3-D OBJECT RECOGNITION USING MULTI-WAVELET AND NEURAL NETWORK

Dr. Tariq Zeyad Ismail   **Zainab Ibrahim Abood**

ABSTRACT

This search has introduced the techniques of multi-wavelet transform and   neural network for recognition 3-D object from 2-D image using patches. The proposed techniques were tested on database of different patches features and the high energy subband of discrete multi-wavelet transform DMWT (gp) of the patches. The test set has two groups, group (1) which contains images, their (gp) patches and patches features of the same images as a part of that in the data set beside other images, (gp) patches and features, and group (2) which contains the (gp) patches and patches features the same as a part of that in the database but after modification such as rotation, scaling and translation. Recognition by back propagation (BP) neural network as compared with matching by minimum distance, gave (94%) and (83%) score by using group (1), (gp) and features respectively, which is much better than the minimum distance. Recognition using (gp) neural network (NN) gave a (94%) and (72%) score by using group (2), (gp) and features respectively, while the minimum distance gave (11%) and (33%) scores. Time consumption through the recognition process using (NN) with (gp) is less than that minimum distance.

الخلاصة

هذا البحث يقدم تقنيات تحويل الموجة المتعددة والشبكة العصبية الاصطناعية لتمييز الأجسام الثلاثية الأبعاد من صور ثنائية الأبعاد باستعمال الشرائح. التقنيات المقترحة تختبر عن طريق بيانات تحوي خصائص الشرائح المختلفة و الجزء ذات الطاقة العالية بعد استخدام اسلوب يعتمد على تحويل الموجة المتعددة للشرائح. مجموعة الاختبار تحتوي على مجموعتين المجموعة الأولى تحتوي على صور وشرائح وخصائص الشرائح و الجزء ذات الطاقة العالية بعد استخدام اسلوب يعتمد على تحويل الموجة المتعددة للشرائح تشابه بعض تلك الموجودة في قاعدة البيانات. أما المجموعة الثانية فتحتوي على صور، شرائح، خصائص الشرائح و الجزء ذات الطاقة العالية بعد استخدام اسلوب يعتمد على تحويل الموجة المتعددة للشرائح تشابه بعض تلك الموجودة في قاعدة البيانات لكن بعد إجراء تحويرات عليها مثل التدوير، التصغير، التكبير والتزحيف. التمييز باستخدام الشبكات العصبية الاصطناعية للمجموعة الأولى أعطت نسبة (94%) و (83%) باستخدام الخصائص و الجزء ذات الطاقة العالية للشرائح على الترتيب والتي هي أفضل مقارنة بالتمييز بطريقة قياس المسافة الأقل للتشابه. أما بالنسبة للتمييز باستخدام الشبكات العصبية الاصطناعية للمجموعة الثانية فقد أعطت نسبة (94%) و (72%) باستخدام الخصائص والجزء ذات الطاقة العالية للشرائح على الترتيب والتي هي أفضل مقارنة بالتمييز بطريقة قياس المسافة الأقل للتشابه والتي أعطت نسبة (11%) و (33%) على الترتيب. كذلك الوقت المستغرق خلال عملية التمييز باستخدام الشبكات العصبية الاصطناعية أقل مما في طريقة قياس المسافة الأقل للتشابه .

KEYWORD: Object recognition, feature extraction, patches, multi-wavelet, neural network.

Dr. Tariq Zeyad Ismail
Zainab Ibrahim Abood

INTRODUCTION

Object recognition is at the top of a visual task hierarchy. In its general form, this is a very difficult computational problem, which will probably play an important role in the eventual building of intelligent machines. A large number of psychological and neurophysiologic studies support the idea that humans represent three-dimensional objects internally as a small set of bi-dimensional images [R. Cesar, 2005]. View-based method has been proposed by some researchers in which the object is described using a set of 2-D characteristic views or aspects. Main disadvantage of this method is the inherent loss of information in the projection from 3D object to 2D image. A single 2D view-based approach may not be appropriate for 3-D object recognition since only one side of an object can be seen from any given viewpoint. [M. Y. Mashor, 2004]. A better alternative is to obtain the features from several 2-D views from a few static cameras as suggested in this proposed approach. An effective representation of 3-D object properties using 2-D images is considered. With multiple views technique enables this technique to be used in 3-D object modeling.

It is a classic difficult problem for a computer to recognize images that is because a computer lacks ability of adaptive learning. The inductive processes embody the universal and efficient means for extracting and encoding the relevant information from the environment, the evolution of intelligence could be seen as a result of interactions of such a learning mechanism with the environment. In consensus with this, any one strongly believe that the pivot of image recognition should be arranged around learning processes at all levels of feature extraction and object recognition [Y. Min, 2005].

THE PROPOSED ALGORITHM

The proposed algorithm illustrated in the following steps:

1. Input all images of all views and I is the number of image.
249 image for generating a training set.
 68 image for generating a test set.
2. Preprocess these images by filtering them using median filter.
3. Edge detection using canny edge detector and select patches to get features from them.
4. Two methods of feature extraction from the patches are used:
   a. (21) features about patches shape and location.
   b. High energy subband results from decompose each patch of each image by using
      DMWT.
5. Store the features of the training set in the data base and the others stored as a test set in   order to be ready to inter to the image recognition stage.
6. Recognition stage contains two methods of recognition, minimum distance and neural network.
Figure (1) shows the block diagram of generation training and testing sets.

Figure (2) shows the flow chart of overall proposed system.

 The block diagram of the Image recognition stage is shown in figure (3).

DATA BASE

The model that is stored in the memory as a data base (reference images) consists of (249) image, each of size (449×267), the (gp) patches and features of each patch in each image in each model. These images are divided into (4) sets. The sets contain car model 1, car model 2, airplane model 1 and airplane model 2. They are named as: c1, c2, a1 and a2 respectively. There are (3) views, izo, side and top view (with rotation) for each model which are named as: i, s and t respectively. The images are named according to the set which they are belong to, their view and their number in this view, i.e. c1s018 means that it is belong to (c1) set , side view and number of image in this view

is 18. Figure (4) shows samples of data base images.

In the test set there are two groups which are group (1) and group (2) each of (34) image. Group (1) means that the test set contains images, their (gp) patches and patches features of the same images as a part of that in the data set beside other images, (gp) patches and features, While group (2) means that the test sets contains the (gp) patches and patches features the same as a part of that in the database but after modification such as rotation, scaling and translation. Figure (5) shows samples of test set images.

FEATURE EXTRACTION

Two methods of feature extraction are used. The first method is the extraction of (21) features from the patches which are represented something about their locations and shapes, it is used in order to compare it with the second method of feature extraction which use the high energy subband results from decompose each patch of each image by using DMWT.

Each patch is a part of object so all rules (in this work) for the object in image is the same as for the patch, i.e. these features are extracted from the patches.

The location of the patch can be determined by calculating the coordinates of the centroid and area of each patch as in the following steps:

*Area of object: The object area given by:

$$A = \sum_i \sum_j f[i,j] \qquad (1)$$

where $f[i,j]$ represents the object pixels, The area is thus computed as the total number of object pixels in the object [S. E. Umbaugh, 1995].

*Location of object: The location of the object is usually given by the center of mass which is given by:

$$\bar{i} = \frac{\sum_i \sum_j [i f(i,j)]}{A} \;,\; \bar{j} = \frac{\sum_i \sum_j [j f(i,j)]}{A} \qquad (2)$$

Where $\bar{i}$ and $\bar{j}$ are the coordinates of the centroid of the object and A is the area of the object [T. Acharya, 2005].

Other features are:
* Orientation of an Object: When the objects have elongated shape, the axis of elongation is the orientation of the object. The axis of elongation is a straight line so that the sum of the squared distances of all the object points from this straight line is minimum. The distance here implies the perpendicular distance from the object point to the line [Y. Amit, 2002]. The axis corresponds to the about which it takes the least amount the energy to spin an object of like shape or the axis of least inertia. If the origin was moved to the center of area (i,j) and θ is the angle between the x-axis and the axis of least second moment counterclockwise, then the axis of least second moment ($\tan(2\theta)$) will be defined as follows [S. E. Umbaugh, 1995]:

$$\tan(2\theta) = 2 \frac{\sum_{i=0}^{N-1}\sum_{j=0}^{N-1} ij f(i,j)}{\sum_{i=0}^{N-1}\sum_{j=0}^{N-1} i^2 f(i,j) - \sum_{i=0}^{N-1}\sum_{j=0}^{N-1} j^2 f(i,j)}$$

$$(3)$$

• **Euler number of an image:** It is defined as the number of object minus the number of holes. For a single object, it tells that how many closed curves the object contains [S. E. Umbaugh, 1995].
• **Projection of an object onto a line:** The projections of an image provide good information about the image. The projections may be computed along horizontal, vertical, or diagonallines. The horizontal projection is obtained by counting the number of object pixels in each column of the image, while the total number of

object pixels in each row yields the vertical projection as follows [T. Acharya, 2005]:

$$P_{hor}(i) = \sum_{j} f(i,j) \qquad (4)$$

$$P_{ver}(j) = \sum_{i} f(i,j) \qquad (5)$$

INVARIANT MOMENTS

M.K.Hu represented the concept of the invariant moments in 1961 firstly. The invariant moments are the highly compressed image features, which meet the invariability of the translation, the ratio and the rotation to the continuous function [Z. Song, 2007].

Basic Theory

i. For the digital image, the discrete invariant moments are used, the geometric moments m_{pq} of the $(p+q)^{th}$ order (p and q are the arbitrary non-negative integer respectively) are:

$$m_{pq} = \sum_{i} \sum_{j} i^p j^q f(i,j) \qquad (6)$$

where $p, q = 0,1,2,...$ and $f(i,j)$ is the function of the image value, i and j are the image coordinates respectively.

ii. Because of the translation invariability of m_{pq}, the central moments of the $(p+q)^{th}$ order are [R. C. Gonzalez, 2002, Z. Song, 2007]:

$$\mu_{pq} = \sum_{i} \sum_{j} (i - \bar{i})^p (j - \bar{j})^q f(i,j) \text{ where}$$

$$\bar{i} = \frac{m_{10}}{m_{00}} \quad , \quad \bar{j} = \frac{m_{01}}{m_{00}} \quad \text{and } \mu_{01} = \mu_{10} = 0$$

$$(7)$$

where $\bar{i}$ and $\bar{j}$ are the image center coordinates [Z. Song, 2007].

The normalized central moments shown below will add scale invariance [R. C. Gonzalez, 2002]:

$$\eta_{pq} = \frac{\mu_{pq}}{\mu^r_{00}} \qquad \text{where} \quad r = \frac{p+q}{2} + 1$$

for $\quad p + q = 2,3,... \qquad (8)$

Therefore Hu made seven invariant moments [Z. Song, 2007].

$$\phi_1 = \eta_{20} + \eta_{02}$$
$$\phi_2 = (\eta_{20} - \eta_{02})^2 + 4\eta_{11}^2$$
$$\phi_3 = (\eta_{30} - 3\eta_{12})^2 + (3\eta_{21} - \eta_{03})^2$$
$$\phi_4 = (\eta_{30} + \eta_{12})^2 + (\eta_{21} + \eta_{03})^2$$
$$\phi_5 = (\eta_{30} - 3\eta_{12})(\eta_{30} + \eta_{12})[(\eta_{30} + \eta_{12})^2 - 3(\eta_{21} + \eta_{03})^2] + (3\eta_{21} - \eta_{03})(\eta_{21} + \eta_{03})$$
$$[3(\eta_{30} + \eta_{12})^2 - (\eta_{21} + \eta_{03})^2]$$
$$\phi_6 = (\eta_{20} - \eta_{02})[(\eta_{30} + \eta_{12})^2 - (\eta_{21} + \eta_{03})^2] + 4\eta_{11}(\eta_{30} + \eta_{12})(\eta_{21} + \eta_{03})$$
$$\phi_7 = (3\eta_{21} - \eta_{03})(\eta_{30} + \eta_{12})[(\eta_{30} + \eta_{12})^2 - 3(\eta_{21} + \eta_{03})^2] - (\eta_{30} - 3\eta_{12})(\eta_{21} + \eta_{03})$$
$$[3(\eta_{30} + \eta_{12})^2 - (\eta_{21} + \eta_{03})^2]$$

$$(9)$$

Invariant Moment's Expansion

The actual invariant moments are:

$$I_k = Log_{10}|\phi_k| \qquad (10)$$

where $\quad k = 1,2,3,...,7$

It is supposed that two images $f_1(i,j)$ and $f_2(i',j')$ have difference in the contrast, the ratio, the translation and the rotation, but their content is same. In order to obtain more general discrete invariant moments, their mutual relationships can be expressed using the following equation:

$$f_2(i',j') = kf_1(i,j) \qquad \text{so}$$

$$\begin{bmatrix} i' \\ j' \end{bmatrix} = c\begin{bmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{bmatrix}\begin{bmatrix} i \\ j \end{bmatrix} + \begin{bmatrix} t_i \\ t_j \end{bmatrix} \qquad (11)$$

where k is the contrast factor; c is the ratio factor; θ the rotation angle; and (t_i, t_j) are the displacement in the i and j direction respectively.

The more general discrete invariant moments can be taken using equations (9), (10) and (11):

$$\beta_1 = \frac{\sqrt{I_2}}{I_1} \quad \beta_2 = \frac{I_3}{I_2 I_1} \quad \beta_3 = \frac{I_4}{I_3} \quad \beta_4 = \frac{\sqrt{I_5}}{I_4}$$

$$\beta_5 = \frac{I_6}{I_4 I_1} \qquad \beta_6 = \frac{I_7}{I_5} \qquad (12)$$

where $I_1, I_2, \dots I_7$ are actual invariant moments [Z. Song, 2007].

MULTI-WAVELET TRANSFORM

Wavelet transforms provide both spatial information about the image and also frequency information [R. William Ross, 1999]. The resulting from the wavelet transform is a set of two signals, each half the length of the original. The overall effect of the lowpass filter is a lower resolution representation of the original signal scaled by some factor. The high-pass filter leaves behind only the high frequency components. Multi-resolution analysis is accomplished by continuing the process on the result of the low-pass filter [H. Chung, 2002].

Until 1999, only wavelets were known. These are wavelets generated by one scaling function. But one can imagine a situation where there is more than one scaling function. This leads to the notation of multi-wavelets [M. Alfaouri, 2008] which are use several scaling functions and mother wavelets [H. Soltanian - Zadeha,2004].

Motivation of Multi-wavelets

Using several scaling functions and mother wavelets adds several degrees of freedom in multi-wavelet design and makes it possible to have several useful properties such as symmetry, orthogonality short support, and a higher number of vanishing moments simultaneously. The usefulness of these properties is well known in wavelet design. Symmetric property allows symmetric extension when dealing with the image boundaries, this prevents discontinuity at the boundaries and therefore a loss of information in these points would be prevented. Orthogonality generates independent sub-images. A higher number of vanishing moments result in a system capable of representing high-degree polynomials with a small number of terms [H. Soltanian - Zadeha, 2004].

Computing discrete multi-wavelet transform, scalar wavelet transform can be written as follows [M. Alfaouri, 2008]:

$$\begin{bmatrix} H_0 & H_1 & H_2 & H_3 & 0 & 0 \\ G_0 & G_1 & G_2 & G_3 & 0 & 0 \\ 0 & 0 & H_0 & H_1 & H_2 & H_3 \\ 0 & 0 & G_0 & G_1 & G_2 & G_3 \end{bmatrix} \qquad (13)$$

where H_i and G_i are low and high pass filter impulse responses, are 2-by-2 matrices which can be written as follows [M. Alfaouri, 2008]:

$$\begin{bmatrix} H_0 & H_0 & H_1 & H_1 & \dots & \dots & \dots & \dots & \dots \\ H_0 & H_0 & H_1 & H_1 & \dots & \dots & \dots & \dots & \dots \\ G_0 & G_0 & G_1 & G_1 & \dots & \dots & \dots & \dots & \dots \\ G_0 & G_0 & G_1 & G_1 & \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & 0 & H_0 & H_0 & H_1 & H_1 & \dots \\ 0 & 0 & 0 & 0 & H_0 & H_0 & H_1 & H_1 & \dots \\ 0 & 0 & 0 & 0 & G_0 & G_0 & G_1 & G_1 & \dots \\ 0 & 0 & 0 & 0 & G_0 & G_0 & G_1 & G_1 & \dots \end{bmatrix} \qquad (14)$$

For computing discrete multi-wavelet transform, scalar wavelet transform matrix must be used as in eq. (13) where a system with H_k for GHM four scaling matrices defined as follows [M. Alfaouri, 2008]:

$$H_0 = \begin{bmatrix} 3/5\sqrt{2} & 4/5 \\ -1/20 & -3/10\sqrt{2} \end{bmatrix}$$

$$H_1 = \begin{bmatrix} 3/5\sqrt{2} & 0 \\ 9/20 & 1/\sqrt{2} \end{bmatrix}$$

$$H_2 = \begin{bmatrix} 0 & 0 \\ 9/20 & -3/10\sqrt{2} \end{bmatrix}$$

$$H_3 = \begin{bmatrix} 0 & 0 \\ -1/20 & 0 \end{bmatrix}$$

(15)

And a system with G_k for GHM four scaling matrices defined as follows [M. Alfaouri, 2008]:

$$G_0 = \begin{bmatrix} -1/20 & -3/10\sqrt{2} \\ 1/10\sqrt{2} & 3/10 \end{bmatrix}$$

$$G_1 = \begin{bmatrix} 9/20 & -1/\sqrt{2} \\ -9/10\sqrt{2} & 0 \end{bmatrix}$$

$$G_2 = \begin{bmatrix} 9/20 & -3/10\sqrt{2} \\ 9/10\sqrt{2} & -3/10 \end{bmatrix}$$

$$G_3 = \begin{bmatrix} -1/20 & 0 \\ -1/10\sqrt{2} & 0 \end{bmatrix}$$

(16)

Computation of 2-D DMWT Algorithm

Repeated row preprocessing (Over-sampling scheme) is used here [Sudhakar. R, 2006], so, for computing a single-level 2-D multi-wavelet transform the next steps should be followed:

1. Checking input dimensions: Input matrix (patch matrix) should be a square matrix of length $N \times N$, where N must be power of two. If the patch is not a square matrix some operation must be done to the patch like resizing the patch or adding rows or column of zeros to get a square matrix.

2. Constructing a transformation matrix: An $N \times N$ transformation matrix should be constructed using GHM low and high pass filters matrices given in eq.'s (15) and (16). The transformation matrix can be written as eq. (13). After substituting GHM matrix filter coefficients values, a $2N \times 2N$ transformation matrix results with the same

dimensions as the input patch matrix dimensions after preprocessing will be obtained.

3. Preprocessing rows: Row preprocessing doubles the number of the input matrix rows. So if the 2-D input is $N \times N$ matrix elements, after row preprocessing the result is $2N \times N$ matrix. The odd rows 1, 3… 2N-1 of this resultant matrix are the same original matrix rows values 1, 2, 3…, N respectively. While the even rows numbers 2, 4…2N are the original rows values multiplied by α, for GHM system functions $\alpha = 1/\sqrt{2}$.

4. Transformation of input rows: can be done by

a. Apply matrix multiplication to the $2N \times 2N$ constructed transformation matrix by the $2N \times N$ preprocessing input matrix.

b. Permute the resulting 2N x 2N matrix rows by arranging the row pairs 1,2 and 5, 6, …., 2N-3, 2N-2 after each other at the upper half of the resulting matrix rows. Then arrange the row pairs 3, 4 and 7, 8, …, 2N-1, 2N below them at the next lower half.

5. Preprocess columns: It can be done by repeating the same procedure used in preprocessing rows:

a. Transpose the $2N \times N$ transformed matrix from step **(4)**.

b. Repeat step **(3)** to the $N \times 2N$ matrix which results in $2N \times 2N$ column preprocessed matrix.

6. Transformation of input columns:

a. Apply matrix multiplication to the 2Nx2N constructed transformation matrix by the 2Nx2N column preprocessed matrix.

b. Permute the resulting 2Nx2N matrix rows by arranging the row pairs 1,2 and 5,6…,2N-3,2N-2 after each other at the upper half of the resulting matrix rows. Then arrange the row pairs 3,4 and 7,8,…,2N-1,2N below them at the next lower half.

7. To get the final transformed matrix the following should be applied:

a. Transpose the resulting matrix from column transformation step.

b. Apply coefficients permutation to the resulting transpose matrix. Coefficient permutation is applied to each of the basic

four subbands of the resulting transpose matrix so that each subband permutes rows then permutes columns.

Finally, a 2N×2N DMWT matrix results from the N×N original patch matrix by using repeated row preprocessing[M. Alfaouri, 2008].

The results of implementing this algorithm is shown in figure (6).

The normalized energy for the DMWT subband is computed and the high energy subband (L1L1) will be taken as a feature and it will be known as (gp) i.e. ghm patch to refer to the patch after transformation by 2-D multi-wavelet transform.

CLASSIFICATION

After generating training and test sets, they should be stored as a database to be used later for testing and evaluation. If a complete set of discriminatory features for each pattern class can be found, classification can be reduced to a simple matching process. However, this assumption is really too quixotic to be achieved in practical pattern recognition problems.

Therefore, only some, or the best discriminatory features are usually adopted. As to classification, its aim is similar to that of feature extraction, which is to find the best class that is the closest to the classified pattern [Y. Kai Wang, 1996].

RECOGNITION METHODS

After the extraction of 21 features from each patch and the high energy subband of each patch in the other side, the minimum distance[T, Zeyad, 2001] and neural network[M, Kantardzic, 2003] methods are used.

In a Minimum distance method, when dealing with a one dimensional vector with more than one element, Euclidean distance is a good measurement for the difference between the two vectors. In this search, the two vectors are two patches vectors which are of the same or different size or rotation angle, translation distance or location (one from the test set and the other from the

training set), i.e. the patches features vectors of the training set and the test set and then the (gp) patches of the training set and the test set. So if the difference is 0, it is surely the best match.

A neural network trained to perform a particular function by adjusting the values of the connections (weights) between elements. Commonly neural networks are trained, so that a particular input leads to a specific target output. There, the network is adjusted, based on a comparison of the output and the target then the error is calculated and the result is fed-back from output layer and the weights are adjusted. These steps represented for all the inputs in the training set and each time the weights are adjusted. The training continues until the mean square error value between the values of the output and the target reaches. Then this net will be used to train an unknown input image that is wanted to recognize.

In this search the parameters of NN training are:

* Performance function is MSE.
* No. of hidden layers is 2 layers, the activation functions used are tan sigmoid in the first hidden layer and purelin in the second hidden layer.
* Epoch 1000 iterations (maximum number of epoch to train)
* Gradient is 1.00e-10

Neural network training is shown in Figure (7).

TESTING AND EVALUATION OF RESULTS

This example will be represented for testing and evaluation the proposed algorithms.

1. Enter the test image c2i03, figure (8).
2. Preprocess the image using median filters.
3. Apply canny edge detection and patch selection and then (21) features are extracted from each patch of this image.

4. Decompose each patch of this image by using DMWT.
5. Extract the high energy subband of each patch of this image.
6. Matching by minimum distance and (BP) neural network respectively.

Recognize the image when the result is for the same image (using above two methods), so it is labeled as true **(T)** but when it is **not recognized**, the result is wrong and it is labeled as false **(F)**.

As shown in the figure, recognize the image except for the gp patches in group (1) and (2) in the matching by minimum distance which is wrong, i.e. recognize the image for features in matching by minimum distance, while it is recognized for features and gp patches in group (1) and (2) in the recognition by neural network.

The results of implementing the algorithms of matching by minimum distance and BP neural network are shown in table (1), where **recognize** the image is labeled as true **(T)** and when it is **not recognized**, the result is wrong and it is labeled as false **(F)**, i.e. when matching the patch features of the test image such as (c1i01) with the patch features of the data base images by using minimum distance, the result is (T) because it is recognized, but when matching the (gp) patches of the same test image (c1i01) with the (gp) patches of the data base images the result is (F) because it is not recognized. while, when matching the patch features of the same test image (c1i01) with the patch features of the data base images by using neural network, the result is (T) because it is recognized and when use the (gp) patches

the result is also (T) because it is recognized. For (a1i09), when matching the patch features using minimum distance, the result is (T) because it is recognized, but when matching the (gp) patches the result is (F) because it is not recognized. While, when matching the patch features using neural network, the result is (F) because it is not recognized and when use the (gp) patches the result is (T) because

it is recognized. For (a2i07), when matching the patch features or (gp) patches using minimum distance, the result is (F) because it is not recognized. While, when matching the patch features using neural network, the result is (T) because it is recognized and when use the (gp) patches the result is (F) because it is not recognized, and so on….

Conclusions

From the above simulation one can be concluded that the proposed techniques are much better performance in comparison with minimum distance for the group (1) or group (2). Recognition by back propagation (BP) neural network as compared with matching by minimum distance, gave (94%) and (83%) score by using group (1), (gp) and features respectively, which is much better than the minimum distance. Recognition using (gp) neural network (NN) gave (94%) and (72%) score by using group (2), (gp) and features respectively, while the minimum distance gave (11%) and (33%) score. Using multi-wavelet transform to gain better feature extractor to each patch and high energy subband of multi-wavelet transform of the patch gave high recognition score than patches features. Time consumption through the recognition process using (NN) with (gp) is less than that when using the minimum distance.

85

References

H. Chung, K. yu Lai, A. Lip, and J. Yip, "Selective Image Compression of Passport Photographs Using Wavelets and Neural Networks", phD Thesis, University of Manitoba, 2002, pp 16.

H. Soltanian - Zadeha, F. Rafiee - Rad, S. Pourabdollah - Nejad, "Comparison of multi-wavelet, Wavelet, Haralick, and Shape Features for micro calcification Classification in Mammograms", Journal, Pattern recognition 37, (2004), E-mail: hamids@rad.hfh.edu

M. Alfaouri, "Image Recognition using Combination of Discrete Multi-wavelet and Wavenet Transform", American journal of applied sciences 5 (4): 418-426, 2008. http://www.scipub.org/fulltext/ajas/ajas5441 8-426.pdf

M. Kantardzic, "Data Mining: Concepts, Models, Methods, and Algorithms", John Wiley and Sons© 2003, ch9.

M. Y. Mashor, M. K. Osman and M. R. Arshad, "3D Object Recognition using 2D Moments and HMLP Network", Proceedings of the International Conference on Computer Graphics, Imaging and Visualization (CGIV'04), IEEE, 0-7695-2178-9/04, 2004.

R. Cesar, A. Rosa, "**3D Object Prototype-Based Recognition from 2D Images using Features net**", 2005, pp. 155-162. http://portal.acm.org/citation.cfm?id=13698 00

R. C. Gonzalez, and R. E. Woods, "Digital Image Processing" Second Edition, 2002, pp. 24, 29, 77, 78, pp. 673-674, Publisher: Tom Robbins.

R. William Ross, "Target Recognition Using Neural Networks for Wind Tunnel Model Deformation Measurements", Msc. Thesis, Christopher Newport University, 1999, pp28.

S. E. Umbaugh, "Computer Vision and Image Processing: A Practical Approach Using CVIP tools", phD, Book, 1995, ch2 (pp61, pp133, pp136).

Sudhakar. R, and Jayaraman. S, "Fingerprint Compression Using Multiwavelets", pp79-80, International Journal of Signal Processing 2; 2 © www.waset.org Spring 2006, pp78-87.

T. Acharya and A. K. Ray, "Image Processing, Principles and Applications**",** 2005, pp84, pp49-52, Published by John Wiley & Sons, Inc., Hoboken, New Jersey.

T. Zeyad, "Recognition of 3-D Objects", PhD Thesis, Baghdad university, 2001.

Y. Amit, "2D Object Detection and Recognition", Models, Algorithms and Networks, 2002, pp12, 13, 190, MIT Press, Cambridge, Massachusetts, London, England.

Y. Kai Wang' and K. Fan', "Applying Genetic Algorithms on Pattern Recognition: An Analysis and Survey", IEEE, 1996, P740. E-mail: ykwang@iis.sinica.edu.tw , kcfan@csie.ncu.edu.tw

Y. Min', J. Zhtwei', Y. Wensheng', and Z. Xiaornin, "Applications of Generalized Learning in Image Recognition", IEEE, First International Conference on Neural Interface and Control Proceedings, 26-28 May 2005.

Z. Song, B. Zhao, Z. Zhu, E. Mao "Research on Traffic Number Recognition Based on Neural Network and Invariant Moments", Proceedings of the Sixth International Conference on Machine Learning and Cybernetics, Hong Kong, 19-22 August 2007, China Agricultural University, E-mail: songzhenghe@cau.edu.cnzb1981@sina.com .

List of Abbreviations

NN:	Neural network
BP:	Back-propagation
DMWT:	Discrete multi-wavelet transforms
GHM:	Filter proposed by Geronimo, Hardian, Masopute
gp:	High energy subband multi-wavelet transform (ghm patch)
HH :	High-high
HL :	High-low
LH :	Low-high
LL :	Low-low
2-D:	Two dimensional
3-D:	Three dimensional

Input reference

Preprocessing

Edge detection
and patch
selection

Patch feature
extraction

Patch 2-D DMWT
coefficients

Training DMWT
coefficients and
features

Training set

(a)

Input test

Preprocessing

Edge detection
and patch
selection

Patch feature
extraction

Patch 2-D DMWT
coefficients

Testing DMWT
coefficients and
features

Test set

(b)
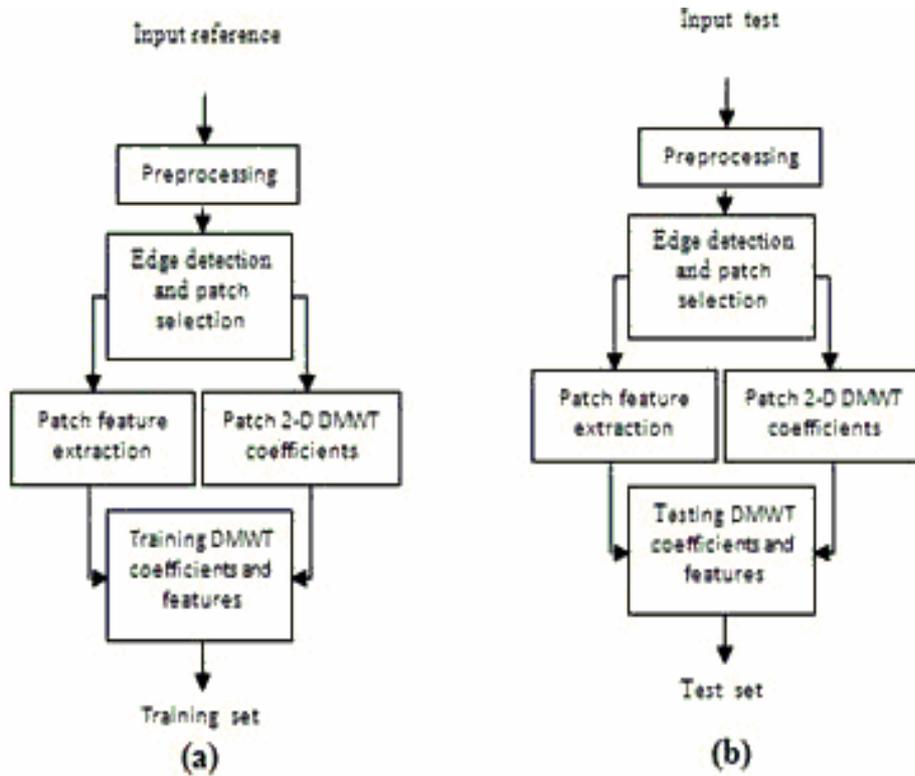
Fig. (1): Block diagrams of: a. Generation training set   b. Generation testing set
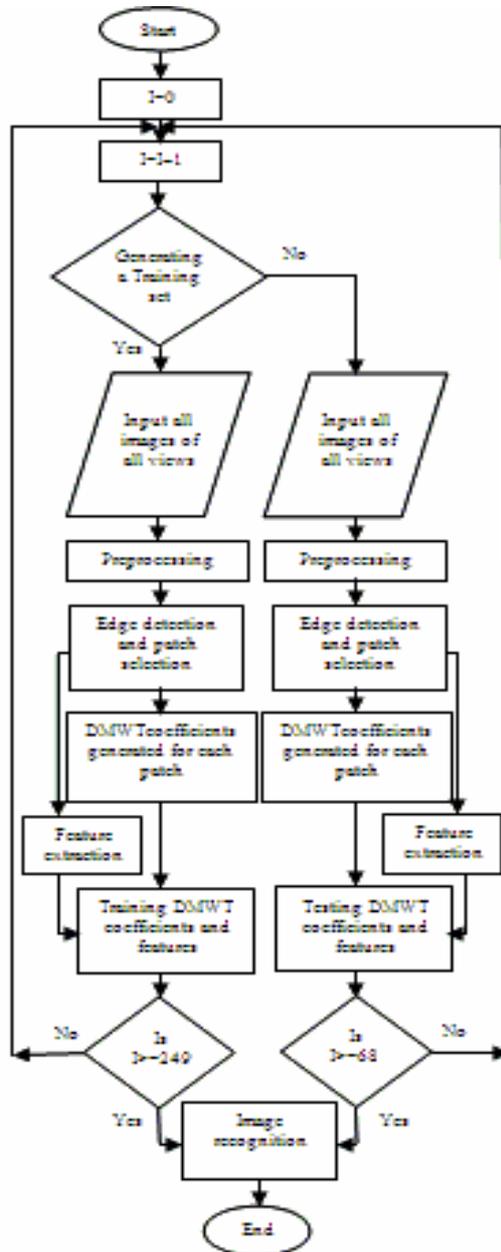
Fig. (2): Flow chart of overall proposed
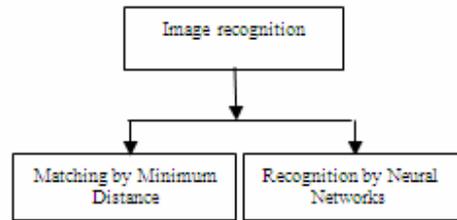system

Fig. (3): Block diagram of Image recognition stage
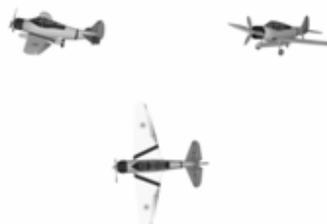
Set (c1)

Set (c2):

Set (a1):

Set (a2):

Fig. (4): Samples of database images
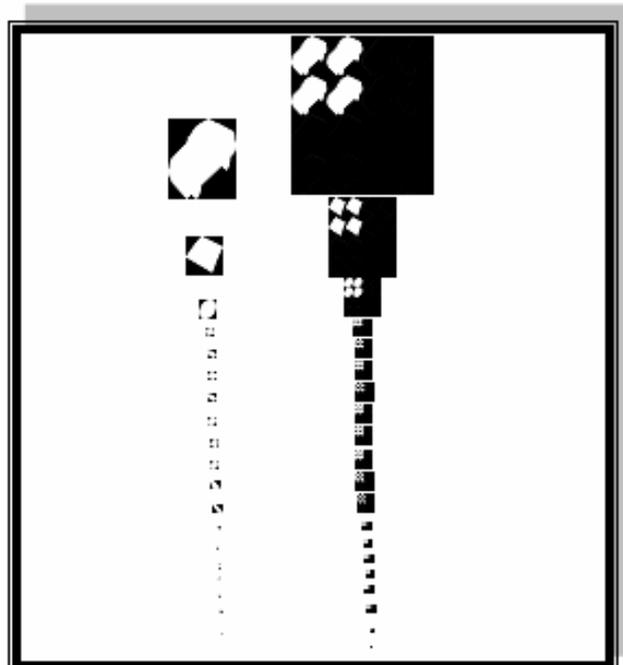
Fig. (5): Samples of test set images.

Fig. (6): DMWT implementation of izo view car
(1) Patches after conversion to square and power of 2
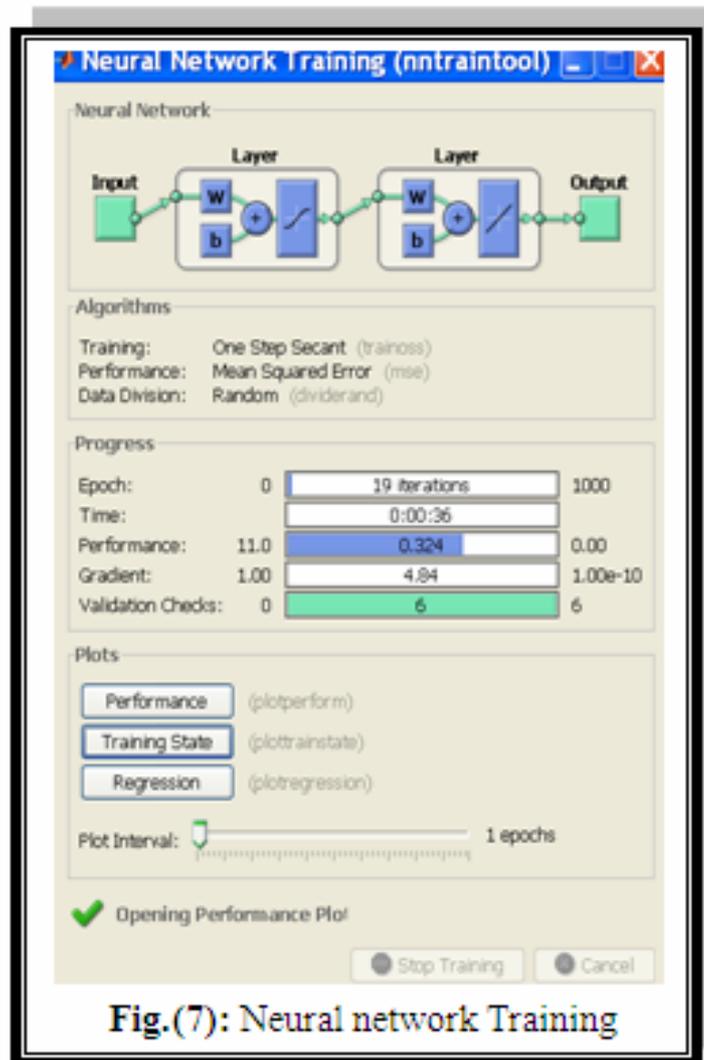(2) Patches after applying DMWT

Fig.(7): Neural network Training

Test image

Matching by **minimum distance:**

features   gp

Group(1)

T   F

Recognition by **Neural network:**

features   gp

Group(1)

T   T

Matching by **minimum distance:**

features   gp

Group (2)

T   F

Recognition by **Neural network**

features   gp

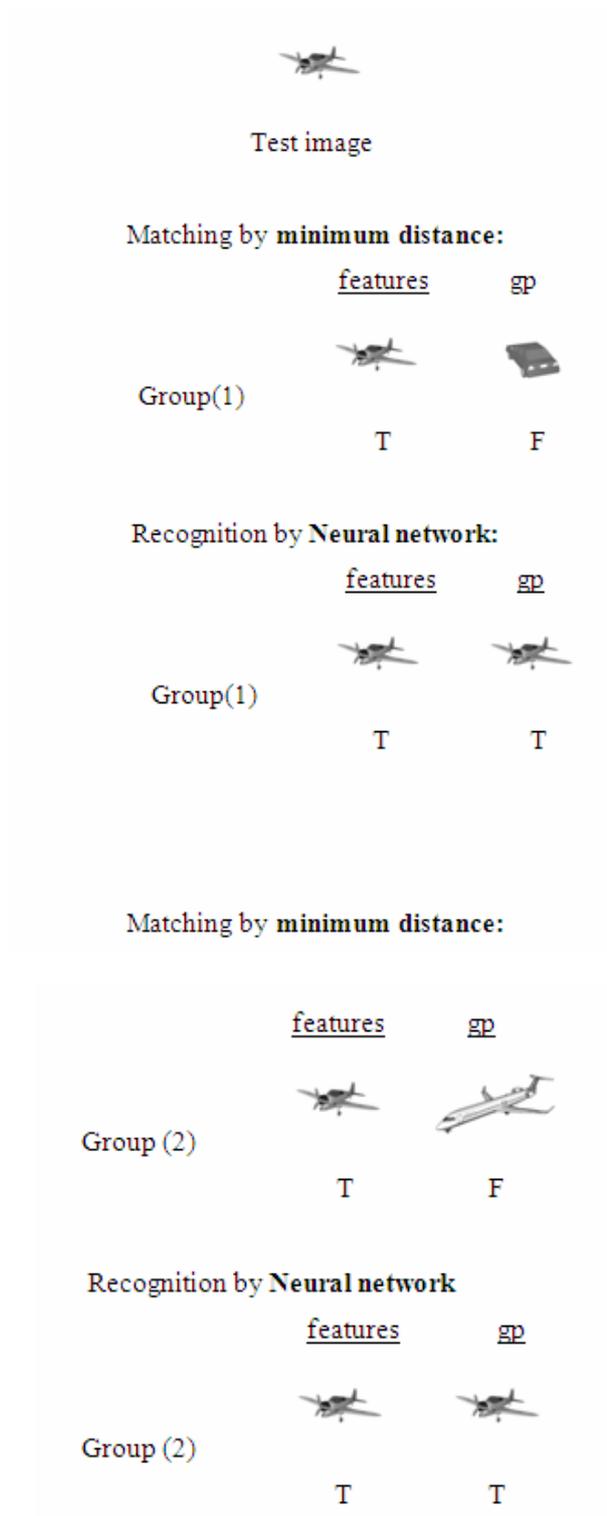
Group (2)

T   T

Fig. (8) Evaluation and testing

Table (1): Matching by minimum distance and BP neural network

	Min. dist. Using feat--ures	Min. dist. Using (gp) Patc--hes	NN. using Feat--ures	NN. using (gp) patc--hes
c1i01	T	F	T	T
c1i08	T	F	T	T
c1s01	T	F	T	T
c1s016	T	F	T	T
c1t014	T	F	T	T
c2i03	T	F	T	T
c2s01	T	F	T	T
c2s019	T	F	T	T
c2t015	T	F	T	T
a1i09	T	F	F	T
a1s011	T	F	F	T
a1t03	T	F	T	T
a1t05	T	F	F	T
a2i03	T	F	T	T
a2i07	F	F	T	F
a2s04	T	F	T	T
a2s05	T	F	T	T
a2t017	T	F	T	T