International Journal of Information, Control and Computer Sciences
ISSN: 2517-9942
Vol:2, No:1, 2008

# Mining Image Features in an Automatic Two-Dimensional Shape Recognition System

R. A. Salam, M.A. Rodrigues

***Abstract*****—**The number of features required to represent an image can be very huge. Using all available features to recognize objects can suffer from curse dimensionality. Feature selection and extraction is the pre-processing step of image mining. Main issues in analyzing images is the effective identification of features and another one is extracting them. The mining problem that has been focused is the grouping of features for different shapes. Experiments have been conducted by using shape outline as the features. Shape outline readings are put through normalization and dimensionality reduction process using an eigenvector based method to produce a new set of readings. After this pre-processing step data will be grouped through their shapes. Through statistical analysis, these readings together with *peak measures* a robust classification and recognition process is achieved. Tests showed that the suggested methods are able to automatically recognize objects through their shapes. Finally, experiments also demonstrate the system invariance to rotation, translation, scale, reflection and to a small degree of distortion**.**

***Keywords*****—** Image mining, feature selection, shape recognition, peak measures.

## I. Introduction

IT becomes extremely easy to obtain and store large quantities of data. However, research progress in image mining still has a big room for improvement, particularly in multimedia images. One of the greatest challenges is devising an effective automatic recognition and categorization. In today's industrial and ever more automated world, there is a strong need for robust and reliable areas, such as medical, manufacturing, autonomous navigation and multimedia applications.

A robust automatic recognition without a priori information has been a primary concern to many researchers in image mining [1]. The core issue is tackling the problem without any human intervention in feeding information from the beginning to the end of the recognition process. In this research paper, an automatic shape recognition is presented. Choosing the right features for object representation is another main issue in this area. Too much information can lead to a slow and inefficient system, whereas too little information can result in misclassification. Therefore, choosing the right features is one of the main problem, especially in developing a robust system. The selection process must be carefully decided, particularly as, once information for an object is discarded, it normally cannot be restored later. This is more challenging when the selected features need to be used for different tasks. Another issue in feature selection is to reduce computational cost. Using all available features to recognize objects can suffer from curse dimensionality. Feature selection and extraction is the pre-processing step of image mining. A lot of information can been reduce if only shape outlines of an object are considered. The use of shape outline is not a new idea and it has shown a significant results for a recognition system [2] – [7]. Others have use shape, color and texture [8], [9].

The first part of this research is based on our early vision system. Early vision system plays an important part of our earlier stage of perception. One of its function at this stage is the edge and bar detection. At this level it processes the visual information necessary for perception and then brings it to the higher order sensory in the brain. In relation to our early vision systems, shape outline was used as the features for this recognition system.

The eigenvector based method are dimensionality reduction schemes and have been investigated for the image mining process in this research. It was used for reducing the amount of data that need to be process. This will produce an effective and robust automatic recognition system. The mining problem that has been focused is the grouping of features for different shapes. Grouping objects according to their shapes can provide a meaningful categories. It provides a hierarchical model for recognition and classification of objects that are defined purely through their shapes. The approach assumes no priori information regarding the geometrical knowledge of the shape in term or scale, rotation, location or particular features.

Through statistical analysis, these readings together with *peak measures* a robust classification and recognition process is achieved. Tests showed that the suggested methods are able to automatically recognize objects through their shapes. Finally, experiments also demonstrate the system invariance to rotation, translation, scale, reflection and to a small degree of distortion.

Manuscript received April 21, 2004.
M. F. Omar is with Universiti Utara Malaysia, Perlis, Malaysia. (ijatz@yahoo.com).
R. A. Salam is with Universiti Sains Malaysia, 11800 Penang, Malaysia. (Tel: +60-4-6532486; Fax: +60-4-6573335; email: rosalina@cs.usm.my).
R. Abdullah is with Universiti Sains Malaysia, 11800 Penang, Malaysia. (email: rosni@cs.usm.my).
N. A. Rashid is with Universiti Sains Malaysia, 11800 Penang, Malaysia. (email: nuraini@cs.usm.my).

## II. Motivation, Methods and assumptions

### A. Vision System

The first area of visual processing is the retina of the eye. Retina does not only collect light through the photoreceptors, but serves as a filter as well. Information from the retina is transmitted through the optic nerve to the lateral geniculate nucleus (LGN). This is where it processes the necessary information for perception. From LGN, neurons that carry the visual information will send it to the primary visual cortex. The primary visual cortex, contains neurons which respond to various features of the image. The neurons respond most strongly to edges of a particular orientation [10]. This edge-detection process is through the connection from LGN to primary cortex.

Our studies was inspired by the front end visual system. At this stage the basic visual information, that is the edges is available for perception. The visual information is then carry for further processing in our extra-striate visual cortex. Recognition and motion processing happen at this stage. Zenon Pylyshyn [11], in his paper concluded that the output of early vision system consists of shape representations involving at least surface layouts, occluding edges, where these are parsed into objects and other details allow parts to be looked up in a shape–indexed memory in order to identify known objects.

In the research conducted, the visual information at the front end visual system, that is the shape representation, namely the shape outline is used as an input to the vision system. Therefore, the visual information from the shape outline, together with the knowledge that we have for an object, this vision system is expected to recognize a given objects if it has seen it before otherwise it will start to learn new views of new objects.

### B. Shape Outlines

Shape outline will be the main feature extracted from the image and it was based on the human vision system. An edge following technique was used for acquiring shape outline readings and storing them in a list format. This technique is used assuming that there is no background information on the image. This is a new method for an outline detection based on edge following technique. The reason why a new method was developed, instead of using an existing method, was that, the outline reading used in the prototype system needs to be in an ordered or sequential list format. Another reason was the need for an automatic boundary detection method. This cannot be obtained from the Snakes [12] active model, even though it has been used in a number of vision application systems. In snakes, the initial point needs to be chosen by external force.

Brownian String [13] seems to be an automatic boundary detection method, but its use is limited to a number of applications and its output is not an order, sequential list of points taken at regular intervals.

The first stage is data acquisition where two-dimensional images that is a raster format were used. Initially the starting point of an object in the image need to be identified. Once the initial position was found, the edge was followed around the object until the initial point was reached. This initial point of the outline is determined by firing a number of simulated range finders sensors from random positions at the border of the display window pointing to its centre until a point on the outline is encountered. This can be seen in Fig. 1. As soon as an object is encountered by at least two nearby simulated sensors the pixel co-ordinates *(x,y)* will be returned. Only one point of these co-ordinates will be used as the initial point.

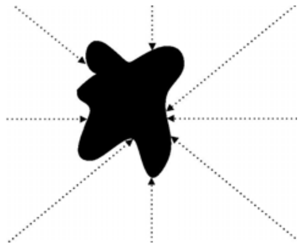

FIGURE 1
INITIAL POINT OF THE SHAPE OUTLINE

In the next stage, three virtual sensors were configured. The reason why three sensors were used is because the average of these three sensors can be used to eliminate noise problem and can produced better results. These three sensors are configured to be at least two pixels apart. These three sensors follow the object's outline, recording a list of *(x,y)* positions. The first thing that these three sensors will do is to rotate until the next reading is obtained. All three sensors must hit the shape for a valid readings.

Rotation angles for all three sensors are recorded. The rotation angle $\theta$ can be computed as:

$$\theta = \frac{1}{3}\int_{0}^{3} \theta_i$$

$$\theta_i = \tan^{-1}(\Delta y_i / \Delta x_i) \tag{1}$$

where

$$\Delta y_i = y_0 - y_1$$

$$\Delta x_i = x_0 - x_1 \tag{2}$$

$(x_0 y_0)$ is the initial point and $i$ is the range of 1…3 and $\theta$ is the average of the three angles.

The above procedure is repeated until the initial point, that is $(x_0 y_0)$ is reached, where this process will be automatically stopped. The ordered list of all the angle rotation values (the average), that is, the $n$ measurements, $\theta$ , is constructed as:

$$\theta = [\theta_1, \theta_2, ..., \theta_n] \tag{3}$$

$d\theta$ the difference from one angle to the next one. The list of the differential angles $d\theta$ is the pre-processing stage of the image mining process. These data will go through a process of dimensionality reduction and normalization before being used for training and testing in the recognition stage.

### *C. Normalization and Dimensionality Reduction*

Outline readings went through a process of transformation which involved normalization and dimensionality reduction. This transformation used eigenvector based methods, which can reduce the computational burden of pattern recognition algorithms and the image mining process. To increase the statistical significance of the used samples, random noise were added to each outline reading creating new equivalent views of the same object.

The list of $d\theta$ described earlier is computed during the feature extraction process, is further filtered by calculating the average of every three readings. Each current value is substituted with the average reading. The reason for this, is that in the taking of outline readings, readings are sometimes affected by noise, and reduces the error created by noise. The new list of $d\theta$ is computed as:

$$d\theta_i = \left( \sum_{i-1}^{i+1} d\theta_i \right) / 3 \tag{4}$$

where $I$ = 2,3,..n-1. Therefore, a new set of $d\theta$ is obtained. Let the list of $d\theta$ be transformed into list of vectors **V**, where **V** = $\{v_1, v_2, \ldots, v_n\}$. A vector $v_i$ is defined as:

$$v_i = \begin{pmatrix} x_i \\ y_i \end{pmatrix} = \begin{pmatrix} m \cos d\theta_i \\ m \sin d\theta_i \end{pmatrix} \tag{5}$$

The new co-ordinates after the transformation can be constructed as follows:

$$C = E^T V \tag{6}$$

Where, $C$ $(C = \{c_1, c_2, ...c_n\})$ is the new set of co-ordinates after the transformation, $V$ $(V = \{v_1, v_2, ...v_n\})$ is the set of vectors computed from rotation angles and $E$ $(E = \{e_1, e_2, ..., e_n\})$ is the set of eigenvectors. The eigenvector $e_i$ is computed as:

$$e_i = \begin{pmatrix} k_x (x_0 + i) \\ k_y \left( \dfrac{d\theta_i}{|d\theta_{\max}|} \right) \end{pmatrix} \tag{7}$$

where $x_0$ is 0 and $|d\theta_{\max}|$ is the largest absolute value in the list. $k_x$ and $k_y$ are arbitrary constant factors in the $x$ and $y$ axis. These constants are determined experimentally and play a very important role in determining the new co-ordinates. The value chosen for $k_y$ was 50 and the value chosen for $k_x$ was 120.

Normalization is carried out on $C$, where three of the values are added up and the average obtained. The new set of readings after normalization is $Z$, that is $Z = \{z_1, z_2, ..., z_n\}$ and represent the new set of vectors. These new set of data is the data that produced in the mining process. The data will then go through the next stage that is the shape categorization process. Data that has been mined can be visualize using a graphical format. An example of the graphical representation of a rectangular shape can be seen in Fig. 2. Peaks in the graph correspond to changes in shape, such as sharp corners.

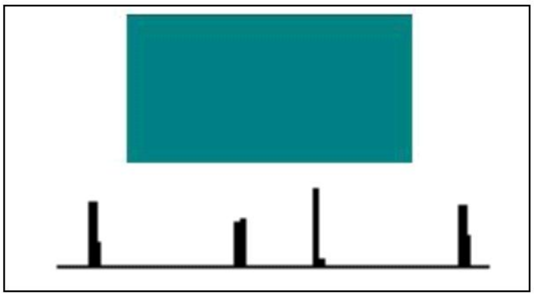

FIGURE 2
A RECTANGLE INITIAL AND ITS GRAPHICAL DATA REPRESENTATION OF THE SHAPE

### *D. Pattern Recognition*

Pattern matching during the classification stage represented another major task in this research. Since there is no priori information of every new images, any new data will have to be trained and saved in the database. This is done by using the unsupervised classification. The first set of data will go through the statistical process, trained and saved in a database. The following sets of data will undergo the same process and saved in the database. The research concentrates on shapes recognition and different shapes have its own representation. Similar shapes will be put in a same category and grouped properly. This is similar on how human brain works where there is a shape-indexed memory [10].

Data obtained from the earlier stage, were subject to statistical analysis, through the use of the *z*-scores method for the classification of each point in the list. Matching was accomplished together with the peaks and distance measures for more accurate results.

Assuming that the list of points of each signature is normally distributed [14]:

$$f(y)=\frac{1}{\sigma\sqrt{2\pi}}e^{-\frac{1}{2}\left(\frac{y-\mu}{\sigma}\right)^2} \tag{8}$$

where $y$ can assume all values from $-\infty$ to $+\infty$ and the parameter $\mu$ and $\sigma$ represent respectively the mean and the standard deviation of the distribution. Since it is a continuous probability density function, the probability that a point $y$ lies between two specified values $a$ and $b$ of a point in the database is given by an integration:

$$\Pr(a\le y\le b)=\int_a^b \frac{1}{\sigma\sqrt{2\pi}}e^{-\frac{1}{2}\left(\frac{y-\mu}{\sigma}\right)^2}dy \tag{9}$$

The above equation can be simplified {14] by carrying out the transformation:

$$z_i=\frac{y_i-\mu}{\sigma} \tag{10}$$

where $z$ is the $z$-scores of observation $y_i$ and is the answer to `how many standard deviations away from the mean the observation is'. The greater the number of standard deviations away from the mean the observation is, the less likely it will have occurred by random chance.

The most suitable value for $z$ was determined based on the results of the experiments. The value of $z$ was between -1.96 and 1.96, that is 5 per cent of the distribution (2.5 percent on each side). If $z$ lies outside this range, then the point is rejected and it does not belong to the list stored in the database.

At least 120 comparisons with the stored lists were made for each of the test signatures. This is repeated for other lists stored in the database. The results may be that all points belong to a list stored in the database, or not. If the results showed that 85 percent belong to the list (also called the confidence interval), that is, at least 100 points are correctly matched with values stored in the database, or it can be said that 15 percent are errors and do not belong or misclassified then we conclude that the test object is the same to the stored object. If the correspondence is less than 85 percent, then the object does not belong to a particular set of lists.

If the results are higher than 85 percent, further tests will be carried out to determine if the object is a complex object or a simple shaped object with straight lines. If the latter is the case, then the number of peaks will be taken into consideration. The number of peaks can roughly determine the type of shape. As an example, a square or rectangle will have four peaks and for a circle, almost all of them are peaks. Complex objects can have any number of peaks. Straight lines will results in the value of $y$ becoming zero. The distance between peaks also provide the internal relationship of a particular shape.

## III. Experimental Results

Experiments were conducted to test the shape outline reading on a set of objects. Raster images were used, with the size of between 300 X 300 pixels and 400 X 400 pixels. Shapes varies from simple to complex objects. Fig. 3 and Fig. 4 shows some examples of the used objects. Each shapes were recreated to 100 images by adding noise before the training process began. This is to allow for a more flexible and robust shape recognition system.

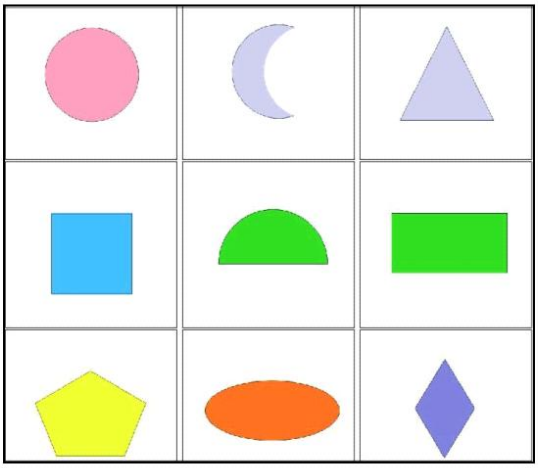

Figure 3
Example of Simple Shapes Used In the Training Process

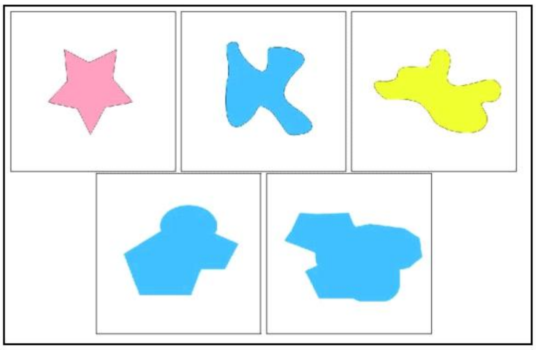

Figure 4
Example of Complex Shapes Used In The Training Process

Fig. 5-8 show the graphical representation of different shapes. It can be seen that the number of peaks showed the sharp corners of each shapes. Straight lines will produce zero readings.

FIGURE 5
GRAPHICAL REPRESENTATION OF THE SHAPE TRIANGLE

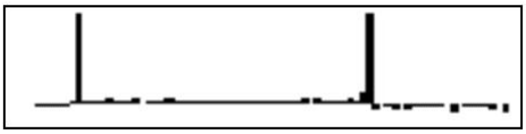

FIGURE 6
GRAPHICAL REPRESENTATION OF THE SHAPE MOON

FIGURE 7
GRAPHICAL REPRESENTATION OF THE SHAPE CIRCLE

FIGURE 8
GRAPHICAL REPRESENTATION OF THE SHAPE STAR

First set of experiments were to test the extraction of the shape outline from an object. Further experiments were conducted to investigate that the system is invariant to rotation, translation, size and reflection and to a certain degree of distortion. Fig. 9 and Fig. 10 show the results of object being rotated and object with different sizes respectively.

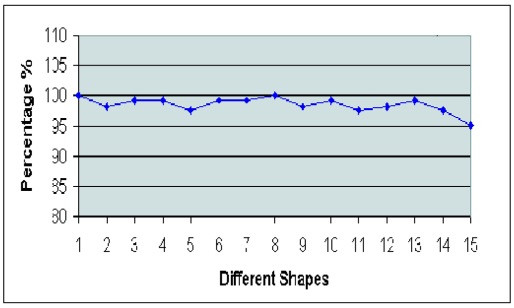


FIGURE 9
TEST RESULTS ON SHAPE ROTATED BY 30 DEGREES

Results obtained from the test for 15 objects rotated at 30 degrees, shows that the method used is invariant to rotation. This can be seen in Fig. 9. The accuracy level for all objects are above 95%.

Translation of objects is similar to rotation, this is because since the difference in angles were used, and there is no problem in identifying a translated object. Results of 15 different shapes that went through that were translated in x, y and xy direction is shown in Table 1.

TABLE
Results On Shapes Translated in *xy* direction

| OBJECT | RESULTS (100%) |
|---|---|
| 1 | 100 |
| 2 | 98.3 |
| 3 | 100 |
| 4 | 100 |
| 5 | 99.2 |
| 6 | 99.2 |
| 7 | 99.2 |
| 8 | 100 |
| 9 | 99.2 |
| 10 | 97.5 |
| 11 | 99.2 |
| 12 | 99.2 |
| 13 | 99.2 |
| 14 | 99.2 |
| 15 | 98.3 |

Experiments conducted for testing the invariance in sizes, shows that, the results for accuracy for 15 objects were above 95%. This can be seen in Fig. 10.

Mirror effect or reflection is another important aspect of object recognition. Readings from the shape outline is stored in a list. Mirror effect can be obtained easily by using the reverse list. Each view of an object were tested through using the following list:

$$view_i = [y_{120}, y_{119}, \ldots, y_1] \tag{11}$$

An object will be not be classified as the same object when it is reflected, however with the use of the reverse list, an object is easily classified. Results obtained for an object, that is the non-reversed list will create a totally new object.

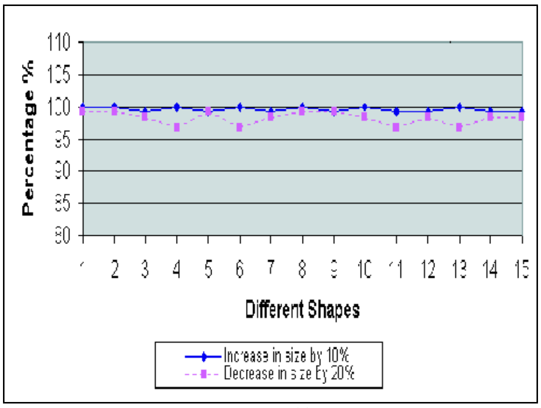


FIGURE 10
TEST RESULTS ON SHAPES INCREASED BY 10 PERCENT AND DECREASED BY 20 PERCENT

Experiments were carried out on a small degree of distortion and translation. Objects were translated into x, y and xy direction for testing the accuracy level of translated objects. The accuracy level are all above 95% for the 15 objects. Objects were distorted by applying the distortion facility provided by Corel Draw, using the displacement map. This was done by altering 20% horizontally and vertically on the displacement map. Results for these tests were above the accuracy level. Further tests were conducted by further distorting all of the objects. Results shows that, accuracy level were achieved for a maximum of 30% of distortion. When the distortion on the displacement map were increased above 30%, the method failed to classify these objects.

Further experiments were conducted on new objects. This is to further test the systems on the shape categorization. Different shapes will be classified differently. Example of five test shapes that were used can be seen in Fig. 11. Results of these five test shapes is shown in Table 2. If the new object does not belong to any existing group of shapes, new group will be automatically created. This new shape will be stored in a new category. Results showed that recognition that based on only matching points were not accurate. Peaks and the distance from peaks are essential to identify whether a shape can be decided to be categorized as a same group or not.

TABLE 2
RESULTS OF THE NEW SHAPES TOWARDS THE TRAINED SHAPES

| Shapes | Shape 1 | Shape 2 | Shape 3 | Shape 4 | Shape 5 |
|---|---|---|---|---|---|
| % match towards the stored objects | 94% (rectangle) | 87.5 (oval) | 92.5 (pentagon) | 96.6 (oval) | 97.5 (rectangle) |
| % peaks match | 30% | 10% | 20% | 20% | 50% |

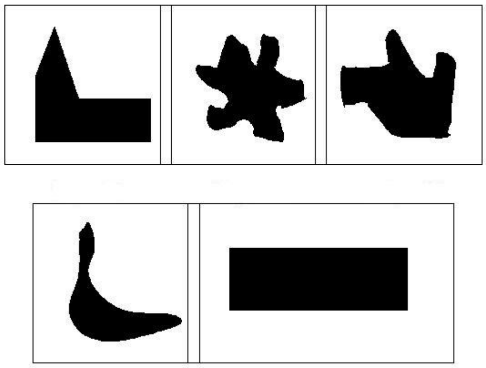

FIGURE 11
EXAMPLE OF FIVE TESTS OBJECTS

Simple objects normally have more straight lines compared to complex objects. When there are straight lines or curves, then the system can easily identify two different objects. Straight lines will have zero angle difference and therefore it is much easier to identify objects with the same shape.

Matching and recognition process will be tougher when there are too many straight lines or curves in an object. The system will try to classify them as the same object even when it is dealing with two different objects such as a rectangle and square. Using the data from the earlier stage will classify them as the same object. However, grouping the data obtained with the number of peaks and the distance from one peak to another can solve the recognition problem. Distance between peaks will show how closely one sharp corner from one to another. As an example the distance of a peak from a rectangle will be different from a square. Therefore these two objects will be classified differently however they will be put in the same category, since both will have four peaks. The rectangle in Fig. 2 compared to the rectangle (shape 5 in Fig. 11) will not be classified as the same object since the number of peaks match is only 50%. However, they will be put under the same category, because of the number of peaks that they have.

Every time a new object was introduced, the system will automatically calculate the shape outline, the number of peaks and the distance measure. In most cases, the system managed to identify the object, if the object is closely matched with existing objects.

## V. CONCLUSIONS

Methods used in this system has shown the capability of an automatic recognition and categorization of shapes. Images were put through the pre-processing process of image mining and data produced were grouped together through their shapes. Results can be visualized in graphical format and different shapes can be seen clearly from this graphical representation. Results which were stored in list went through a statistical method, using $z$ - scores, *peak measure* and were used to classify and recognize simple and complex objects. Experiments showed that the system is invariant to rotation, translation, size, mirror effect and to a certain degree of distortion.

Peaks measure are essential for recognizing objects of different shapes, as the shape outline itself is insufficient. A larger number of peaks or higher readings occur where there is a significant change in shape, such as a sharp corner or a curve. Different shapes will be grouped accordingly. The system is capable of grouping new objects, that are objects that do not belong to any existing category by putting it into a new category. This is achieved without any human interference.

The research were carried out to test the capability of producing an automatic shape recognition system by mining relevant image features. From the experiments and the results it showed that the method is capable of producing a generic automatic shape recognition system that is invariant to rotation, translation, size and to a certain degree of distortion.

The method can be extended to three-dimensional objects, which is currently under investigation. Color, depth and texture can be grouped together to form a set of new features. Selecting and grouping these data can be another part of the data mining process. The current method will be tested with a much larger scale of images. The current system is limited in classifying and recognizing objects with a greater distortion. This will be look into in the next coming project.

In comparison with other methods such as neural networks, the next stage of the research could carry out a real comparison with the same data for both methods. Another possibility is the combination of both methods, and this would be a very useful area of investigation.

## REFERENCES

[1] J. Zhang, W. Hsu, and M. L. Lee, An Information-driven Framework for Image Mining, in *Proceedings of the 12th International Conference on Database and Expert Systems Applications (DEXA)*, Munich, German, 2001.

[2] I. Bierderman, and G. Ju, Surface vs. Edge-based Determinants of Visual Recognition. *Cognitive Psychology,* 20, 38-64, 1988.

[3] W. G. Hayward, Effects of Outline Shape in Object Recognition. *Journal of Experimental psychology: Human Perception and Performance*, 24(2), 427-440, 1988.

[4] I. Taylor and M. M. Taylor, *The Psychology of Reading*. London and New York Academic Press, 1983.

[5] I. Rock, F. Halper, T. Clayton, *The* Perception and Recognition of Complex Figures. *Cognitive Psychology*, 3, 655-673, 1972.

[6] R. N. Haber, R. Haber, Visual components of the Reading Process. *Visible Language*, 15, 147-182, 1981.

[7] R. G. Crowder, *The Psychology of Reading*. Oxford University Press, 1982.

[8] A. Jain, A. Vailaya, Image Retrieval using Color and Shape, *Pattern Recognition,* 29(8), 1233-1244, 1996.

[9] W. Ma, Y. Deng, and B. S, Manjunath, Tools for Texture/Color Based Search of Images, *SPIE International Conference*, Human Vision and Electronic Imaging, 497-507, 1997.

[10] K. Schulten, *The Development of the Primary Visual Cortex*. Theoretical Biophysics Group, Beckman Institute, University of Ilionis, USA, Available : http://www.ks.uiuc.edu/Research/Neural/development. html, (16th September 2002).

[11] Z. Pylyshyn, *Is Vision Continuous with Cognition? - The Case for Cognitive Impenetrability of Visual Perception*. Technical Report TR-38, 1998, Rutgers Center for Cognitive Science, Rutgers University, New Brunswick, NJ, Available: http://ruccs.rutgers.edu/ publicationsreports.html

[12] M. Kass, A. Witkin, and D. Terzopoulos, Snakes: Active Models. *International Journal of Computer Vision*, 321-331, 1988.

[13] R. P. Grzeszczuk and D. N. Levin, Brownian Strings: Segmenting Images with Stochastically Deformable s. IEEE *Transactions on Pattern Analysis and Machine Intelligence*, 19, 1100-1114, 1997.

[14] H. Mulholand and C. R. Jones, *Fundamental of Statistics*. London Butterworths, London, 1968.

**Rosalina Abdul Salam** was born in Penang, Malaysia on 13th of January 1968. She received her Bachelors degree in Computer Science in 1992 from Leeds Metropolitan University, United Kingdom. She received a scholarship to pursue her Masters degree and PhD. She received her Masters degree in Software Engineering from Sheffield University, United Kingdom in 1997. She completed her PhD in 2001 from Hull University in the area of artificial intelligence and image processing.

She was a system analyst in Intel Penang, from 1992 to 1995. She is currently a lecturer of the School of Computer Sciences, Universiti Sains Malaysia and a member of Artificial Intelligence Research Group. She has published more than 25 papers in journals and conferences. Her current research area is in the area of artificial intelligence, image processing and bioinformatics applications.

**Marcos Aurelio Rodrigues** received his BEng in Mechanical Engineering from the Federal University of Santa Catarina (Brazil) in 1983. He was awarded an MSc in Computer Science in 1989 and a PhD in Computer Science in 1991, both from the University of Wales, Aberystwyth.

He has been appointed a Reader in Intelligent Systems within the School of Computing and Management Sciences at Sheffield Hallam University in January 2000 and awarded a Personal Chair in Computer Science in February 2003. He has published over 100 technical papers in international journals and conferences on the subjects of robotics, computer vision, pattern recognition, systems modelling and artificial intelligence. His main current research interests include 2D and 3D machine vision, machine learning, and pattern recognition