A Hierarchical Classification Structure based on Trainable Bayesian Classifier for Logo Detection and Recognition

Hossein Pourghassem
Young Research Club-Islamic Azad University- Najafabad Branch, Iran.
Email: h_pourghasem@iaun.ac.ir

Received: March 2010 Revised: June 2010 Accepted: July 2010

ABSTRACT:
The ever-increasing number of logo (trademark) in official automation systems for information management, archiving and retrieval applications has created greater demand for an automatic detection and recognition logo. In this paper, a hierarchical classification structure based on Bayesian classifier is proposed to logo detection and recognition. In this hierarchical structure, using two measures false accept rate (FAR) and false reject rate (FRR), a novel and straightforward training scheme is presented to extract optimum parameters of the trained Bayesian classifier. In each level of the hierarchical structure, a separable feature set of shape and texture features is used to train and test classifier based on complexity of the logo pattern. The logo candidate regions are extracted from document images by a wavelet-based segmentation algorithm, and then recognized in the proposed structure. The proposed structure is evaluated on a vast database consisting of the document and non-document images with Persian and international logos. The obtained results show efficiency of the proposed structure in the real and operational conditions.

KEYWORDS: Logo detection and recognition, trainable Bayesian classifier, training scheme, shape and texture features and wavelet-based segmentation algorithm.

1. INTRODUCTION

Due to the increasing content of multimedia databases (such as text, image, audio, and video), the demands for archiving and retrieving in search and data mining applications is a vital requirement. In this domain, document image analysis and understanding have received a great deal of interests in the last few years for many diverse applications such as, digital library, Internet publishing and searching, on-line shopping and official automation systems. Along with logo detection and recognition is an important requirement in the document image analysis and shape-matching domain as it enables us to identify the source of documents based on the organization where a document originates. For example, in official automation systems application, content of scanned official letters (as a document image) should be recognized and classified by their logos. Therefore, logos can act as a valuable means in identifying sources of documents [1].
The major previously research related to logo in document images have focused in logo recognition [2]-[8] and rare investigation have consisted of both logo detection and recognition [9], [10]. In [11], a modified line segment Hausdorff distance has proposed that incorporates structural and spatial information to compute dissimilarity between two sets of line segments rather than two sets of points. In this paper, logo is first generated to line segments and represented with feature vector. In [9], a logo detection system is presented based on segmentation the document image into smaller images using a top-down X-Y cut algorithm [12]. In this paper, sixteen features of the connected components in each segment are extracted and used by a rule-based classification scheme. In [13], an approach to logo detection and extraction in the document images using a multi-scale boosting strategy has presented. An initial two-class Fisher classifier at a coarse image scale on each connected component is used. Each detected logo candidate region is then classified at finer image scales by a cascade of simple classifiers [13]. In [10] a simple logo detection method has presented based on the assumption that the spatial density of foreground pixels in a logo region is greater than that in non-logo regions. A document image is first binarized into foreground and background pixels. Then, the spatial density within each fixed size window is computed and the region with the highest density is hypothesized as a logo region. In [14], a method for such a system based on the image content, using a

shape feature has presented. Zernike moments of an image are used as a feature set. In this method, to recognize the detected logo, a similarly measure based on shape feature (Zernike moments) has defined. In [15], an automatic content-based logo retrieval method has proposed. The proposed method in [15] automatically selects appropriate features (such as area, deviation, symmetry, centralization, complexity and 2-level contour representation strings) based on feature selection principles to discriminate logo. In this method, the user can submit a query through logo examples to get a list of database trademarks ordered by similarity ranks. In [4], a color image retrieval system based on multiple classifiers has presented. In this approach, a region-growing technique for segmentation of the input image into logo candidate regions and three complementary region-based classifiers (color, shape and relational classifiers) have applied to logo recognition. In each classifier, a virtue probability representing the probability that an image is similar to the query image is defined. A set of virtue probabilities is calculated to define similarity measure in each classifier. In [16], a shape-based similarity retrieval system has developed based on database classification, which exploits the contour and interior region of a shape efficiently. In this paper, angular radial transform (ART) region feature is employed to compare the query with the candidate sets according to the priority order.

In this paper, a hierarchical classification structure based on Bayesian classifier for logo detection and recognition is proposed. In this structure, a novel training approach is used to extract the optimum parameters of Bayesian classifier by false accept rate and false reject rate measures. In each level of this structure, an isolated set of shape features consisting of Seiden features [9], Yin features [15], ART features [17], and texture features consisting of energy, homogeneity, correlation extracted from co-occurrence matrix [18] are extracted based on logo complexity. The used features in different decision levels; from down levels to up levels are more intricate. In other word, simple and intricate features are used to simple and intricate logos, respectively.

The proposed structure in this paper has two differences with the previously works. Firstly, as we previously imply that the major previously research related to logo in the document images have focused in logo recognition and rare investigation have consisted of both logo detection and recognition. While, in this paper, these two stages have merged in a primitive segmentation stage and a logo candidate region classification stage. Secondly, the used images in the previously research are scanned official letters. While, the used database for evaluation of the proposed algorithm consists of a large number of images such as personal, journal and newspaper images and a few document images. Certainly, evaluation of the proposed algorithm with this database not only obtains real conditions of tests but also asks more accuracy in algorithm design.

This paper is organized as follows: Section 2 describes the used segmentation algorithm. In section 3, hierarchical structure of logo detection and recognition is presented. In this section, details of this structure consisting of the extracted features in each level, training approach of classifiers and logo classification and recognition procedure are described. Experimental results are shown in Section 4. Section 5 provides a conclusion to the work.

2. SEGMENTATION

The first stage in logo detection and recognition procedure is the detection and extraction of candidate logo region from document images. The previous used approaches are divided to two general categories. Approaches of the first category use morphological operators for logo detection purpose [19].

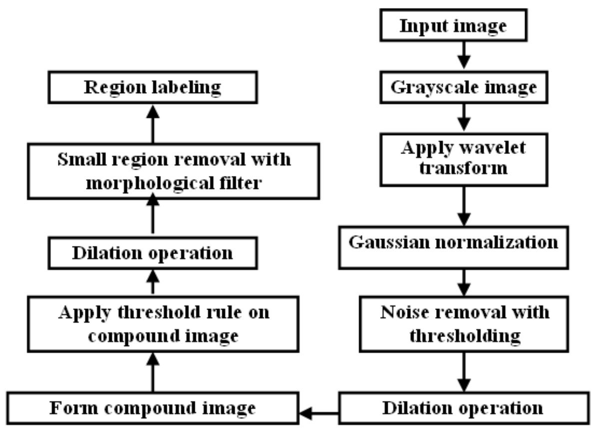

Fig. 1. Block diagram of wavelet-based segmentation algorithm [1].

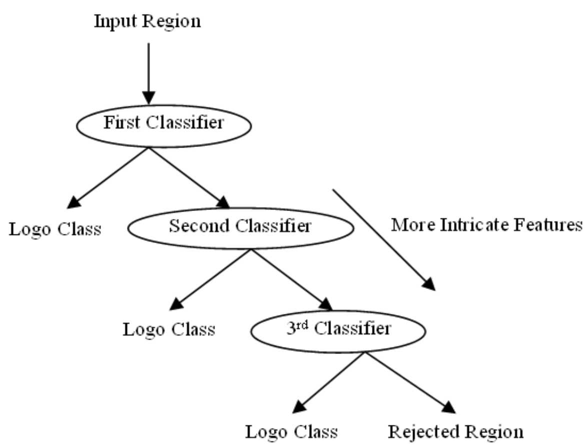

Fig. 2. Block diagram of hierarchical classification structure of the logo candidate region.

Approaches of the second category define and use measures such as spatial density of the foreground pixel in the image, relative and spatial information of the objects in the image and color features of logo [4], [10], [13], [20]- [22]. In this paper, a segmentation algorithm of the document image in page layout analysis application is used [1]. In [1], a two-stage segmentation algorithm based on wavelet transform and thresholding has proposed. Figure 1 shows block diagram of the wavelet-based segmentation algorithm. In the wavelet-based segmentation algorithm, segmentation is carried out on the wavelet transform subbands of the grayscale document image. Seven sets of the wavelet transform subbands (approximation and details in three directions, vertical, horizontal and diagonal) are formed in two levels of the wavelet transform. Small coefficients of subbands are replaced with zero value by applying a threshold value as a de-noising processing step. Threshold value of de-noising process is dependent to energy of each subband. It is determined by below equation:

$$Th = Min(coeff) + \alpha(Max(coeff) - Min(coeff)) \quad (1)$$

where $Min(coeff)$ and $Max(coeff)$ are the minimum and maximum values of the wavelet transform coefficients in each subband, respectively. α is a constant that we set to 0.05. Neighbor pixels in the wavelet subbands are amplified by dilation morphological operator, until segmented regions of the document image are formed. The detected segments by the wavelet-based segmentation algorithm are removed from the document image and the rest of the image is segmented by the threshold-based segmentation algorithm. In this stage, non-segmented regions are divided to foreground and background segments by applying a threshold that is determined from gray-level histogram based on Otsu's method [23]. Then foreground segments are labeled and extracted from image as text or picture segments.

3. HIERARCHICAL STRUCTURE OF LOGO DETECTION AND RECOGNITION

In this section, logo detection and recognition procedure is described based on the candidate logo regions. The decision-making procedure of the logo candidate regions in pre-defined logo classes is performed in a hierarchical structure. Figure 2 shows block diagram of the hierarchical classification structure of logo candidate region. In this structure, decision is carried out in the three level or three classifiers. It does not mean that the three classifiers classify each input region. However, in the worse case, the tertiary classifier determines the type of the input region. In other word, more complicated pattern of the logo causes more complex feature set and training stage. Details of the hierarchical structure consist of Bayesian classifier; the used features in each classifier and training procedure of the classifiers are described in the following.

3.1. Bayesian Classifier (Minimum Distance)

In the proposed structure, Bayesian classifier (Minimum distance) is used. A Bayesian classifier with c discriminate functions corresponds to c classes and $p(x \mid w_i) \sim N(\mu_i, \Sigma_i)$, $\Sigma_i = \sigma^2 I$ are defined as

$$g_i(x) = -\frac{\|x-\mu_i\|^2}{2\sigma^2} + \ln P(\omega_i) \quad i = 1,2,..,c \quad (2)$$

where $\|x-\mu_i\|^2 = (x-\mu_i)^t(x-\mu_i)$ .If prior probabilities $P(w_i)$ are the same for all c classes. Then the $\ln P(w_i)$ term becomes unimportant additive constant that can be ignored. Therefore, Bayesian discrimination function simplifies as below,

$$g_i(x) = -\|x-\mu_i\|^2 \quad i = 1,2,..,c \quad (3)$$

The above equation is called a minimum distance classifier.

3.2. Feature Extraction

In this paper, three different feature sets corresponds to three levels of the hierarchical logo detection and recognition structure are used. The first feature set consists of Seiden features [9] and Yin features such as area, deviation, symmetry, centralization and complexity [15]. The second feature set consists of energy, homogeneity, correlation extracted from co-occurrence matrix [18] and the tertiary feature set consists of ART features [17]. In the following, ART feature is described in details.

3.2.1. Angular Radial Transform features

ART is the 2-D complex transform defined on a unit disk that consists of the complete orthonormal sinusoidal basis functions in polar coordinates. The transformation is defined as [17]

$$F_{nm} = \langle V_{nm}(\rho,\theta), f(\rho,\theta)\rangle = \int_0^{2\pi}\int_0^1 V_{nm}^*(\rho,\theta), f(\rho,\theta)\rho\, d\rho\, d\theta \quad (4)$$

Here, F_{nm} is an ART coefficient of order n and m, $f(\rho,\theta)$ is an image function in polar coordinates, and $V_{nm}(\rho,\theta)$ is the ART basis function that are separable along the angular and radial directions, i.e.,

$$V_{nm}(\rho,\theta) = A_m(\theta)R_n(\rho) \quad (5)$$

The angular and radial basis functions are defined as follows:

$$A_m(\theta) = \frac{1}{2\pi}\exp(jm\theta) \quad (6)$$

$$R_n(\rho) = \begin{cases} 1 & n = 0 \\ 2\cos(\pi n\rho) & n \neq 0 \end{cases} \quad (7)$$

To describe a shape, all pixels constituting the shape are transformed with ART, and the transformed coefficients are formed into the ART descriptor. Twelve angular and three radial functions are used. By discarding the DC coefficient, 35 AC components form the descriptor vector.

3.3. Training of Classifiers

In this paper, three Bayesian classifiers in the three levels of the hierarchical structure determine class type of unknown logo region while false accept and false reject rates of the classifiers are less than two determined thresholds. Block diagram of the training procedure of the hierarchical classifiers is shown in Figure 3. In this block diagram, using train and test sets and segmented regions, optimum parameters of the classifiers are determined. In the following, details of this block diagram are described.

3.3.1. Train and Test Samples

To determine the optimum parameters of the classifiers in the hierarchical structure, three sets of data are used. The first data set is the segmented regions from document images by segmentation algorithm that is called database of detected regions in the block diagram of Figure 3. The other two sets are train and test samples of desirable logos of user. Train and test samples are created of the scaled and rotated versions of the desired logo

Train set consists of the original logo, rotated samples with rotation angels of $\pm 7^0$ and resized samples with scale coefficients of 0.7 and 1.2. Test set consists of five resized samples with scale coefficients of 0.65, 0.85, 0.95, 1.15, 1.35 and 6 rotated samples with rotation angles of $\pm 3^0$, $\pm 6^0$ and $\pm 9^0$ and 4 rotated and resized samples with rotation angles and scale coefficients of $\pm 2^0$, $\pm 5^0$ and 0.8, 0.9, 1.2 and 1.3, respectively. Ultimately, 15 test samples and 5 test samples are formed.

3.3.2. Determination of optimum classification parameters

Feature vector of an unknown logo x is belong to class i if prototype of class i (μ_i) is the nearest prototype in all prototypes to sample x based on equation (3). But, we use this classifier in another application, i.e. to aim to false reject rate less than threshold β, in each classifier of hierarchical structure. It may that an input logo belongs to many predefined classes of logos. It causes that false accept rate increases. Now if false accept rate of the first classifier is more than threshold α, and then the second classifier classifies the input logo by features that are more intricate. Now if false accept rate of the second classifier is not less than threshold α. Ultimately, the tertiary classifier carries out the final classification. In each classifier, the maximum value of discriminate function $g_i(x)$ for acceptance of the logos in the class i is saved as optimum parameters of the training stage when false accept and false reject rates are less than thresholds α and β ($FAR \le \alpha\%$ and $FRR \le \beta\%$), respectively.

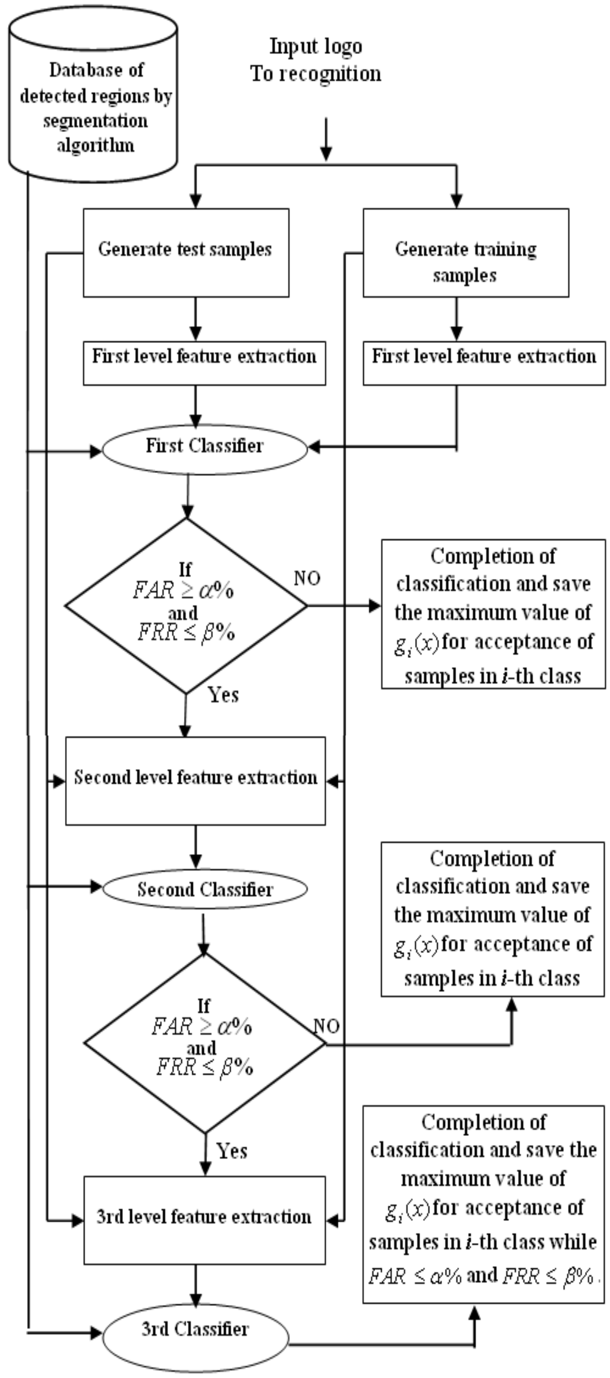

Fig.3. Block diagram of the training procedure of the hierarchical classifiers.

Figure 4 shows graph of false accept rate versus false reject rate and desired and undesired regions for a classifier. In this graph, performance of classifier is accept that fall into regions with $FAR \le \alpha\%$ and $FRR \le \beta\%$. Whereas, the other regions of graph with $FRR \le \beta\%$ and $FAR > \alpha\%$ are not accept that they are called undesired regions. In these cases, other classifiers with more complicated features are used unavoidably.

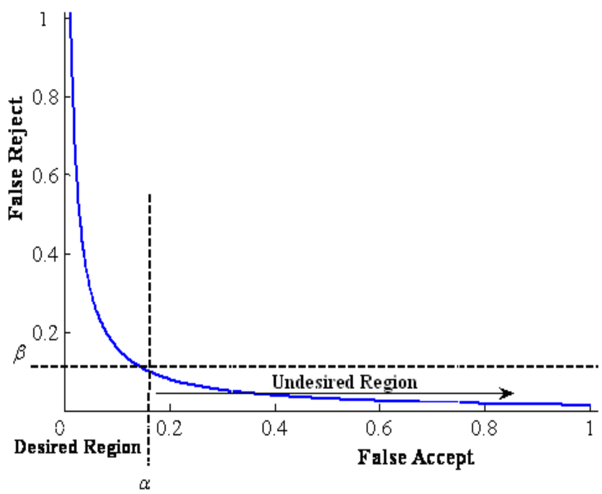

Fig. 4. Graph of false accept rate versus false reject rate and desired and undesired regions for a classifier.

4. RESULTS

The used document image database is a collection of 1600 images comprising three classes, document image (images only with text regions), pure picture (images without text regions) and combined images (images with text and picture regions). This database is used to evaluate the two-stage segmentation algorithm. The used logo database is a collection of 1980 images comprising international and Persian logos. To evaluate the proposed logo detection and recognition algorithm, these logos have inserted to document images manually. Image sample of the document image database are shown in Figure (5). In the following, results of the wavelet-based segmentation algorithm and training procedure of Bayesian classifier are described.

4.1. Wavelet-based Segmentation Algorithm

In this paper, the wavelet-based segmentation algorithm is applied to detect the logo candidate regions. The wavelet-based segmentation algorithm uses morphological operators in the wavelet subbands to detect text and shape regions [1]. Image sample of the segmented regions by the wavelet-based segmentation algorithm is shown in Figure (6). In Figure (6), the segmented regions consist of shape logos, a section of text regions and a section of pure picture.

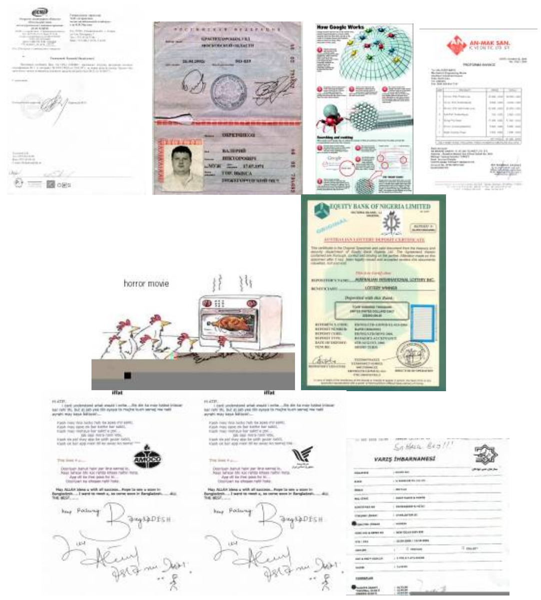

Fig.5. Image sample of document images database.

Fig.6. Image sample of the extracted regions by the wavelet-based segmentation algorithm.

Fig. 7. Image sample of desired logos for classification problem (from left to right corresponding to class 1 through 5, respectively).

4.2. Extraction of optimum classification parameters

Predefined classes of the extracted regions by the wavelet-based segmentation algorithm are determined using hierarchical structure. In this hierarchical structure, using three Bayesian classifiers and test and train samples, the optimum parameters of training such as mean value of each class as prototype sample and the maximum value of $g_i(x)$ for acceptance of samples are calculated.

As you can see in Figure 4, thresholds and parameters of classifiers are designed that the rates of false accept and false reject are less than two thresholds α and β, respectively. In each level of hierarchical classification, if false accept rate is more than α while false reject rate less than β, then class of the unknown region will determine in the next level. This strategy is designed that the rejected logos are decreased while false accept rate of irrelevant logos be in an acceptable limit. 15 samples of the test samples have very large variance and accumulation of all 15 samples in a class is irrational (notice that rotated and scaled logos with large values form new logos.). Therefore, it is accepted, if 6 samples of the test samples reject (i.e. $\beta = 40\%$). Threshold of false accept rate is set less than 1% (i.e. $\alpha = 1\%$). Table 1 shows classification results of the logo candidate region with shown logos in Figure 7. The selected logos in this test have complex pattern and considerable similarity with other logos. These problems cause acceptance of 23, 11 and 5 logos of irrelevant logos in the first, second and fifth classes, respectively. It is happened while the total of logo candidate region correctly accepted. In other word, accuracy rate of logo detection is 100%. Of course, a number of regions are detected incorrectly as logo candidate region. Figure 8 shows image sample of false accept cases of classification problem corresponding to logos of Figure 7. False accepted logos in Figure 8 are similar to original logo in each class. Therefore, error of classification is justifiable. In Table 1, false rejected logos in each class are resized and rescaled versions of original logo with very large rotation angles and scale coefficients. Therefore, the rejection of these logos is admissible. Ultimately, accuracy rate of the proposed algorithm is 97.6% for a 5-class classification problem.

5. CONCLUSION AND COMPARISON

In this paper, a hierarchical classification structure based on trainable Bayesian classifier for logo detection and recognition is proposed. In this structure, a novel training approach is used to extract the optimum parameters of Bayesian classifier such as mean of samples and the maximum value of discriminate function for acceptance of a sample in a class by false accept and false reject rates. In each level of this structure, an isolated set of shape and texture features are extracted based on logo complexity.

The used features in different decision levels; from down levels to up levels are more intricate. The proposed structure has two main advantages, logo detection and recognition stages using shape and morphological features simultaneously and generalization of training using the proposed training procedure of Bayesian classifier.

Fig. 8. Image sample of false accept cases of classification problem corresponding to logos of Figure 7. (a) Class 1, (b) Class 2, (c) Class 5 and (d) Class rejection.

The proposed structure evaluated on a very large database consisting of the document and non-document images with Persian and international logos. Tests have

carried out in the worse case conditions. False accept and false reject rates of less than 1% and 40% have obtained respectively for a 5-class classification problem. These results are very good in the worse case of test conditions. The segmentation algorithm has obtained false reject rate of zero for logo candidate region detection stage.
An exact comparison across the presented algorithms in the literature is a complex task (especially as no standard dataset is available at this time). However, for example in [15], a retrieval logo structure has presented in China logo application. Recall rate of 76% has reported for retrieval of the first ten relevant logos. In [15], a shape feature set and relevance feedback mechanism are used to improve the performance.

6. ACKNOWLEDGMENT

This work was supported by Young Research Club-Islamic Azad University Najafabad Branch (IAUN) based on approved research proposal.

Table 1. Classification Results of logo candidate regions.

Class	The number of logos	True Accept	False Accept	False Reject	False Accept Rate (%)
1	15	13	23	2	0.92
2	15	9	11	6	0.44
3	15	9	0	6	0
4	15	10	0	5	0
5	15	13	5	2	0.2
Reject class	2431	2392	21	39	-
Total	2506	2446	60	60	1.59

7. REFERENCES

[1] H. Pourghassem, "**Page Layout Analysis of the Document Image based on Region Classification in a Hierarchical Decision Structure**", *Journal of Electronics and Power Engineering*, Vol. 2, pp. 37-44, Spring, (2009)

[2] M. Goria, M. Magginia, S. Marinaib, J.Q. Shengc, G. Sodab, "**Edge-backpropagation for noisy logo recognition**", *Pattern Recognition*, Vol. 36, pp. 103-110, (2003)

[3] D. Doermann, E. Rivlin, and I. Weiss, "**Applying algebraic and differential invariants for logo recognition**", *Machine Vision and Application*, Vol. 9, No. 2, pp.73-86, (1996)

[4] I. S. Hsieh, K. C. Fan, "**Multiple Classifiers for Color Flag and Trademark Image Retrieval**", *IEEE Transactions on Image Processing*, Vol. 10, No. 6, pp. 938-950, (2001)

[5] J. Neumann, H. Samet, and A. Soffer, "**Integration of local and global shape analysis for logo classification**", *Pattern Recognition Letters*, Vol. 23, No. 12, pp. 1449-1457, (2002)

[6] P. Suda, C. Bridoux, B. Kammerer, and G. Maderlechner,"**Logo and word matching using a general approach to signal registration**", *In Proc. Int. Conf. Document Analysis and Recognition*, pp. 61-65, (1997)

[7] K. Zyga, J. Schroeder, and R. Price. "**Logo recognition using retinal coding**". *In Proc. 38th Asilomar Conf. Signals, Systems and Computers,* Vol. 2, pp. 1549–1553, (2004)

[8] E. Francesconi, P. Frasconi, M. Gori, S. Marinai, J. Q. Sheng, G. Soda, A. Sperduti, "**Logo Recognition by Recursive Neural Networks**", *Graphics Recognition Algorithms and Systems*, Vol. 1389, pp. 104-117, (1998)

[9] S. Seiden, M. Dillencourt, S. Irani, R. Borrey, T. Murphy, "**Logo detection in document images**", *Proc. of the Int. Conf. on Imaging Science, Systems, and Technology*, Las Vegas, Nevada, pp. 446–449, (1997)

[10] T. Pham, "**Unconstrained logo detection in document images**", *Pattern Recognitio*n, Vol. 36, No. 12, pp. 3023-3025, (2003)

[11] Jingying Chen, Maylor K. Leung ,Yong sheng Gao, "**Noisy logo recognition using line segment Hausdorff distance**", *Pattern recognition*, Vol. 36, pp. 943-955, (2003)

[12] G. Nagy, S. Seth, and M. Viswanathan, "**A prototype document image analysis system for technical journals**", *Computer*, Vol. 7, pp. 10-22, (1992)

[13] G. Zhu, D. Doermann, "**Automatic Document Logo Detection**", *9th Int. Conf. on Document Analysis and Recognition,* Vol. 2, pp. 864-868, (2007)

[14] Y.S. Kim, W.Y. Kim, "**Content-based trademark retrieval system using a visually salient feature**", *Image and Vision Computing*, Vol. 16, pp. 931-939, (1998)

[15] P. Y. Yin, C. C. Yeh, "**Content-based retrieval from trademark databases**", *Pattern Recognition Lett*ers, Vol. 23, pp. 113-126, (2002)

[16] M. H. Hung, C. H. Hsieh, C. M. Kuo, "**Similarity retrieval of shape images based on database classification**", J. *Vis. Communication and Image Representation,* Vol. 17, pp. 970-985, (2006)

[17] W. Y. Kim, Y. S. Kim, "**A new region-based shape descriptor**", *ISO/IEC MPEG99/M5472*, TR15-01, Maui, Hawaii, (1999)

[18] R. M. Haralick, K. Shanmugan, I. Dinstein, "**Textural features for image classification**", *IEEE Trans on Sys Man and Cybern*, Vol. 3, No. 6, pp. 610-621, (1973)

[19] W.Q. Yan, J. Wang, M.S. Kankanhalli, "**Automatic video logo detection and removal**", *Multimedia Systems*, Vol. 10, No. 5, pp. 379-391, (2005)

[20] H. Wang, Y. Chen, "**Logo Detection in Document Images Based on Boundary Extension of Feature Rectangles**", *10th Conf. of Document Analysis and Recognition*, pp. 1335-1339, (2009)

[21] K. Gao, Sh. Lin, Y. Zhang, Sh. Tang; D. Zhang, "**Logo detection based on spatial-spectral saliency and**

partial spatial context", *IEEE Int. Conf. on Multimedia and Expo.,* pp. 322-329, (2009)

[22] Z. Ahmed, H. Fella, "**Logos extraction on picture documents using shape and color density**", *Int. IEEE Conf. on Industrial Electronics,* pp. 2492-2496, (2008)

[23] N. Otsu, "**A Threshold Selection Method from Gray-Level Histograms**", *IEEE Transactions on Systems, Man, and Cybernetics*, Vol. 9, No. 1, pp. 62-66, (1979)