Multilayer Neural Networks and Nearest Neighbor Classifier Performances for Image Annotation

Mustapha OUJAOURA

Laboratory of Information Processing and Telecommunications, Computer Science Department, Faculty of Science and Technology, Sultan Moulay Slimane University, PO Box. 523, Béni Mellal, Morocco.

Brahim MINAOUI

Laboratory of Information Processing and Telecommunications, Computer Science Department, Faculty of Science and Technology, Sultan Moulay Slimane University, PO Box. 523, Béni Mellal, Morocco.

Mohammed FAKIR

Laboratory of Information Processing and Telecommunications, Computer Science Department, Faculty of Science and Technology, Sultan Moulay Slimane University, PO Box. 523, Béni Mellal, Morocco.

Abstract—**The explosive growth of image data leads to the research and development of image content searching and indexing systems. Image annotation systems aim at annotating automatically animage with some controlled keywords that can be used for indexing and retrieval of images. This paper presents a comparative evaluation of the image content annotation system by using the multilayer neural networks and the nearest neighbour classifier. The region growing segmentation is used to separate objects, the Hu moments, Legendre moments and Zernike moments which are used in as feature descriptors for the image content characterization and annotation.The ETH-80 database image is used in the experiments here. The best annotation rate is achieved by using Legendre moments as feature extraction method and the multilayer neural network as a classifier.**

Keywords-Image annotation; region growing segmentation; multilayer neural network classifier; nearest neighbour classifier; Zernike moments; Legendre moments; Hu moments; ETH-80 database.

I. INTRODUCTION

As the online resources are still a vital resource in everyday life, developing automatic methods for managing large volumes of digital information is increasingly important. Among these methods, automatic indexing multimedia data remain an important challenge, especially the image annotation [1]. Annotated images play a very important role in information processing. They are useful for an image retrieval based on keywords and image content [2]. Manual annotation is not only boring but also not practical in many cases due to the abundance of information. Most images are, therefore, available without adequate annotation. Automatic image content annotation becomes a recent research interest [3]. It attempts to explore the visual characteristics of images and associate them with image contents and semantics.

Many classifiers have been used for the image classification and annotation without any performance's evaluation. In this paper, we use the neural network classifier for image annotation, and we compare its performance with the nearest neighbour classifier. We use the same type of moments as a method of features extraction for each classifier in order to have an objective comparison between performances of the considered classifiers. In such case, we used a system that tends to extract objects from each image and find the annotation terms that describe its individual content.

The rest of the paper is organised as follows. The Section 2 presents the adopted annotation system while the Section 3 discusses the primordial tasks of any annotation and recognition system. They are the image segmentation and features extraction problems in addition to a brief formulation for Hu, Legendre and Zernike moments as features extraction method. The Section 4 is reserved for the image classification and annotation using neural network or the nearest neighbour classifier. Finally, the Section 5 presents the experimental annotation results and the comparing performance of each used classifier.

II. ANNOTATION SYSTEM

The automatic annotation's techniques of image's content attempt to explore visual features of images that describe an image content and associate them with its semantics. It's an effective technology to improve the image indexing and searching in the large volume of information that is available in the media. The algorithms and systems, used for image annotation, are commonly divided into those tasks [4]:

- Segmentation and Features extraction;
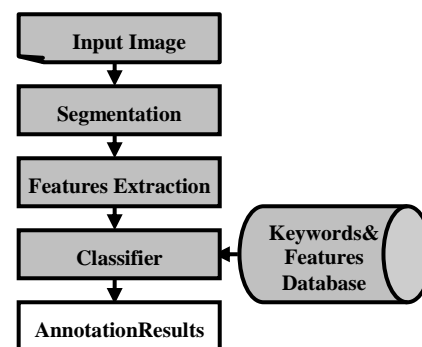
- Classification and Annotation.

Figure 1. The block diagram of the image annotation system.

The annotation system adopted in this work is shown in Fig.1. The system has a reference database that contains keywords and features descriptor of images that are already annotated by experts (manual offline annotation).This database is used for modelling and training the classifier in order to choose the appropriate keywords. To achieve this goal, by using region growing segmentation, the input image is firstly segmented into regions that represent objects in the image. The feature vectors of each region are secondly computed and extracted from the image. Those features are finally feed into the input of the classifier, that can be the multilayer neural networks or the nearest neighbor classifier, in order to decide and choose the appropriated keywords for annotation tasks. Therefore, the image content is annotated and the performance of each classifier can be evaluated.

III. IMAGE SEGMENTATION AND FEATURES EXTRACTION

The features' vector, which is extracted from the entire image, loses local information. So, it is necessary to segment an image into regions or objects of interest and use of local characteristics. The image segmentation is the process of partitioning a digital image into multiple segments. It is very important in many applications for any image processing, and it still remains a challenge for scientists and researchers. So far, the efforts and attempts are still being made to improve the segmentation techniques. With the improvement of computer processing capabilities, there are several possible segmentation techniques of an image: threshold, region growing, active contours, level sets, etc... [5]. Among these methods, the region growing is well suited because of its simplicity and has been successfully used several times as a segmentation technique of digital images.

The regions are iteratively grown by comparing all unallocated neighbouring pixels to the regions. The difference between a pixel's intensity and the pixel's mean in one region is used as a measure of similarity, it's a predicate that control the evolution of segmentation process. The pixel, with the smallest difference measured this way, is allocated to the respective region. This process continues until all pixels are allocated to a region.

While all the pixels in image are not visited;

  1. Choose an unlabeled pixel p_k;

  2. Set the region's mean to intensity of pixel p_k;

  3. Consider unlabeled neighboring pixels p_{kj};

  If (pixel's intensity - region's mean) < threshold;

    a. Affect the pixel p_{kj} to the region labeled by k.

    b. Update the region's mean and go back to ③;

  Else k = k + 1 and go back to①.

  End If

End While

Figure 2. Region growing segmentation algorithm.

By assuming that the objects are localised at the center of images, the region growing segmentation is started from the corner of the image to isolate the objects in the center of the image. The region growing segmentation algorithm used in this paper is presented in Fig.2 and an example of image segmentation is in Fig.3.
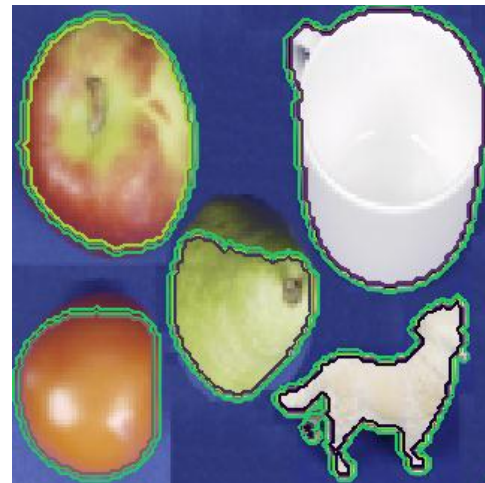

Figure 3. Example of image segmentation result.

After dividing the original image into several distinct regions that correspond to objects in a scene, the feature vector can be extracted from each region and can be considered as a representation of an object in the entire image.

The feature extraction task transforms carefully the rich content and large input data of images into a reduced representation set of features in order to decrease the processing time. Not only it enhances the retrieval and annotation accuracy, but the annotation speed as well, since a large image database can be organized according to the classification rule and, therefore, search can be performed [6].

In the feature extraction method, the representation of the image content must be considered in some situations such as: translation, rotation and change of scale. This is the reason that justifies the use of moments for feature extraction method from the segmented image.

The use of moments for image analysis and pattern recognition was inspired by Hu [7] and Alt [10]. In this paper, the moments used are:

- Hu moments.
- Legendre moments.
- Zernike moments.

A. Hu moments

For a discrete image of M x N pixels with intensity function f(x, y), Hu [7] defined the following seven moments that are invariant to the change of scale, translation and rotation:

$$\phi_1 = \mu_{20} + \mu_{02} \qquad (1)$$

$$\phi_2 = \left(\mu_{20} + \mu_{02}\right)^2 + 4\mu_{11}^2 \qquad (2)$$

$$\phi_3 = (\mu_{30} - 3\mu_{12})^2 + (3\mu_{21} - \mu_{03})^2 \qquad (3)$$

$$\phi_4 = (\mu_{30} - \mu_{12})^2 + (\mu_{21} - \mu_{03})^2 \qquad (4)$$

$$\phi_5 = (\mu_{30} - 3\mu_{12})(\mu_{30} + \mu_{12}) \times$$
$$\left[(\mu_{30} + \mu_{12})^2 - 3(\mu_{21} + \mu_{03})^2 \right] +$$
$$(3\mu_{21} - \mu_{03})(\mu_{21} + \mu_{03}) \times \qquad (5)$$
$$\left[3(\mu_{30} + \mu_{12})^2 - (\mu_{21} + \mu_{03})^2 \right]$$

$$\phi_6 = (\mu_{20} - \mu_{02})\left[(\mu_{30} + \mu_{12})^2 - (\mu_{21} + \mu_{03})^2 \right]$$
$$+ 4\mu_{11}(\mu_{30} + \mu_{12})(\mu_{21} + \mu_{03}) \qquad (6)$$

$$\phi_7 = (3\mu_{21} - \mu_{03})(\mu_{30} + \mu_{12}) \times$$
$$\left[(\mu_{30} + \mu_{12})^2 - 3(\mu_{21} + \mu_{03})^2 \right] -$$
$$(\mu_{30} - 3\mu_{12})(\mu_{21} + \mu_{03}) \times \qquad (7)$$
$$\left[3(\mu_{30} + \mu_{12})^2 - (\mu_{21} + \mu_{03})^2 \right]$$

Where,

$$\mu_{pq} = \frac{\displaystyle\sum_{x=0}^{N-1}\sum_{y=0}^{M-1}(x - \bar{x})^p (y - \bar{y})^q f(x,y)}{\left(\displaystyle\sum_{x=0}^{N-1}\sum_{y=0}^{M-1} f(x,y) \right)^{\frac{p+q}{2}+1}} \qquad (8)$$

And,

$$\begin{cases} \bar{x} = \dfrac{\displaystyle\sum_{x=0}^{N-1}\sum_{y=0}^{M-1} x^p f(x,y)}{\displaystyle\sum_{x=0}^{N-1}\sum_{y=0}^{M-1} f(x,y)} \\[20pt] \bar{y} = \dfrac{\displaystyle\sum_{x=0}^{N-1}\sum_{y=0}^{M-1} y^q f(x,y)}{\displaystyle\sum_{x=0}^{N-1}\sum_{y=0}^{M-1} f(x,y)} \end{cases} \qquad (9)$$

B. Legendre moments

Legendre moments were first introduced by Teague [8]. They were used in several pattern recognition applications [9]. The orthogonal property of Legendre polynomials implies no redundancy or overlap of information between the moments with different orders. This property enables the contribution of each moment to be unique and independent of the information in an image [10]. The Legendre moments for a discrete image

of M x N pixels with intensity function f(x, y) is the following [13]:

$$L_{pq} = \lambda_{pq} \sum_{x=0}^{M-1}\sum_{y=0}^{N-1} P_p(x_i)\, P_q(y_j)\, f(x,y) \qquad (10)$$

Where $\lambda_{pq} = \dfrac{(2p+1)(2q+1)}{M \times N}$, xi and yj denote the normalized pixel coordinates in the range of [-1, +1], which are given by:

$$\begin{cases} x_i = \dfrac{2x - (M-1)}{M-1} \\[12pt] y_j = \dfrac{2y - (N-1)}{N-1} \end{cases} \qquad (11)$$

$P_p(x)$ is the p$^{\text{th}}$-order Legendre polynomial defined by:

$$P_p(x) = \sum_{k=0}^{p} \left\{ \frac{(-1)^{\frac{p-k}{2}} x^k (p+k)!}{2^p\, k! \left(\frac{p-k}{2} \right)! \left(\frac{p+k}{2} \right)!} \right\}_{p-k=even} \qquad (12)$$

And, the recurrent formula of the Legendre polynomials is:

$$\begin{cases} P_p(x) = \dfrac{(2p-1)x}{p} P_{p-1}(x) - \dfrac{(p-1)}{p} P_{p-2}(x) \\[12pt] P_1(x) = x \quad , \quad P_0(x) = 1 \end{cases} \qquad (13)$$

In this work the recurrent formula is used for calculating Legendre polynomials in order to increase the computation speed.

C. Zernike moments

Zernike moments are the mapping of an image onto a set of complex Zernike polynomials. As these Zernike polynomials are orthogonal to each other, Zernike moments can represent the properties of an image with no redundancy or overlap of information between the moments [11]. Due to these characteristics, Zernike moments have been utilized as feature sets in many applications [12].

The discrete form of the Zernike moments of an image size $M \times N$ represented by f(x, y) is expressed, in the unit disk $(x^2 + y^2) = 1$, as follows [13]:

$$Z_{pq} = \frac{p+1}{\lambda} \sum_{x=0}^{M-1}\sum_{y=0}^{N-1} R_{p,q}(r_{xy}) e^{-jq\theta_{xy}} f(x,y) \qquad (14)$$

Where

$$\begin{cases} R_{p,q}(r) = \sum_{s=0}^{(p-|q|)/2} \dfrac{(-1)^s (p-s)! \, r^{p-2s}}{s! \left(\dfrac{p+|q|}{2} - s \right)! \left(\dfrac{p-|q|}{2} - s \right)!} \\ p - |q| = (even), \ |q| \le p, \ p \ge 0 \end{cases} \quad (15)$$

λ is the number of pixels located in the unit circle, and the transformed phase θ_{xy} and the distance r_{xy} at the pixel of coordinates (x, y) are [13]:

$$\begin{cases} \theta_{xy} = \tan^{-1} \left(\dfrac{(2y-(N-1))/(N-1)}{(2x-(M-1))/(M-1)} \right) \\ r_{xy} = \sqrt{ \left(\dfrac{2x-(M-1)}{M-1} \right)^2 + \left(\dfrac{2y-(N-1)}{N-1} \right)^2 } \end{cases} \quad (16)$$

Most of the time taken for the computation of Zernike moments is due to the computation of radial polynomials. Therefore, researchers have proposed faster methods that reduce the factorial terms by utilizing the recurrence's relations on the radial polynomials. In this paper, we obtained Zernike moments using the direct method.

IV. IMAGE'S CLASSIFICATION AND ANNOTATION

As the features are extracted, a suitable classifier must be chosen. Many classifiers are used and each classifier is found suitable to classify a particular kind of feature vectors depending on their characteristics.

Some of these classifiers are based on supervised learning which require an intensive learning and a training phase of the classifier parameters (parameters of Support Vector Machines [14], Boosting [15], parametric generative models [16], decision trees[17], and Neural Network[18]). They are also known as parametric classifiers. Other classifiers base their classification decision directly on the data, and require no learning and training of parameters. These methods are also known as nonparametric classifiers. The most common nonparametric classifier, based on distance estimation, is the Nearest Neighbour classifier.

There are several types of the classifier. Based on their ability to detect complex nonlinear relationships between dependent and independent variables, the multilayer neural networks are used in this paper to classify and annotate image content. Their performances are compared to those of the Nearest Neighbour classifier.

A. Multilayer Neural Network classifier

A multilayer neural network consists of an input layer including a set of input nodes, one or more hidden layers of nodes, and an output layer of nodes [19]. Fig. 4 shows an example of a three layer network, used in this paper, having input layer formed by m nodes, one hidden layer formed by 20 nodes, and output layer formed by n nodes. This neural network is trained to classify inputs according to target classes.

The training input data are loaded from the reference database while the target data should consist of a vector of all zero values except for the element i that represents the appropriate class.
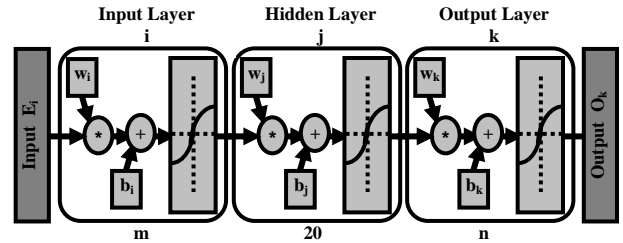

Figure 4. The three layer neural network.

After random initialisation of bias and connection weights, the training principle of the neural network is based on a loop of steps. It starts by propagating inputs from the input layer to the output layer. The error is calculated and back propagated for each layer of the neural network in order to update the bias and connection weights. When the bias and connection weights are changed, the propagation of the inputs is repeated until having the minimum error between outputs and targets. This is the criterion stop of the neural network training.

The inputs Y_i are presented to the input layer and propagated to the hidden layer using the following formula:

$$Y_j = f\left(\sum_{i=1}^{m} Y_i w_i + b_i \right) \quad (17)$$

Then from the hidden layer to the output layer:

$$Y_k = f\left(\sum_{j=1}^{20} Y_j w_j + b_j \right) \quad (18)$$

Finally, the outputs are:

$$O_k = f\left(\sum_{k=1}^{n} Y_k w_k + b_k \right) \quad (19)$$

Where f is the activation function (hyperbolic tangent sigmoid) used in the tree layer neural network, defined by:

$$f(x) = \frac{2}{1 + e^{-2x}} - 1 \quad (20)$$

At the output layer, the error between the desired output T_k and the actual output O_k is calculated by:

$$E_k = \left(1 - O_k^2\right)\left(T_k - O_k\right) \quad (21)$$

The calculated error is propagated to the hidden layer using the following formula:

$$E_j = \left(1 - Y_k^2\right)\sum_{k=1}^{n} w_k . E_k \qquad (22)$$

Then the calculated back propagation of error from the hidden layer to the input layer is:

$$E_i = \left(1 - Y_j^2\right)\sum_{j=1}^{20} w_j . E_j \qquad (23)$$

The bias and the connections weights of the input layer i, the hidden layer j and the output layer k are adjusted by:

$$\begin{cases} \Delta w_l = Y_l . E_l \\ \Delta b_l = E_l \end{cases} \quad for \ l = i, j, k \qquad (24)$$

B. Nearest Neighbor classifier

The nearest neighbour classifier is used to compare the feature vector of the input image and the feature vectors stored in the database. It is obtained by finding the distance between the prototype image and the database. The class is found by measuring the distance between a feature vector of input image and feature vectors of images in reference database. The Euclidean distance measurement is used in this paper, but other distance measurement can be also used [14].

Let X_1, X_2, …, X_k be the k class features vectors in the database and X_q the feature vector of the query image. The feature vector with the minimum distance is found to be the closest matching vector. It is given by:

$$d(X_q, X_j) = \min_{j\in\{1,2,...k\}}\left\{ \sqrt{\sum_i \left(x_q(i) - x_j(i)\right)^2} \right\} \qquad (25)$$

C. Image Annotation

For the image annotation, low-level feature vectors are calculated iteratively for each region in the image, either by using Hu moments or Legendre moments or Zernike moments. These features vector are presented either to the nearest neighbour classifier or to the Multilayer neural network that was already trained.

When features vector are presented to the nearest neighbour classifier to test matching with the feature values in reference database, the label or keyword of image class with the minimum distance is selected.

When features vectors are feed to the input layer of the Multilayer neural network, where each of the input neurons or nodes correspond to each element of the features vector, The output neurons of the neural network represent the class labels of images to be classified and annotated. Then, each region is annotated by the corresponding label that is found by the classifier.

The input layer of the neural network has a variable number of input nodes. It has seven input nodes in the case of

adoption of Hu moments, nine input nodes when adopting Zernike moments and ten input nodes when used Legendre moment as feature extraction method. However, the number of input nodes for the neural networks can be changed and increased when using Zernike and Legendre moments, as a feature extraction method, to increase the accuracy of the annotation system. The features can be also combined together and feed to the Multilayer neural network or the nearest neighbour classifier.

V. EXPERIMENTS AND RESULTS

A. Experiments

In our experiments, for each region that represent an object from the input image, the number of input features extracted using Hu invariants features extraction method is 7 (hu1, hu2, hu3, hu4, hu5, hu6, hu7) while the number of input features extracted using the order 4 of Zernike moments is 9 (Z00, Z11, Z20, Z22, Z31, Z33, Z40, Z42, Z44) and the number of input features extracted using the order 3 of Legendre moments is 10 (L00, L01, L02, L03, L10, L11, L12, L20, L21, L30). These inputs are feed to the input of the multilayer neural network or the nearest neighbour classifier in order to do matching with the feature values in the reference database. Then, the appropriate keywords are selected and used for annotation of the input image.

The fig. 5 shows some examples of image objects from the ETH-80 image database used in our experiments. The experiments are based on eight classes of objects (Apple, Car, Cow, Cup, Dog, Horse, Pears, and Tomato). The number of prototypes per class is 5.
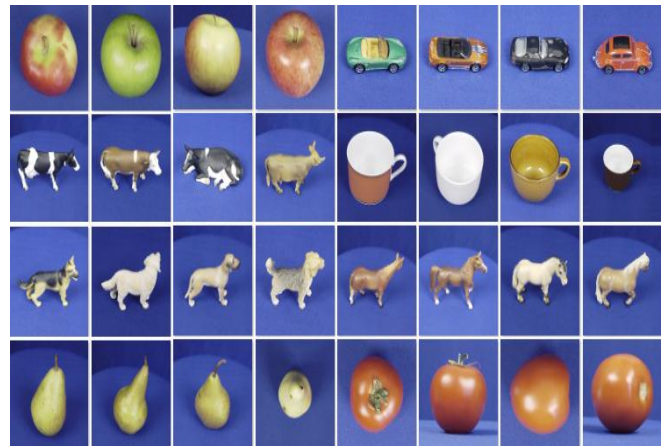

Figure 5. Some examples of objects from the ETH-80 database.

The accuracy of image annotation system is evaluated by the precision rate which is the number of the correct results divided by the number of all returned results.

All the experiments are conducted using the ETH-80 database containing a set of 8 different object images [20]. The proposed system has been implemented and tested on a Core 2 Duo personal computer using Matlab Software.

B. Results

For both of the used classifiers (the neural network classifier and Nearest Neighbour classifier), the annotation

rates of Hu, Zernike and Legendre Moments, as the feature extraction method, are given for each object in Table 1.

TABLE I.   OJECTS ANNOTATION RATE OF HU, ZERNIKE AND LEGENDRE MOMENTS USING NEURAL NETWORK AND NEAREST NEIGHBOUR CLASSIFIER

Object	Neural networkclassifier			Nearest neighbourclassifier		
	Hu	*Zernike*	*Legendre*	*Hu*	*Zernike*	*Legendre*
Apple	71.20%	78.40%	92.13%	69.12%	76.84%	90.93%
Car	68.53%	74.37%	83.11%	67.03%	72.87%	81.91%
Cow	41.32%	50.00%	61.37%	39.82%	49.13%	60.10%
Cup	63.27%	71.20%	81.20%	62.07%	70.37%	80.02%
Dog	62.57%	72.30%	80.41%	60.87%	70.93%	79.10%
Horse	48.60%	56.40%	65.28%	46.80%	55.10%	63.80%
Pears	69.29%	78.40%	89.38%	66.80%	76.94%	87.85%
Tomato	67.49%	76.52%	86.40%	65.90%	75.10%	84.70%
Average	61.53%	69.70%	79.91%	59.80%	68.41%	78.55%

The general annotation rates and error rates of the neural network classifier or the Nearest Neighbour classifier based on Hu moments, Zernike moments and Legendre Moments as feature extraction method are given in Table 2. The experimental results showed that the annotation rate of the neural network classifier based on Legendre moments, Zernike moments and Hu moments is higher than the annotation rate of the Nearest Neighbour classifier based on Legendre moments, Zernike moments and Hu moments.

TABLE II.   Annotation and Error Rate of Hu, Zernike and Legendre Moments using Neural Network and Nearest Neighbour classifier

Classifier	Features Extraction method	Annotation Rate	Error Rate
Neural network	Hu moments	61.53%	38.47%
	Zernike moments	69.70%	30.30%
	Legendre moments	79.91%	20.09%
Nearest neighbour	Hu moments	59.80%	40.20%
	Zernike moments	68.41%	31.59%
	Legendre moments	78.55%	21.45%

The best annotation rate is achieved when we use the Legendre moments as the features extraction method and the multilayer neural networks as a classifier. However, for the computation time, the Nearest Neighbour classifier based on Hu moments is the best.

The annotation rates of Zernike moments is lower than the annotation rate of Legendre moments because Zernike moments are computed only from pixels of an image inside a unit circle obtained by the mapping transformation over a unit circle. In one hand, those pixels outside a unit circle are not considered. On the other hand, only the absolute values of extracted complex Zernike moments are feed into the classifier.

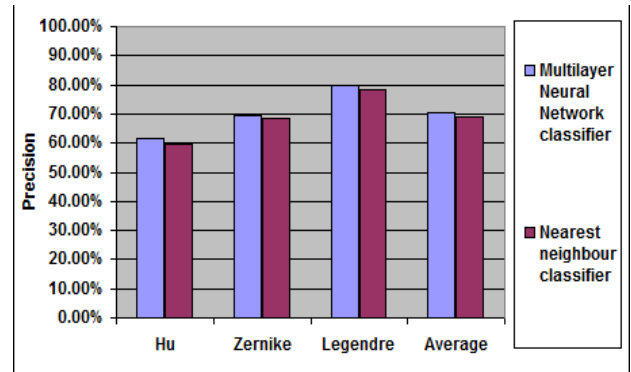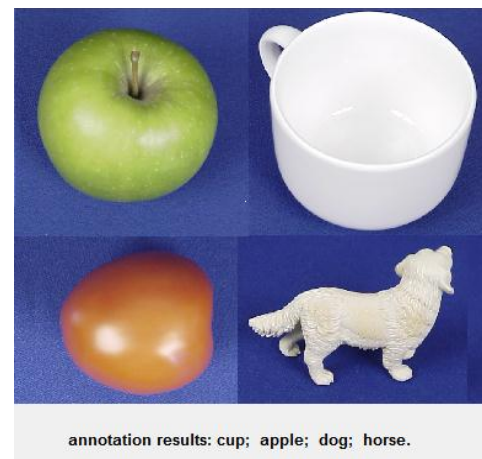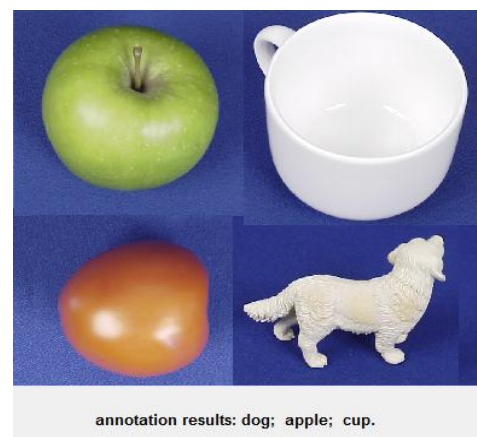This is the main reason for the modest results obtained by Zernike moments.

Figure 6.   Comparison of general annotation rates.

The fig 6 gives the comparison of general annotation rate for each extraction method and for each classifier.

annotation results: cup; apple; dog; horse.

(a)

annotation results: dog; apple; cup.

(b)

Figure 7.   Example of the annotation results using Zernike moment features and (a) the multilayer neural network classifier (b) the Nearest Neighbour classifier.

The Fig. 7 presents an example of annotation results obtained using the presented system and Zernike moment as a

feature extraction method. The Graphical User Interface is illustrated in Fig. 8.
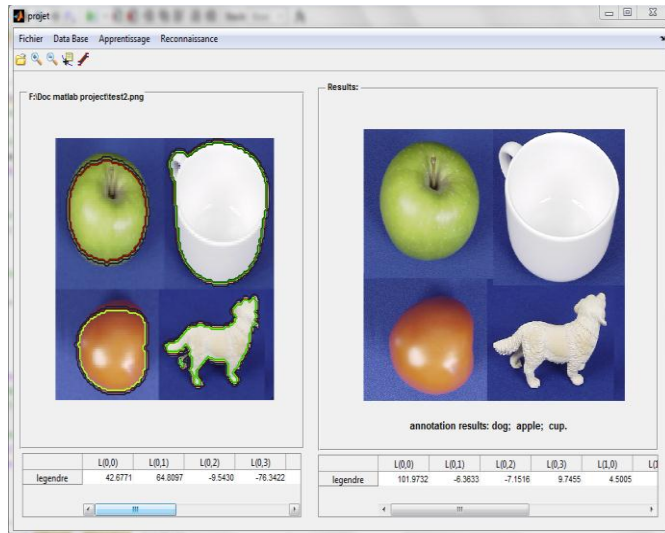

Figure 8.   Graphical User Interface (GUI).

The results are also affected by the accuracy of the image segmentation method. In most cases, it is very difficult to have an automatic ideal segmentation. This problem decreases the annotation rates. Therefore, any annotation attempt must consider the image segmentation as an important step, not only for automatic image annotation system, but also for the other systems which require its use. The Multilayer neural network classifier based on Legendre moments and Zernike moments is very expensive regarding to the computation time of the features in addition to the training time of the classifier. So, any use of them in a real time for an online image annotation system will be difficult and impracticable.

VI.   CONCLUSION

In this paper, we evaluated the image annotation performance of the neural network classifier and the nearest neighbour classifier based on Hu moments, Legendre moments and Zernike moments as feature extraction methods. The performance of each classifier and each feature extraction method has been experimentally analyzed. The successful experimental results proved that the proposed image annotation system based on the multilayer neural network classifier gives the best results for some images that are well and properly segmented. However, the processing time and the Image segmentation remain a challenge that needs more attention in order to increase precision and accuracy of the image annotation system. Also, the gap between the low-level features and the semantic content of an image must be reduced and considered for more accuracy of any image annotation system. Other Image segmentation and other classifiers may be considered in the future works for other image databases.

REFERENCES

[1] N. Vasconcelos and M. Kunt, Content-Based Retrieval from Image Databases: Current Solutions and Future Directions, Proc. Int'l Conf. Image Processing, 2001.

[2] A. Smeulders, M. Worring, S. Santini, A. Gupta, and R. Jain, Content-Based Image Retrieval: The End of the Early Years, IEEE Trans. Pattern. Analysis and Machine Intelligence, vol. 22, no. 12, pp. 1349-1380, Dec. 2000.

[3] Lei Ye, Philip Ogunbona and Jianqiang Wang, Image Content Annotation Based on Visual Features, Proceedings of the Eighth IEEE International Symposium on Multimedia (ISM'06), IEEE computer society, San Diego, USA, 11-13 December 2006.

[4] Ryszard S. Chora´s, Image Feature Extraction Techniques and Their Applications for CBIR and Biometrics Systems, International Journal Of Biology And Biomedical Engineering, Issue 1, Vol. 1, pp. 6-16, 2007.

[5] Frank Y. Shih, Shouxian Cheng, Automatic seeded region growing for color image segmentation, Image and Vision Computing 23, pp. 877–886, 2005.

[6] Zijun Yang and C.-C. Jay Kuo, Survey on Image Content Analysis, Indexing, and Retrieval Techniques and Status Report of MPEG-7, Tamkang Journal of Science and Engineering, Vol. 2, No. 3, pp. 101-118, 1999.

[7] M. K. Hu, Visual Problem Recognition by Moment Invariant, IRE. Trans. Inform. Theory, Vol. IT-8, pp. 179-187, Feb. 1962.

[8] M.R. Teague, Image analysis via the general theory of moments, J. Opt. Soc. Amer. 70 , pp. 920–930, 1980.

[9] Chee-Way Chonga, P. Raveendranb and R. Mukundan, Translation and scale invariants of Legendre moments, Pattern Recognition 37, pp. 119 – 129, 2004.

[10] F. L. Alt, Digital Pattern Recognition by Moments, J. Assoc. Computing Machinery, Vol. 9, pp. 240-258, 1962.

[11] Sun-Kyoo Hwang, Whoi-Yul Kim, Anovel approach to the fast computation of Zernike moments, Pattern Recognition 39, pp. 2065 – 2076, 2006.

[12] R. Sinan Tumen, M. Emre Acer and T. Metin Sezgin, Feature Extraction and Classifier Combination for Image-based Sketch Recognition, Eurographics Symposium on Sketch-Based Interfaces and Modeling, pp. 1–8, 2010.

[13] Mustapha Oujaoura, Brahim Minaoui and Mohammed Fakir. Article: Image Annotation using Moments and Multilayer Neural Networks. IJCA Special Issue on Software Engineering, Databases and Expert Systems SEDEX(1):46-55, September 2012. Published by Foundation of Computer Science, New York, USA.

[14] Oren Boiman, Eli Shechtman and Michal Irani, In Defense of Nearest-Neighbor Based Image Classification, IEEE Conference on Computer Vision and Pattern Recognition (CVPR), June 2008.

[15] A. Opelt, M. Fussenegger, A. Pinz, and P. Auer. Weak hypotheses and boosting for generic object detection and recognition, In ECCV, 2004.

[16] G. Wang, Y. Zhang, and L. Fei-Fei. Using dependent regions for object categorization in a generative framework. In CVPR, 2006.

[17] A. Bosch, A. Zisserman, and X. Munoz. Image classification using random forests and ferns. In ICCV, 2007.

[18] P.Simard, D. Steinkraus, J. C. Platt, Best Practices for Convolutional Neural Networks Applied to Visual Document Analysis, ICDAR, 2003, pp. 958-962.

[19] S. Haykin, Neural networks: a comprehensive foundation, 2nd Edition. Prentice Hall, New Jersey. 1999.

[20] ETH-80 database image. [Online]. Available: http://www.d2.mpi-inf.mpg.de/Datasets/ETH80