# Analysis of Visual Features in Local Descriptor for Multi-Modality Medical Image

Hizmawati Madzin[1], Roziati Zainuddin[2], and Nur-Sabirin Mohamed[3]
[1]Multimedia Department, University Putra Malaysia, Malaysia
[2]Artificial Intelligence Department, University of Malaya, Malaysia
[3]Centre for Foundation Studies in Science, University of Malaya, Malaysia

**Abstract:** *In medical application, the usage of multiple medical images generated by computer tomography such as x-ray, Magnetic Resonance Imaging (MRI) and CT-scan images is a standard tool of medical procedure for physicians. The major problems in analyzing various modality of medical image are the inconsistent orientation and position of the body-parts of interest. In this research, local descriptor of texture, shape and color are used to extract features from multi-modality medical image in patches and interest point's descriptor. The main advantage of using local descriptor is that these features do not need preprocessing method of segmentation and also robust to local changes These features are then will be classified based on its modality using Support Vector Machine (SVM) and k-Nearest Neighborhood (k-NN) classifiers. It shows that different modality have different characteristic and the importance of selecting significance features.*

**Keywords:** *Multi-modality medical images, texture, shape and color features analysis, local patches, local interest points.*

*Received December 7, 2011; accepted February 16, 2014; published online February 26, 2014*

## 1. Introduction

Digital media archives are increasing to colossal proportions in the world today, which includes audio, video [6] and images [5]. Medical images are not exempt where thousands of different medical image modalities are dealt with by medical professionals each day [7]. Medical image analysis is one of the most important and challenging task in order to develop computerized system that can analyze different types of medical images and extract useful information for clinical purposes [11]. Medical image in its widest sense is a part of biological imaging and incorporate radiology, medical photography and microscopy. The major objectives of image analysis in biomedical engineering are to gather information, screening or investigating, to diagnose, therapy and control, monitoring and evaluation [3].

The ongoing developments of medical imaging instrumentation and techniques have created an enormous growth in the quantity of data produced including large amount of medical images. Multimodality of medical images constitutes an important source of anatomical and functional information for diagnosis of disease, medical research and education. The potential of multimodality imaging in providing information can be very useful and significant in biomedical research and clinical investigations.

The capabilities of this application fields can be extended to provide valuable teaching, training and enhanced image interpretation support, by developing techniques supporting the automated archiving and the retrieval images by content. Daily, thousands of medical images produced in various of modality at the radiology department. In Geneva University Hospital alone the number of images produced by Radiology department has increased up to 70,000 images per day in 2007 [12]. Therefore the needed of extract the structure and content of medical images are significant and the important of efficient image retrieval system for patient care, education and research [8].

Content-Based Image Retrieval (CBIR) is a system to extract features to represent the image itself. The visual features used can be classified into primitive, logical and abstract features [24]. Most available system in medical image indexing and retrieval used primitive features which based on texture, shape and color. Nevertheless different medical imaging modalities reveal different characteristics of the human body. The quality of a medical image is determined by the imaging method and equipment characteristics. There are at least 5 factors need to be considered in quality characteristics which include contrast, blur, noise, artifacts and distortion [13]. The combination of primitive features and quality characteristics may increase the performance in CBIR system.

Texture feature generally capture the information of image characteristic with respect to the changes in certain direction and scale of the image. This information gives benefit for regions or images with homogeneous texture. Among popular texture descriptor methods that have been used for medical image indexing and retrieval are co-occurrence matrices [16, 17], wavelets [10, 2] and Fourier transform [18].

Shape features had been often to describe as visual information that based on two classes: Region and contour of the image. Contour shape-based techniques in global level has firstly introduced in [18]. The application is easy to compute and robust to noise but limit in discriminatory power. The study of shape-based spectral descriptor of Fourier and wavelet has been performed by [5]. Though these methodologies are robustness to noise and easy normalization, these descriptors are not inherently rotations invariant. In region-based techniques, moment-based shape features provide a numerical shape-preserving representation that is invariant to translation, rotation and scale [1]. However the drawback of using this technique is it relied on image segmentation. In medical image application automated segmentation of images is completely an unsolved problem since it depends on the content of the medical image [24].

Color features can be the most effective features for system that employ colors image. The color descriptors consist of a number of histogram descriptors, a dominant color descriptor, and a color layout descriptor [16]. This set of descriptors was done to serve different application domains while in this research paper is concentrate in specialized fields' namely medical domain. Color histogram is one of the most frequently used color descriptors that characterize the color distribution in an image. Color histogram descriptor was defined that would be able to capture the color distribution with reasonable accuracy for image search and retrieval applications. Conventionally, Red, Green, Blue (RGB) color space is used for indexing and querying. However this color space does not correspond very well in human perception. Therefore other color space is used as alternative such as Hue, Saturation, Value (HSV) [10], Luv [17] for better respect to human perceive.

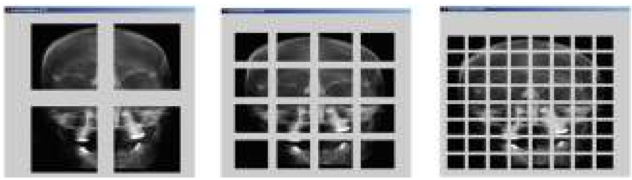
These expressive visual features can be extracted in global and local level. However, it shown in [18], shape feature in global level is weak to discriminate different shape. Methods based on local presentation have shown promise good result for image indexing and retrieval task. Local presentation is an image pattern which has different value or characteristic from its instantaneous neighborhood and associated with a change of image properties (shape, color and texture) simultaneously [22]. The advantages of local presentation are it robust with respect to noise, variability in object shape and partial occlusions [19]. Considering medical images have a specific composition for each modality and anatomic region, we proposed to used local level application and skip the segmentation part in order to speed up analysis and increase the accuracy of the performance.

This research paper is fold to five sections. Initially will be the section 1. The section 2 will describe the methodology of the experimentation for visual features of texture, shape and color in local descriptor followed by the experimentation setup in Section 3. Section 4 presents the results based on the experimentation followed by the conclusion and discussion in section 5.

## 2. Methodology

Experiments are performed using 2000 medical images from various modality inclusive of X-Ray (XR), Computed Tomography (CT) Scan, Ultrasound (US), Nuclear Medicine (NM), Positron Emission Tomography (PET), Magnetic Resonance Image (MRI), Optical Imaging (PX) and Graphic (GX). These images are taken from Image CLEF 2010[1] dataset. We follow the standard image indexing paradigm, in which we initially extract meaningful visual features of texture, shape and color. The major steps involves in the extracting the visual features can be summarized as follows:

1. *Preprocessing*: For texture and shape features extraction process, convert the image to grayscale if needed and resize the image to 256×256 pixels. As for color feature the RGB image need to convert to HSV color space.
2. *Local Descriptor*: Partition each image to 2×2, 4×4 and 8×8 non-overlap patches and extract interest points from the whole image as depicted in Figure 1.



a) 2×2. b) 4×4. c) 8×8.

Figure 1. Examples of local patches.

3. *Feature Extraction Process*: Evaluate each patch and blocks from each interest point with texture and shape descriptor. As for texture we analyze feature of contrast, correlation, energy and homogeneity of each medical image using GLCM method. For shape descriptor we use Hu moment invariant analysis. Finally for color feature we used color histogram descriptor.
4. *Feature Vector for Classification*: We used Support Vector Machine (SVM) and k-Nearest Neighbourhood (k-NN) to train the feature vectors. The results of the evaluation will be the precision and recall value for each feature and also percentage of correctly classified in each modality of medical image.

[1] www.imageclef.org

### 2.1. Local Image Presentation

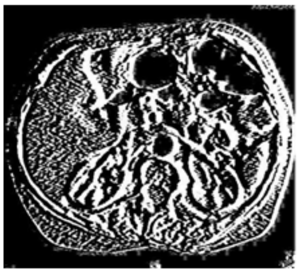
Invariant features stay unchanged when a transformation (object's position and/or orientation changes) is applied. Therefore, local invariant feature is a new image representation that allows describing the objects/parts in any different transformation (translation, rotation and scale). The features can be in points, edges or small image patches [1]. Therefore it is suitable to select interesting locations in the image and to speed up analysis since there is no need of segmentation.

The easiest way to describe local level feature extraction is to use patches of fixed size and location which known as partitioning of the image [11]. Each medical image will be partitioned into 2×2, 4×4 and 8×8 non overlap patches as depicted in Figure 1. The purposed to have variety of patches size is to evaluate which size is suitable for local feature extraction.

As for interest points, we focused on extracting features using interest points that based on corner detector. Harris detector is a corner detector created by Harris and Stephens is based on second moment matrix, M. The method developed is to cater image regions with texture and isolated feature. The matrix calculation is describing the gradient distribution in the local neighbourhood of a point as shown in Equation 1:

$$M = \sum w(x, y)\begin{bmatrix} i_x^2(x) & i_x\, i_y(y) \\ i_x\, i_y(x) & i_y^2(y) \end{bmatrix} \quad (1)$$

Initially the first order local image derivatives, $i_x$ and $i_y$ are computed. Then take the product of these gradient images. Figure 1 shows the initial step of Harris Detector.



a) The local derivative of an image in the x-direction.

b) The local derivative of an image in the y-direction.

Figure 1. The initial step of the Harris detector.

Next is to smooth the image with Gaussian kernels, *w(x,y)* in different scale value, ($\sigma$) as in Equation 2:

$$w(x,y) = g(x,y,\sigma) = ½\,(2\pi\sigma^{e\,[-x2+y2/\,2\pi]}) \quad (2)$$

The eigenvalues represent the significant signal changes in two orthogonal directions in neighbourhood around the point. Different values in Gaussian Kernel may produce different number of interest points [9].

Finally, Harris measure for cornerness that can be defined by positive local extrema in Equation 3:

$$cornerness = det(M) - \lambda\; trace(M) \quad (3)$$

Where typically, $\lambda = 0.04$.

Once the interest points are generated, a block of 20×20 pixels is generated for each point and interest point as centre point as depicted in Figure 2. In this research, the maximum number of generated interest points is 20. Each block will be extracted with shape and texture features. Color will not be used since HSV color histogram is not suitable in extracting a small size block.

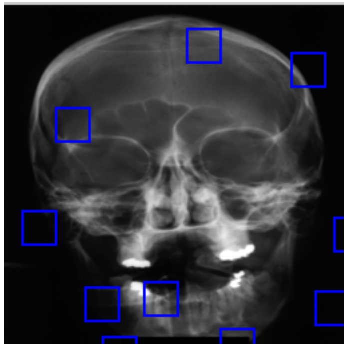


Figure 2. Harris interest point feature extraction.

### 2.2. Texture Feature

In this experimentation we analyse texture features based on contrast, correlation, energy and homogeneity. These texture measures try to capture the characteristics of the image parts with respect to changes in certain directions and the scale of the changes. Gray-Level Co-occurrence Matrix (GLCM) is a statistical method of examining texture that considers the spatial relationship of pixels which also known as the gray-level spatial dependence matrix. The GLCM functions characterize the texture of an image by calculating how often pairs of pixel with specific values and in a specified spatial relationship occur in an image, creating a GLCM, and then extracting statistical measures from this matrix. The texture features information is derived from GLCM by using the following formula:

1. *Contrast*: Measures the local variations in the GLCM:

$$\sum_{i,j=0}^{N-1} P_{i,j}\,(i-j)^2 \quad (4)$$

2. *Correlation*: Measures the joint probability occurrence of the specified pixel pairs:

$$\sum_{i,j=0}^{N-1} P_{i,j}\left[\frac{(i-\mu_i)\,(i-\mu_\mu)}{\sqrt{(\sigma_i^2)\,(\sigma_j^2)}}\right] \quad (5)$$

3. *Energy*: Provides the sum of squared elements in the GLCM. Also, known as uniformity or the angular second moment:

$$\sum_{i,j=0}^{N-1} P_{i,j}\,(-ln\, P_{i,j}) \quad (6)$$

4. *Homogeneity:* Measures the closeness of the distribution of elements in the GLCM to the GLCM diagonal:

$$\sum_{i,j=0}^{N-1} \frac{P_{i,j}}{1+(i-j)^2} \quad (7)$$

Where $i$ is the row number, $j$ the column number, $P_{i,j}$ the normalized value in the cell $i,j$ and $N$ the number of rows or columns.

## 2.3. Shape Feature

In this research paper we analyse shape descriptor using Hu moment invariant method. Medical image is usually complex to high variability and only have subtle differences from other image in the context of visual appearance, it seems inappropriate to compute in global level. Alternatively, we calculate the shape feature in local patches to access more detail information of the medical image without any segmentation.

For each patch, we used Hu moment function $m_{pq}$ with $(p+q)$ order moment is given as:

$$m_{pq} = \sum_y x^p y^q f(x,y) \quad (8)$$

Computing the coordinates of centers of mass for each patch as below:

$$x' = \frac{M_{10}}{M_{00}} \quad y' = \frac{M_{10}}{M_{00}} \quad (9)$$

Therefore the central moments can be defined as:

$$\mu_{pq} = \sum_x \sum_y (x-x')^p (y-y')^q f(x,y) \quad (10)$$

When a scaling normalization is applied the central moments change as:

$$\eta_{pq} = \frac{M_{pq}}{M_{00\gamma}} \quad (11)$$

Where $\gamma = (p + q / 2) + 1$ is the normalization factor. In particular, Hu moment invariant method described a set of six moments that are rotation, scaling, translation invariant and the seventh invariant is skew invariant.

## 2.4. Color Feature

The color histogram describes the proportion of pixels of each color in an image with simple and computationally effective manner. The color histogram is obtained by quantizing image colors into discrete levels and then counting the number of times each discrete color occurs in the image. In this research paper, generic color histogram is used. Initially the image color needs to be in HSV format. Then the HSV histogram is created. Generic color histogram descriptor was defined that would be able to capture the color distribution with reasonable accuracy for image search and retrieval applications. Later is to do normalization of color histogram to produce better result.

## 2.5. Classifier

Two classifiers will be used in this research paper namely SVM and k-NN.

### 2.5.1. Support Vector Machine

SVM is a binary classifier that finds the optimal linear decision surface based on the concept of structural risk minimization. The input to a SVM algorithm is a set $\{x_i, y_i\}$ of labeled training data where $(x_i, \ldots, x_n)$ is the data and $y_i$ ε (+1, -1) is the label. SVM tries to discover a hyperplane to separate the two classes by selecting the maximized distance value of support vectors. The hyperplane has the form of:

$$f(x) = \sum_{i-1}^{n} \alpha_i \; y_i \; x_i \; .x + b \quad (12)$$

The coefficient $\alpha_i$ and $b$ in equation 12 are the solutions of quadratic programming problem [19]:

$$maximize \sum_{i-1}^{l} \alpha_i - \frac{1}{2} \sum_{i,j-1}^{l} \alpha_i \alpha_j (x_i, y_i) y_i y_j \quad (13)$$

With:

$$\sum_{i=1}^{l} \alpha_i y_i = 0 \quad 0 < \alpha_i < C \quad (14)$$

Where $C$ is the regularization parameter, selected by the user. For multi-class classification which has to classify more than two classes, there are two general approaches for classifier which are one-against-one, which the calculation is made from each pair of classes and the final classification is based on majority voting scheme all classifiers, while one-against all is the calculation from each class versus all classes and then the first object that is classified as a single class is the class of the unseen data point [20].

### 2.5.2. k-Nearest Neighbor

k-NN classifier is a direct concept and amongst the simplest method in pattern classifier. It classified an object by majority votes of its neighbour. This means, the object being assigned to the class most common amongst its K nearest neighbour.

## 3. Experiment Setups

We conducted three types of experiments. The experiments involved of 2000 medical images from eight different modality which include XR, CT, US, NM, PET, MRI, PX and GX. We used SVM and k-NN classifiers. The training and testing dataset is based on 10-fold cross validation.

- *Experiment 1. Classifiers Experiment with k-NN and SVM*: The performance evaluation of k-NN and SVM are performed. Prior to determined which K value is the best for this research study we implement K = {1, 2, 3, 4, 5} for the experiment. As for SVM classifier, we compare the kernel performance between polykernel and RBF in SVM. Later the comparison between k-NN and SVM classifier is performed.

- *Experiment 2. Local Level Evaluation with 2×2, 4×4 and 8×8 Patches and Interest Points*: Experiments setup for local level performance involves patches and interest points. For patches experimental, there will be three subcategories. Each subcategory the medical image will be divided into 2×2, 4×4 and 8×8 patches. The experiments will use the combination of texture and shape features. The evaluation is based on which size of patches performs the best performance in this experiment. The next experiment in local level is to extract visual features based on interest points. Prior to that Haris detector method is used to extract interest points in each medical image. For each interest point, a square block of size 20×20 pixels is created as interest point will be the centre point. Then texture and shape features will be extracted in these boxes and combine all features in one feature vector. We limit to produce the maximum of 20 interest points for each medical image. This is because to avoid huge large dimension of feature vector. The final experiment for this section is to compare the performance between patches, interest points and combination of patches and interest points.
- *Experiment 3. Multi-modality Medical Images Evaluation*: The experiment was to extract texture and shape visual features using local level and color feature. The performance will be evaluated based on the formula correctness rate, which is the percentage of correctly classified image divided by total number of images and also the rate of precision and recall of each modality.

## 4. Results

Before we can evaluate the performance of visual features, we need to select optimum value for k-NN and SVM classifiers. These values will be used for other upcoming experiments. The purposed of experiment is to identify which is the best K value in k-NN classifier for K = {1, 2, 3, 4, 5} and for SVM classifier we compared the performance of polykernel and RBF kernel. Figure 4 illustrates the MAP value of texture features using k-NN classifier. There is only subtle difference from each K, nevertheless K=1 has higher value compare to others.

Comparison between polykernel and RBF kernel has resulted to polykernel with MAP value of 0.0638 compared to RBF kernel is 0.0072 as depicted in Figure 5. Therefore we decided to used K-NN classifier with K=1 and polykernel for SVM classifier for other experiments.

Different size of patches will produce different dimensional of feature vector as depicted in Table 1. The more patches are created the higher number of dimensions produce in a feature vector. Figure 6 shows the experiment result. The chart represents percentage of accurately classified and the MAP value produce by each size of patches. From the figure it shows that 4×4 patches has performed the highest percentage of correctly classified with 43.17% and the MAP value is 0.121. However, it is only 2.12% different from 8×8 patches and 3.3% different from 2×2 patches.

Table 1. List of dimensions produced based on patches.

| Patches | Dimensions |
|---|---|
| 2×2 | 44 |
| 4×4 | 108 |
| 8×8 | 166 |

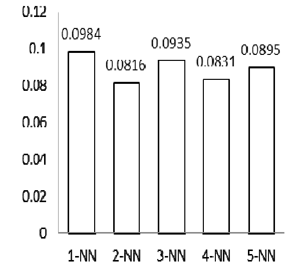



Figure 3. MAP value for K = {1, 2, 3, 4, 5}.

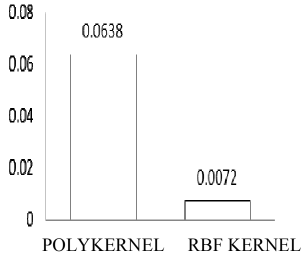



Figure 4. MAP value for SVM using poly kernel and RBF kernel.

The result shows that using 4×4 patches has produced better result compared to 2×2 and 8×8. Applying 2×2 patches in medical image may course too general information extract in each patch since the patches is large and more likely extract the whole image (global level) instead. In contrast to 8×8 the patches are too small and many of the patches contain plain dark pixel which surrounding at the edges of the medical image. Therefore the 4×4 patches is the most suitable size for local level in extracting visual features.

Next is to extract features from interest blocks created based on interest points. Initially interest points are produced in medical image. A block size of 20×20 pixels is created in each interest point produced. As previously mentioned, 20 interest points will be generated for each medical image. Therefore the dimension of feature vector for each medical image is 460 dimensions. Each feature involves of shape and texture features. The system will extract these features on each block. The result of 43.93% of correctly classified and MAP value of 0.128 are obtained from the experiment. It clearly shows that the interest point's result is better than 4×4 patches.

The final experiment for this section is to combine 4×4 patches and interest blocks which produce 568 dimensions (108 dimensions of 4×4 patches, 460 dimensions of interest points). As depicted in Figure 7 the combination result is 0.45% higher than interest block's result which is 44.38% of correctly classified and the MAP value is 0.135.

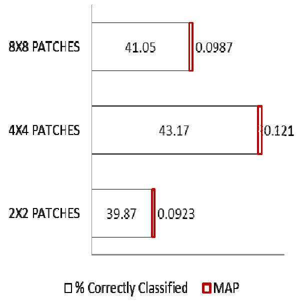



Figure 5. Comparison between 2×2, 4×4 and 8×8 patches.

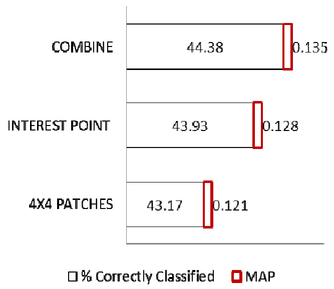



Figure 6. Comparison between 4×4 patches, interest points and combined both.

The final result is the result from multi-modality medical image evaluation. Table 2 show the percentage of accuracy classified based on each modality for visual features of texture, shape and color. From the table we can explain that GX has the higher value in each visual feature. This is because GX only represent chart and graph of medical image which it is not a complex image. In contrast XR and MR have low value in all features due to the complexity of the image. Nevertheless the difference between the values from each modality is subtle.

Table 2. Percentage of multi-modality medical image correctly classified.

| Modality | Texture | Shape | Color |
|---|---|---|---|
| **CT** | 56 | 49 | 30 |
| **GX** | 91 | 89 | 85 |
| **XR** | 49 | 13 | 29 |
| **MR** | 17 | 25 | 42 |
| **NM** | 26 | 52 | 69 |
| **US** | 38 | 64 | 41 |
| **PET** | 54 | 27 | 38 |
| **PX** | 18 | 46 | 52 |

Figures 8, 9 and 10 show the accuracy value of precision and recall for texture, shape and color features. It shows that CT, XR, US, PET and PX are suitable to be classified using texture descriptor as depicted in Figure 8. As for moment shape descriptor, the result almost similar with texture, but the value of accuracy is higher as shown in Figure 9. It explains that shape descriptor using local level has better presentation compare to texture descriptor in analysing multi-modality medical images. Finally, the color descriptor in Figure 10 shows that GX, PX, NM and PET have higher value. This is due to those modalities have used more colors compare to other modalities which concentrate only on grey-scale image.

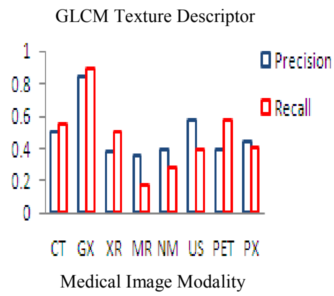



Figure 7. Precision and recall for texture feature.

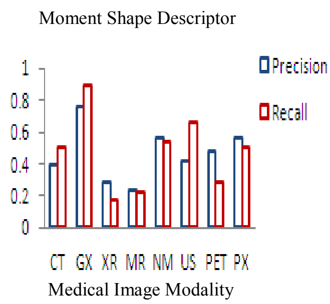



Figure 8. Precision and recall for shape feature.

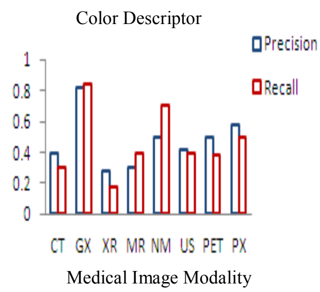



Figure 9. Precision and recall for color feature.

## 5. Conclusions

This research presents the evaluation of local descriptor of visual features for multi-modality medical image. Prior to that an experiment of classifier comparison has been done and it shows that k-NN classifier is suitable in classification of multi-modality medical image. Furthermore, we also evaluate the effectiveness between patches and interest points for local descriptor. It shows that interest block is better than patches; nevertheless combination of both descriptors can perform better result but will lead to high value of dimensional. For multi-modality medical image experiment, it shows that different modality of medical image has different value of features. From the explanation in result section, it shows that different modality of medical images has different characteristic in term of selecting significance features value. As for example, the GX mostly has higher value in precision and recall. This is because this image is not complex compare to other modalities. It also shown that shape descriptor in local level has better performance compare to texture descriptor in the application of medical images. It may be due the texture of medical images have only subtle difference in each modality especially in CT, MR, XR and US. For future work we will further experiment for more specific in medical image such as anatomy and pathology.

## References

[1] Burkhardt H. and Siggelkow S., "Invariant Features in Pattern Recognition Fundamentals and Application," Kotropoulos C. and Pitas I., Nonlinear Model-Based Image/Video Processing and Analysis, John Wiley and Sons, 2001.

[2] Chen Y., Tseng D., and Chang P., "Wavelet-Based Medical Image Compression with Adaptive Prediction," *in Proceedings of International Symposium on Intelligent Signal Processing and Communication Systems*, Hong Kong, China, pp. 825-828, 2005.

[3] Costrarido L., *Medical Image Analysis Method: Medical-Image Processing and Analysis for CAD Systems*, Taylor & Francis, USA, 2005.

[4] Eakins J. and Graham M., "Content-Based Image Retrieval," *Joint Information Systems Committee*, 1999.

[5] Flickner M., Sawhney H., Niblack W., Ashley J., Huang Q., Dom B., Gorkani M., Hafner J., Lee D., Petkovic D., Steele D., and Yanker P., "Query by Image and Video Content: The QBIC System," *IEEE Computer Society* , vol. 28, no. 9, pp. 23-32, 2001.

[6] Halin A., Rajeswari M., and Abbasnejad M., "Soccer Event Detection via Collaborative Multimodal Feature Analysis and Candidate Ranking," *the International Arab Journal of Information Technology*, vol. 10, no. 5, pp. 493 - 502, 2013.

[7] Haux R., "Health Information System: Past, Present and Future," *International Journal Med Inform*, vol. 75, no. 3-4, pp. 268-281, 2006.

[8] Kuo W., Chang R., Lee C., and Moon W., "Retrieval Technique for the Diagnosis of Solid Breast Tumors on Sonogram," *Ultrasound in Medicine & Biology*, vol. 28, no. 7, pp. 903-909, 2002.

[9] Madzin H. and Zainuddin R., "Medical Image Analysis Using Local Invariant Features-Interest Points," *in Proceedings of International Conference of Software Engineering & Computer Systems*, Kuantan Malaysia, pp. 303 - 307, 2009.

[10] Manjunath B., Ohm J., Vasudevan V., Yamada A., "Color and Texture Descriptors," *IEEE Transaction CSVT*, vol. 11, no. 6, pp. 703-715, 2001.

[11] Mueen A., Zainuddin R., and Baba M., "Automatic Multilevel Medical Image Annotation and Retrieval," *Journal of Digital Imaging*, vol. 21, no. 3, pp. 290 - 295, 2008.

[12] Muller H., Michoux N., Bandon D., and Geissbuhler A., "A Review of Content-Based Image Retrieval Systems in Medical Applications - Clinical Benefits and Future Directions" *International Journal of Medical Informatics*, vol. 73, no. 1, pp. 1 - 23, 2003.

[13] Perry S., *The Physical Principles of Medical Imaging*, Wolters Kluwer Law & Business, 1987.

[14] Pizurica A., Philips W., Lemahieu I., and Acheroy M., "A Versatile Wavelet Domain Noise Filtration Technique for Medical Imaging," *IEEE Transactions on Medical Imaging*, vol. 22, no. 3, pp. 323-331, 2003.

[15] Sclaroff S., Taycher L., La Cascia M., "ImageRover: A Content-Based Browser for the World Wide Web," *in Proceedings of IEEE Workshop on Content Based Access of Image and Video Libraries*, Puerto Rico, San Juan, pp. 2-9,1997.

[16] Setia L., Teynor A., Halawani A., and Burkhardt H., "Grayscale Medical Image Annotation Using Local Relational Features," *Pattern Recognition Letters*, vol. 29, no. 15, pp. 2039-2045, 2008.

[17] Smith J. and Chang S., "Visual SEEK: A Fully Automated Content-Based Image Query System," *in Proceedings of the 4th ACM International Multimedia Conference and Exhibition*, Boston, USA, pp. 1-12, 1996.

[18] Sumanaweera T. and Liu D., "Medical Image Reconstruction with the FFT," available at: http://http.developer.nvidia.com/GPUGems2/gpugems2_chapter48.html, last visited 2005.

[19] Tuytelaars T. and Mikolajczyk K., "Local Invariant Feature Detectors: A Survey," *Foundations and Trends® in Computer Graphics and Vision*, vol. 3, no. 3, pp. 177-280, 2008.

[20] Vapnik V., "Structure of Statistical Learning Theory," *Computational Learning and Probabilistic Reasoning*, pp. 1-3, 1996.

[21] William H., Antani S., Long L., Neve L., and Thoma G., "SPIRS: A Web-Based Image Retrieval System for Large Biomedical Databases," *International Journal Medical Inform*, vol. 78, no. 1, pp. S13-S24, 2009.

[22] Zhang D. and Lu G., "A Comparative Study on Shape Retrieval Using Fourier Descriptors with Different Shape Signatures," *in Proceedings of the International Conference on Intelligent Multimedia and Distance Education*, Fargo, USA, pp. 1 - 9, 2001.

[23] Zheng X., Zhou M., and Wang X., "Interest Point Based Medical Image Retrieval," *in Proceedings of the 2nd International Conference Medical Imaging and Informatics*, Beijing, China, pp. 118-124, 2008.

[24] Zhu Y., De Silva L., and Ko C., "Using Moment Invariants and HMM in Facial Expression Recognition," *Pattern Recognition Letters*, vol. 23, no. 1-3, pp. 83-91, 2002.


**Hizmawati Madzin** is a PhD student specializing in multi-modal medical image anaysis and retrieval at University of Malaya (UM), Malaysia. She obtained her Bachelor of Computer Science (Hons.) from University Putra Malaysia (UPM) and Master of Software Engineering from UM. She is now working at the Faculty of Computer Science & Information Technology, Universiti Putra Malaysia.

**Roziati Zainuddin** is a Professor in Department of Artificial Intelligence, Faculty of Computer Science & Information Technology, University of Malaya, Malaysia. Her current research interests include Scientific Multimedia Visualization of Medical Process, Computer Vision and Image Processing.

**Nur-Sabirin Mohamed** is a Professor in Physics Division, University of Malaya, Malaysia. She received MSc degree in 1987 and PhD degree from UM in 2004. Her research interest is in the area of computer vision and medical image processing, advanced materials which include polymer electrolytes, ceramic electrolytes and electrode materials.