KIDS's evaluation in medical image retrieval task at ImageCLEF 2004

Pei-Cheng Cheng[a], Been-Chian Chien[be], Hao-Ren Ke[c], and Wei-Pang Yang[ad]

[a]Department of Computer & Information Science, National Chiao Tung University,
1001 Ta Hsueh Rd., Hsinchu, TAIWAN 30050, R.O.C.
{ cpc, wpyang}@cis.nctu.edu.tw

[b]Department of Information Engineering, I-Shou University
1, Section 1, Hsueh-Cheng Rd., Ta-Hsu Hsiang, Kaohsiung, TAIWAN 840, R.O.C.
cbc@isu.edu.tw

[c]University Library, National Chiao Tung University,
1001 Ta Hsueh Rd., Hsinchu, TAIWAN 30050, R.O.C.
claven@lib.nctu.edu.tw

[d]Department of Information Management, National Dong-Hwa University
1, Sec. 2,Da Hsueh Rd., Shou-Feng, Hualien,Taiwan, R.O.C.

[e]Department of Computer Science and Information Engineering, National University of TAINAN
33, Sec. 2, Su-Lin Street, Tainan, 700, Taiwan, R.O.C.
bcchien@ipx.ntntc.edu.tw

Abstract

We describe our participation in the medical retrieval task of the ImageCLEF 2004. This task aims at finding images that are similar with respect to modality (CT, radiograph, MRI, and so on). We propose several image features, including *color histogram, gray-spatial histogram*, *coherence moment*, and *gray correlogram*, to facilitate the retrieval of similar images. The initial retrieval results are obtained via visual feature analysis. An automatic feedback mechanism clusters visually and textually similar images among these initial results to help refine the query. In this paper, we present the system used, focusing on novel and newly developed aspects. The evaluated result shows that the automatic feedback mechanism improves the precision by 15%.

Keywords: Medical image retrieval; Color histogram; Relevance feedback;

1. Introduction

The importance of digital image retrieval techniques increases in the emerging fields of medical imaging and picture archiving and communication systems (PACS). The increasing reliance of modern medicine on diagnostic techniques such as radiology, histopathology, and computerized tomography has led to the explosion in the number and importance of medical images now stored by most hospitals. While the prime requirement for medical imaging systems is to be able to display images relating to a named patient, there is increasing interest in the use of CBIR (content-based image retrieval) techniques to aid diagnosis by identifying similar past cases.

In the past years, content-based image retrieval has been one of the most hot research areas in the field of computer vision. The commercial QBIC [Flickner95] system is definitely the most well known system. Another commercial system for image and video retrieval is Virage [Bach 96] [Hampapur97] that has well known commercial customers such as CNN. In the academia, some systems including Candid [Kelly95], Photobook [Pentland96], and Netra [Ma97] use simple color and texture characteristics to describe image content. The Blobword system [Carson99][Belongie98] exploits higher-level information, such as segmented objects of images, for queries. A system that is available free of charge is the GNU Image Finding Tool (GIFT) [Squire00]. Some systems are available as demonstration version on the Web such as Viper, WIPE or Compass. Most of the available systems are hard to compare.

Imaging systems and image archives have often been described as an important economic and clinical factor in the hospital environment [Greenes00]. Several methods from computer vision and image processing have already been proposed for the use in medicine [Pun94]. Medical images have often been used for retrieval

systems, and the medical domain is often cited as one of the principal application domains for content-based access technologies [Smeul00] [Kelly95] [Beretti01] [Orphan94] in terms of potential impact. Still, there has rarely been an evaluation of the performance.

One of the most significant problems in content-based image retrieval results from the lack of a common test-bed for researchers. Although many published articles report on content-based retrieval results using color photographs, there has been little effort in establishing a benchmark set of images and queries. It is very important that image databases are made available free of charge for the comparison and verification of algorithms. Only such reference databases allow comparing systems and to have a reference for the evaluation that is done based on the same images. ImageCLEF offers numerous medical images for evaluation that has many benefits in advancing the technology and utilization of content-based image retrieval systems.

In this year's ImageCLEF evaluation, we participated in the medical retrieval task. In the following sections, we detail the approach taken for the medical retrieval task. We analyze the results of the various evaluations, and have a discussion about the relative performance of our systems. In this task, we need to find the relevant images from approximate 9000 medical images that are similar to the 26 query images respectively. The data of this task contains cross language text (French and English) and visual medical images (CT, radiograph, MRI, and so on). We submitted two runs in this task. In the first run we use the visual features to retrieve similar images. In the second run we analyze the results of visual example queries and exploit a relevance feedback mechanism to improve the result. This paper is organized as follows. . In Section 2, we describe the features we use to represent the images. The similarity metric is proposed in Section 3. In Section 4, we explicate the automatic feedback mechanism. The submit runs are discussed in Section 5. Section 6 concludes this paper.

2. Feature Extraction

The medical image collection of ImageCLEF 2004 contains gray and color images. In color images, users are usually attracted by the change of colors more than the positions of objects. Thus, we use *color histogram* as the feature of color images to retrieve similar color images. Color histogram is suitable to compare images in many applications. Color histogram is computationally efficient, and generally insensitive to small changes in the camera position.

Color histogram has some drawbacks. Color histogram provides less spatial information; it merely describes which colors are present in an image, and in what quantities. Because gray images encompass few colors (usually 256 gray levels), directly using color histogram in gray images will get bad retrieval results. For gray images, we must emphasize spatial relationship analysis; furthermore, object and contrast analysis is important for medical images; therefore, three kinds of features that can indicate the spatial, coherence, and shape characteristics, *gray-spatial histogram*, *coherence moment*, and *gray correlogram*, are employed as the features of gray images.

In the following we describe the four kinds of features, one for color images and three for gray images, used in this paper.

2.1. Color image features

Color histogram [Swain91] is a basic method and has good performance for representing image content. The color histogram method gathers statistics about the proportion of each color as the signature of an image. Let C be a set of colors, $(c_1, c_2 \ldots c_m) \in C$, that can occur in an image. Let I be an image that consists of pixels $p(x,y)$[1]. The color histogram $H(I)$ of image I is a vector $(h_1, h_2, \ldots, h_i, \ldots, h_m)$, in which each bucket h_i counts the ratio of pixels of color c_i in I. Suppose that p is the color level of a pixel. Then the histogram of I for color c_i is defined as Eq. (1):

$$h_{c_i}(I) = \Pr_{p \in I}\{p \in c_i\} \qquad (1)$$

In other words, $h_{c_i}(I)$ corresponds to the probability of any pixel in I being of the color c_i. For comparing the similarity of two images I and I', the distance between the histograms of I and I' can be calculated using a standard method (such as the L_1 distance or L_2 distance). Then, the image in the image database most similar to a query image I is the one having the smallest histogram distance with I.

---

[1] p(x, y) indicates the color of the corresponding pixel as well.

LL	HL
LH	HH

(a)   (b)   (c)
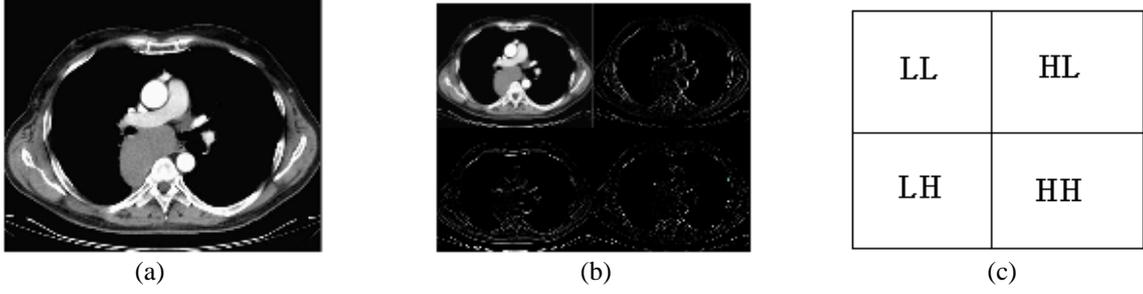
Figure 1 An example of wavelet translation.

Any two colors have a degree of similarity. Color histogram is hard to express the similar characteristic. In this paper, each pixel does not only assign a single color. We set an interval range δ to extend the color of each pixel. Then the histogram of image *I* is redefined as the Eq. (2):

$$h_{c_i}(I) = \frac{\sum_{j=1}^{m} \frac{[p_j - \frac{\delta}{2}, p_j + \frac{\delta}{2}] \cap c_i}{\delta}}{m} \qquad (2)$$

where p_j is a pixel of image, and *m* is the total number of pixels.

The colors of an image are represented in HSV (Hue, Saturation, Value) space, which is closer to human perception than spaces such as RGB (Red, Green, Blue) or CMY (Cyan, Magenta, Yellow). In implementation, we quantize HSV space into 18 hues, 2 saturations and, 4 values, with additional 4 levels of gray values; as a result, there are a total of 148 bins.

Using the modified color histogram, the similarity of two color images *q* and *d* is defined as Eq. (3):

$$SIMcolor(H(q), H(d)) = \frac{H(q) \cap H(d)}{|H(q)|} = \frac{\sum_{i=1}^{n} \min(h_i(q), h_i(d))}{\sum_{i=1}^{n} h_i(q)} \qquad (3)$$

2.2. Gray image features

Gray images are different from color images in human's perception. Gray images have fewer colors than color images, only 256 gray levels in each gray image. Human's visual perception is influenced by the contrast of an image. The contrast of an image from the viewpoint of human is relative rather than absolute. To emphasize the contrast of an image and handle images with less illuminative influence, we normalize the value of pixels before quantization. In this paper we propose a relative normalization method. First, we cluster the whole image into four clusters by the K-means cluster method [Han01]. We sort the four clusters ascendently according to their mean values. We shift the mean of the first cluster to value 50 and the fourth cluster to value 200; then each pixel in a cluster is multiplied by a relative weight to normalize. Let m_{c1} is the mean value of cluster 1 and m_{c4} is the mean value of cluster 4. The normalization formula of pixel p(x,y) is defined as Eq. (4).

$$p(x, y)_{normal} = (p(x, y) - (m_{c1} - 50)) \times \frac{200}{(m_{c4} - m_{c1})} \qquad (4)$$

After normalization, we resize each image into 128*128 pixels, anduse one level wavelet with Haar wavelet function [Stollnitz96] to generate the low frequency and high frequency sub-images. Process an image using the low pass filter will obtain an image that is more consistent than the original one; on the contrary, processing an image using the high pass filter will obtain an image that has high variation. The high-frequency part keeps the contour of the image. Figure 1 is an example of wavelet translation. Figure 1 (a) is the original image, (b) is one-level wavelet transformed image; there are four sub-bands denoted by Low_Low (LL), Low_Height (LH), Height_Low (HL), Height_Height (HH), as shown in Figure 1 (c). High-frequency pixels may be important in medical images for doctor diagnoses. By performing the OR operation for LH, HL, and HH bands, we get the contour of a medical image.
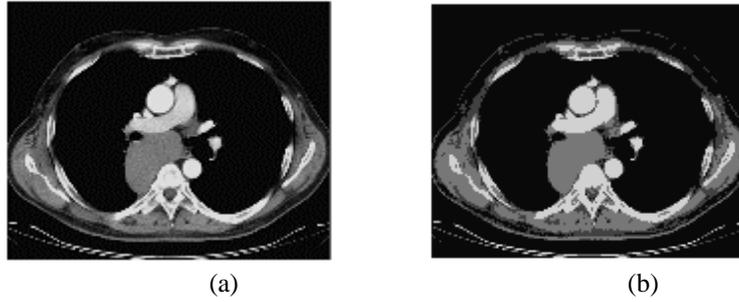
Figure 2 (a) original image with 256 levels; (b) new image after clustering with only 4 levels
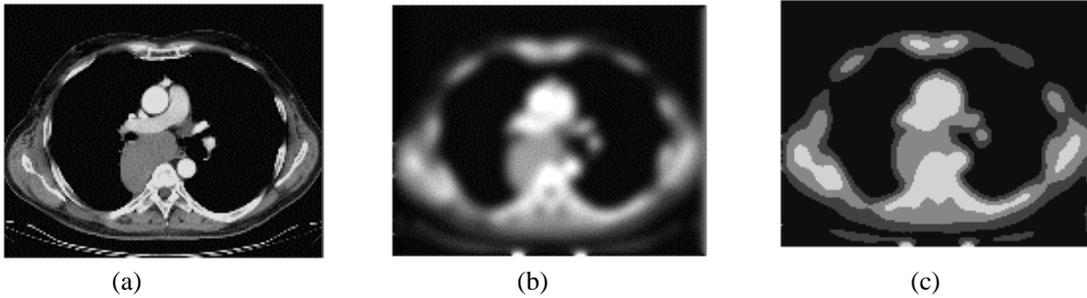
Figure 3 (a) original image; (b) image after smoothing; (c) image after clustering into four classes.

2.2.1. Gray-spatial histogram

In a gray image the spatial relationship is very important especially in medical images. Medical images always contain particular anatomic regions (lung, liver, head, and so on); therefore, similar images have similar spatial structures. We add spatial information into the histogram so we call this representation as gray-spatial histogram in order to distinguish from color histogram. We use the LL band for gray-spatial histogram and coherence analysis. To get the gray-spatial histogram, we divide the LL band image into nine areas. The gray values are quantized into 16 levels for computational efficiency.

The gray-spatial feature estimates the probability of each gray level that appears in a particular area. The probability equation is defined in Eq. (2), where δ is set to 10. The gray-spatial histogram of an image has a total of 144 bins.

2.2.2. Coherence moment

One of the problems to design a image representation is the semantic gap. The state-of-the-art technology still cannot reliably identify objects. The coherence moment feature attempts to describe the features from the human's viewpoint in order to reduce the semantic gap.

We cluster an image into four classes by the K-means algorithm. Figure 2 is an example. Figure 2 (a) is the original image and Figure 2 (b) is four-level gray image. We almost can not visually find the difference between the two images. After clustering an image into four classes, we calculate the number of pixels (COH_κ), mean value of gray value (COH_μ) and standard variance of gray value (COH_ρ) in each class. For each class, we group connected pixels in eight directions as an object. If an object is bigger than 5% of the whole image, we denote it as a big object; otherwise it is a small object. We count how many big objects (COH_o) and small objects (COH_v) in each class, and use COH_o and COH_v as parts of image features.

Since we intend to know the reciprocal effects among classes, so we smooth the original image. If two images are similar, they will also be similar after smoothing. If their spatial distributions are quite different, they may have different result after smoothing. After smoothing, we cluster an image into four classes and calculate the number of big objects (COH_τ) and small objects (COH_ω). Figure 3 is an example. Each pixel will be influenced by its neighboring pixels. Two close objects of the same class may be merged into one object. Then, we can analyze the variation between the two images before and after smoothing. The coherence moment of each class is a seven-feature vector, (COH_κ, COH_μ, COH_ρ, COH_o, COH_v, COH_τ, COH_ω). The coherence moment of an image is a 28-feature vector that combines the coherence moments of the four classes.

2.2.3. Gray correlogram

The contour of a medical image contains rich information. Diseases can be easily detected in the high frequency domain. But, in this task we are going to find similar medical images, not to detect the affected part. A broken bone in the contour may be different from the healthy one. Thus we choose a representation that can estimate the partial similarity of two images and can be easy to calculate their global similarity.

We analyze the high frequency part by our modified correlogram algorithm. The definition of the correlogram [Huang97][Ojala01] is as Eq. (5). Let D denote a set of fixed distances $\{d_1, d_2, d_3, \ldots, d_n\}$. The correlogram of an image I is defined as the probability of a color pair (ci, cj) at a distance d.

$$\gamma_{c_i, c_j}^{d}(I) = \Pr_{p_1 \in c_i, p_2 \in I}\{p_2 \in c_j \| p_1 - p_2 \| = d\} \qquad (5)$$

For computational efficiency, the autocorrelogram is defined as Eq. (6)

$$\lambda_{c_i}^{d}(I) = \Pr_{p_1 \in c_i, p_2 \in I}\{p_2 \in c_i \| p_1 - p_2 \| = d\} \qquad (6)$$

The contrast of a gray image dominates human's perception. If two images have different gray levels they still may be visually similar. Thus the coorelogram method cannot be used directly.

Our modified correlogram algorithm works as follows. First we sort the pixels of the high frequency part descendently.. Then we order the results of the preceding sorting by the ascendant distances of pixels to the center of the image. The distance of a pixel to the image center is measured by the L2 distance. After sorting by gray value and distance to the image center, we select the top 20 percent of pixels and the gray values higher than a threshold to estimate the autocorrelogram histogram. We set the threshold zero in this task. Any two pixels have a distance, and we estimate the probability that the distance falls within an interval. The distance intervals we set are {[0,2], [2,4], [4,6], [6,8], [8,12], [12,16], [16,26], [26,36], [36,46], [46,56], [56,66], [76,100]}. The high frequent part comprises 64*64 pixels, thus the maximum distance will be smaller than 100. The first n pixels will have n*(n+1)/2 numbers of distances. We calculate the probability of each interval to form the correlogram vector.

3. Similarity metric

While an image has features to represent it, we need a metric to measure the similarity between two feature vectors (and consequently, the similarity between two images). The similarity metric of color histogram is defined as Eq. (3) and that of gray-spatial histogram is defined as Eq. (7):

$$\text{SIM}_{gray\_spatial}(H(q), H(d)) = \frac{H(q) \cap H(d)}{|H(q)|} = \frac{\sum_{i=1}^{n} \min(h_i(q), h_i(d))}{\sum_{i=1}^{n} h_i(q)} \qquad (7)$$

The similarity metric of the coherence moment is defined as Eq. (8)

$$DIS_{coh}(COH(q), COH(d)) = \sum_{i=1}^{4classes}(|COH_{\kappa}^{q_i} - COH_{\kappa}^{d_i}| + |COH_{\mu}^{q_i} - COH_{\mu}^{d_i}| \times |COH_{\rho}^{q_i} - COH_{\rho}^{d_i}| +$$
$$|COH_{o}^{q_i} - COH_{o}^{d_i}| + |(COH_{v}^{q_i})^{1/2} - (COH_{v}^{d_i})^{1/2}| + \qquad (8)$$
$$|COH_{\tau}^{q_i} - COH_{\tau}^{d_i}| \times 2 + |(COH_{\omega}^{q_i})^{1/2} - (COH_{\omega}^{d_i})^{1/2}|)$$

The correlogram metric is defined as Eq. (9):

$$DIS_{hf}(H(q), H(d)) = \frac{\sum_{i=1}^{n} |h_i(q) - h_i(d)|}{\sum_{i=1}^{n} |h_i(q) + h_i(d)|} \qquad (9)$$

The similarity of two images Q and D is measured by Eq. (10):

$$\text{SIM}_{image}(Q, D) = W_1 \times \text{SIM}_{color}(H(Q), H(D) + W_2 \times \text{SIM}_{gray\text{-}spatial}(H(Q), H(D) +$$
$$W_3 \times 1/(1 + DIS_{coh}(COH(Q), COH(D))) + W_4 \times 1/(1 + DIS_{hf}(COH(Q), COH(D))) \qquad (10)$$
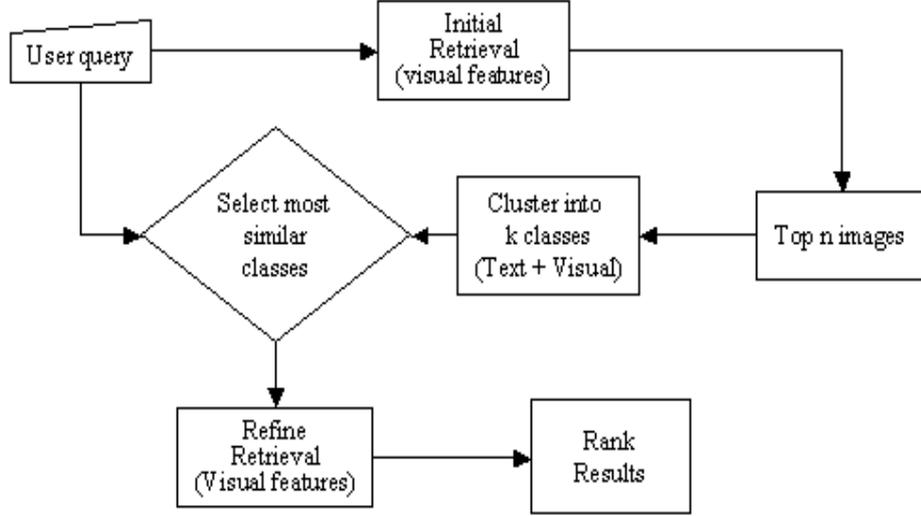
Figure 4 The feedback mechanism

, where W_i is the weight of each feature. In this task the database contains color and gray images. When the user queries an image by example, we first determine whether the example is color or gray. We calculate the color histogram, if the four bins of gray values occupy more than 80% of the whole image, we decide the query image is gray; otherwise it is color. If the input is a color image, then we set $W_1=10$, $W_2=0.1$, $W_1=10$, and $W_1=10$; otherwise we set $W_1=0.1$, $W_2=1$, $W_1=100$, and $W_1=100$.

4. Feedback Mechanism

When the user inputs the visual query example, the system first employs visual features to retrieve relevant images from the database. After the initial retrieval, the system selects the top-n relevant images as candidate positive images. The similarity between the visual query example and each of the top-n images must also be greater than a threshold. In the next step, we cluster the top-n images into k classes. Figure 4 illustrates the feedback mechanism. In addition to images, the database of ImageCLEF 2004 contains a diagnosis text for each image. However, A patient case contains a variety of visual images. The images of the same case are sometimes not all similar in visually. So, in this paper while doing the relevance feedback, the weight of textual is slighter.

We first translate the diagnosis texts in French into possible English. The vector space model [Salton88] is used to create a vector representation of a diagnosis text. Each entry of the vector represents a term of the text and the value of the entry is the term frequency (tf) * inverse document frequency (idf) value. The similarity between two diagnoses is computed as the cosine between their vector representations, as shown in Eq. (10)

$$\cos ine(q,d) = \frac{\sum_{i=1}^{n} w_i^q \times w_i^d}{\sqrt{\sum_{i=1}^{n} (w_i^q)^2} \times \sqrt{\sum_{i=1}^{n} (w_i^d)^2}} \qquad (10)$$

where W_i^q is the weight of term i in text q and n is the numbers of terms.

The similarity between two images consists of visual similarity and textual similarity. We set the weight of textual part as 0.1 and visual features part as 0.9. In our implementation, we cluster the top-20 images into 6 classes by the minimum distance hierarchical cluster algorithm [Han01]. The class most similar to the query example in vision becomes the next-generation query example images. We use the OR operation among exemplary images to measure the similarity of database images; in other words, we use the maximum similarity between positive query images and an image in the database to measure the similarity of the latter image to the query.

(a) First query step	(b) With feedback
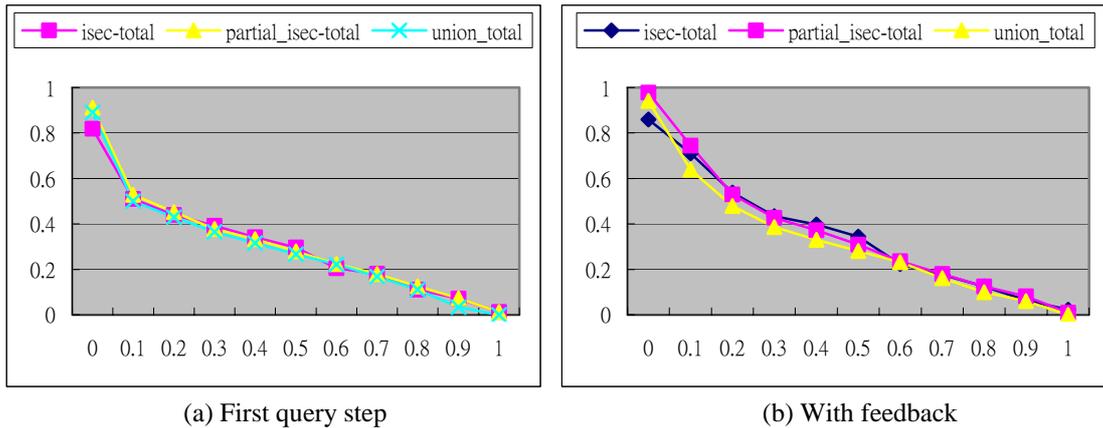
Figure 5 Precision Vs. Recall graphs without and with feedback

5.   Experimental Results

In the ImageCLEF 2004, the process of evaluation and the format of results employs the trec_eval tool. The evaluation procedure did as following:

> Extracted the top-60 runs from each submission (we received 43 submissions in total). Computed the union of runs to create a document pool for each topic. Manually assessed images in the document pool using three assessors (images judged as relevant and partially relevant). Created 9 sets of relevant images for each topic (9 *qrels* sets). Compared each system run against the qrels. Computed uninterpolated mean average precision across all topics using trec_eval.

Three "expert" assessors judged the image pools generated from pooling the submissions. They created 9 sets of qrels based on the overlap of relevant images between assessors, and whether partially relevant images were included in the qrels set. The partially relevant judgment was used to pick up image where the judge thought it was in some way relevant, but could not be entirely confident. The 9 relevance sets are listed here:

> ➢   isec-rel: images judged as relevant by all three assessors.
>
> ➢   isec-partial: images judged as partially relevant by all three assessors.
>
> ➢   isec-total: images judged as either relevant or partially relevant by all three assessors.
>
> ➢   partial_isec-rel: images judged as relevant by at least 2 assessors.
>
> ➢   partial_isec-partial: images judged as partially relevant by at least 2 assessors.
>
> ➢   partial_isec-total: images judged as either relevant or partially relevant by at least 2 assessors.
>
> ➢   union_rel: images judged as relevant by all at least 1 assessor.
>
> ➢   union_partial: images judged as partially relevant by at least 1 assessor.
>
> ➢   union_total: images judged as either relevant or partially relevant by at least 1 assessor.

In this task, we submit three runs. We forgot to sort descendently the third run's output and get wrong results; thus we omit to discuss it. The first run uses the visual feature of the query example image to query the databse. The second run is the result of the automatic feedback mechanism, which uses the images of the most similar class as the positive query examples to query the image database. The test result shows that the auto-feedback mechanism, run2, has better result than the first run. In the partial_isec-total results summary, the mean average precision of the first run of our system is 0.2960. The mean average precision of run2 (with feedback) is 0.3457.   Figure 5 shows the precision and recall graphs. Figure 5 (a) is the run1 we submitted and Figure 5 (b) is the run2 with the automatic feedback mechanism.

In this paper we consider that the contrast of a gray image dominates human's perception. We use a relative normalization method to reduce the impact of illumination. Figure 6is the result of an example query returned by our system. It can be observed from Figure 6 that 9719.jpg and 16870.jpg are darker than the query image (def_queries/1.jpg), but our system still can find them out.

Figure 6 Result of an example query

The first run has accuracy above 50% in the first 20 images. The really similar images may have similar features in some aspect and similar to each other. The miss judged images are always less consistent. So we try to refine the initial result by the automatic feed back mechanism. We cluster the first 20 images into six classes. If the class contains diverse images, the center of the class will become farther, and consequently more different, from the query image. Thus we can improve the result by our feedback method.

6.   Conclusion

In this paper we propose several methods to represent medical images. Although the color histogram of content-based image retrieval methods has good performance in general-propose color images, unlike general-propose color images the X-ray images only contain gray-level pixels. Thus, we concentrate on the contrast representation of images.

The image representations we propose have obtained good results in this task. Our representation is immune in defective illumination. A total of 322 features is used. It is very efficient in computation. The auto feedback mechanism also provides a good result in medical images.

An image contains thousands of words. An image can be viewed from various aspects; furthermore, different people may have different interpretation of the same image. This will cause too many parameters need to be tuned. In the future, we will try to learn the user behavior and tune those parameters by artificial learning methods.

Reference

[Smeul 00] A. W. M. Smeulders, M. Worring, S. Santini, A. Gupta, R. Jain, Content-based image retrieval at the end of the early years, *IEEE Transactions on Pattern Analysis and Machine Intelligence* 22 No 12 (2000) 1349-1380.

[Swain91] Swain M.J. and D. H. Ballard, "Color Indexing", *International Journal of Computer Vision,* Vol. 7, pp.11-32, 1991.

[Flickner95] M. Flickner, H. Sawhney, W. Niblack, J. Ashley, Q. Huang, B. Dom, M. Gorkani, J. Hafner, D. Lee, D. Petkovic, D. Steele, P. Yanker, Query by Image and Video Content: The QBIC system, *IEEE Computer* 28 (9) (1995) 23-32.

[Bach 96] J. R. Bach, C. Fuller, A. Gupta, A. Hampapur, B. Horowitz, R. Humphrey, R. Jain, C.-F. Shu, The Virage image search engine: An open framework for image management, in: I. K. Sethi, R. C. Jain (Eds.), *Storage & Retrieval for Image and Video Databases IV*, Vol. 2670 of IS&T/SPIE Proceedings, San Jose, CA, USA, 1996, pp. 76-87.

[Hampapur97] A. Hampapur, A. Gupta, B. Horowitz, C.-F. Shu, C. Fuller, J. Bach, M. Gorkani, R. Jain, Virage video engine, in: I. K. Sethi, R. C. Jain (Eds.), *Storage and Retrieval for Image and Video Databases V*, Vol. 3022 of SPIE Proceedings, 1997, pp. 352-360.

[Kelly95] P. M., M. Cannon, D. R. Hush, Query by image example: the CANDID approach, in: W. Niblack, R. C. Jain (Eds.), *Storage and Retrieval for Image and Video Databases III*, Vol. 2420 of SPIE Proceedings, 1995, pp. 238-248.

[Pentland96] A. Pentland, R. W. Picard, S. Sclaro_, Photobook: Tools for content-based manipulation of image databases, *International Journal of Computer Vision* 18 (3) (1996) 233-254.

[Ma97] W. Y. Ma, Y. Deng, B. S. Manjunath, Tools for texture- and color-based search of images, in: B. E. Rogowitz, T. N. Pappas (Eds.), *Human Vision and Electronic Imaging II*, Vol. 3016 of SPIE Proceedings, San Jose, CA, 1997, pp. 496-507.

[Carson 99] C. Carson, M. Thomas, S. Belongie, J. M. Hellerstein, J. Malik, Blobworld: A system for region-based image indexing and retrieval, in: D. P. Huijsmans, A. W. M. Smeulders (Eds.), *Third International Conference On Visual Information Systems* (VISUAL' 99), no. 1614 in Lecture Notes in Computer Science, Springer{Verlag, Amsterdam, The Netherlands, 1999, pp. 509-516.

[Belongie 98]S. Belongie, C. Carson, H. Greenspan, J. Malik, Color and texture based image segmentation using EM and its application to content-based image retrieval, in: *Proceedings of the International Conference on Computer Vision (ICCV'98)*, Bombay, India, 1998, pp. 675-682.

[Squire 00] D. M. Squire, W. Muller, H. Muller, T. Pun, Content-based query of image databases: in- spirations from text retrieval, *Pattern Recognition Letters (Selected Papers from The 11th Scandinavian Conference on Image Analysis SCIA '99)* 21 (13-14) (2000) 1193-1198, b.K. Ersboll, P. Johansen, Eds.

[Greenes 00] R. A. Greenes, J. F. Brinkley, Imaging systems, in: Medical Informatics: *Computer Applications in Healthcare (2nd edition), Springer,* New York, 2000, Ch. 14, pp. 485- 538.

[Pun 94] T. Pun, G. Gerig, O. Ratib, Image analysis and computer vision in medicine, *Computerized Medical Imaging and Graphics* 18 (2) (1994) 85-96.

[Beretti 01] S. Beretti, A. Del Bimbo, P. Pala, content- based retrieval of 3D cellular structures, in: *Proceedings of the second International Conference on Multimedia and Exposition* (ICME'2001), IEEE Computer Society, Tokyo, Japan, 2001, pp. 1096-1099.

[Orphan 94] S. C. Orphanoudakis, C. E. Chronaki, S. Kostomanolakis, I2Cnet: A system for the indexing. *storage and retrieval of medical images by content*, *Medical Informatics* 19 (2) (1994) 109-122.

[Huang 97] J. Huang, S. R. Kumar, M. Mitra, W. –J. Zhu, R. Zabih, Image indexing Using Color Correlograms, *Proceedings of IEEE Conference on Computer Vision and Pattern Recognition*, San Juan, Puerto Rico 1997.

[Ojala 01] T. Ojala, M. Rautiainen, E. Matinmikko and M. Aittola, Semantic Image Retrieval with HSV Correlograms, *Proceedings of 12th Scandinavian Conference on Image Analysis*, Bergen, Norway 2001.

[Han01] J. Han, M. Kamber, "Data mining: concepts and techniques", *Academic press*, San Diego, CA, USA 2001.

[Stollnitz96] E. J. Stollnitz, T. D. DeRose, and D. H. Salesin, Wavelets for computer Graphics – Theory and Applications. San Francisco, CA: Morgan Kaufmann Publilshers, Inc, 1996.

[Salton88] G. Salton, Automatic Text Processing. Addison-Wesley Publishing Company (1988).