximum synergism between different RF techniques.
2. Our system is the first one that can automatically choose the optimal RF approach for a given query at a particular feedback iteration.
3. We use a shared long-term memory to accumulate the relevance knowledge acquired from multiple users' experiences. The long-term relevance knowledge significantly improves the retrieval performance.
4. The efficacy of the proposed system is validated using a real image database. The results show that our integrated approach is better than any single

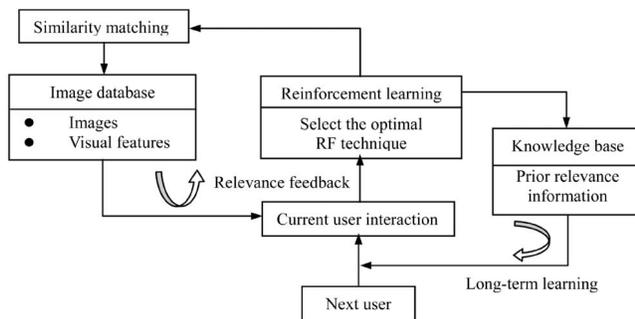


Fig. 4. The system diagram for the image relevance reinforcement learning (IRRL) model.

relevance feedback approach with or without long-term learning.

3 THE PROPOSED APPROACH

In this section, we propose our model (see Fig. 4) for image relevance reinforcement learning (IRRL). When a user starts a new query session, the prior relevance information about this query is first retrieved. With the relevance information derived by long-term learning, the model performs retrievals based on previous learning state. When entering the session, the reinforcement learning navigates the model to select the optimal RF technique for the query at every feedback iteration, and the most probable relevant images are searched and returned to the user for feedback interactions. When the user terminates the session, the latest relevance information is captured in the knowledge base for updating the corresponding entry.

3.1 Integration Schemes for Multiple RF Approaches

The integration scheme addresses the issue of how the multiple RF techniques applied within a query session interact with each other. We present two schemes named *combination* and *hybridization*. Let us assume that the retrieval system has a set of the three major existing RF techniques, namely, the QVM, FRE, and BI, and also the system is capable to improve the retrieval performance of a query session by executing several RF iterations. We define an RF *strategy* as a sequence of selected RF techniques to be applied at various feedback iterations for a given query. The assumptions yield many possible strategies as shown in Fig. 5. Each strategy corresponds to a path originating from the leftmost node (initial state of a query session) and ending at any of the rightmost nodes (terminating states of a query session). For the target search, it is possible that the target image is not found within the first few feedback iterations but can be found by experiencing more iterations. Since the performance of a target search is evaluated in bilevel setting, i.e., whether the target image is found or not, the number of experienced feedback iterations would be a critical factor and should be large enough to guarantee a fair analysis. However, the purpose of this paper is for category search and the performance is evaluated by calculating the precision rate (the ratio of the number of positive retrievals divided by the number of total retrievals in the same display). The precision rate is not a binary value (success or not) as used in target search, but rather it is a real number between 0 and 1. In our experiments with a large real image database (more than 10,000 images), the precision rate of a typical query session

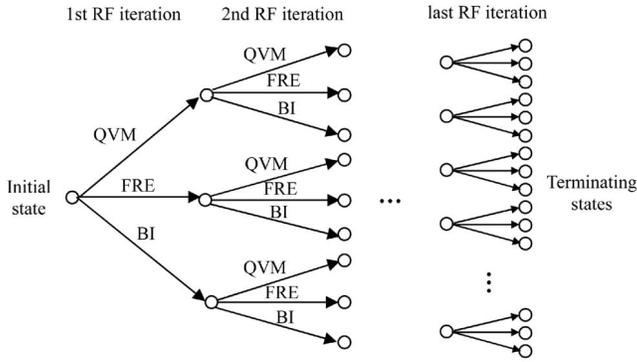


Fig. 5. Many possible RF strategies in a query session for selecting RF techniques at different iterations.

improves significantly during the first two feedback iterations, and won't change much afterwards. Hence, in the following, we focus our discussion on the integration of RF techniques applied at two feedback iterations.

For the strategies that apply the same RF technique at both iterations, the RF process can be performed by the corresponding formulation. While for those strategies that apply distinct RF techniques at various iterations, for example, apply the QVM at the first iteration and FRE at the second, there are two means for integrating them. The first type of integration, called *combination*, simply applies one RF technique at the first iteration and applies the other RF technique at the second. However, the second type of integration, called *hybridization*, applies one RF technique at the first iteration and when the second iteration is entered, both RF techniques are applied simultaneously to strengthen the synergetic effect. In particular, we embed the three most popular RF approaches, namely, the QVM, FRE, and BI, in our IRRL model and propose three specific ways to hybridize two of them. It behaves algebraically the same at the second iteration if we reverse the order of the two RF approaches applied at consecutive iterations, although the RF applied at the first iteration is different if we reverse the order. So, there are exactly three ways to implement the hybridization scheme, each of which corresponds to hybridization of every two RF techniques. We do not make the hybridization scheme different when the order of

applied RF techniques is reversed because each RF technique has its unique formulation and complements the other in a particular way. The details of the two types of within-session integration schemes are described as follows:

- Integration between QVM and FRE.** Without loss of generality, we assume QVM is applied at the first iteration and FRE at the second. Fig. 6 gives an illustration. Let the original query be the i th database image and be denoted by $X_i^{(0)}$. The system returns the closest images to $X_i^{(0)}$ with feature weights $w_1 = w_2$, i.e., returns those images residing at the interior of the circle centered at $X_i^{(0)}$ (see Fig. 6a), and requests for relevance feedback. The user identifies relevant and nonrelevant images from the retrievals. By performing QVM at the first RF iteration, the new query vector $X_i^{(1)}$ is derived by $X_i^{(1)} = \alpha X_i^{(0)} + \beta \sum_{Y_j \in R} Y_j / |R| - \gamma \sum_{Y_j \in N} Y_j / |N|$, while the feature weights remain unchanged. Thus, the new query vector is moved to a location closer to the mass centroid of relevant images (see Fig. 6b). At the second RF iteration where FRE is applied, there are two integration schemes. For the combination scheme (see Fig. 6c), the query vector is not changed ($X_i^{(2)} = X_i^{(1)}$) because the QVM is not applied at this iteration, only the feature weights are updated ($w_1 > w_2$) according to the FRE so as to stretch the boundary of the query's neighborhood to involve more relevant images. For the hybridization scheme (see Fig. 6d), in addition to updating the feature weights ($w_1 > w_2$) based on FRE, the query vector $X_i^{(2)}$ is reformulated by $X_i^{(2)} = \alpha X_i^{(1)} + \beta \sum_{Y_j \in R} Y_j / |R| - \gamma \sum_{Y_j \in N} Y_j / |N|$. As such, the FRE is hybridized with the QVM at the second iteration. It is observed that both types of integration schemes for QVM and FRE preserve the advantages of each approach, and improve the retrieval performance than using the same approach at all feedback iterations (with and without long-term learning).

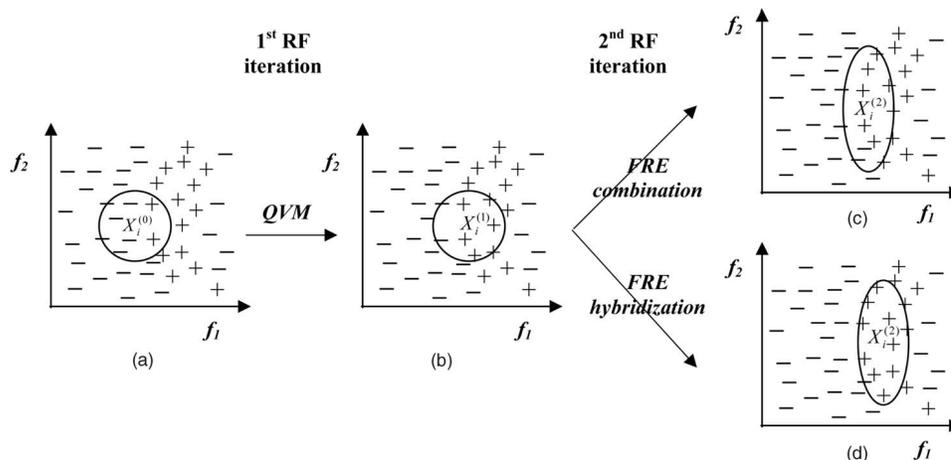


Fig. 6. Illustration of combination and hybridization schemes of QVM and FRE. (a) The system returns the initial retrievals to the original query. (b) The query vector is reformulated by QVM and is moved closer to the mass centroid of relevant images. (c) The FRE is applied to stretch the neighborhood while the query vector is unchanged. (d) In addition to stretch the neighborhood by FRE, the query vector is further reformulated by QVM.

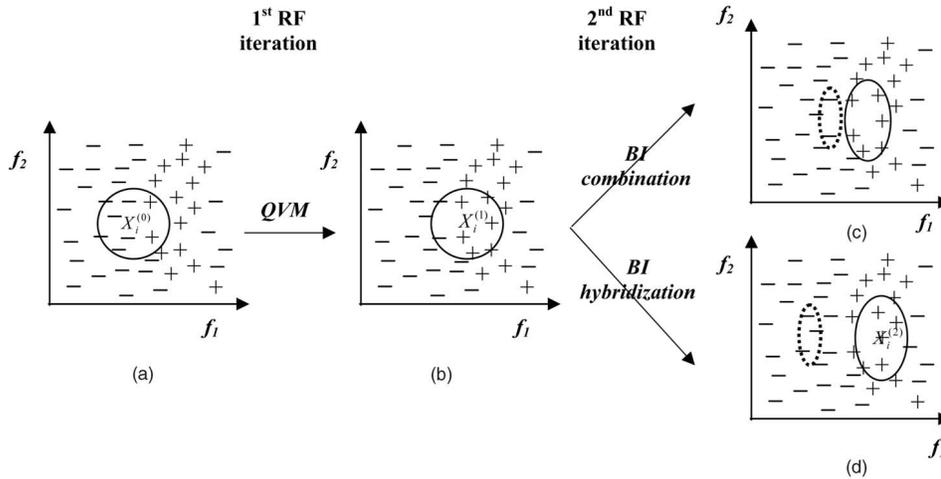


Fig. 7. Illustration of combination and hybridization schemes of QVM and BI. (a) The system returns the initial retrievals to the original query. (b) The query vector is reformulated by QVM and is moved closer to the mass centroid of relevant images. (c) The BI is applied to estimate the a posteriori probabilities. (d) The BI is applied with the mean vectors estimated using QVM.

- Integration between QVM and BI.** Again, without loss of generality, we assume QVM is applied at the first iteration and BI at the second. Fig. 7 gives an illustration. The situations of the first retrievals and the retrievals after the first RF iteration (see Figs. 7a and 7b) are the same as those in the previous case. At the second RF iteration where BI is applied, there are two possibilities. For the combination scheme (see Fig. 7c), the a priori probabilities of $p(Y|R)$ and $p(Y|N)$ are estimated using the observed samples in R and N , respectively, identified at the second RF iteration. If we assume the relevant images form a Gaussian density, then $p(Y|R) \equiv N(\mu^R, \sigma^R)$ with $\mu^R = \mu(\{Y|\forall Y \in R\})$ and $\sigma^R = \sigma(\{Y|\forall Y \in R\})$, where $\mu(\cdot)$ and $\sigma(\cdot)$ denote the vectors of mean and standard deviation of the observed samples. Although the nonrelevant images may belong to multiple classes, we intend, for the problem discussed here, to make a 2-class discrimination between relevant and nonrelevant classes. So, we use a Gaussian density to model nonrelevant images and let $p(Y|N) \equiv N(\mu^N, \sigma^N)$ with $\mu^N = \mu(\{Y|\forall Y \in N\})$ and $\sigma^N = \sigma(\{Y|\forall Y \in N\})$. Fig. 7c sketches the estimations of $p(Y|R)$ and $p(Y|N)$ using solid and dashed borders, respectively. The most relevant images are determined using the Bayesian classifier. On the other hand, for the hybridization scheme (see Fig. 7d), the query vector $X_i^{(2)}$ is reformulated, according to QVM, by $X_i^{(2)} = \alpha X_i^{(1)} + \beta \sum_{Y_j \in R} Y_j / |R| - \gamma \sum_{Y_j \in N} Y_j / |N|$. Since $X_i^{(2)}$ is an estimate for the mass centroid of all possible relevant images based on both relevant and nonrelevant observed samples, it is also an estimate for the mean vector of the assumed Gaussian density of relevant images. Hence, we set $\mu^R = X_i^{(2)}$. Similarly, the mean vector of nonrelevant images can be determined based on the QVM formulation, we get $\mu^N = \rho \sum_{Y_j \in N} Y_j / |N| - \eta \sum_{Y_j \in R} Y_j / |R|$, where ρ and η are parameters controlling the relative contribution of

each component. The standard deviation vectors are derived as in the combination scheme. The estimations of $p(Y|R)$ and $p(Y|N)$ for the hybridization scheme are sketched in Fig. 7d. As such, the BI is hybridized with the QVM by replacing the estimates for the mean vectors with those obtained using the QVM.

- Integration between FRE and BI.** Without loss of generality, we assume FRE is applied at the first iteration and BI at the second. Fig. 8 gives an illustration. Following the same notations used above, the system first retrieves the closest images to $X_i^{(0)}$ with $w_1 = w_2$ (see Fig. 8a), and requests for relevance feedback. The user then identifies relevant and nonrelevant images. By performing FRE at the first RF iteration, the weights are updated as $w_1 > w_2$ and the query vector remains unchanged ($X_i^{(1)} = X_i^{(0)}$), as illustrated in Fig. 8b. At the second RF iteration, the BI is to be applied. For the combination scheme (see Fig. 8c), the a priori probabilities of $p(Y|R)$ and $p(Y|N)$ are simply estimated using observed samples in R and N identified at the second RF iteration. However, for the hybridization scheme (see Fig. 8d), the weights are further updated due to FRE (say, w_1 is now much larger than w_2). In the sense of FRE, a larger feature weight has resulted from a denser distribution of images on the corresponding feature component. Hence, the standard deviation of the Gaussian density is inversely proportional to the corresponding feature weight for the feature component. We, thus, set $\sigma_j^R = (1 - w_j) \sum_{k=1}^d \sigma_k^R / \sum_{k=1}^d (1 - w_k)$, where σ_j^R denotes the standard deviation of the Gaussian density of relevant images on the j th feature component. Also, the standard deviation vector of nonrelevant images is similarly derived as $\sigma_j^N = (1 - w_j) \sum_{k=1}^d \sigma_k^N / \sum_{k=1}^d (1 - w_k)$. The mean vectors μ^R and μ^N are computed as in the combination scheme. Fig. 8d sketches the estimations of $p(Y|R)$ and $p(Y|N)$ for the

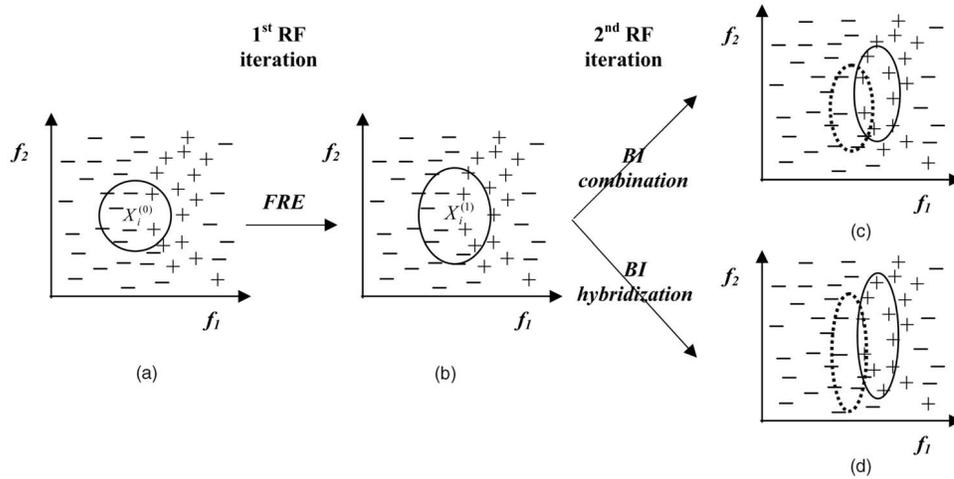


Fig. 8. Illustration of combination and hybridization schemes of FRE and BI. (a) The system returns the initial retrievals to the original query. (b) The neighborhood of the query vector is stretched by FRE to involve more relevant images. (c) The BI is applied to estimate the a posteriori probabilities. (d) The BI is applied with the standard deviation vectors estimated using FRE.

hybridization scheme. As such, the BI is hybridized with the FRE by replacing the estimates for the standard deviation vectors with those derived from the FRE.

Let us consider again an RF process with two iterations, there are three strategies applying the same RF approach at both iterations and six strategies applying distinct RF approaches at different iterations. Since each of the six strategies can be facilitated with combination or hybridization schemes, there are 15 ways of performing the RF process for a given query.

We have conducted extensive experiments on a testing database of 10,038 images to evaluate the comparative performances between the 15 RF processes. There is no statistically significant evidence emerged from our experiment that says that any particular RF process is superior to others for all query images. There are some queries for which a particular RF process gives much better retrieval performance than the others, and there also exist some queries for which the same RF process produces worse retrievals. Therefore, the choice for an optimal strategy for selecting RF techniques at various feedback iterations and the choice for the best within-session integration scheme of the selected RF techniques are both query-dependent.

3.2 Reinforcement Learning for Image Relevance

To learn the optimal strategy and the best within-session integration scheme for using multiple RF techniques, we propose an image relevance reinforcement learning (IRRL) model as sketched in Fig. 9. A user initializes a query session by submitting a query image to an agent which is a CBIR system with multiple RF mechanisms. We define that an RF *mechanism* is an existing RF approach facilitated with a particular within-session integration scheme, for example, the RF mechanism FRE/hybridization instructs the agent to apply FRE at the current feedback iteration and hybridize FRE with the RF technique that is applied at the preceding iteration, if it exists. The agent applies an action selection rule to perform an action from the set of possible RF mechanisms. The nearest t images to the query are computed by calculating the similarity measure which is

defined by the selected RF mechanism, these images are then returned to the environment (the end user) for requesting a relevance feedback. The user identifies relevant and non-relevant images from the retrieved result, and a precision rate about the retrievals can be computed. The state of the environment is, therefore, changed to another state. The precision rate is also provided to the agent as a reward that reveals the desirability about the state transition. The agent observes the new state and repeats the cycle again. This produces a sequence of states s_i , actions a_i , and rewards r_i as shown at the bottom of Fig. 9. The agent's goal is to learn an optimal strategy for selecting an action in a given state that maximizes the expected sum of total rewards.

3.2.1 IRRL Model

Let the image database contain a collection of n images, and let the agent be allowed to perform an RF process on a specific query for at most m iterations. Assume the agent is provided a set of u possible RF mechanisms to choose from. Some notations of the IRRL model are defined as follows:

- A set of states, $S = \{s_{i,j,k}\}_{1 \leq i \leq n, 0 \leq j \leq m, 0 \leq k \leq u}$. A state is characterized by three elements, namely, the query image i , feedback iteration j , and the last RF mechanism k performed to this query. Note that $j = 0$ indicates the state before any relevance feedback is performed within the current query session, and $k = 0$ describes that no RF mechanism has ever been performed to this query image. The index structure is designed to express that the underlying task of learning the optimal RF strategy depends upon the query image, the number of current iteration, and the last performed RF mechanism.
- A set of actions, $A = \{a_h\}_{1 \leq h \leq u}$. Performing an action corresponds to executing an existing RF mechanism to the query image. The possible RF mechanisms are precoded in the agent and, once they are called upon, they will replace the similarity metric with their own definitive formulations.

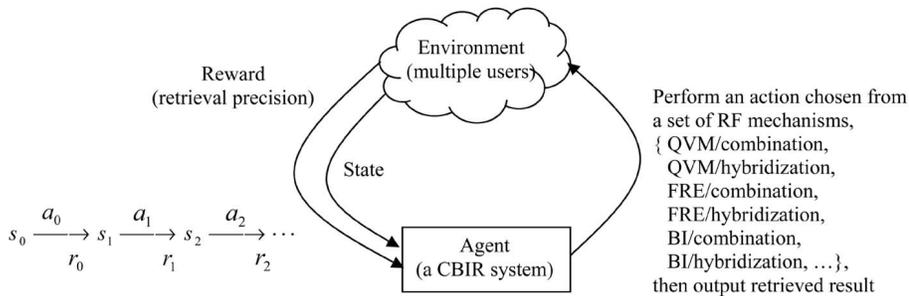


Fig. 9. The image relevance reinforcement learning (IRRL) model.

- *Positive real-valued rewards*, $r \in R^+$, where R^+ is the set of positive real numbers. The reward can be described by the precision rate regarding the user's desirability about the current retrievals and it is given by $r = \text{Positive Retrievals} / \text{Total Retrievals}$.
- *A state transition function*, $\delta : S \times A \rightarrow S$. In particular, by the above definitions, we have $\delta(s_{i,j,k}, a_h) = s_{i,j+1,h}$, that is, when the agent performs an RF action, it will sense a new state describing the transition to the next feedback iteration and the action just performed.
- *A reward function*, $\tau : S \times A \rightarrow R^+$. In particular, $\tau(s_{i,j,k}, a_h)$ will return the precision rate on the current retrievals obtained when the agent performs action a_h in state $s_{i,j,k}$.

The IRRL model learns the optimal strategy, $\pi^* : S \rightarrow A$, that maximizes the cumulative rewards received over time (or the expected sum of the precision rates obtained at all feedback iterations), that is, $\pi^* = \arg \max_{\pi} \{r_0 + \gamma r_1 + \gamma^2 r_2 + \dots\} = \arg \max_{\pi} \sum_{v=0}^{\infty} \gamma^v r_v$, where r_v is the reward received v steps into the future using strategy π to select actions, and $\gamma \in [0, 1]$ is the discounting factor that determines the relative value of immediate and delayed rewards.

3.2.2 Methodology

We apply the Q -learning algorithm [26], which is the most popular method for conducting a reinforcement learning task, to learn the optimal RF strategy. The Q -learning algorithm could have a slow convergence speed for real-time applications, but it is applicable in our system for three reasons. First, most other systems which employ reinforcement learning should wait for the completion of the training process then use the output strategy for deriving the final solution. While our system trains, the CBIR agent in the background when users are using it. The users initially meet a level of performance no worse than those of existing RF techniques, and then achieve a gradual performance improvement in the future. Second, the proposed system is highly adaptive to users' intention by updating the \hat{Q} estimates. This is an important property in a multiuser environment where a mechanism handling the transition between different intentions is necessary. Third, there is currently no similar system which can be used for comparison on convergence speed. Our system is the first one that integrates multiple RF techniques and maximizes their synergetic performance. In Section 4, we provide a comparison with a single RF technique with and without long-term learning. Let $Q(s_{i,j,k}, a_h)$ be the maximum cumulative reward which can be received by performing action a_h in state $s_{i,j,k}$ and, then, proceeding optimally using π^* . The Q -learning

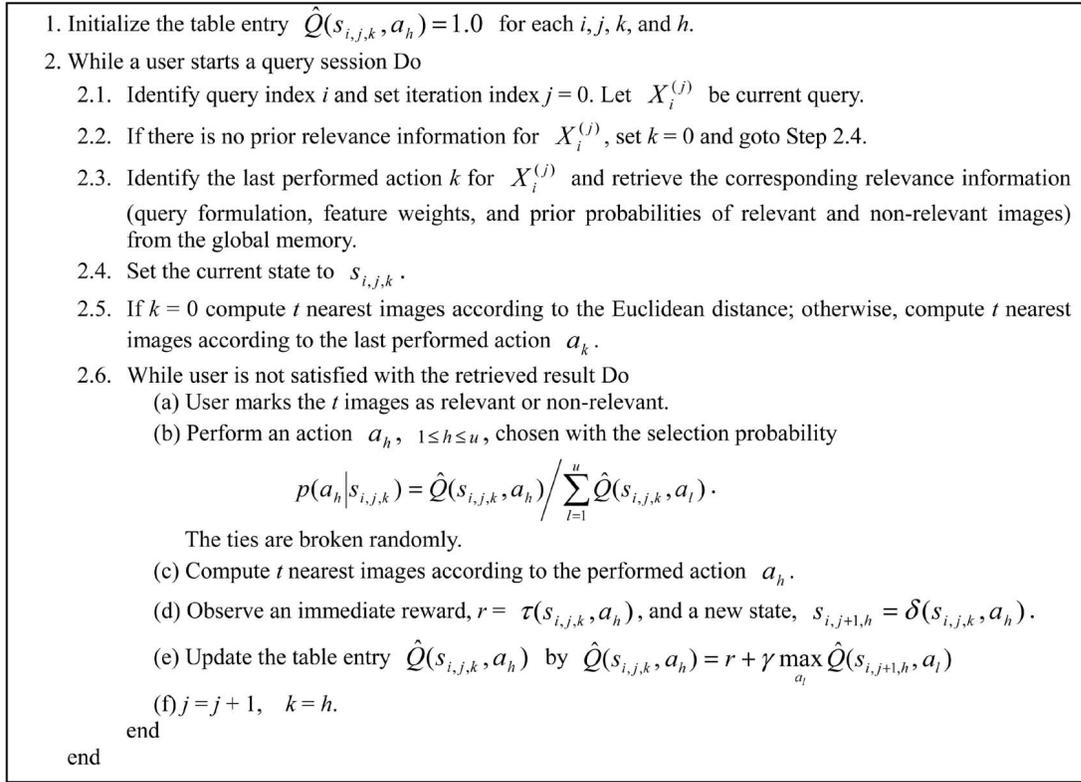
algorithm iteratively approximates the Q function by the following recursive definition:

$$\begin{aligned} Q(s_{i,j,k}, a_h) &= \tau(s_{i,j,k}, a_h) + \gamma \max_{a_l} Q(\delta(s_{i,j,k}, a_h), a_l) \\ &= r + \gamma \max_{a_l} Q(s_{i,j+1,h}, a_l), \end{aligned}$$

where r is the observed immediate reward and $s_{i,j+1,h}$ is the observed new state.

The Q -learning algorithm for the IRRL model is presented in Fig. 10 and is explained as follows: First, the algorithm initializes a table of estimate of the Q function for each possible state-action pair. When a user starts a new query session by submitting a query, say, image i , if the image was never used as a query before, the algorithm computes the t nearest images according to the Euclidean distance; otherwise, the algorithm retrieves all relevance knowledge (involving query formulation, feature weights, and prior probabilities of relevant and nonrelevant images) about this query from a global shared memory, and then computes the t nearest images according to the last performed action when image i was previously used as a query. Next, if the user is not satisfied with the retrieved results, he/she can perform an RF process to improve the retrievals. At each RF iteration, the user marks the retrievals as relevant or nonrelevant. The algorithm performs an action based on the selection probability. We have experimented with a probabilistic action selection rule according to which actions with higher \hat{Q} values will be assigned higher selection probabilities. Let the agent in state $s_{i,j,k}$ be going to choose an action from a set of u alternatives. The probability with which the agent chooses action a_h is given by $p(a_h | s_{i,j,k}) = \hat{Q}(s_{i,j,k}, a_h) / \sum_{l=1}^u \hat{Q}(s_{i,j,k}, a_l)$. As such, the probability with which an action is chosen is linearly proportional to the corresponding \hat{Q} estimate, and every action is assigned a nonzero probability. Note that the ties with respect to $p(a_h | s_{i,j,k})$ are broken randomly.

Then, the algorithm computes t nearest images according to the performed action, observes an immediate reward and a new state, then updates the corresponding \hat{Q} table entry. The algorithm iteratively approximates the optimal strategy for choosing an action (one of existing RF mechanisms) to perform in a given state, and guide the agent to maximize the expected sum of total rewards (retrieval precisions obtained at all feedback iterations).

Fig. 10. Q -learning algorithm for the IRRL model.

3.3 Convergence Analysis and Storage Reduction

When a database image is used as a query, one may ask a question about whether the CBIR agent has learned a near-optimal RF strategy with proper within-session integration schemes for this query. This can be answered by examining the current values of the estimates for the Q function. Since the estimate approximates the maximal expected sum of precision rates, a near-optimal selection for an RF strategy is learned if the value of the corresponding estimate is significantly larger than those of other possible choices. Consider that the agent is provided three RF techniques and two within-session integration schemes, this totally results in six available RF mechanisms. Let the agent be in state $s_{i,j,k}$ and have six choices of possible actions, each of which is assigned a selection probability, $p(a_h | s_{i,j,k}), h = 1, 2, \dots, 6$. We compute the information entropy regarding to action selection probabilities in state $s_{i,j,k}$ by

$$E(s_{i,j,k}) = - \sum_{h=1}^6 p(a_h | s_{i,j,k}) \log_2 p(a_h | s_{i,j,k}).$$

The smaller the value of $E(s_{i,j,k})$, the more deterministic the action selection in state $s_{i,j,k}$. Thus, the CBIR agent has learned a dominant strategy $\Phi(X_i)$ for query X_i if the entropy values in the initial state and those states sensed during all subsequent feedback iterations using this strategy are less than a small real-valued threshold e , that is, $\Phi(X_i)$ is *determined* if $E(s_{i,0,0}) < e$ and

$$E(\delta(s_{i,j,k}, \arg \max_{a_l} \{\hat{Q}(s_{i,j,k}, a_l)\})) < e,$$

for $j = 0, 1, 2, \dots$

Thus, the Q -learning algorithm has reached a “mature” convergence for query X_i if the dominant strategy $\Phi(X_i)$ is determined. To save the storage demand incurred by storing the \hat{Q} estimate for each state-action pair, we present a concept digesting method as follows: Recall that, if an image has determined a dominant strategy, the selection probabilities for the actions along the dominant path are significantly higher than others. Thus, if two images determine the same dominant strategy, the distributions of selection probabilities are very similar, so are the relative entry values of the two \hat{Q} tables. It would have little impact on computing selection probabilities if we merge the two \hat{Q} tables by averaging the corresponding entry values. Accordingly, we let the IRRL agent merge, by averaging, the \hat{Q} tables of those images that determine the same dominant strategy. The \hat{Q} entry update of these images is then operated on the same table. Nevertheless, for those images that have not determined their dominant strategy yet, each of them should still be prepared for a separate \hat{Q} table for learning the optimal strategy. So, there are two types of concepts existing in our system. A *determined concept* consists of those images that determine the same dominant strategy, while a *nondetermined concept* has only one element which is an image that has not determined a dominant strategy. Note that here the concept should not be referred to images with high similarity in visual content; instead, the concept is a class of query images that are optimally treated with the same RF strategy. The CBIR agent keeps a separate \hat{Q} table for every concept of both types. Hence, the storage demand is proportional to the number of concepts existing at that moment. When the database is just created, the number of concepts is equivalent to the number of database images since there is no determined concept

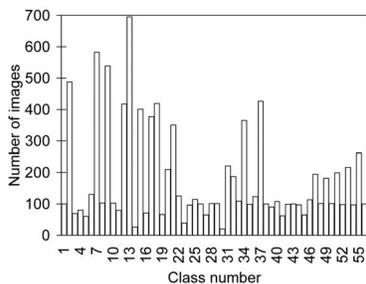


Fig. 11. Real image database: histogram of the distribution of the numbers of images in each class.

digested and each database image corresponds to a non-determined concept. As the agent experiences more query sessions, it digests a determined concept by merging many nondetermined concepts and the number of total concepts decreases. The storage demand is reduced and it provides a room for adding new images to the database. Consequently, the CBIR agent is scalable against a dynamic database with growing size.

4 EXPERIMENTAL RESULTS AND COMPARATIVE PERFORMANCE EVALUATION

In this section, we conduct extensive experiments to evaluate the performance. We have implemented the three major RF techniques, namely, the QVM, FRE, and BI, and integrated them with the IRRL model with various schemes (combination and hybridization). The programs are coded in C++ language and executed on a PC Web server with a Pentium IV 2.4 GHz CPU and 248 MB RAM. To simulate the practical situation of online users, the sequence of query images used in our experiments is generated at random. Each query session is allowed to refine its retrievals by executing a relevance feedback process for two iterations. The average precision rates over all query images obtained at three different stages, namely, the one before any relevance feedback (PR0), the one after the first iteration of relevance feedback (PR1), and the one after the second iteration of relevance feedback (PR2), are computed. The

UCR database is chosen for the experiments and its content is briefly described as follows:

- Real Image Database.** The database is obtained from the UCR Visualization and Intelligent Systems Laboratory (VISLab) [27]. There are 10,038 images covering a variety of real-world scenes such as castles, cars, humans, animals, etc. In order to generate the ground-truth for each of the images, three persons are asked to make independent judgment regarding the class assignment with 56 possible class names given to them. Each image is assigned to the appropriate class when the majority of these three persons agreed with this assignment. For the images which are assigned to three distinct classes by the three persons, the prelabeling provided by the image processing and computer vision expert is used as the ground-truth. The number of images in each class varies from 20 to 695. The histogram of this distribution is shown in Fig. 11. The sample images from each class are shown in Fig. 12. Each image in the database is represented by a 22-dimensional feature vector having 16 Gabor features [10] (the mean and standard deviation of filtered images at 4 orientation and 2 scales) and six color features (the mean and standard deviation from the HSV color domain [28]).
- Experiment 1: Integration of reinforcement learning with relevance feedback.** The first experiment is conducted to show that the integration of multiple existing RF approaches using the proposed IRRL model provides better retrieval performance than using the same RF approach at all iterations. We experiment with 200,000 random queries and compute the average precision of each competing method. The first three rows of Table 3 report the average precision using each existing short-term RF approach alone. It is seen that the three RF approaches have comparable performances. The fourth row gives the average retrieval performances of the three methods and the results will be referred to as baseline short-term precisions for assessing the



Fig. 12. Real image database: sample images from each class (arranged from left to right and from top to bottom).

TABLE 3
Comparative Performances of the Proposed IRRL Model with the Existing RF Approaches

Approaches and improvement ratios	PR0 (CPU time)	PR1 (CPU time)	PR2 (CPU time)
QVM with short-term learning	25.41% (0.015)	40.42% (0.049)	41.95% (0.049)
FRE with short-term learning	25.41% (0.016)	40.40% (0.016)	42.68% (0.016)
BI with short-term learning	25.41% (0.015)	41.25% (0.140)	41.67% (0.146)
Average (as baseline short-term)	25.41%	40.69%	42.10%
QVM with long-term learning	28.70% (0.053)	41.85% (0.052)	42.14% (0.053)
FRE with long-term learning	30.61% (0.016)	45.29% (0.017)	45.88% (0.017)
BI with long-term learning	27.62% (0.147)	44.25% (0.293)	44.36% (0.307)
Average (as baseline long-term)	28.98%	43.80%	44.13%
IRRL with short-term learning	25.41% (0.057)	53.57% (0.074)	59.61% (0.128)
Improvement ratio to baseline short-term	0%	31.65%	41.59%
IRRL with long-term learning	48.72% (0.055)	60.90% (0.074)	61.96% (0.124)
Improvement ratio to baseline long-term	68.12%	39.04%	40.40%



Fig. 13. Retrieval result without relevance feedback using IRRL with short-term learning scheme, three images are identified as relevant (images 1, 2, and 4), the others are identified as nonrelevant. The precision is 3/10.

performance of the proposed model. Also, we add long-term mechanism to each existing RF approach by storing the parameter values derived from historic sessions. As seen from the next three rows in Table 3, the long-term learning can improve the retrieval performance even if the researcher works on making the single favorite RF technique better. The baseline long-term retrieval performances are also given. Further, we apply separately the short-term learning scheme and the long-term learning scheme for cross-session integration in the proposed IRRL model. Since the relevance concepts of a particular user rarely change during the same query session, we skip the nonrelevant images which have been annotated in the same query session. The average precision rates and the improvement ratios to the baseline precisions are shown at the bottom of Table 3. The improvement ratio is defined as the increment on the precision rate divided by the corresponding baseline precision. It is seen that the proposed model using either cross-session integration schemes obtains higher retrieval precision than the baseline performance of using a single RF approach at all feedback iterations. This hints us that the development of a good RF integration model is as important as that of a new RF approach. The improvement ratio of the proposed model ranges from 31.65 percent to 41.59 percent for the short-term learning, and from 39.04 percent to 68.12 percent for the long-term learning. Note that there is no improvement ratio gain on PR0 for the short-term learning since no relevance knowledge is shared between multiple query sessions. However, the long-term learning is useful for improving the retrieval performances at all iterations, and is

particularly more significant in improving PR0. This is because the previous relevance knowledge is retrieved and used when the algorithm computes the first retrievals of a query session. Further, there is no RF strategy facilitated with either combination or hybridization schemes is superior to the others for all queries. An RF strategy is best suited to only a certain class of images. This implies that the within-session relevance learning is image-dependent.

Table 3 also shows the processing times (in seconds) of all the approaches used. It is observed that BI is the most computationally expensive method as expected, FRE is the fastest, while the IRRL model has moderate CPU time cost because it conducts alternative RF techniques and will not resort to a single approach. The results also reveal that the additional processing time for Q value updating used in the IRRL model is negligible.

Fig. 13 shows the first retrieval result of a particular query session obtained by IRRL with short-term learning, that is, no relevance information is retrieved at the beginning of this session. All the retrieved images are sorted in increasing order of their Euclidean distances from the query, and the first retrieved image is also the query image itself. With the human-labeled ground-truth, three images are identified as relevant (a precision rate of 30 percent), and the others as nonrelevant. Fig. 14 shows the retrieval result using the short-term IRRL based on the feedbacks given in Fig. 13, a retrieval precision of 50 percent (five images are identified as relevant) is achieved. On the other hand, we also submit the same query image to the long-term IRRL agent, that is, the cross-session relevance knowledge is retrieved at the beginning of this query session. Fig. 15 shows the retrieval result after the first feedback iteration, a



Fig. 14. Retrieval result after the first relevance feedback using IRRL with short-term learning scheme, five images are identified as relevant (images 1, 2, 3, 4, and 6), the others are identified as nonrelevant. The precision is 5/10.



Fig. 15. Retrieval result after the first relevance feedback using IRRL with long-term learning scheme, seven images are identified as relevant (images 1, 2, 3, 4, 5, 6, and 9), the others are identified as nonrelevant. The precision is 7/10.



Fig. 16. Retrieval result without relevance feedback using IRRL with long-term learning scheme, five images are identified as relevant (images 1, 2, 3, 6, and 8), the others are identified as nonrelevant. The precision is 5/10.



Fig. 17. Retrieval result after the first relevance feedback using IRRL with long-term learning scheme, seven images are identified as relevant (images 1, 2, 3, 4, 5, 8, and 10), the others are identified as nonrelevant. The precision is 7/10.

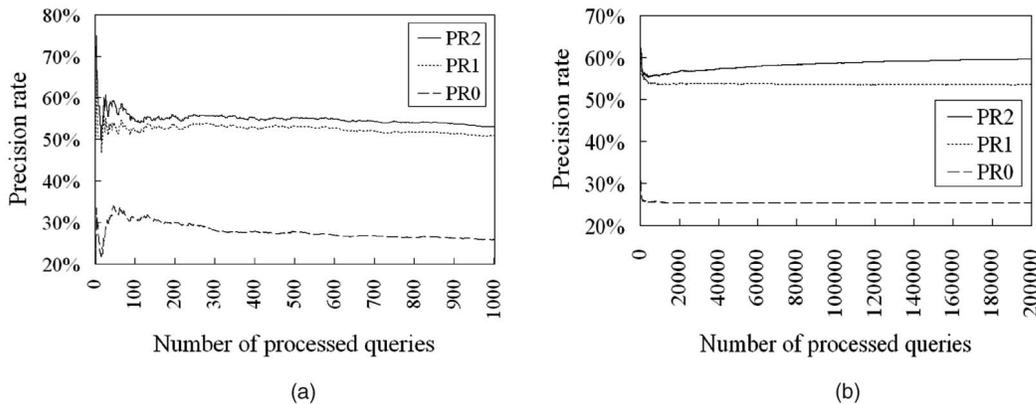


Fig. 18. Accumulated average precision rates versus the number of processed queries using the proposed IRRL model with short-term learning scheme. (a) The accumulated average precision rates fluctuate during early query sessions from 0 to 1,000. (b) The accumulated average precision rates improve with the number of processed queries except PR0.

higher retrieval precision of 70 percent (seven images are identified as relevant) is achieved as compared to that obtained by short-term IRRL.

Fig. 16 shows the first retrieval result of another query session obtained by IRRL with long-term learning scheme. In this case, a relatively high precision rate of 50 percent (five relevant images) has been already achieved in the first retrieval result since long-term relevance learning provides valuable information for the initial query formulation. Fig. 17 shows the retrieval result based on the relevance feedback given in Fig. 16, it is observed that two additional relevant images are included at this iteration (a precision rate of 70 percent).

- **Experiment 2: Convergence analysis.** This experiment consists of two parts: convergence analysis of the retrieval precision and convergence analysis of

the probabilistic action selection rule. Fig. 18a shows the accumulated average precision rates of the proposed model versus the number of already processed queries from 0 to 1,000 using the short-term learning scheme, and Fig. 18b corresponds to that for all processed query sessions from 0 to 200,000. It is observed that, at the initial learning stage as shown in Fig. 18a, the accumulated average precision rates fluctuate depending on which images are initially selected as query images since some images may involve extremely high precision rates and some others may cause extremely low precision rates when the feedback history is too short. However, if we look at a longer period as shown in Fig. 18b, the plotted curves for the accumulated average PR1 and PR2 climb up as the number of queries that are already processed increases. This is because the IRRL agent

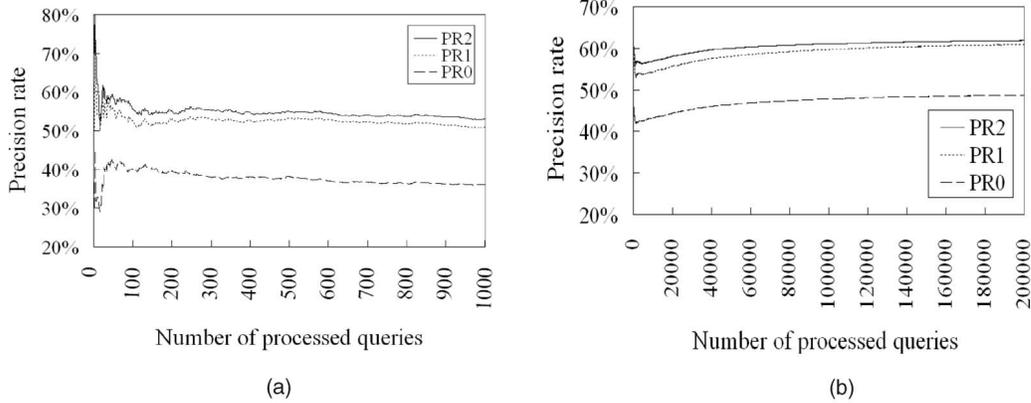


Fig. 19. Accumulated average precision rates versus the number of processed queries using the proposed IRRL model with long-term learning scheme. (a) The accumulated average precision rates fluctuate during early query sessions from 0 to 1,000. (b) The accumulated average precision rates improve with the number of processed queries. The PR0 is also improved due to the use of prior relevance knowledge.

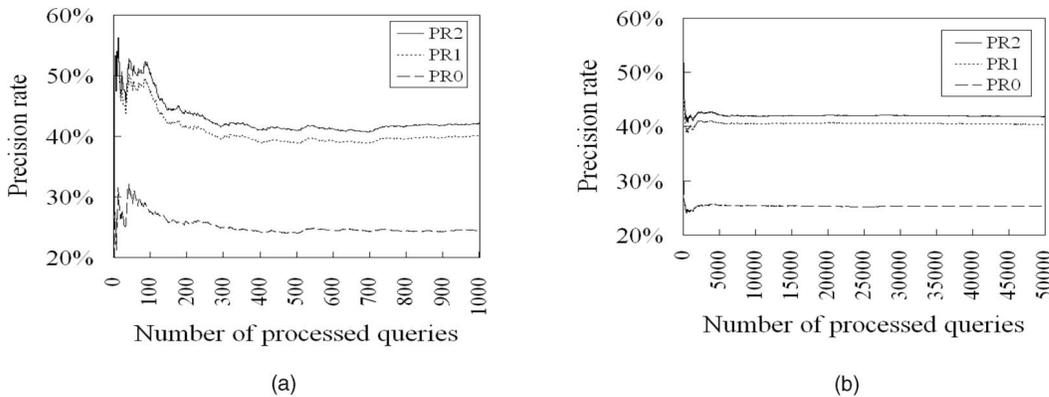


Fig. 20. Accumulated average precision rates versus the number of processed queries using the traditional QVM approach. (a) The accumulated average precision rates fluctuate during early query sessions from 0 to 1,000. (b) The accumulated average precision rates stay at fixed values since no relevance knowledge is exploited across sessions.

can approximate the optimal RF strategy and find the best within-session integration scheme for most query images. The slopes of these curves then become gentle and the precision rates finally converge to some upper bounds when little new relevance information can be learned and the performance is hardly improved further. Note that the accumulated average PR0 is not improved with the number of processed queries since no prior relevance information across sessions has ever been derived and utilized at the beginning of a query session. Fig. 19 shows the accumulated average precision rates of the proposed model versus the number of already processed queries using the long-term learning scheme. It reveals similar phenomena as those shown in Fig. 18 except that the accumulated average PR0 is also improved with the number of processed queries due to the use of prior relevance knowledge derived across sessions. On the other hand, Fig. 20 corresponds to those obtained using the traditional QVM approach at all iterations. It is seen that all the accumulated average precision rates are almost fixed no matter how many query images have been processed. This is because that no experiences from processed queries are utilized and that the same QVM approach is used at various feedback iterations for every query image.

For the convergence analysis of the probabilistic action selection rule, we compute the probability for choosing the actions with maximum Q values as the number of experienced query sessions increases. As shown in Fig. 21, the convergence probability of the action selection rule is increasing as the system experiences more queries. The curve reaches 55 percent after experiencing 50,000 query sessions, which means more than half of the users have consensus behavior with the action selection rule performed by the IRRL model.

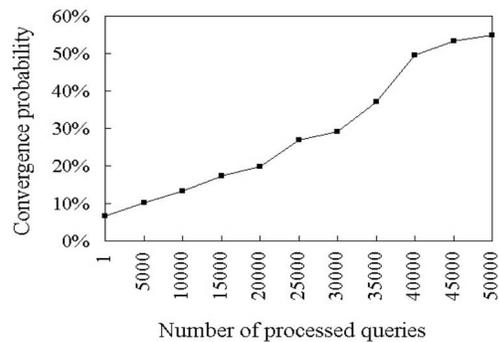


Fig. 21. Convergence probability of the action selection rule.

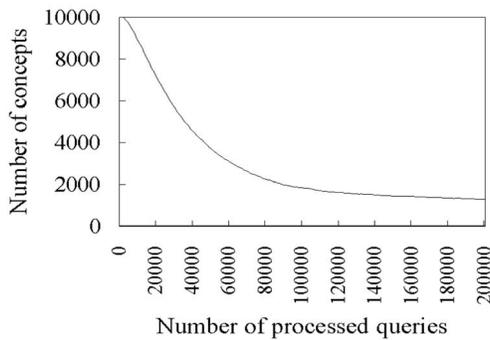


Fig. 22. Concepts learned versus processed queries.

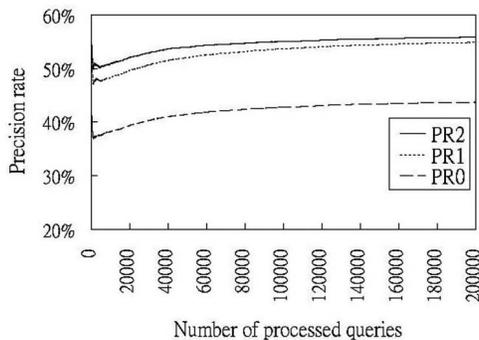


Fig. 23. Accumulated average precision rates versus the number of processed queries using the long-term IRRL model with the concept digesting method.

- Experiment 3: Demonstration of concept digesting method.** The third experiment is intended to show that the concept digesting method can be applied to reduce the storage demand incurred by the within-session relevance learning. Fig. 22 shows the number of concepts digested with an entropy threshold $e = 1.0$ versus the number of already processed queries. When the database is just created, the number of concepts is equivalent to the number of images (10,038) in the database since every image corresponds to a nondetermined concept. As the IRRL agent experiences with more query sessions and digests determined concepts by merging many nondetermined concepts with the same dominant strategy, the number of total concepts decreases. Finally, when the system has dealt with 200,000 query sessions, the number of concepts is reduced to 1,296 which is only about 13 percent of the original number. Since the IRRL agent keeps a separate Q table for every concept, the storage demand of within-session relevance learning is also reduced to 13 percent of its original complexity. Fig. 23 shows the accumulated average precision rates of the long-term IRRL model with the concept digesting method applied. It is seen that the precision rates converge to PR0 = 43.71 percent, PR1 = 54.87 percent, and PR2 = 55.86, which are very close to the precision rates obtained without applying the concept digesting method (PR0 = 48.72 percent, PR1 = 60.90 percent, and PR2 = 61.96, as shown in Table 3). Thus, the IRRL agent is suited to work with a dynamic database and is able to perform relevance learning for newly added images.

5 CONCLUSIONS

Most researchers in the relevance learning community strive to develop a new relevance feedback approach. However, they ignore the possible synergetic contribution due to the integration of multiple existing techniques. In this paper, we have proposed an image relevance reinforcement learning model that learns the optimal strategy for selecting the right relevance feedback technique, at the right iteration, for a given query image. A long-term learning scheme has been presented to derive the prior relevance knowledge about a query, such that, the relevance learning can start from the preceding state. The average precision rates obtained using the proposed model are significantly higher than those obtained using the traditional methods with and without long-term learning. The improvement ratio ranges from 31.65 percent to 68.12 percent at various feedback iterations. To save the storage demand, we also proposed a concept digesting method. Experimental results manifest that the storage demand can be reduced to 13 percent of its original complexity. Hence, the proposed model is scalable to a large dynamic database.

REFERENCES

- M. Flickner and H. Sawhney, "Query by Image and Video Content: The QBIC System," *Computer*, vol. 28, no. 9, pp. 23-32, Sept. 1995.
- Pattern Recognition*, J.C.M. Lee and A.K. Jain eds., special issue on image database, vol. 30, no. 4, 1997.
- Y. Rui, T.S. Huang, and S.F. Chang, "Image Retrieval: Current Techniques, Promising Directions, and Open Issues," *J. Visual Comm. and Image Representation*, vol. 10, no. 1, pp. 39-62, 1999.
- 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, Dec. 2000.
- B. Moghaddam, H. Biermann, and D. Margaritis, "Regions-of-Interest and Spatial Layout for Content-Based Image Retrieval," *Multimedia Tools and Applications*, vol. 14, no. 2, pp. 201-210, 2001.
- J.J. Rocchio Jr., "Relevance Feedback in Information Retrieval," *The SMART System*, G. Salton, ed., pp. 313-323, Prentice Hall, 1971.
- E. Ide and G. Salton, "Interactive Search Strategies and Dynamic File Organization in Information Retrieval," *The SMART System*, G. Salton, ed., pp. 373-393, Prentice Hall, 1971.
- Y. Rui, T.S. Huang, and S. Mehrotra, "Relevance Feedback Techniques in Interactive Content-Based Image Retrieval," *Proc. SPIE Conf. Storage and Retrieval for Image and Video Databases*, pp. 25-36, 1998.
- 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.
- J. Peng, B. Bhanu, and S. Qing, "Probabilistic Feature Relevance Learning for Content-Based Image Retrieval," *Computer Vision and Image Understanding*, vol. 75, nos. 1-2, pp. 150-164, 1999.
- S. Aksoy, R.M. Haralick, F.A. Cheikh, and M. Gabbouj, "Weighted Distance Approach to Relevance Feedback," *Proc. Int'l Conf. Pattern Recognition*, pp. 812-815, 2000.
- I. Cox, M. Miller, T. Minka, S. Omohundro, and P. Yianilos, "Target Testing and the Pichunter Bayesian Multimedia Retrieval System," *Proc. Research and Technology Advances in Digital Libraries*, pp. 66-75, 1996.
- C. Meilhac and C. Nastar, "Relevance Feedback and Category Search In Image Database," *Proc. Int'l Conf. Multimedia Computing and Systems*, pp. 512-517, 1999.
- I. Cox, M. Miller, T. Minka, T. Papathomas, and P. Yianilos, "The Bayesian Image Retrieval System, Pichunter: Theory, Implementation, and Psychophysical Experiments," *IEEE Trans. Image Processing*, vol. 9, no. 1, pp. 20-37, 2000.
- B. Bhanu and A. Dong, "Exploitation of Meta Knowledge for Learning Visual Concepts," *Proc. IEEE Workshop Content-Based Access of Image and Video Libraries*, pp. 81-88, 2001.

- [16] B. Bhanu and A. Dong, "Concept Learning with Fuzzy Clustering and Relevance Feedback," *Eng. Applications and Artificial Intelligence*, vol. 15, pp. 123-138, 2002.
- [17] P.Y. Yin, B. Bhanu, K.C. Chang, and A. Dong, "Improving Retrieval Performance by Long-Term Relevance Information," *Proc. Int'l Conf. Pattern Recognition*, vol. 3, pp. 533-536, 2002.
- [18] R.S. Sutton and A.G. Barto, *Reinforcement Learning: An Introduction*. MIT Press, 1998.
- [19] L.P. Kaelbling and A.W. Moore, "Reinforcement Learning: A Survey," *J. Artificial Intelligence Research*, vol. 4, pp. 237-285, 1996.
- [20] J. Peng and B. Bhanu, "Delayed Reinforcement Learning for Adaptive Image Segmentation and Feature Extraction," *IEEE Trans. System, Man, and Cybernetics—Part C*, vol. 28, no. 3, pp. 482-488, 1998.
- [21] J. Peng and B. Bhanu, "Closed-Loop Object Recognition Using Reinforcement Learning," *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol. 20, no. 2, pp. 139-154, Feb. 1998.
- [22] D. Heesch and S. Ruger, "Relevance Feedback for Content-Based Image Retrieval: What Can Three Mouse Clicks Achieve," *Proc. European Conf. Information Retrieval Research*, 2003.
- [23] J. Lim, J. Wu, S. Singh, and D. Narasimhalu, "Learning Similarity Matching in Multimedia Content-Based Retrieval," *IEEE Trans. Knowledge and Data Eng.*, vol. 13, pp. 846-850, 2001.
- [24] J. Smith, S. Srinivasan, A. Amir, S. Basu, G. Iyengar, C. Lin, M. Naphade, D. Ponceleon, and B. Tseng, "Integrating Features, Models, and Semantics for TREC Video Retrieval," *Proc. Text REtrieval Conf. (TREC 2001)*, pp. 240-249, 2001.
- [25] S. Aksoy and R. Haralick, "Classification Framework for Content-Based Image Retrieval," *Proc. Int'l Conf. Pattern Recognition*, pp. 240-249, 2002.
- [26] C. Watkins and P. Dayan, "Q-Learning," *Machine Learning*, vol. 8, pp. 279-292, 1992.
- [27] Visualization and Intelligent Systems Lab (VISLab), Univ. of California, Riverside, <http://www.vislab.ucr.edu>, 2005.
- [28] K.R. Castleman, *Digital Image Processing*. Prentice-Hall, 1996.



Bir Bhanu (S'72-M'82-SM'87-F'95) received the SM and EE degrees in electrical engineering and computer science from the Massachusetts Institute of Technology, Cambridge, the PhD degree in electrical engineering from the Image Processing Institute, University of Southern California, Los Angeles, and the MBA degree from the University of California, Irvine. He was the founding professor of electrical engineering and served its first Chair at the University of California, Riverside (UCR). He has been the Cooperative Professor of Computer Science and Engineering and Director of Visualization and Intelligent Systems Laboratory (VISLab) since 1991. Currently, Dr. Bhanu also serves as the founding director of an interdisciplinary center for research in intelligent systems (CRIS) at UCR. Previously, he was a Senior Honeywell Fellow at Honeywell Inc. in Minneapolis, Minnesota. He has been on the faculty of the Department of Computer Science at the University of Utah, Salt Lake City, and has worked at Ford Aerospace and Communications Corporation, California, INRIA-France and IBM San Jose Research Laboratory, California. He has been the principal investigator of various programs for DARPA, NASA, NSF, AFOSR, ARO, and other agencies and industries in the areas of learning and vision, image understanding, pattern recognition, target recognition, biometrics, navigation, image databases, and machine vision applications. He has received two outstanding paper awards from the Pattern Recognition Society and has received industrial and university awards for research excellence, outstanding contributions and team efforts. He has been on the editorial board of various journals and has edited special issues of several IEEE transactions (*PAMI*, *SMC*, *R&A*, *IP*) and other journals. He holds 11 US and international patents and over 230 reviewed technical publications in the areas of his interest. He is a fellow of the IEEE, AAAS, IAPR, and SPIE.



Kuang-Cheng Chang received the BS degree in information management from Ming Chuan University, Taoyuan, Taiwan, and is currently pursuing the MBA degree in information management at National Chi Nan University, Nantou, Taiwan. His research interests include pattern recognition, machine learning, software engineering, and bioinformatics.



Anlei Dong received the PhD degree in electrical engineering from the University of California at Riverside (UCR) in 2004. Previously, he received the BS and MS degrees in automation from the University of Science and Technology of China. His research interests are machine learning and computer vision, including object detection, image retrieval, representation, understanding, indexing, and recognition.

▷ For more information on this or any other computing topic, please visit our Digital Library at www.computer.org/publications/dlib.



Peng-Yeng Yin received the BS, MS, and PhD degrees in computer science from National Chiao Tung University, Hsinchu, Taiwan. From 1993 to 1994, he was a visiting scholar at the Department of Electrical Engineering, University of Maryland, College Park, and the Department of Radiology, Georgetown University, Washington DC. In 2000, he was a visiting professor in the Visualization and Intelligent Systems Laboratory (VISLab) at the Department of Electrical Engineering, University of California, Riverside (UCR). From 2001 to 2003, he was a professor at the Department of Computer Science and Information Engineering, Ming Chuan University, Taoyuan, Taiwan. Since 2003, he has been a professor of the Department of Information Management, National Chi Nan University, Nantou, Taiwan, and is currently the chairman of the Department there. He received the Overseas Research Fellowship from Ministry of Education in 1993, Overseas Research Fellowship from National Science Council in 2000. He has received the best paper award from the Image Processing and Pattern Recognition Society of Taiwan. He is a member of the Phi Tau Phi Scholastic Honor Society and listed in *Who's Who in the World*. His current research interests include pattern recognition, content-based image retrieval, relevance feedback, machine learning, computational intelligence, and computational biology.