An Image Retrieval Method Using Homogeneous Region and Relevance Feedback

Vu Van Hieu
Information Technology Faculty, Haiphong University, Haiphong, Vietnam
Email: hieuvv@fithpu.edu.vn

Nguyen Huu Quynh
Information Technology Faculty, Electric Power University, Hanoi, Vietnam
Email: quynhnh@epu.edu.vn

***Abstract*—Relevance feedback and region based image retrieval are two effective ways to improve accuracy in content-based image retrieval. In this paper, we propose a content-based image retrieval method using relevance feedback and homogeneous region. By extracting a number of homogeneous color regions from the image and calculating the occurrence frequency of regions, we convert image feature vectors to weighted vectors. On the basis of the weighted vectors, we calculate the similarity between two weighted vectors and using relevant feedback technique. Our experimental results on a Wang database of over 10,000 images suggest that the technique results in which is close to user's intention better than the CBsIR and CCH methods.**

***Index Terms*—content based image retrieval, weighted vectors, feature vectors, machine learning.**

I. INTRODUCTION

In digital decades, millions of images are stored in the huge database and on the Internet, find interesting images within a large collection of images requires a new approach. Most of the images not owned by us, therefore, we do not have the knowledge to support finding images of interest. If we search for images in the database is manually, we can search for the most desirable because image content recognition capabilities of the human is great (nothing can compare). However, the most challenges when retrieval big database with human is speed.

Many content-based image retrieval systems (CBIR) executable retrieval based on major feature global. Many time users access to CBIR system for search object but these system seem failure, by an indication is computed for the whole image that can't collect enough all the important properties of the separate objects. Content-based image retrieval systems based on region (RBIR) [1]-[4] trying to overcoming limitations of the global feature through representing images at object level closer to human perception [2].

However, pixels are sparsely scattered in the image so not became homogeneous regions (homogeneous regions contain adjacent pixels that have similar color values). The systems above use all the pixels in the image (include pixels are scattered) in retrieval process, so high computational complexity and sometimes the retrieval quality is not improved. That is why we are considered only to some homogeneous regions in matching process similar two images.

One of the interactive learning techniques is relevance feedback (RF) was developed for text retrieval [5]. Since the mid-1990s, relevance feedback (RF) has been proposed to CBIR to improve performance for retrieval systems [6]-[9]. The main idea of RF is to let users guide the system. In retrieval process, the users interact with the system and appreciate associated of images retrieved (according to subjective of users). With this additional information, system will learn on user attention and propose best results.

Field of Information Retrieval has been developed in a long time and many efficient information retrieval techniques were developed with relevance feedback. However, we did not take advantage of this effective technique in content-based image retrieval. That's why we apply integrated information retrieval technique to image retrieval. To do this, in addition to using the homogeneous regions of image, our proposed method conversion feature vector representing image to weights vector, then based on the method of cosine measures to compute the similarity of two weight vectors and use the model relevance feedback information retrieval.

The remaining of the paper is organized as follows: In Section 2, weight vector and relevance feedback techniques will be briefly reviewed. Feature extraction and image feature representation are described in Section 3. In Section 4, content-based image retrieval with relevance feedback is addressed in detail. Experimental results and conclusions will be given in Section 5 and Section 6 respectively.

II. RELEVANCE FEEDBACK AND WEIGHT VECTOR

A model most popular models for information retrieval is the vector space model [10]-[12], therefore, in this section we will present the weighted vectors to represent documents and compute similarity between two

Manuscript received July 12, 2014; accepted September 19, 2014.

doi: 10.12720/ijsps.3.1.14-18

documents. After obtaining the weight vector and computed similarity, relevance feedback technique will be combined to retrieval process to increase system performance.

A. Weighting Vector

Term weighting is a technique of assigning different weights for different keywords (different terms) according to their relative importance to the document [10], [12].

If we define w_{ik} to be the weight for term t_k, $k = 1, 2, ..., N$ in document i, where N is the number of terms, document i can be represented as a weight vector:

$$\vec{V}(d_i) = [w_{i1}; ...; w_{ik}; ...; w_{ik}] \quad (1)$$

To correctly estimate the weights, we need to consider two aspects. First, if term k is frequently occurred in the document i then w_{ik} should be assigned high value. This intuition suggests that a term frequency (tf) factor should be included in the estimation of w_{ik}. Second, tf alone can't ensure an acceptable estimation. When the high frequency term is not concentrated in a few documents but instead spreading over all documents, we should give this term low weight. This leads to the inverse document frequency (idf), which varies inversely with the number of documents in which a term appears.

$$idf_k = \log_2 \frac{M}{df_k} + 1 \quad (2)$$

where df_k is the document frequency for term k and M is the total number of documents in the collection. Experiments have shown that the product of tf and idf is a good estimation of the weights [10]-[12].

The query q has the same model as that of document d mean it is a weight vector in the term space:

$$\vec{V}(q) = [w_{q1}; ...; w_{qk}; ...; w_{qN}] \quad (3)$$

The similarity between $\vec{V}(d)$ and $\vec{V}(q)$ is defined as the Cosine distance

$$Sim(d,q) = \frac{\vec{V}(d)\vec{V}(q)}{\|\vec{V}(d)\|\,\|\vec{V}(q)\|} \quad (4)$$

where numerator is the product of two vectors $\vec{V}(d)$ and $\vec{V}(q)$, while the denominator is the product Euclidean length.

B. Relevance Feedback

As we can see from the subsection A, in the vector model, the specification w_{qk} in q is very critical since the similarity values Sim(d, q) are computed based on them. However, it is usually difficult for a user to map his information need into a sets of terms precisely. To overcome this difficulty, the technique of relevance feedback has been proposed [10]-[12]. Relevance feedback is the process of automatically adjusting an existing query using information feedback by the user about the relevance of previously retrieved documents.

The mechanism of this method can be described elegantly in the vector space. If the sets of relevant documents (D_R) and non-relevant documents (D_N) are known, the optimal query can be proven to be [10]-[12]:

$$\vec{V}(q_{opt}) = \frac{1}{N_R}\sum_{i \in D_R}\vec{V}(d_i) - \frac{1}{N_T - N_R}\sum_{i \in D_N}\vec{V}(d_i) \quad (5)$$

where N_R is the number of documents in D_R and N_T the number of the total documents.

In practice, D_R and D_N are not known in advance. However, the relevance feedback obtained from the user furnishes approximations to D_R and D_N which are referred as D'_R and D'_N.

The original query q can be modified by putting more weights on the relevant terms and less weight on the non-relevant terms.

$$\vec{V}(q') = \alpha\vec{V}(q) + \beta(\frac{1}{N_{R'}}\sum_{i \in D'_R}\vec{V}(d_i) - \gamma(\frac{1}{N_{N'}}\sum_{i \in D'_N}\vec{V}(d_i) \quad (6)$$

where α, β and γ are suitable constants [3], [13]; $N_{R'}$ and $N_{N'}$ are the numbers of documents in D'_R and D'_N. The q' approaches q_{opt} as the relevance feedback iteration moves on. Experiments indicate that the retrieval performance can be improved considerably by using relevance feedback [10]-[12].

III. Image Feature Extraction and Representation

We will use color feature based image retrieval in this paper. To demonstrate the validity of the proposed approach, we will convert color feature representations to the weighted vector model. We briefly describe the two representations about extract and feature represented in this section.

We design an algorithm that extracts the feature vector. The algorithm works as follows: The first, pixel values with the average value in a small local neighborhood (currently including the 8 adjacent pixels), then discretize the color space, such that there are only N distinct colors in the image. Then computed the regional groups color similarity through classify the color similarity pixel in the bin color and each bin color, built collection regions (each region consists of pixels belonging to the same bin color and size is predefined). Finally, building vector representing the image consists of N components (N colors respectively), each component is a list of regions same color.

Following is the LoHR algorithm. The algorithm returns a representing image vector consists of N components (N colors respectively), each component is a list of regions same color.

LoHR Algorithm (List of Homogeneous Region)

Input: *I* - image consists n pixels

T - size threshold of region

Output: *R* - collection regions of image I

1. $r_{i,j} \leftarrow \varnothing$
2. NumberRegion←0
3. N←*getcolor* (I);
4. for i←1 to N do
   - 4.1 $bin_i \leftarrow$ *getcolorpixel* (I, i)
   - 4.2 for j←1 to *size* (bin_i) do

     {

     $r_{i,j}$←*getneightborpixel* (bin_i, T);

     size (bin_i)←*size* (bin_i)-*size* ($r_{i,j}$)

     if ($r_{i,j<>} \varnothing$)

     NumberRegion++

     }
5. for i←1 to NumberRegion do
   - 5.1 R←R $\cup$ {ri, j}

LoHR algorithm has parameters: I, T, $r_{i,j}$, NumberRegion, N, bin_i and R. There is I input image, T is threshold for determining the region, $r_{i,j}$ is a region, Numberregion is number of regions in image, N is the number of colors to be quantified, bin_j is color bucket corresponds to ith color, R is the list of regions in the image. The algorithm uses functions getcolor(), getcolorpixel(), size() and getneightborpixel(), getcolor() function return color number from image, getcolorpixel() function get pixels have i color from image, size() function get amount pixels of a color bucket, getneightborpixel() get connected pixels in bin_i with amount connected pixels in bin_i larger threshold T indicated.

After receiving the list of R regions with the corresponding colors, we will get feature vector of I image as follows:

$$\mathrm{R_i} = [r_{i,1};...;r_{i,k};...r_{i,N}] \tag{7}$$

where *N* is the number of specific colors and also known as dimensional of feature vector and $r_{i,j}$ with *j*=1, 2, …, N is understood as a list of *j* regions color in the I image.

IV. Image Retrieval with Relevance Feedback

Related techniques are described in Section 2 is a powerful technique, but it can only application in information retrieval vector model. To use this technique we need to develop technique that can convert the image feature vector to the weight vector in the vector model.

For retrieval, we proposed image retrieval algorithm using IRuRF relevant feedback. Algorithm is performed as follows: The first, building collective of dictionary of region DR. This dictionary collection consists of N regions, is represented by: $DR = [r_{dr,1};...;r_{dr,k};...r_{dr,N}]$. The next, images extract features in the database according to the Lohr algorithm we are obtained collection of feature vectors $\mathrm{R_i} = [r_{i,1};...;r_{i,k};...r_{i,N}]$. Extracted of query q image feature according to Lohr algorithm we are obtained feature vector $\mathrm{q} = [r_{q,1};...;r_{q,k};...r_{q,N}]$. In each this feature vector, each component $r_{i,j}$ ($r_{q,j}$) is a list of regions with j color. After that, determine the weight vector of query image $\mathrm{q} = [w_{q,1};...;w_{q,k};...w_{q,N}]$ and weight vector of database images $\mathrm{I_i} = [w_{i,1};...;w_{i,k};...w_{i,N}]$ through the region frequency $r_{dr,j}$ in $r_{q,j}$ and $r_{dr,j}$ in $r_{i,j}$ respectively and inverse images collection frequency. The computing is based on determining two regions is similar (two regions are similar if the same color and equal in size or different a threshold θ). The similarity between the weight vector of the query image $\mathrm{q} = [w_{q,1};...;w_{q,k};...w_{q,N}]$ and weight vector $\mathrm{R_i} = [w_{i,1};...;w_{i,k};...w_{i,N}]$ of each the image database according to Cosine similarity $Sim(d,q) = \frac{\vec{V}(d)\vec{V}(q)}{\|\vec{V}(d)\|\|\vec{V}(q)\|}$. The last, apply formula (6) to get information from user and improve the quality result collection. Following is the algorithm image retrieval to use IRuRF relevance feedback. The algorithm returns an images result collection.

IRuRF algorithm (Image Retrieval using Relevance Feedback)

Input:

q - query image

DBI - image set in database

DR - dictionary regions

T - size threshold of region

θ - threshold similar region

Output:

S - result images set

1. $T_R \leftarrow \phi;\ T_N \leftarrow \phi$
2. q ←LoHR (q, T, R)
3. for each image i $\in$ DBI do
   - 3.1 I_i ←LoHR (I_i, T, R)
4. for i←1 to N do
   - 4.1 $w_{q,i} \leftarrow count(r_{dr,i}, r_{q,i}, \theta)$
   - 4.2 $q \leftarrow q \cup \{w_{q,i}\}$
5. for each image i $\in$ DBI do
   - 5.1 for j←1 to N do

     $w_{i,j} \leftarrow count(r_{dr,i}, r_{i,j}, \theta)$

     $I_i \leftarrow I_i \cup \{w_{i,j}\}$
6. for each image i $\in$ DBI do
   - 6.1 Computing $Sim(q, I_i) \leftarrow \frac{\vec{V}(q)\vec{V}(I_i)}{\|\vec{V}(q)\|\|\vec{V}(I_i)\|}$
7. Repeat
   - 7.1 User select I'_R relevance images and I'_N non-relevant images
   - 7.2 for each image i $\in I'_R$ do

     $T_R \leftarrow T_R + I_i$
   - 7.3 for each image i $\in I'_N$ do

     $T_N \leftarrow T_N + I_i$
   - 7.4 $q' \leftarrow \alpha q + \beta(\frac{1}{N_{R'}} T_R) - \gamma(\frac{1}{N_{N'}} T_N)$

7.5 Readln (Answer);
8. Until (Answer = "*No*");
9. Return S

IRuRF algorithm has parameters: q, DBI, DR, S, T_R, T_N, T, R, Ii, $w_{q,i}$, $r_{dr,i}$, $r_{q,i}$, θ, I'_R, I'_N, T_R, T_N, $N_{R'}$, $N_{N'}$, α, β and γ. Where q is input query image, DBI is image collection database, DR is dictionary collection region, S is image result collection, T_R is sum of image relevance vectors with query image, T_N is sum of image non-relevant vectors with query image, T is threshold region, I_i is ith image in database, $w_{q,i}$ is ith weight of query image vector, $r_{dr,i}$ is dictionary region has i color, $r_{q,i}$ is list regions has i color in query image, θ is similarity threshold region, I'_R is images relevance collection that user to select, I'_N is images non-relevance collection that user to select, $N_{R'}$ is amount images relevance that user to select, $N_{N'}$ is amount images non-relevance that user to select, α, β and γ appropriate constants. The algorithm also uses Lohr() algorithm, function $count(r_{dr,i}, r_{q,i}, \theta)$. LoHR() algorithm return list of regions and $count(r_{dr,i}, r_{q,i}, \theta)$ computed frequently occurred and inverse frequently in the image $r_{dr,i}$ in $r_{q,i}$ with similar threshold θ.

V. Experimental Results

The system is implemented on PC Pentium 2.8GHz processors and 1GB of main memory, running Windows 8, using a Wang database of 10,000 images[1]. The images are stored in JPEG format with size 126×85 and quantified in 15 colors. The database consists of 19 topics: The Sea, card, horse, butterfly, flower, sports athletes, windsurfing, sailing, fruits, flags, birds, house, waterfall, bear, wildebeest, cars, mountain, sunset, and forest. This database will be used to prove the accuracy of technique. The 100 images in the first image retrieval are categorized into positive and negative samples (according to users).

In this research we compare the results with CBsIR [13] and CCH [14]. To provide reliable results, random selection 10 images from each of nineteenth categories above. Recall precision graph [15] is used to compare IRuRF, CBsIR and CCH.

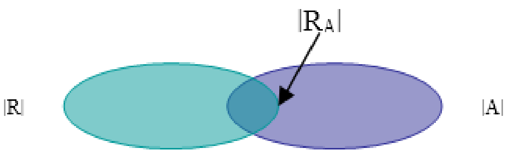

Figure 1. Recall and precision for query results

The R is a collection of relevant images in a database, A is a collection of images retrieval returned, R_A is a collection of relevant images in collection A (Fig. 1).

[1]http://wang.ist.psu.edu/IMAGE

Recall is the ratio of relevant images in database that been retrieved with a query. Precision is the ratio of relevant images that relevant to query image.

$$recall = \frac{area(R_A)}{area(R)}; \quad precision = \frac{area(R_A)}{area(A)}$$

Precision is the average precision of the total 187 queries and the results are shown in Fig. 2. The results indicate that IRuRF is better than CBsIR and CCH.

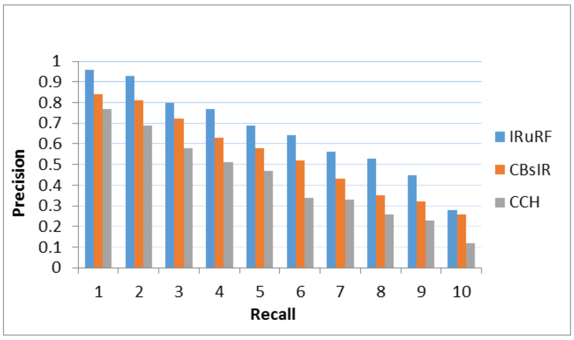

Figure 2. Compare precision - recall of IRuRF with CBsIR and CCH.

Table I summarizes the average query results. The retrieval results are summary as precision and recall. We performed some experiments, the first IRuRF technique is used to retrieved, CBsIR technique [13] is used in the second experiment, and the final CCH technique [14].

TABLE I. The Average Results of the Query

call	*Precision*		
	IRuRF	***CBsIR***	***CCH***
0.1	0.96	0.84	0.77
0.2	0.93	0.81	0.69
0.3	0.8	0.72	0.58
0.4	0.77	0.63	0.51
0.5	0.69	0.58	0.47
0.6	0.64	0.52	0.34
0.7	0.56	0.43	0.33
0.8	0.53	0.35	0.26
0.9	0.45	0.32	0.23
1	0.28	0.26	0.12

Horse image is used for query image for IRuRF, CBsIR and CCH to indicate the effectiveness of IRuRF. Group 30 images are found by IRuRF, CBsIR and CCH that are indicated in Fig. 3, Fig. 4 and Fig. 5.

VI. Conclusions

We have developed IRuRF that is a novel image retrieval method based on homogeneous regions using relevance feedback technique. IRuRF method has two advantage: We haven't process all pixels in image (processing some homogeneous regions only) and benefit advantage of relevance feedback in information retrieval and lead to the collection result closer to human perception.

The experimental results on a database consisting of 10,000 images indicates semantic accuracy of propose method. The experimental also indicate performance of IRuRF is higher than CBsIR and CCH methods.

Figure 3. These images found by IRuRF

Figure 4. These images found by CBsIR

Figure 5. These images found by CCH

REFERENCES

[1] J. Li, J. Z. Wang, and G. Wiederhold, "IRM: Integrated region matching for image retrieval," in *Proc. the 8th ACM Intl. Conf. on Multimedia*, 2000, pp. 147-156.

[2] J. Z. Wang, J. Li, and G. Wiederhold, "Simplicity: Semantics sensitive integrated matching for picture libraries," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 23, no. 9, pp. 947-963, 2001.

[3] R. O. Stehling, M. A. Nascimento, and A. X. Falc˜ao, "An adaptive and efficient clustering-based approach for content based image retrieval in image databases," in *Proc. Intl. Data Engineering and Application Symposium*, 2001, pp. 356-365.

[4] J. J. Rocchio, "Document retrieval systems-optimization and evaluation," PhD thesis, Harvard Computational Laboratory, Cambridge, MA, 1966.

[5] G. Salton, *Automatic Text Processing: The Transformation, Analysis, and Retrieval of Information by Computer*, Boston, MA, USA: Addison-Wesley Longman Publishing Co., Inc., 1989.

[6] T. S. Huang, *et al.*, "Learning in content-based image retrieval," in *Proc. the 2nd International Conference on Development and Learning*, 2002.

[7] Y. Rui and T. S. Huang, "Optimizing learning in image retrieval," in *Proc. IEEE Int. Conf. on Computer Vision and Pattern Recognition*, Jun. 2000

[8] S. Tong and E. Chang, "Support vector machine active leaning for image retrieval," in *Proc. the Ninth ACM International Conference on Multimedia*, Ottawa, Canada, 2001, pp. 107-118.

[9] L. Zhang, F. Z. Lin, and B. Zhang, "Support vector machine learning for image retrieval," in *Proc. IEEE International Conference on Image Processing*, Oct. 2001.

[10] G. Salton and M. J. McGill, *Introduction to Modern Information Retrieval*, McGraw-Hill Book Company, 1983.

[11] C. Buckley and G. Salton, "Optimization of relevance feedback weights," in *Proc. SIGIR*, 1995.

[12] W. M. Shaw, "Term-Relevance computations and perfect retrieval performance," *Inf. Process. Manage*, vol. 31, no. 4, pp. 491-498, Jul. 1995.

[13] R. A. B. Yates and B. R. Neto, *Modern Information Retrieval*, Boston, MA, USA: Addison-Wesley Longman Publishing Co., Inc., 1999.

[14] R. O. Stehling, M. A. Nascimento, and A. X. Falcao, "Cell histograms versus color histograms for image representation and retrieval," *Knowledge and Information Systems Journal (KAIS)*, vol. 5, no. 3, pp. 315-336, Jul. 2003.

[15] J. Luo and M. A. Nascimento, "Content-Based sub-image retrieval using relevance feedback," in *Proc. 2nd ACM International Workshop on Multimedia Databases*, New York, 2004, pp. 2-9.

Vu Van Hieu received the BS degree in computer science from the Haiphong Private University of Vietnam in 2003 and the MS degree in computer science from the Vietnam Military University of Science and Technology in 2007. He has been working for Information Technology Faculty at Haiphong University from 2003. He is currently a PhD candidate in the Institute of Information Technology, Vietnam Academy of Science and Technology. His research interests include multimedia retrieval, Pattern Recognition and Image Processing, Content based image retrieval, data mining, and machine learning.

Quynh Nguyen Huu received the B. S. degree in informatics, M.S. and Ph. D. degrees in computer science, all from Hanoi University of Viet Nam, in 1998, 2004 and 2010, respectively.

Dr. Quynh has been working for Information Technology Faculty at Electric Power University from 1999. His current research interests include content based image retrieval, intelligent image processing, GIS, image features extracting, and multimedia systems, Image Database Management Systems, Natural Language processing. Dr. Quynh has published over 20 international journal articles and international conference papers.

