World Academy of Science, Engineering and Technology
International Journal of Computer and Information Engineering
Vol:1, No:7, 2007

Research on the Relevance Feedback-based Image Retrieval in Digital Library

Rongtao Ding, Xinhao Ji, and Linting Zhu

***Abstract*—**In recent years, the relevance feedback technology is regarded in content-based image retrieval. This paper suggests a neural networks feedback algorithm based on the radial basis function, coming to extract the semantic character of image. The results of experiment indicated that the performance of this relevance feedback is better than the feedback algorithm based on Single-RBF.

***Keywords*—**Image retrieval; Relevance feedback; Radial basis function.

I. INTRODUCTION

THERE are massive information resources in the digital library, contains many multimedia information such as the words, the images, the sound, the video, and so on. Therefore, how to search for the information needed in the digital library quickly and effectively is an important research direction of content-based image retrieval system. In the retrieval process, user first choose a character vector or some character vectors according to the aim and the request of retrieval, then determines the weight of each character based on their different importance in the retrieval. This requests user has rich knowledge to the image character expression, but the average user certainly does not have the specialized image knowledge, and can bring some difficulties to the retrieval.

Relevance feedback [1] is a powerful method to enhance the system search effect. It studies from the real interactive process of the user and the search system, then discovers and captures user's actual search intention, and modifies the search strategy of system, thus obtains the search result which tallies as precise as possible with the user's actual demand. Image retrieval based on relevance feedback is an unceasingly repeated and gradually advanced processes, the interaction between the system and the user enables the retrieval to approach the user's expectation, and finally achieves the requests.

Rongtao Ding (1973), male, is a lecturer in Zhejiang Vocational College of Commerce. His current research interests include intelligent teaching, multimedia technology and E-commerce. His research takes place in several projects (phone: 13018961726, e-mail: rongtaoding@gmail.com).

Xinhao Ji (1979), male, is an engineering master of Zhejiang University of Technology and a lecturer in Zhejiang Vocational College of Commerce. His current research interests include computer information processing and image processing. His research takes place in several projects (phone: 13305715009, e-mail: jxh@zjvcc.cn).

Linting Zhu (1982), female, is an engineering master candidate at Zhejiang University, and a lecturer in Zhejiang Vocational College of Commerce. Her current research interests include E-commerce and computer information processing. Her research takes place in several projects (phone: 13588818168, e-mail: jannetty@163.com).

II. RELEVANCE FEEDBACK TECHNOLOGY

A. Main Idea of Relevance Feedback

The relevance feedback mechanism had been introduced into the image retrieval system, and the interactive search technology based on the relevance feedback appeared. The relevance feedback technology adjusts the search automatically according to the user's relevant feedback of the preceding retrieval result, enables the adjusted retrieval to approach the user's expectation much more.

The basic steps of user's relevance feedback are as follows [2]:

1) System search for the demonstration image given by user;
2) The user compares the retrieval result returned from the system with own demands, and appraises their pertinence.
3) System analyzes the character which can indicate user's retrieval aim best automatically from the feedback information produced by the user, adjusts the similarity method, then carries on the retrieval again, repeats the step 2).

Relevance feedback technology is the most active domain in current image retrieval research, the relevance feedback method is connected knowledge that joins the substrate character and the high level semantics into user's feedback, only this knowledge is provided by user. The aim of relevance feedback is to study from the real interaction between the user and the retrieval system, discover and capture user's actual search intention, and modify the search strategy of system, thus obtains the search result which tallies as precise as possible with the user's actual demand. The relevance feedback can real-time revise the inquiry strategies, thus can increase the auto-adapted function for the image retrieval system.

B. Image Retrieval Mode of Relevance Feedback l Stage

To analyze and extract the characteristic needed in the content-based image retrieval, we also consider the relevant problem and the user relevance feedback problem therefore must use a set of reasonable retrieval models to carry on the quantification to it, then make the retrieval results may rely on. We explain the image retrieval model based on relevance feedback as follows:

First, we should define the object image models; an object image model I may be represented as:

$$I = I(D, F, R) \tag{1}$$

D is the primitive image data, such as the image in JPEG form, etc. $\mathrm{F}=\{f_t\}$ is the substrate characteristic set of this object image, these characteristics including the color, the texture and the shape, etc. $R=\{r_y\}$ is the expression of a certain characteristic f_t, the color histogram and the color matrix are the expression way of color characteristic, each characteristic expression r_y is possibly a vector which is composed by many components, can be written in the following form:

$$r_y = \left\{r_{y1}, r_{y2}, \ldots, r_{yk}\right\} \tag{2}$$

K is the vector dimension.

For expressing the abundant content of the image fully, this object model permit carrying on many characteristics expression to the description of the picture, and each characteristic has a corresponded dynamic weight value. The image characteristic weight value exists in each level of the model, W_i, W_{ij} and W_{ijk} separately corresponds to the image characteristic f_i, the characteristic expression r_{ij}, and each component of the characteristic expression r_{ijk}. The aim of relevance feedback is to search for the proper weight value that can manifest user's information exactly. An object image model $I\left(D,F,R\right)$ and a group of similarity measure algorithms $M=\left\{m_{ij}\right\}$ constitute $CBIR$ model $\left(D,F,R,M\right)$, and the similarity algorithm M is used for calculating the similarity between image objects.

The image retrieval process based on relevance feedback can be described as follows:

1) Initialize all weight values $W=\left\{W_i,W_y,W_{yk}\right\}$ into WO, WO is a group of agonic weight values, it makes all characteristics and the characteristic component has the same weight.

$$\begin{aligned} W_i &= WO_i = \frac{1}{L} \\ W_{ij} &= WO_{ij} = \frac{1}{J_i} \\ W_{ijk} &= WO_{ijk} = \frac{1}{K_{ij}} \end{aligned} \tag{3}$$

L is the number of image characteristic, J_i is the number of expression form of the characteristic f_i, and K_{ij} is the dimension of characteristic vector r_{ij}.

2) Divide the search object Q provided by the user into a group of image characteristics f_i according to the weight W_i, and each characteristic f_i can be divided into the corresponding characteristic expression r_{ij} according to the weight W_{ij}.

3) In a certain characteristic expression r_{ij}, the similarity between the image I and the demonstrative image Q is calculated according to the corresponding similarity algorithm m_{ij} and the weight value W_{ijk}.

$$S\left(r_{ij}\right)=m_{ij}\left\{r_{ij},W_{ijk}\right\} \tag{4}$$

The similarity between the image I and the demonstrative image Q in a certain image characteristic f_i is obtained by merging the similarity of each characteristic expression:

$$S\left(f_{ij}\right)=\sum_j W_{ij} S\left(r_{ij}\right) \tag{5}$$

4) The total similarity S between the image I and the demonstrative image Q can be given by merging each $S\left(f_{ij}\right)$:

$$S=\sum_i W_{ij} S\left(f_i\right) \tag{6}$$

5) Calculate the total similarity between the demonstrative image Q and all images in the database, and arrange images according to the similarity, then return the first N images which are most similar to the image user needed, N is the number of images user needs.

6) The user judges the relativity between each image returned from the system and the query according to own search demands and the subjective opinion. The relativity can be divided into five kinds, extremely related, related, no judgment, not related and extremely unrelated.

7) The system adjusts the weight value according to user's feedback opinion, and skips to the step 2).

C. *Update the Weight Value according to User's Feedback*

There are some key processes affecting the performance of the content-based image retrieval system, such as: the choice of the vision characteristic, the computational method of weight, the expression form and the adjustment of the search vector, as well as similarity measure method. After the system had determined the characteristic index form and the visual characteristic vector, the enhancement of the performance mainly relies on the adjustment of search and the renewal method of the weight. In the computer-based retrieval system, the expression and the weight value of image vision characteristic is definite, and in relevance feedback-based retrieval system, it is necessary to update the weight value dynamically, and express the rich image content by many kinds of characteristics expressions.

W_{ij} is corresponded to the weight of different characteristic vector, and it reflects the different attention to various characteristics in total similarity, the adjustment to W_{ij} may follow the formula hereinafter according to user's relevance feedback.

Suppose that T is the aggregate of the first N images, which are the most similar ones that determined by the total similarity S , S_i is the grade of image I appraised by user.

$$S_i = \begin{cases} 3 & Extremely-related \\ 1 & Related \\ 0 & \text{No judgment} \\ -1 & Not-related \\ -3 & \text{Extremely unrelated} \end{cases} \tag{7}$$

For each r_{ij}, set T_{ijo} be the most similar image to the search image, which determined according to $S\left(r_{ij}\right)$, set $W_{ij}=0$, and then adjusts the weight as follows:

$$W_{ij} = \begin{cases} W_{ij}+s_i & \text{if } T_{ijo} \in T_M \\ W_{ij} & Others \end{cases} \tag{8}$$

The weight related to r_{ijk} has reflected the contribution to expression vector r_{ij} from r_{ijk}. Adjustment to the weight W_{ijk} can base on the standard deviation.

Suppose that there are M images in the database, put the expression vectors r_{ij} of extremely related images together into a $M \times K$ matrix, and each column of this matrix is a r_{ijk} sequence. The reciprocal of standard deviation of this sequence is preferable estimation to weight, $W_{ijk} = 1/\partial_{ijk}$.

III. NN Feedback Algorithm based on the Radial Basis Function

A. Single-RBF Feedback Algorithm

The structure of RBF nerve network is simple, and it can approach to any discretional nonlinear function. In order to simulate how people distinguishing the similar image or not, Muneesawang used this merit of RBF nerve network [3], took the network as a nonlinear mapping function to scale the similarity between images. The network output is:

$$f\left(X\right)=\sum_{i=1}^{p} G_i\left(x_i, z_i\right)$$

$$=\sum_{i=1}^{p} \exp\left(-\frac{\left(x_i - z_i\right)}{2\partial_i^2}\right) \tag{9}$$

$Z=\left[z_1,\ldots,z_i,\ldots,z_p\right]^T$ is the central vector of RBF network, also is the inquiry characteristic vector of the demonstration image, $\partial_i, i=1,\ldots,p$ is the variance, $X=\left[x_1,\ldots,x_i,\ldots,x_p\right]^T$ is the characteristic vector in the characteristic database that corresponds to an image. The kernel function $G_I\left(x_i, z_i\right)$ adopts the Gauss function, and its focus is a single component of inquiry characteristic vector, this structure can be called the Single-RBF network. To the output of the function, $f\left(x\right)$ express the similarity between the input characteristic vector X and the inquiry characteristic vector Z , each RBF concealed cell react partial similarity estimation on the non-linear input-output mapping of the entire network. Obviously, the value is the biggest when $X=Z$.

In formula (9), the Single-RBF network needs to determine two parameters: central vector Z and variance ∂_i. According to user's feedback, the incessant revision to these two parameters can improve the retrieval performance of system.

Because the inquiry characteristic vector Z reflect the user expected image, the inquiry characteristic vector Z can be revised by user's feedback, and cause this characteristic vector approach to similar image [4]. It can be revised in following formula.

$$Z_q\left(t+1\right)=\overline{X}^{'}-\partial_N\left(\overline{X}^{''}-Z_q\left(t\right)\right) \tag{10}$$

$$\overline{X}^{'}=\frac{1}{M}\sum_{m=1}^{M} X_m^{'} \tag{11}$$

$$\overline{X}^{''}=\frac{1}{N}\sum_{n=1}^{N} X_n^{'} \tag{12}$$

$\left\{X_m^{'}\right\}_{m=1}^{M}$ is the collection of negative feedback example and $\left\{X_n^{''}\right\}_{n=1}^{N}$ is the collection of positive feedback example, $\overline{X}^{'}$ and $\overline{X}^{''}$ are the average value vectors of these two sample collections, reflected the distributed situation in the characteristic space. Formula (10) makes new inquiry vector $Z_q\left(t+1\right)$ toward to the positive feedback example, and departure from the negative feedback example, ϑ^{N} controls the moving speed.

The variance ∂_i emphasizes the characteristic vector that is highly related to the inquiry vector, and neglects slightly related characteristic vector. After obtaining a new inquiry vector, the variance ∂_i to be revised by the collection of positive feedback example:

$$\partial_i = \eta \max_m \left[x_{mi}^{w} - z_i \right] \tag{13}$$

x_{mi}^{w} is the ith component of the mth characteristic of the collection of positive feedback example, z_i is the ith component of the inquiry vector. Therefore, when some component of the characteristic vector deviates the central inquiry vector, the output of corresponded concealed cell become smaller, and the influence on the output of total similarity to the entire network become smaller, even may be neglected. η makes the output of concealed cell that used the Gauss function in a reasonable range, generally is 3.

B. The RBF NN Feedback Algorithm based on the Gauss Function

The Single-RBF feedback algorithm simulates the people's discretion about the image similarity by using a kind of simplification of radial basis nerve network structure based on the Gauss function, but it has two limitations in the experimental process: 1, the choice of parameter σ_N brings tremendous influence to the final result; 2, this structure adopts the one-dimensional Gauss function to be the kernel function, and it can not simulate the people's discretion about the image similarity well. We advanced the RBF nerve network feedback algorithm based on the Gauss function. The experiments have indicated that, the performance of this algorithm has been enhanced. This algorithm also adopted a kind of simplification of radial basis nerve network structure, and used the multi-dimensional Gauss function to be the kernel function. The network output is:

$$f(x) = \sum_{m=1}^{M} G_m(X, V_m, \partial_m) \tag{14}$$

$$G_m(X, V_m, \partial_m) = \exp\left(-\frac{(X - V_m)^T (X - V_m)}{2\partial_m^2} \right) \tag{15}$$

$G_m(X, V_m, \partial_m)$ is the output of mth Gauss kernel function, V_m and ∂_m are the central vector and the variance. We can modify these parameters by using user's feedback information, and improve the performance.

In formula (15), each component of the characteristic vector has the same weight, and has the consistent importance in the similarity comparison process. We may know from the adjustment to characteristics weight in section II-B, if the value of a certain component of all the corresponded images is extremely approach to each other; the image characteristic corresponded to this component reflected user's inquiry well, and should be given a higher weight. On the contrary, if the value differs from each other widely, the image characteristic corresponded to this component unable to reflect user's inquiry well, and should be given a lower weight, and take the reciprocal of standard deviation of a certain component sequence of this characteristic vector to be the estimation of weight. Formula (15) is unable to satisfy requests, and adopt the ellipse radial function proposed by Rui:

$$\phi(X, V_m) = \sum_{p=1}^{p} \alpha_p \left(x_p - v_p \right)^2 \tag{16}$$

α_p is the correlated weight, it can be obtained from the standard deviation ξ_p of the collection of positive feedback example $\left\{ X_m^{w} \right\}_{m=1}^{M}$.

$$\alpha_p = \begin{cases} 1 & \xi_p = 0 \\ \dfrac{1}{\xi_p} & \xi_p > 0 \end{cases} \tag{17}$$

The output of the mth Gauss kernel function $G_m(X, V_m, \partial_m)$ is:

$$G_m(X, V_m, \partial_m) = \exp\left(-\frac{\phi(X, V_m)}{2\partial_m^2} \right) \tag{18}$$

After the user has submitted the feedback information first time, the system must initialize the network structure, and initialize the M examples of the collection of positive feedback example $\left\{ X_m^{w} \right\}_{m=1}^{M}$ into the central vector of concealed neuron, i.e. the number of concealed neuron is M. In the following feedback process, we profited from the study vector quantification in the pattern recognition, modify each neuron center by using the negative feedback example. Because in the characteristic database, the negative feedback example is the closest characteristic vector to the inquiry the vector and the characteristic vector of positive feedback example. Suppose that a certain vector $X_n^{w}(t)$ of the collection of negative feedback $\left\{ X_n^{w} \right\}_{n=1}^{N}$ is closest to the central vector of the mth neuron V_i in the characteristic space, we can modify the central vector as follows:

$$V_i(t+1) = V_i(t) - \eta(t)\left[X_n^w(t) - V_i(t)\right] \tag{19}$$

$\eta(t)\left(0 \le \eta(t) \le 1\right)$ is the factor controls the study speed, and $\eta(t)$ is decreased progressively along with the iteration times. According to the CBIR, this paper adopts formula (20) to determine the value of $\eta(t)$, and determine the network study speed dynamically:

$$\eta(t) = \frac{p}{\text{p+n}} \tag{20}$$

P is the number of the positive feedback example, and n is the number of the negative feedback example. After had determined each central vector of the network concealed unit, we calculates the variance by formula (21):

$$\partial_m = 3 \bullet \max_i \left\| X_i^w - V_m \right\|, i = 1, \ldots, M \tag{21}$$

IV. Experiment Analysis

This article designed two systems to compare these two algorithms, system 1 adopts the Single-RBF algorithm (algorithm 1), and system 2 adopts the RBF nerve network feedback algorithm based on the Gauss function (algorithm 2), the steps of retrieval as follows:

1) User chooses the demonstrating image; the algorithm picks up its characteristic, and sets the values of the biggest correlated feedback times T and the biggest number of showed images.

2) Calculate the distances between the demonstrating image and each image in the characteristic database according to the Euclidean distance, then returns the first Q pictures that the distance is the smallest.

3) Set the feedback times t=0.

4) System 1 obtains the new values of the inquiry central vector and the variance using the formula (10) and formula (11); and system 2 obtains the new values of the inquiry central vector and the variance using the formula (19) and formula (21).

5) Then, system 1 calculates the similarity of modified inquiry vector and each image in the characteristic database using the formula (9); and system 2 using the formula (12) and formula (13).

6) Return the first Q images that have the high similarity.

7) The feedback times adds 1, i.e. t=t+1.

8) If t >T, change to step (9), otherwise, change to step (4).

9) End the retrieval.

In order to maintain the consistency of the experiments data, these two systems use the identical test image database, which has 1000 images. There are 10 categories of images, and stochastically chooses 5 images from each kind to be the inquiry image (altogether 50 images), and search for these images by these two systems separately. Set the iteration times T=5, and the number of returned image Q=20 and Q=40.

The retrieval stipulated that, whether two images similar to each other base on whether belongs to the identical category. This article used the precision to scale the performance of the algorithm, as (22) showed.

$$\text{Precision} = \frac{s}{r} \times 100\% \tag{22}$$

s is the number of similar images, and r is the number of returned images. Take the average of these 50 precisions as the final result. And count the results returned in 20and 40 images separately, showed in Table I and Table II.

Summarize from these two tables that, the performance of system 2 that adopts the RBF nerve network feedback algorithm based on the Gauss function is better than system 1 which adopts the Single-RBF algorithm under the condition of the same quantity images. Compare the results of returned in 20and 40 images, the retrieval performance of system 1 falls vastly along with the addition of returned images (from 84.6% to 78.95%), The retrieval performance of system 2 falls less (from 92.3% to 88.95%), and it supports the better retrieval performance.

TABLE I
The Precision Returned in 20 Images

Feedback method	0 Iter.	1 Iter.	2 Iter.	3 Iter.	4 Iter.	5 Iter.
System 1	65.2	79.2	81.9	82.3	83.1	84.6
System 2	65.2	86.5	88.4	90.4	91.5	92.3

TABLE II
The Precision Returned in 40 Images

Feedback method	0 Iter.	1 Iter.	2 Iter.	3 Iter.	4 Iter.	5 Iter.
System 1	61.3	76.44	77.65	78.2	78.5	78.95
System 2	61.3	83.6	85.4	86.5	88.6	88.95

V. Conclusion

The relevance feedback used for solving the “semantic gap” between the substrate characteristic of image and the high level semantics. This article emphatically studied the neural networks relevance feedback algorithm based on the radial basis function, and carried on the confirmation through the experiment. The performance of the neural networks relevance feedback algorithm based on the radial basis function is better than the feedback algorithm based on the Single-RBF. There is a limitation in the current relevance feedback algorithm that, the feedback information user submitted just influences the next retrieval. If the user wants to carry on the same retrieval again, the feedback information must be submitted to the system.

References

[1] Chen Wan- mi, Xie Zhi- bing, Fei Min- rui.Research on majorization strategies of vision system of RoboCup.IEEE, 2004: 372- 375.

[2] Feng Jing, Mingjing Li, Hong-Jiang Zhang, Bo Zhang. Support Vector Machines For Region-based Image Retrieval. Multimedia and Expo, 2003. ICME '03. Proceedings. 2003 International Conference on ,Volume: 2,6-9 July 2003: 21-24

[3] Paisarn Muneesawang,Ling Guan.An Interative Approach for CBIR Using a Network of Radial Basis Function,IEEE Trans on Multimedia.6(5),Oct.2004

[4] Tao Da- cheng, Tang Xiao- ou, Li Xue- long, et al.Asymmetric bagging and random subspace for Support Vector Machines-based relevance feedback in image retrieval[J].IEEE Transaction on Pattern Analysis and Machine Intelligence, 2006, 28(7).