

# A General Framework for Content-Based Medical Image Retrieval with its Application to Mammograms

Chia-Hung Wei<sup>\*</sup>, Chang-Tsun Li<sup>†</sup>, and Roland Wilson<sup>#</sup>

Department of Computer Science  
University of Warwick, Coventry, CV4 7AL, UK

## ABSTRACT

In the medical field, content-based image retrieval (CBIR) is used to aid radiologists in the retrieval of images with similar contents. CBIR methods are usually developed for specific features of images, so that those methods are not readily applicable across different kinds of medical images. This study proposes a sound methodology for CBIR of mammograms, which is applicable to various formats of medical image. The methodology is divided into two parts—image analysis and image retrieval. In the image analysis part, 19 abnormal regions of interest (ROI) and 20 normal ROIs are selected as samples for the whole ROI dataset. These two groups of ROIs are used to analyze 11 textural features based on gray level co-occurrence matrices. The multivariate  $t$  test is then applied to examine the significance of the differences for these 11 textural features from normal and abnormal ROIs. The discriminating features are incorporated into a feature descriptor for the ROI. This descriptor is embedded into the CBIR system. In the image retrieval part, a CBIR system for mammograms is developed. For normalization of feature vectors, a novel technique is proposed to clip the values of feature elements of the top 5%, and then project each image feature onto the unit sphere. To determine the similarity between query image and each ROI in the dataset, the  $L_2$  norm is used to measure the similarity between two images. This system was designed by query-by-example (QBE). Query images were selected from different classes of abnormal ROIs. To evaluate the performance of the CBIR system, the precision and recall were measured. A maximum precision of 51% and recall of 19% were obtained using the gray level co-occurrence matrices and a distance of 5. The averages of precision and recall are 49% and 18% in this experiment.

**Keywords:** Content-Based Image Retrieval (CBIR), Mammogram, Gray Level Co-occurrence Matrices

## 1. INTRODUCTION

The huge volume of medical images generated in hospitals creates a need to develop new tools to retrieve such visual information. The task of content-based image retrieval (CBIR) in the medical field is to help radiologists to retrieve images with similar contents. However, CBIR methods are usually developed for specific features of images, so that methods are not always applicable between different kinds of medical images.

Among visual features, texture is widely used for content-based access to medical images [1]. Through textural analysis, it is possible to discover the texture signature of a medical image relevant to the diagnostic problem. The effectiveness of textural analysis depends on the methods used to extract meaningful features [2]. There have been several methods of textural feature extraction, such as gray level co-occurrence matrices [3] and Tamura's textural features [4].

Human intervention is a necessary aid in the CBIR process because the existing algorithms cannot segment pathology-bearing regions in most medical images automatically. In particular, there are often several pathology-bearing regions in an image. One CBIR technique in medical imaging applications is the use of a region of interest (ROI). CBIR systems have definite targets to search for using delineation of ROIs.

Breast cancer continues to be a serious disease across the world. Mammography is a reliable method for detection of breast cancer. There are an enormous number of mammograms generated in hospitals. How to effectively retrieve a

---

\* E-mail: [rogerwei@dcs.warwick.ac.uk](mailto:rogerwei@dcs.warwick.ac.uk)

† E-mail: [ctli@dcs.warwick.ac.uk](mailto:ctli@dcs.warwick.ac.uk)

# E-mail: [rgw@dcs.warwick.ac.uk](mailto:rgw@dcs.warwick.ac.uk)

desired image from mammography databases is a challenging problem. This study concentrates on textural analysis based on gray-level co-occurrence matrices for the content-based retrieval of mammograms. The objectives of this study are as follows:

- 1) To analyze the textural features presented in the ROI of abnormal breast tissue as compared to the same information presented in normal tissue;
- 2) To develop the optimal mammographic descriptors generated from gray level co-occurrence matrices.
- 3) To evaluate the effectiveness of the CBIR system using descriptors with different unit pixel distances.

The rest of the paper is organized as follows: Section 2 reviews previous literature on medical content-based image retrieval. Section 3 describes the methodology for this experiment. Section 4 presents the results of the experiment and discusses the results. Section 5 concludes the current work and suggests possible future work.

## 2. PREVIOUS STUDIES

In recent years, several studies for content-based retrieval of mammograms have been published. Honda et al. [5] developed a CBIR system for mammograms to aid the diagnosis of breast lesions. They used spatial gray level dependence matrices to extract textural features of mammograms from 136 clinical cases. Selected ROIs and whole images were used for the performance evaluation. Precisions for selected ROIs ranged from 22% to 100%. Precisions for whole images were from 31% to 100%. El-Naqa et al. [6] proposed a relevance feedback approach, based on incremental learning, for mammogram retrieval. They adapted support vector machines (SVM) to develop an online learning procedure for similarity learning. The approach they proposed is implemented on clustered microcalcification images. They reported that the approach significantly improves the effectiveness of retrieval.

Felipe et al. [7] also extracted textural features of medical images from gray level co-occurrence matrices. Their results show that a descriptor combining gradient, entropy, and homogeneity performs better than any other descriptors with single features. The retrieval accuracy of the best descriptor is over 90%. Antani et al. [8] propose a partial shape matching method for spine X-ray image retrieval. The method developed from the procrustes distance model is to allow users to define a ROI along the vertebra boundary. The delineated partial shape is then matched with every other region in the shape. The similarity of the shape is measured by the procrustes metric.

Liu et al. [9] presented a content-based scheme for retrieving lung computed tomographic (CT) images. In their system, users are allowed to enter a query image by selecting the ROI. Liu et al. applied neural networks for classification, obtaining a set of candidate classes. At the retrieval stage, the system first identifies which class the query is likely to belong to. Then candidate images in the class are ranked, based on geometrical characteristics. Shyu et al. [10] implemented a physician-in-the-loop approach to retrieving images of high-resolution computed tomography (HRCT). The approach requires experts to delineate the pathology-bearing regions and anatomical landmarks for each image. The pathology-bearing regions are used for feature extraction. A multi-dimensional index is developed based on the values of those features.

## 3. METHODOLOGY

The method described in this work contains two major stages—*image analysis* and *image retrieval*. The objective of the image analysis stage is to examine the textural features of mammograms, and then test the statistical significance of the differences between normal and abnormal mammograms. These discriminating features are selected to construct a textural descriptor of mammograms. The descriptor constructed in the image analysis stage is embedded into the CBIR system. The feature descriptor is extracted from the query image in order to retrieve the mammograms relevant to the query image. The performance of the CBIR system is then evaluated. The detailed steps and components of the experiment are described in the following sections.

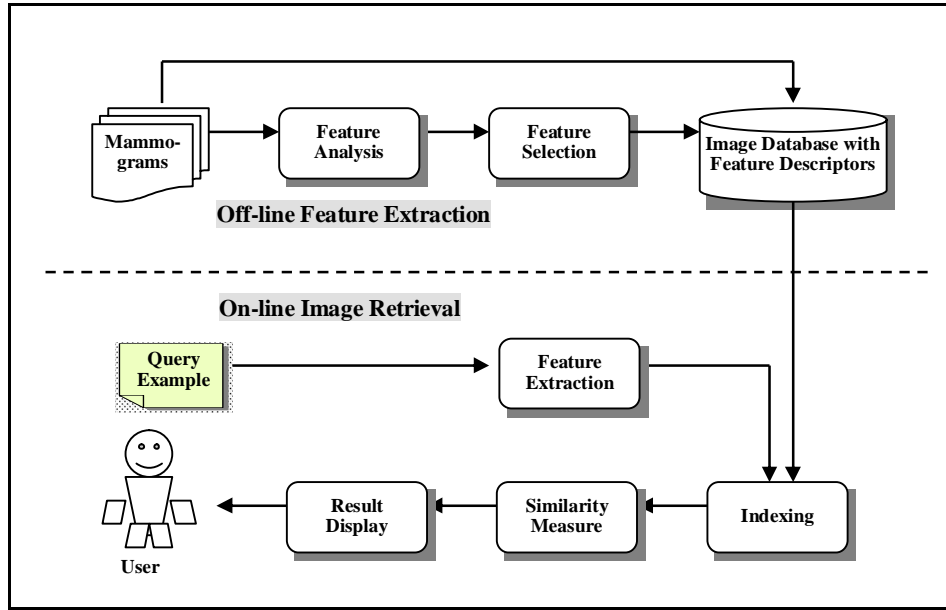


Figure 1. Methodology of content-based image retrieval for mammograms.

### 3.1. Mammogram Dataset

Mammograms were obtained from the database of the Mammographic Images Analysis Society (MIAS) [11]. The size of each image was  $1024 \times 1024$  pixels. All of the images have been annotated for class, severity and location of abnormality, character of background tissue, and radius of circle enclosing the abnormality. Abnormalities are classified into calcifications, architectural distortions, asymmetries, circumscribed masses, speculated masses, and ill-defined masses. Sub-images of size  $200 \times 200$  pixels were cropped as ROIs from each mammogram. 122 sample ROIs (including 29 images in calcification class, 19 in architectural distortion class, 15 in asymmetry class, 25 in circumscribed masses class, 19 in speculated masses class, and 15 in other or ill-defined masses class) were selected deliberately from abnormal tissues. Another 207 ROIs were obtained arbitrarily from normal tissues. These 329 ROIs were used to analyze their textural features based on gray level co-occurrence matrices.

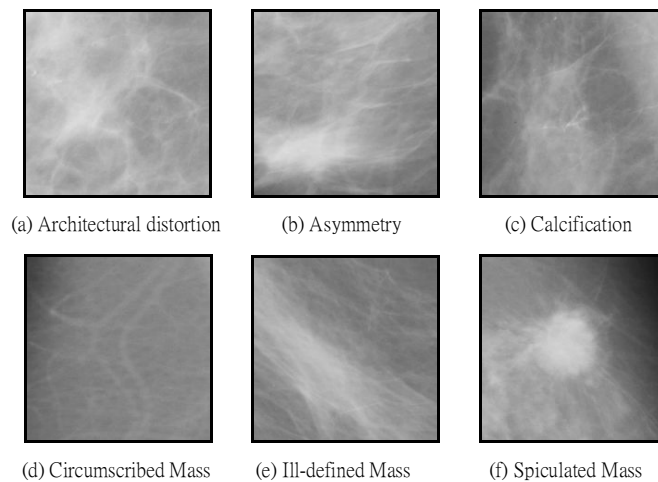


Figure 2. Abnormal mammograms are classified into calcification, architectural distortion, asymmetry, circumscribed masses, speculated masses, and ill-defined masses.

### 3.2. Feature Analysis

The presence of a breast lesion may cause a disturbance in the homogeneity of tissues, and result in architectural distortions in the surrounding parenchyma [12]. Therefore, the textures of digital images contain a lot of valuable information for further research and application. This study applies gray level co-occurrence matrices, a statistical textural method, to analyze the textural features of mammograms and develop descriptors for content-based image retrieval. Gray level co-occurrence matrices will be introduced in the following section.

#### 3.2.1. Gray level co-occurrence matrices

Gray level co-occurrence matrix (GLCM) is a statistical method for computing the co-occurrence probability of textural features [3]. Given an image  $f(x, y)$  of size  $L_r \times L_c$  with a set of  $N_g$  gray levels, define the matrix  $p(i, j, d, \theta)$  as

$$\begin{aligned} P(i, j, d, \theta) = \text{card}\{ & ((x_1, y_1), (x_2, y_2)) \in (L_r \times L_c) \times (L_r \times L_c) | \\ & (x_2, y_2) = (x_1, y_1) + (d \cos \theta, d \sin \theta), \\ & f(x_1, y_1) = i, f(x_2, y_2) = j, 0 \leq i, j < N_g \} \end{aligned} \quad (1)$$

where  $d$  denotes the distance between pixels  $(x_1, y_1)$  and  $(x_2, y_2)$  in the image,  $\theta$  denotes the orientation aligning  $(x_1, y_1)$  and  $(x_2, y_2)$ , and  $\text{card}\{\cdot\}$  denotes the number of elements in the set. Texture features which can be extracted from gray level co-occurrence matrices are as follows [13]:

$$\text{Angular Second Moment (ASM)} = \sum_i \sum_j \{p(i, j)\}^2 \quad (2)$$

$$\text{Contrast} = \sum_{n=0}^{N_g-1} n^2 \left\{ \sum_{i=1}^{N_g} \sum_{\substack{j=1 \\ |i-j|=n}}^{N_g} p(i, j) \right\} \quad (3)$$

$$\text{Correlation} = \frac{\sum_i \sum_j (ij) p(i, j) - \mu_x \mu_y}{\sigma_x \sigma_y} \quad (4)$$

$$\text{Variance} = \sum_i \sum_j (i-j)^2 p(i, j) \quad (5)$$

$$\text{Inverse Difference Moment (ID_Mom)} = \sum_i \sum_j \frac{1}{1+(i-j)^2} p(i, j) \quad (6)$$

$$\text{Sum Average (Sum_Aver)} = \sum_{i=2}^{2N_g} i p_{x+y}(i) \quad (7)$$

$$\text{Sum Variance (Sum_Var)} = \sum_{i=2}^{2N_g} (i - \text{Sum\_Entro})^2 p_{x+y}(i) \quad (8)$$

$$\text{Sum Entropy (Sum_Entro)} = - \sum_{i=2}^{2N_g} p_{x+y}(i) \log\{p_{x+y}(i)\} \quad (9)$$

$$\text{Entropy} = - \sum_i \sum_j p(i, j) \log p(i, j) \quad (10)$$

$$\text{Different Variance (Diff_Vari)} = \text{variance of } p_{x-y} \quad (11)$$

$$\text{Different Entropy (Diff_Entro)} = - \sum_{i=0}^{N_g-1} p_{x-y}(i) \log\{p_{x-y}(i)\} \quad (12)$$

### 3.2.2. Mammogram analysis using GLCM

In order to develop the tailored descriptors described in the next section, it is necessary to analyze the features of the mammogram. In this study, 12 GLCMs are constructed in order to compute each ROI in the  $0^\circ$ ,  $45^\circ$ ,  $90^\circ$ , and  $135^\circ$  directions, each with unit pixel distances of 1, 3, and 5, respectively. The 11 features described above are computed for the 12 GLCMs, thus, resulting in a total of 132 texture features for each ROI.

### 3.3. Feature Selection for Image retrieval

At the stage of feature analysis, 132 texture features are generated for each ROI. We obtain 5,148 texture feature sample images from 20 normal and 19 abnormal images. To select the most discriminant features, a statistical multivariate  $t$ -test is used to assess the significance of the difference between the means of two sample set  $A$  and  $B$ , which are independent of each other in the obvious sense, i.e., the individual measures in set  $A$  are in no way related to any of the individual measures in set  $B$ . The value of  $t$ -test is obtained as follows [14]:

$$D_a = \sum (A_i - \mu_a)^2 \quad (13)$$

$$D_b = \sum (B_i - \mu_b)^2 \quad (14)$$

$$V = \frac{D_a + D_b}{(n_a - 1) + (n_b - 1)} \quad (15)$$

$$\sigma = \sqrt{\frac{V}{n_a} + \frac{V}{n_b}} \quad (16)$$

$$t = \frac{\mu_a - \mu_b}{\sigma} \quad (17)$$

Where  $A_i$  and  $B_i$  in equation (13) and (14) are the  $i$ th element of the set  $A$  and  $B$ , while  $\mu_a$  and  $\mu_b$  are the means of the set  $A$  and  $B$ , respectively.  $D_a$  and  $D_b$  in the equation (13) and (14) are the sum of squared deviates of the set  $A$  and  $B$ .  $V$  in equation (15) is the estimated variance of the source population.  $\sigma$  in equation (16) is the standard deviation of the sampling distribution of sample-mean differences.  $t$  in equation (17) is the value of the  $t$ -test. The degree of freedom (d.f.) is  $(n_a - 1) + (n_b - 1)$ .

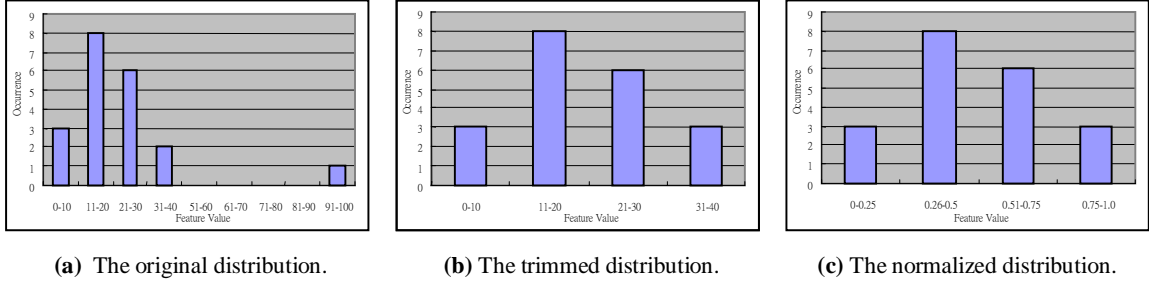
In our case study with 20 normal images (set  $A$ ) and 19 abnormal images (set  $B$ ), the degree of freedom (d.f.) is 37. According to the Table of Critical Values of  $t$  [15], the  $t$  value for 37 degrees of freedom (d.f.) is 1.305. When the value  $t$  obtained in this experiment is greater than 1.305, it means that there is a significant mean difference between normal and abnormal mammograms with regard to the given feature.

The descriptor is composed of features with significant differences in the  $t$  statistic. Individual descriptors were developed for three distances ( $d = 1, 3, \text{ and } 5$ ) in the gray level co-occurrence matrices.

### 3.4. Data Normalization

The purpose of normalization in this experiment is to assign a weight to all features in order to measure their similarity on the same basis. The technique used was to project each feature onto a unit sphere. However, a potential problem exists—if a few elements in a feature space are extremely large, other elements may be dominated by the large ones after normalization.

To solve this problem, the value located at the point of the top 95% of the distribution is taken as the nominal maximum. All features greater than the nominal minimum in the feature space were clipped to the nominal maximal value, i.e. the top 5% of distribution are trimmed. Then all values are divided by the maximal values. An example is given below to illustrate the use of this approach to normalization. Figure 2(a) shows the original distribution of feature values in a feature vector. The results of trimming the top 5% and normalization are illustrated in Figure 2(b) and Figure 2(c).



**Figure 3.** Process of normalization.

### 3.5. Similarity Measure

The similarity measure of two images  $I_a$  and  $I_b$  is the distance between their descriptors  $f_a$  and  $f_b$ . In this work,  $L_2$  norm was adopted to measure the similarity between the query image and each ROI.  $L_2$  is defined as follows.

$$\|d_{ab}\|_2 = \|f_a - f_b\|_2 = \sqrt{\sum_{i=1}^n |f_{a,i} - f_{b,i}|^2}, \quad (17)$$

where  $d_{ab}$  is the similarity distance between descriptors  $f_a$  and  $f_b$ ,  $f_{a,i}$  and  $f_{b,i}$  are the  $i$ th element of  $f_a$  and  $f_b$ , respectively, and  $n$  is the number of elements of the descriptors. The smaller the distance is, the more similar the two images are. After calculating the distance, our CBIR system ranks similarity in descending order and then returns the top five images that are most similar to the query image.

### 3.6. Performance Evaluation

Relevance judgment is a vital part of performance evaluation. The relevance criteria described in Table 1 were developed and used in this work. For example, suppose the query image belongs to the calcification class, the retrieved image would score 0.5 if it belongs to any of the following abnormal classes: ill-defined masses, circumscribed masses, speculated masses, architectural distortion, and asymmetry.

**Table 1.** Criteria for measurement of performance evaluation of CBIR.

Score	Criteria
1.0	The retrieved image belongs to the class of query image.
0.5	The retrieved image belongs to one of the abnormal classes, but not the class of query image.
0	The retrieved image does not belong to any abnormal class.

Precision and recall are basic measures used in evaluating the effectiveness of an information retrieval system. Precision is the ratio of the number of relevant records retrieved to the total number of irrelevant and relevant records retrieved [16]. It indicates the subject score assigned to each of the top five images in this experiment. The formula is expressed as follow:

$$p = \frac{\sum_{i=1}^n S_i}{N} \quad (18)$$

where,  $S_i$  is the score assigned to the  $i$ th hit,  $N$  is the number of top hits retrieved.

Recall is the ratio of the number of relevant records retrieved to the total number of relevant records in the database [16]. It is defined as follow:

$$R = \frac{R_n}{T_n} \quad (19)$$

where  $R_n$  is the number of retrieved relevant hits, and  $T_n$  is the total number of relevant images in the database.

## 4. RESULTS AND DISCUSSION

### 4.1. Results of $t$ Test

**Table 2.** Comparison of mean values obtained by co-occurrence matrices with the distance of 1.

Feature	Normal		Abnormal		t (37 .d.f.)
	$\mu_a$	$D_a$	$\mu_b$	$D_b$	
ASM	1.7721	0.6127	1.3890	0.5542	1.6042
Contrast	5.8209	2.2742	5.6187	3.7038	0.3755
Correlation	-0.3392	0.0742	-0.6135	0.5345	1.6071
Variance	0.2413	0.0768	1.5446	1.9193	-4.2257
ID_Mom	2.9460	0.0671	3.0654	0.1442	-1.1812
Sum_Aver	2.1943	0.3391	2.6984	1.2137	-1.8444
Sum_Var	3.2636	2.2970	0.7699	1.3261	5.9097
Sum_Entro	-6.7253	2.7305	-0.7922	0.1327	-15.6895
Entropy	-4.0729	0.1476	-3.8611	0.3262	-1.3995
Diff_Vari	5.2007	2.3020	6.7532	1.4977	-3.5953
Diff_Entro	-1.7865	0.0332	-1.9603	0.0141	3.5986

Table 2 presents the results of the  $t$  statistic for  $d=1$  of gray level co-occurrence matrices. From the results, it can be seen that the differences between the mean values of ASM, Correlation, Sum\_Var (sum variance), and Diff\_Entro (difference entropy) of the normal and abnormal ROIs are significant ( $t > 1.305$ ). As a result, these four features are selected to construct the descriptor. Table 3 shows that ASM is the only feature with significant discriminating power for two groups of ROIs when  $d=3$ . The descriptor with  $d=3$  contains only ASM.

**Table 3.** Comparison of mean values obtained by co-occurrence matrices with the distance of 3.

Feature	Normal		Abnormal		t (37 .d.f.)
	$\mu_a$	$D_a$	$\mu_b$	$D_b$	
ASM	1.2846	0.3212	1.0075	0.2905	1.6046
Contrast	1.3822	0.2398	1.6247	0.4604	-1.3191
Correlation	0.8311	0.0402	3.6871	15.2704	-3.3534
Variance	1.5943	3.4598	1.1012	0.9300	1.0584
ID_Mom	1.465	0.0518	1.5281	0.1695	-0.6123
Sum_Aver	1.6135	0.1835	1.9774	0.6539	-1.8161
Sum_Entro	-4.9365	1.4720	-0.5789	0.0708	-15.7084
Entropy	-2.6801	0.1052	-2.4949	0.4599	-1.1262
Diff_Vari	4.1858	1.2969	5.3510	0.8446	-3.5984
Diff_Entro	-1.6118	0.0184	1.7424	0.0088	-91.7379

Table 4 shows that Sum\_Var (sum-variance) is the only feature with significant discriminating power for two groups of ROIs when  $d=5$ . The descriptor with  $d=5$  contains only Sum\_Var.

**Table 4.** Comparison of mean values obtained by co-occurrence matrices with the distance of 5.

Feature	Normal		Abnormal		t (37 .d.f.)
	$\mu_a$	$D_a$	$\mu_b$	$D_b$	
ASM	1.2466	0.3019	0.9784	0.2730	0.7924
Contrast	2.1675	0.8227	2.9725	2.0665	-1.5708
Correlation	0.8822	0.0387	4.6453	14.6955	-7.1316
Variance	1.4357	2.8461	1.0654	0.8467	1.0316
ID_Mom	1.2008	0.0497	1.227	0.1625	-0.0742
Sum_Aver	1.5855	0.1775	1.9376	0.6293	-0.8292
Sum_Var	2.3006	2.6245	0.5359	0.6438	4.5870
Sum_Entro	-4.8432	1.4197	-0.5657	0.0677	-8.0337
Entropy	-2.5128	0.1179	-2.3037	0.2179	-0.4203
Diff_Vari	4.1858	1.2969	5.3510	0.8446	-1.6684
Diff_Entro	-1.6118	0.0184	-1.7424	0.0088	0.3150

## 4.2. Results of Performance Evaluation

Precision can be used to describe the accuracy of the proposed CBIR system in finding only relevant images on a search for query images. Table 5 shows that precision rates for the three descriptors ( $d=1$ , 3, and 5) are 47%, 50%, and 51%, respectively. The descriptor with  $d=5$  obtained the highest value in precision and the smallest values in standard deviation. This indicates that the performance of the descriptor ( $d=5$ ) is more stable. On the whole, about half of the results retrieved by these three descriptors are relevant to the query images

**Table 5.** Precision for the 3 descriptors

	CALC	CIRC	SPIC	MISC	ARCH	ASYM	Mean	Std
<b>d=1</b>	42%	44%	54%	44%	46%	54%	47%	5.32%
<b>d=3</b>	57%	43%	54%	42%	54%	51%	50%	6.24%
<b>d=5</b>	48%	53%	54%	50%	50%	48%	51%	2.51%

(Notes: CALC = calcification; CIRC = circumscribed masses, SPIC = speculated masses; ARCH = architectural distortion; ASYM = asymmetry; MISC = other or ill-defined masses.)

Recall measures how well the CBIR system finds all relevant images in a search for a query image. Table 6 indicates that the descriptor with  $d=5$  outperforms the other two. However, the recall values are very close. The largest difference is only 1.5%. The three descriptors can retrieve, on average, about 18% of relevant images in the database. In theory, as precision goes up, recall goes down. The relationship explains why the three recall values are low.

**Table 6.** Recall for the 3 descriptors

	CALC	CIRC	SPIC	MISC	ARCH	ASYM	Mean	Std
<b>d=1</b>	10.69%	12.80%	20.53%	18.67%	16.84%	26.00%	17.59%	5.51%
<b>d=3</b>	13.45%	11.60%	20.00%	20.00%	18.42%	25.33%	18.13%	4.97%
<b>d=5</b>	12.07%	15.20%	21.05%	24.00%	20.00%	22.00%	19.05%	4.51%

The experimental results also show that the descriptor with the largest distance ( $d=5$ ) has the best performance in both precision and recall. The descriptor with  $d=3$  outperforms the descriptor with  $d=1$  in both measures. Although the larger distance has better performance in this experiment, it is still too early to make any conclusions.

## 5. CONCLUSIONS

The main contribution of this work is to present a sound CBIR methodology for mammograms. The methodology is divided into image analysis and image retrieval stages. The purpose of the image analysis is to collect samples from the database, obtain the image signature, and then apply it for the feature extraction in the image retrieval stage. A complete CBIR system based on gray level co-occurrence matrices was implemented. A technique was also proposed to improve the effectiveness of normalization. Three descriptors were evaluated by query images to retrieve the ROIs for the mammogram dataset consisting of 122 images of six sub-classes from abnormal class, and 207 images from normal class. The best precision rate of 51% and recall rate of 19% were achieved with the descriptor using gray level co-occurrence matrices with the pixel distance of 5.

Future work on medical CBIR includes pre-processing of the mammograms (e.g. image enhancement and noise reduction) before conducting image retrieval. More features in conjunction with other feature extraction methods will also be investigated. Weighting factors with regard to mammogram features will be considered. Classification will be applied to find a group of candidates. In addition, relevance feedback will be embedded into the system.

## APPENDIX

**Table 7.** 20 normal images and 19 abnormal images were used to analyze their textural features.

	Normal d=1	Normal d=3	Normal d=5	Abnormal d=1	Abnormal d=3	Abnormal d=3	Mean
ASM (Mean/Std)	1.7721/ 0.7827	1.2846/ 0.5668	1.2466/ 0.5494	1.3890/ 0.7444	1.0075/ 0.5390	0.9784/ 0.5225	1.2797
Contrast (Mean/Std)	5.8209/ 1.5080	1.3822/ 0.4896	2.1675/ 0.9070	5.6187/ 1.9245	1.6247/ 0.6785	2.9725/ 1.4375	3.2644
Correlation (Mean/Std)	-0.3392/ 0.2725	0.8311/ 0.2006	0.8822/ 0.1968	-0.6135/ 0.5345	3.6871/ 3.9077	4.6453/ 3.8335	1.5155
Variance (Mean/Std)	0.2413/ 0.2766	1.5943/ 1.8600	1.4357/ 1.6870	1.5446/ 1.3854	1.1012/ 0.9644	1.0654/ 0.9202	1.1638
ID_Mom (Mean/Std)	2.946/ 0.2582	1.465/ 0.2275	1.2008/ 0.2228	3.0654/ 0.3797	1.5281/ 0.4117	1.227/ 0.4031	1.9054
Sum_Aver (Mean/Std)	2.1943/ 0.5824	1.6135/ 0.4283	1.5855/ 0.4213	2.6984/ 1.1017	1.9774/ 0.8086	1.9376/ 0.7933	2.0011
Sum_Var (Mean/Std)	3.2636/ 2.2970	2.3442/ 1.6498	2.3006/ 1.6200	0.7699/ 1.1516	0.5484/ 0.8222	0.5359/ 0.8024	1.6271
Sum_Entro (Mean/Std)	-6.7253/ 1.6524	-4.9365/ 1.2133	-4.8432/ 1.1915	-0.7922/ 0.3640	-0.5789/ 0.2662	-0.5657/ 0.2602	-3.0736
Entropy (Mean/Std)	-4.0729/ 0.3844	-2.6801/ 0.3244	-2.5128/ 0.3433	-3.8611/ 0.5711	-2.4949/ 0.4599	-2.3037/ 0.4668	-2.9876
Diff_Vari (Mean/Std)	5.2007/ 1.5172	4.1858/ 1.1388	4.1858/ 1.1388	6.7532/ 1.2238	5.3510/ 0.9190	5.3510/ 0.9190	5.1712
Diff_Entro (Mean/Std)	-1.7865/ 0.1805	-1.6118/ 0.1356	-1.6118/ 0.1356	-1.9603/ 0.1248	-1.7424/ 0.0940	-1.7424/ 0.0940	-1.7425

## REFERENCES

1. T. M. Lehmann, M. O. Guld, D. Keysers, T. Deselaers, H. Schubert, B. Wein, and K. Spitzer, "Similarity of medical images computed from global feature vectors for content-based retrieval," *Proceedings of the 8<sup>th</sup> International Conference on Knowledge-Based Intelligent Information and Engineering Systems*, 2004.
2. G. D. Tourassi, "Journey toward computer-aided diagnosis: role of image texture analysis," *Radiology*, **213**(2), pp. 317-320, 1999.
3. R. M. Haralick, "Statistical and structural approaches to texture," *Proceedings of the IEEE*, **67**(5), pp. 786-304, 1979.
4. H. Tamura, S. Mori, and T. Yamawaki, "Texture features corresponding to visual perception," *IEEE Transactions on Systems, Man and Cybernetics*, **6**(4), pp. 460-473, 1976.
5. M. O. Honda, P. M. A. Marques, and J. A. H. Rodrigues, "Content-based image retrieval in mammography: using texture features for correlation with BI-RADS categories," *Proceedings of the 6<sup>th</sup> International Workshop on Digital Mammography*, pp. 401-403, 2002.
6. I. El-Naqa, Y. Yang, N. P. Galatsanos, and M. N. Wernick, "Relevance feedback based on incremental learning for mammogram retrieval," *Proceedings of the International Conference on Image Processing 2003*, pp.729-732, 2003.
7. J. C. Felipe, A. J. M. Traina, and C. Traina, Jr., "Retrieval by content of medical images using texture for tissue identification," *Proceedings of the 16<sup>th</sup> IEEE Symposium on Computer-Based Medical Systems*, pp.26-27, 2003.
8. S. Antani, X. Xu, L. R. Long, G. R. and Thoma, "Partial shape matching for CBIR of spine X-ray images," *Proceedings of SPIE Electronic Imaging*, pp. 1-8, 2004.
9. C. T. Liu, P. L. Tai, A. Y. Chen, C. H. Peng, and J. S. Wang, "A content-based medical teaching file assistant for CT lung image retrieval," *Proceedings of the 2<sup>nd</sup> International Conference on Multimedia and Exposition*, pp. 241-244, 2001.
10. C. R. Shyu, C. E. Brodley, A. C. Kak, A. Kosaka, A. M. Aisen, and L. S. Broderick, "ASSERT: A physician-in-the-loop content based retrieval system for HRCT image databases," *Computer Vision and Image Understanding*, **75**(1/2), pp. 111-132, 1999.
11. J. Suckling, J. Parker, D. R. Dance, S. Astley, I. Hutt, C. R. M. Boggis, I. Ricketts, E. Stamatakis, N. Cerneaz, S.-L. Kok, P. Taylor, D. Betal, and J. Savage, "The mammographic image analysis society digital mammogram database," *Proceedings of the 2<sup>nd</sup> International Workshop on Digital Mammography*, 1994.
12. H. D. Cheng and M. Cui, "Mass lesion detection with a fuzzy neural network," *Pattern Recognition*, **37**(6), pp. 1189-1200, 2004.
13. R. M. Haralick, K. Shanmugan, and I. Dinstein, "Textural features for image classification," *IEEE Transactions on Systems, Man and Cybernetics*, **3**(6), pp. 610-621, 1973.
14. V. Serdobolskii, *Multivariate statistical analysis: A high-dimensional approach*, Kluwer Academic Publishers, London, 2000.
15. A. C. Rencher, *Multivariate Statistical Inference and Applications*, John Wiley & Sons, New York, 1998.
16. R. Baeza-Yates and B. Ribeiro-Neto (Eds.), *Modern Information Retrieval*, Addison-Wesley, Boston, 1999.