International Journal of Multimedia and Ubiquitous Engineering
Vol. 8, No. 1, January, 2013

A Novel Optimized Image Feature Selection Algorithm Using Pairwise Classifiers

Mahdi Bazarganigilani, Ben Arko, Ali Syed, Camille Price, Sandid Burki, Belinda Fridey and Michael Baron

Faculty of Business and Management, Charles Sturt University, Australia
mahdi62b@yahoo.com, benarko@yahoo.com.au, asyed@studygroup.com, cprice@studygroup.com, sburki@studygroup.com, bfridey@studygroup.com, mbaron@studygroup.com

Abstract

This paper introduces an optimized method to improve the accuracy of content based image retrieval systems (CBIR). CBIR systems classify the images according to low and higher features. This research improves both feature selection and classifier partition of a CBIR system. This paper normalizes and descretizes the feature vectors. A pair-wise classifier is employed to improve the accuracy of classification. Results show high accuracy of proposed classifier in an image database.

Keywords: *Efficient Feature Selection; Content Based Image retrieval System; Pairwise Classifier*

1. Introduction

By increasing amount of multimedia data on information systems, there is a huge demand for more accurate and effective image retrieval systems. The similar semantically images likely have similar vision features. This is the main idea behind CBIR systems [1]. Such systems use positive feedback of the users to improve accuracy of image classification [2].

By improving either classification algorithm or retrieval approaches of the system, the overall accuracy is improved. There is plenty of research on classification algorithms [1, 3] and retrieval methodologies [4,6]. This research focuses on choosing efficient features of the images [7]. Consequently, the accuracy of the classifier is improved by employing pair wise classifiers [8].

One of technique to increase the accuracy of the classifier is to use ensemble of classifiers [9]. In this method we use multiple pair-wise classifiers instead of one multiple classifier. Pair-wise classifier needs to be employed on most inaccurate classification groups. In this manner, the overall accuracy of the CBIR system is improved. Improving the accuracy of CBIR systems is very vital due to the high number of feature vectors of images. This paper introduces multiple effective steps to obtain the most accurate classifiers for an CBIR systems.

2. Image Feature Vector Selection

To classify the images, the low level features of images are needed to be extracted. Low level feature include color and shape properties of the images [10]. In case of low-level image features, a feature vector which describes various visual cues, such as color, shape, or texture, is computed for each image in the database. These include features based on color histograms [11, 12,13, 14, 15, 16, 17,18, 19], color moments [20, 21, 22, 18], statistical shape and texture features such as edge direction histograms [23], Tamura texture features [24], and wavelet-based texture features [22, 25]. Given a query image, its feature vector is calculated and those images which are most similar to this query image based on an appropriate distance measure in the feature space are retrieved [26].

Color histograms are generally invariant to translation and rotation of the images and normalizing the histograms leads to scale invariance [27]. However, color histograms do not incorporate spatial adjacency of pixels in the image and may lead to inaccuracies in the retrieval. Therefore, retrieval of images based on color histograms are prone to yielding a large number of false hits, i.e., images with completely different content which just happen to have a similar color composition as the query image [28].

To overcome the quantization effects as in color histogram, [20] proposed to use color moments approach. Color moments of an image are a simple yet effective feature for color-based image retrieval [20, 21, 22]. Most of the color distribution information can be obtained by the low-order moments, using only the first three moments: mean, variance and skewness, these moments have been proven to be efficient and effective in representing color distribution of images [20]. The first three moments are defined as:

$$u_i = \frac{1}{N}\sum_{j=1}^{N} p_{ij}$$

$$\sigma_i = \sqrt{\frac{1}{N}\sum_{j=1}^{N}\left(p_{ij} - u_i\right)^2}$$

$$s_i = \left(\frac{1}{N}\sum_{j=1}^{N}\left(p_{ij} - u_i\right)^3\right)^{\frac{1}{3}}$$

Where p_{ij} is the value of the ith color channel of the jth image pixel.

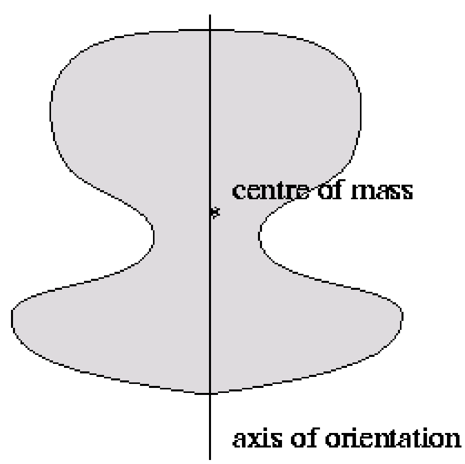

Figure 1. The 0th Moment is the Area of the Object; the 1st Moment Gives the Centre of Mass; and the 2nd Moments Give the Axes of Orientation

Shape is another feature that can be used in content-based image retrieval. In the absence of color information or in the presence of images with similar colors, it becomes imperative to use additional image attributes for an efficient retrieval [26].

Various schemes have been proposed in the literature for shape-based retrieval. Shape representation techniques are mainly included in two classses: boundary-based and region-based. The boundary-based uses only the outer boundary of the shape while the latter uses the entire shape region [29]. The most successful representatives for these two categories are Fourier descriptors and moment invariants respectively. The main idea of Fourier descriptors is to use the Fourier transformed boundary as the shape feature [30].

The main idea of moment invariants is to use region-based moments, which are invariant to transformations, as the shape feature. In [31], the authors represent the shape of an image in terms of seven invariant moments based on the 2nd and 3rd order moments. The idea of using moments in shape recognition gained prominence because of its effectiveness. The central moment of a digitally sampled image in order of $(p+q)$ that has gray function $f(x,y),(x,y=U,...M-1)$ is given by,

$$u_{pq}=\sum_x\sum_y\left(x-\bar{x}\right)^p\left(y-\bar{y}\right)^q f(x,y)$$

Seven moment invariant based on 2d and 3d order moments gained as follow.

$$\begin{aligned}
&M_1=(u_{20},u_{02}),\\
&M_2=(u_{20}-u_{02})^2+4u_{11}{}^2,\\
&M_3=(u_{30}-3u_{12})^2+(3u_{21}-u_{03})^2,\\
&M_4=(u_{30}+u_{12})^2+(u_{21}+u_{03})^2,\\
&M_5=(u_{30}+u_{12})(u_{30}-3u_{12})\left[(u_{30}+u_{12})^2-3(u_{21}-u_{03})^2\right]\\
&+(3u_{21}-u_{03})(u_{21}+u_{03})\left[3(u_{21}+u_{03})^2-(u_{21}-u_{03})^2\right],
\end{aligned}$$

$$\begin{aligned}
&M_6=(u_{20}-u_{02})\left[(u_{30}+u_{12})^2-(u_{21}+u_{03})^2\right]+\\
&4u_{11}(u_{30}+u_{12})(u_{21}+u_{03}),\\
&M_7=(3u_{21}-u_{03})(u_{30}+u_{12})\left[(u_{30}+u_{12})^2-3(u_{21}+u_{03})^2\right]\\
&-(u_{30}-3u_{12})(u_{21}+u_{03})\left[(u_{30}+u_{12})^2-3(u_{21}-u_{03})^2\right]
\end{aligned}$$

M_1 through M_6 are invariant under rotation and reflection. Scale invariance is achieved through following formula.

$$M_1'=M_1/n\ \ M_2'=M_2/r^4\ \ M_3'=M_3/r^6$$
$$M_4'=M_4/r^6\ \ M_5'=M_5/r^{12}\ \ M_6'=M_6/r^8\ \ M_7'=M_7/r^{12}$$

Where r is the radius of gyration of the object.

$$r=(u_{20}+u_{02})^{1/2}$$

Low-level texture features can provide both global and local information, but are hard to define. Moreover, they are limited in their ability in describing the semantic content in the image. Another limitation of texture features is the high computational complexity of matching based on these features [26].

3. Normalization of Feature Vector

Normalization means putting the real value between two predefined values. Mostly, the real values are mapped between 0 and 1 [11]. Following formula is used to normalize the feature i from vector k.

$$x'_{i,k} = \frac{x_{i,k}}{\max(x_{i,1},\ldots x_{i,n})}$$

Normalization is important in learning classifiers. Therefore, sigmoid function is employed in neural network classifiers predominately. Such networks have very small first derivative for the numbers outside the range -5 and 5. Figure 2, shows this range.

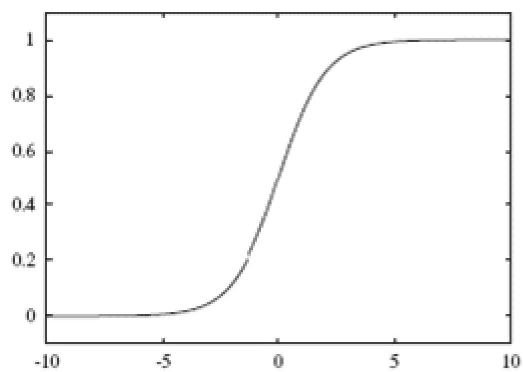

Figure 2. Sigmoid Function

If the network is feeded with the values outside the normalized values, the network converges to 0. Thus, the learning is impaired due to non normalized inputs .By normalizing the inputs, such problem is impeded [7].

4. Descretization of Inputs

The main advantage of descretization is to decrease the number of inputs without losing the accuracy of the classifier [7]. Value sets of a feature are considered as $f_k = \{x_1, x_2, \ldots, x_n\}$. These values are sorted to $x_1 \le x_2 \le \ldots \le x_n$. The set is divided from the point x_p .The suitable point for x_p to divide feature set of f_k in feature vectors U is the position minimizes the following formula [11].

$$E(f, x_p) = \frac{S_1}{U} Entropy(S_1) + \frac{S_2}{U} Entropy(S_2)$$

$$Entropy(S) = -\sum_{c' \subseteq C} \frac{Sc'}{S} \log_2 \frac{Sc'}{S}$$

In which,

U : Is the set of feature vectors.
S_1 : Is the feature vectors which $f_k \leq x_p$.
S_2 : Is the feature vectors which $f_k \succ x_p$.
C : Is the set of semantic concepts in image inputs.

The division of x_p is continued to obtain the best point for each feature set.

5. Feature Selection

There are many features which is not suitable for semantic classification. For instance, brightness is not a good feature. Since every picture from different classes may have different brightness. The main idea behind feature selection is to identify the most discrepancies among various classes [32].

Information Gain is employed to select best features for classifiers. The most accurate features for differentiating the instances is obtained by Information Gain [33]. Information Gain of feature A in feature sets of S is defined as follow:

$$Gain(S, A) = Entropy(S) - \sum_{v \in Value(A)} \frac{|S_v|}{|S|} Entropy(S_v)$$

In which,

$Value(A)$: is the value set of feature A.

S_v : Subset of S in which feature A has value v.

The best features are selected and the most ineffective features are discarded for improving the accuracy of the system.

6. Pair-Wise Classifier

This paper employs pair-wise classifiers instead of one multiple classifier. In such manner the accuracy of the system is effectively improved. The learning in multiple classifiers is to distinguish one category from other categories. While in pair-wise one, it is based on two categories. A secondary network is used to combine the results of multiple pair-wise classifiers [8].

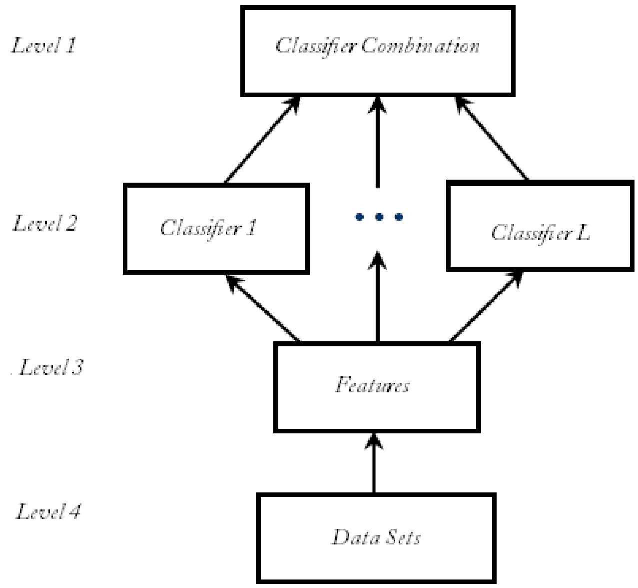

Figure 3. Different Levels in Making Ensemble of Classifers

Optimized feature vectors are inputted to par wise classifiers. In this algorithm, the combination of results is obtained by a multilayer perceptron network.

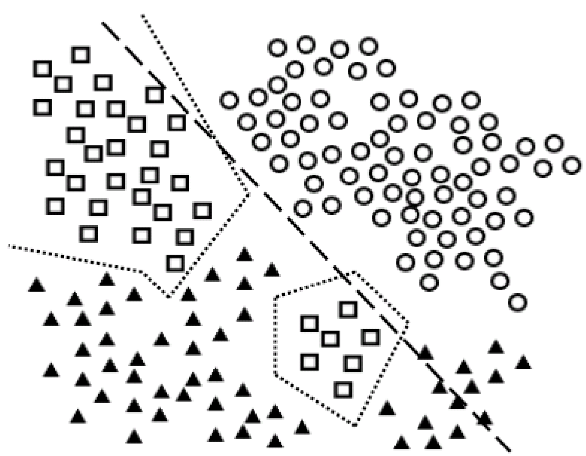

Figure 4. Simplicity of Classifier in Pairwise Boundries (the boundry of two pices of square with circles has shown with dashed line and with others with dotted line) [8]

In Figure 4, we have shown the accuracy of using pair-wise classifier. In the figure distinguishing the rectangles from other shapes such as circles is more accurate. On the other hand, discriminating the rectangles among all other shapes leads to less accuracy. In this approach, first, a few numbers of classifiers with maximum error rate of classification are determined. Second, the most inaccurate categories are classified using pair-wise classifiers. Consequently, other remaining categories are classified using one multiple classifier which results in more speed of convergence. Finally, the outputs of pair-wise and multiple classifiers are combined by employing an Artificial Neural Network [8].

7. Determining the Most Erroneous Pair-Wise Categories

The most erroneous pair-wise categories which have most intersection with each other are defined to establish multiple pair-wise classifiers. Such couple categories are feeded to pair-wise classifiers. In this manner, the effectiveness of the classifier is improved.

A multiple neural network classifier is employed to determine error rate of couple categories. First, feature vectors are optimized using Normalization and Information Gains. Second, such vectors are inputted to the multiple classifier. Consequently, a confliction matrix is created to define the categories which are essential to be classified. Each element of such matrix, determines the number of incorrect instances from the classes corresponding to the row and column numbers [8]. The number of error instances indicates the error rate of couple classifications.

8. Combination of Pair-Wise Classifiers

The final step for classifying the feature vectors is to combine the result of primary network and pair-wise networks. A multilayer perceptron network is employed to combine the results. The output of multi classifier network and pair-wise networks are feeded to combination network. Figure 5, the entire structure of the classifier [8].

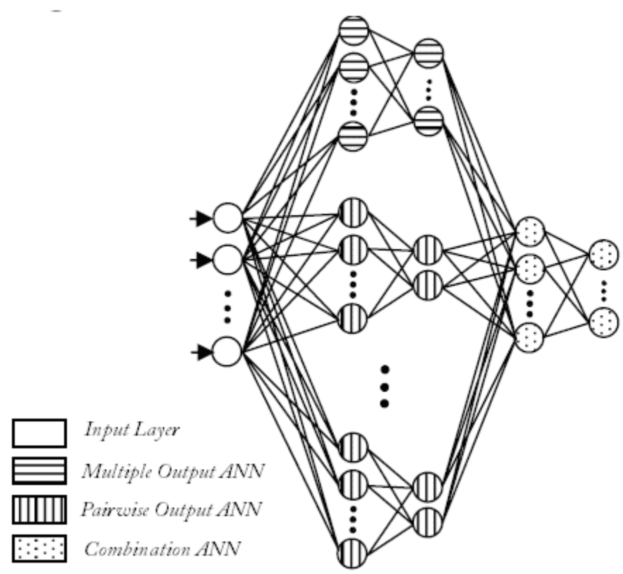

Figure 5. Structure of Artificial Neural Network Classifiers

In this network the input feature vectors are feeded to the system. In next layer, there are a number of pair-wise classifiers along the multiple-neural network. The pair wise classifiers were employed due to the high classification error rate of couple categories. In combination layer, the outputs of pair-wise classifiers and the multiple-neural network is merged. The output layer consists of the number of categories.

5. Experiments and Results

To evaluate the algorithm, Several experiments were conducted. In the experiment I, 300 pictures were chosen from LabelMe database [34]. Such groups are categorized in 10 concept groups including, tree, seashore, sky, sun, hand, cloud, library, window, road and wall. 13

features of pictures are obtained including color and shape properties. Information Gain was applied on the feature vectors and 3 of feature with the lowest values were eliminated. The remaining features were feeded to one multiple Neural Network. Consequently, 4 couple categories with the highest error rates were determined. Therefore, one 4 pair wise networks with 10 feature inputs were employed. The combination network consisted of 18 inputs and ten outputs to determine the concept class of the images. Table 1, shows the results of applying each step.

Table 1. Algorithm Performance

Algorithm	Number of Feature	Precision
Raw Feature	13	74%
Raw Feature	7	68%
Normalized Features	13	77%
Normalized Features	7	73%
Normalized-Descretized Features	13	81%
Normalized Descretized Features	7	75%
Pair-Wise NNs	13	88%
Pair-Wise NNs	7	80%

In the experiment II, the speed of convergence without considering pair-wise networks were measured. Figure 6, shows each step were resulted in faster convergence rate and lower number of Epochs.

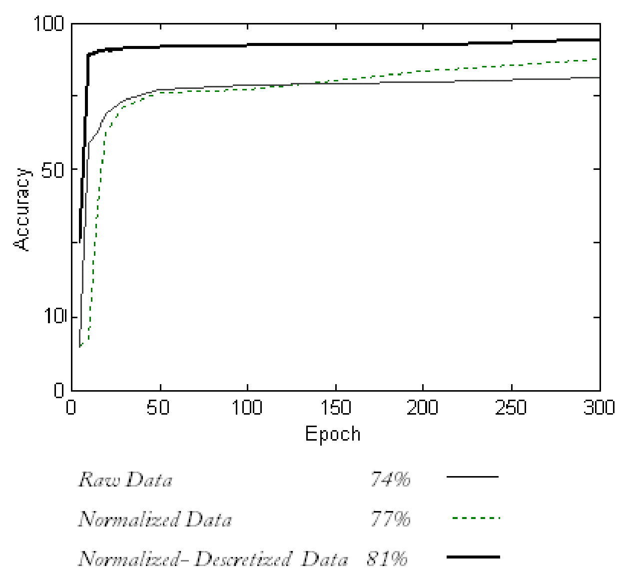

Figure 6. Convergence Speed in Normalized and Descretized Inputs

As results presented, the proposed networks resulted better accuracy than general multi layer neural network. In this approach , normalization along with descretization leads to more effective input vectors and improved overall accuracy of the structure. This efficient input results in better performance and speed of convergence. Furthermore, it is concluded, using pair-wise classifiers improved the accuracy of categorization.

6. Conclusions

This paper introduced an efficient method to optimize feature vectors obtained from images. Feature vectors were normalized to avoid impairment of learning algorithms. Consequently, the inputs were descretized to speed up the convergence and accuracy of the system. Information Gain formula is applied to select the best features. Finally, a new multiple-neural network was created to improve the accuracy of classification. Multiple pair-wise networks were employed to classify the worst features in pairs. Results showed the improved performance and convergence rate of the classifier in each step.

References

[1] J. Z. Wang, J. Li and G. Wiederhold, “SIMPLIcity: Semantics-Sensitive Integrated Matching for Picture Libraries”, IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 23, no. 9, **(2001)**, pp. 947–963.

[2] Y. Liu, D. Zhang, G. Lu and W.-Y. Ma, "A survey of content-based image retrieval with high-level semantics", Journal of Pattern Recognition, vol. 40, **(2007)**, pp. 262– 282.

[3] W. Wang and A. Zhang, "Extracting semantic concepts from images: a decisive feature pattern mining approach", Multimedia Systems, vol. 11, no. 4, **(2006)**, pp. 352–366.

[4] Y. Wang, M. Ding, C. Zhou and T. Zhang, "A Hybrid Method for Relevance Feedback in Image Retrieval Using Rough Sets and Neural Networks", international journal of computational cognition, vol. 3, no. 1, **(2005)**, pp. 78-87.

[5] J. -H. Han, D. S. Huang, T. -M. Lok and M. R. Lyu, "A novel image retrieval system based on BP neural network", IJCNN 2005, vol. 4, no. 31, **(2005)**, pp. 2561-2564.

[6] H. y. Ko and S. I. Yoo, "Intelligent Image Retrieval Using Neural Network", IEICE TRANS. INF. & SYST, vol. E48-D, no. 12, **(2001)**, pp. 1810-1819.

[7] A. Tazaree and K. Faez, "Features Selection to Enforce Contents Based Image Retrieval Systems Accuracy", Proceedings of CSICC2008, Kish, Iran, **(2008)** March 9-11.

[8] M. Analoyi and M. Meshki, “A sclable Method For Improving Retrival Systems’s Accuracy Using Pairwise Classifers”, First Joint Congress on Fuzzy and Intelligent Systems Ferdowsi University of Mashhad, Iran, (In persian), **(2007)**, http://confbank.um.ac.ir/modules/conf_display/isfs2007/article/i357.pdf .

[9] L. I. Kuncheva, “Combining Pattern Classifiers: Methods and Algorithms”, John Wiley & Sons, New Jersey, **(2004)**.

[10] M. K. Hu, "Visual Pattern Recognition by Moment Invariants", IRE Trans. Info. Theory, vol. IT-8, **(1962)**, pp. 179–187.

[11] M. Ioka, “A method of defining the similarity of images on the basis of color information”, Technical Report RT-0030, IBM Research, Tokyo Research Laboratory, **(1989)**.

[12] M. J. Swain and D. H. Ballard, “Similarity of color images”, Journal of Computer Vision, vol. 7, no. 1, **(1991)**, pp. 11-32.

[13] B. V. Funt and G. D. Finlayson, “Color constant color indexing”, Technical report 91-09, School of Computing Science, Simon Fraser University, Vancouver, B.C., Canada, **(1991)**.

[14] D. Lee, R. Barber, W. Niblack, M. Flickner, J. Hafner and D. Petkovic, “Indexing for complex queries on a query-by-content image database”, Proceedings of 12th International Conference on Pattern Recognition (ICPR'94), Jerusalem, Israel, **(1994)**, pp. 142-146.

[15] J. R. Smith and S. -F. Chang, “Local color and texture extraction and spatial query”, Proceedings of IEEE International Conference on Image Processing, Lausanne, Switzerland, vol. 3, **(1996)**, pp. 1011-1014.

[16] G. Pass and R. Zabih, “Comparing images using joint histograms”, ACM Journal of Multimedia Systems, vol. 7, no. 3, **(1999)**, pp. 234-240.

[17] A. Vailaya, M. Figueiredo, A. Jain and H. J. Zhang, “A Bayesian framework for semantic classification of outdoor vacation images”, Proceedings of SPIE: Storage and Retrieval for Image and Video Databases VII, San Jose, CA, vol. 3656, **(1999)**, pp. 415-426.

[18] A. Vailaya, M. Figueiredo, A. Jain and H. J. Zhang, “Content-based hierarchical classification of vacation

images", Proceedings of IEEE Multimedia Systems'99, Florence, Italy, vol. 1, **(1999)**, pp. 518-523.
[19] R. O. Stehling, M. A. Nascimento and A. X. Falcao, "On "Shapes" of colors for content-based image retrieval", Proceedings of ACM International Conference on Multimedia, Los Angeles, CA, USA, **(2000)**, pp. 171-174.
[20] M. Stricker and M. Orengo, "Similarity of color images", Proceedings of SPIE Conference on Storage and Retrieval for Image and Video Databases III, San Jose, CA, vol. 2420, **(1995)**, pp. 381-392.
[21] S. Ghosal and R. Mehrota, "A moment-based unified approach to image feature detection", IEEE Transactions on Image Processing, vol. 6, no. 6, **(1997)**, pp. 781-793.
[22] M. K. Mandal, "Wavelet based coding and indexing of images and video", Ph.D. thesis, University of Ottawa, Canada, **(1998)**.
[23] A. Vailaya, A. Jain and H. J. Zhang, "On image classification: city images vs. landscapes", Journal of Pattern Recognition, vol. 31, no. 12, **(1998)**, pp. 1921-1935.
[24] Q. Li, J. Yang and Y. Zhuang, "Web-based multimedia retrieval: balancing out between common knowledge and personalized views", Proceedings of 2nd International Conference on Web Information System Engineering, Kyoto, Japan, **(2001)**, pp. 92-101.
[25] J. Li, J. Z. Wang and G. Wiederhold, "Classification of textured and nontextured images using region segmentation", Proceedings of Seventh International Conference on Image Processing, Vancouver, BC, Canada, **(2000a)**.
[26] J. Lakshma, "Enhancing retrieval of images on the web through effective use of associated text and semantics from low-level image features", Ph.D. thesis, School of Computing and Mathematics; College of Health and Science; University of Western Sydney, **(2006)**.
[27] A. Vailaya, Y. Zhong and A. K. Jain, "A Hierarchical system for efficient image retrieval", Proceedings of 13th International Conference on Pattern Recognition, Vienna, Austria, **(1996)**, pp. C356-360.
[28] A. Vailaya, "Semantic classification in image databases", Ph.D. thesis, Michigan State University, USA, **(2000)**.
[29] Y. Rui, A. C. She and T. S. Huang, "Modified Fourier descriptors for shape representation – a practical approach", Proceedings of First International Workshop on Image Databases and Multimedia Search, Amsterdam, The Netherlands, **(1996)**.
[30] M. Bouet, A. Khenchaf and H. Briand, "Shape representation for image retrieval", Journal of ACM Multimedia '99, **(1999)**, pp. 1-4.
[31] A. K. Jain and A. Vailaya, "Shape-based retrieval: a case study with trademark image databases", Pattern Recognition, vol. 31, no. 9, **(1998)**, pp. 1369-1390.
[32] I. H. Witten and E. Frank, "Data Mining: Practical Machine Learning Tools and Techniques with JAVA Implementations", Morgan Kaufman, San Francisco, **(2005)**.
[33] T. Mitchell, "Machine Learning", McGraw-Hill, **(1997)**.
[34] B. C. Russell, A. Torralba, K. P. Murphy and W. T.Freeman, "LabelMe: a database and web-based tool for image annotation", MIT AI Lab Memo AIM-2005-025, **(2005)**, www.labelme.csail.mit.edu.