

**Offline Handwriting Arabic Words features
Extractions and classifications using Hybrid
Transform and Self-Organizing Feature Map**

Dr Anmar Abdul Wahab

Al-Turath University College

Dept. Of Computer Science

ABSTRACT

In this paper images of Arabic handwritten words have been introduced to sequences of transforms, to obtain the final features of the given image. The hybrid transform has been used which transforms images of Arabic handwritten words into another form such that it is partially invariant to scale and rotation. The transforms used in this research, such as 2-D Fast Fourier Transform, Radon Transform, 1-D Inverse Fast Fourier Transform and 1-D Discrete Multiwavelete Transform have been considered for feature extraction. In the hybrid transform, images of the Arabic handwritten words have been introduced to the sequences of transforms, to obtain the final coefficients matrix of the given image. The obtained transform coefficients can be as affine invariant pattern features. The experiments showed that a small subset of these coefficients is enough for reliable recognition of complex patterns. A Comprehensive Database of Handwritten Arabic Words [1] had been used in the implementation of the hybrid transform. Then Kohonen Self Organizing Feature Map (SOFM) for pattern clustering has been used to cluster the discriminated information vectors extracted from the hybrid transforms matrix's coefficients. The variations of the transforms used in order to improve generalization, and perform with 89% accuracy on a 7-class lexicon. The

technique suggested in this paper can serve as a pre-processing step in computer vision applications.

Author Keywords: Offline Arabic handwritten recognition, Two Dimensional Fast Fourier Transform, Radon transform, One Dimensional Inverse Fast Fourier Transform, One Dimensional Discrete Multiwavelet Transform, and Self Organizing Feature Map.

الخلاصة

في هذا البحث تم استخدام الكلمات العربية المكتوبة بخط اليد في متوايلات من التحويلات وذلك للحصول على الخصائص النهائية لكلمة عربية معينة. التحويل الهجين الذي استخدم ليقوم بتحويل الكلمات العربية المكتوبة بخط اليد إلى شكل آخر يكون غير متأثر جزئياً بالتدوير والتقييس. التحويلات في هذا البحث، مثل تحويل 2-D فورييه السريع، تحويل الرادون، تحويل 1-D معكوس فورييه السريع، و تحويل 1-D المتقطع للموجات المتعددة استخدمت لاستخراج الخصائص الخاصة بالكلمات العربية المكتوبة بخط اليد. في التحويل الهجين، أدخلت صور من الكلمات العربية المكتوبة بخط اليد إلى متوايلات من التحويلات، للحصول على معاملات المصفوفة النهائية لصورة معينة. إن المعاملات التي تم الحصول عليها ذات ميزات نمطية ثابتة. وأظهرت التجارب أن مجموعة صغيرة من هذه المعاملات تكون كافية في عملية تعرف موثوقة للأسماء المعددة منها. وقد استخدمت قاعدة بيانات شاملة من الكلمات العربية المكتوبة بخط اليد [1] في تفزيذ التحويل الهجين. ثم استخدمت طريقة خريطة التنظيم الذاتي للخصائص لكتلتين خاصة بفصل المجموعات النمطية وذلك لفصل الخصائص المميزة والمتتشابهة لكل نمط والمستخرج من مصفوفة المعاملات الخاصة بالتحويل الهجين. الأشكال المختلفة من التحويلات تستخدم من أجل تحسين التعلم، مع نتائج أداء مقداره 89%. التقنية المقترنة في هذا البحث يمكن أن تكون بمثابة خطوة في المعالجة التمهيدية لتطبيقات الحاسوب القائمة على فهم محتوى الصور.

1-Introduction

Recently commercial products such as computers, mobiles, tablet PC, PADs and others incorporating the process of handwriting recognition to be replaced with the traditional keyboard and mouse. Handwriting recognition is defined as “the ability of a computer to receive and interpret intelligible handwritten input from sources such as paper documents, photographs, touch-screens and other devices”[2]. In principle there are two ways to recognize handwritten words, online and offline, the online method considered to be simpler than the offline

method because its depend on previously collected information regarding the tracking of the movement of the tip of the pen, while the offline method attempted to recognize the images of the actual words after long process of scanning the document , scanned image preprocessing, words segmentation, character extraction, feature extraction , classifications and recognition.

The Preprocessing usually consists of binarization, normalization, sampling, smoothing and denoising, for character extraction, segmentation methods may be used to separate the characters in the words, detection of lines of text, detection of connected components. Feature extraction methods used to extract most distinguishable features in the image to be used as a measure of similarity between objects of similar group. For the classification and segmentation process in which similar group of objects (words) can be classified or clustered in the same class, Kohonen Self Organizing Feature Map (SOFM) for pattern clustering has been used to cluster the discriminated information vectors extracted from the hybrid transforms matrix's coefficients.

Arabic language consists of 28 characters, words are written from right to left in horizontal lines. Each character has more than one form depending on its position in the word [3]. The Arabic handwritten words are cursive, where the characters are joined together, this make the task of separating the characters more difficult task.

In this research the attempt is made to overcome some of the above mentioned difficulties in which face the task of recognition, such the problem of partial rotation and scaling of the handwritten words, the characters segmentation and separation problem. The suggested approach is to transform the image from the special domain into the frequency domain using a sequence of transforms to extract affine invariant features to be used in the process of handwritten recognition process.

In general transforms provide information regarding the spatial frequency content of an image. The transform maps image data into a different mathematical space via transformation equations. It maps the image data from the spatial domain to the frequency domain (also called the spectral domain). These transforms are used as tools in computer imaging. The goal in pattern recognition is to extract information useful for the recognition task and this is done by reducing the amount of image data with the transformations. The images then are modified from the lowest level of pixel data into higher-level representation. Now the feature extraction is considered to solve the recognition problem. The image transforms provides features based on spatial frequency information-spectral features. The method suggested here applies affine invariant feature extraction to the images; each image is described by a subset of hybrid transform Coefficients [4].

These coefficients characterize the image features which allow form compact and meaningful feature vector. Then Self-Organizing Feature Map neural network [5] is used to arrange the feature vectors on the map according to their topological relationship order. Non-normalized images causes distortion in which plays an important role in determining the end result of pattern recognition process (face recognition, fingerprint matching, signature matching etc.).

The technique suggested in this paper can serve as a pre-processing step in computer vision applications, without the need to perform the tedious work for image normalization (scaling, rotation and alignment). The focus here is on a particular method for handling in variances, which has the capability of extracting all of the invariant features from an image using only a small amount of information about the image (such as a few low of low coefficients). The invariant features extraction separate the problem of finding the image viewing transformations, from the problem of deciding which features of the image are

needed for a classification process. Intuitively, affine invariant features are simply a systematic method for transforming from observer-based to image-based coordinates; in the former the image depends on the view, whereas in the latter the image is viewing transformation independent.

The experiments are carried out on a set of handwritten words taken from database of Handwritten Arabic Words[1], for seven different classes image sets. These word images are transformed using the suggested hybrid (Matlab is used to implement the hybrid transforms individually on each transform), and the resulting feature vectors are classified using Self-Organizing Map neural network. Three separable classes were presented one for each class of images. The experiment results showed that the method were highly capable of preserving the actual image features invariant to the scaling process.

2- The Hybrid.

Recently Walidlet Hybrid Transform is used in image enhancement, so an interest is rising in this topic. The main reason is that a complete framework has been recently built [6]. Particularly for what concerns the construction of hybrid transforms for efficient feature representation of image signals. The main characteristic of Walidlet is the possibility to provide a multiresolution analysis of the image in the form of coefficient matrices. The Walidlet transform starts by applying the two dimensions Fast Fourier Transform (2-D FFT) to the two dimensions signals (image).The next step is to map a line sampling scheme into a point sampling scheme using the Radon transform. Hence, it is requires taking the one dimension inverse Fast Fourier Transform (1-D IFFT) for each column of the produced two dimensions signal. Finally, the Multiwavelet transform is performed on each row of the resultant two dimension signal. The structure of the Hybrid transform is given in Figure (1).

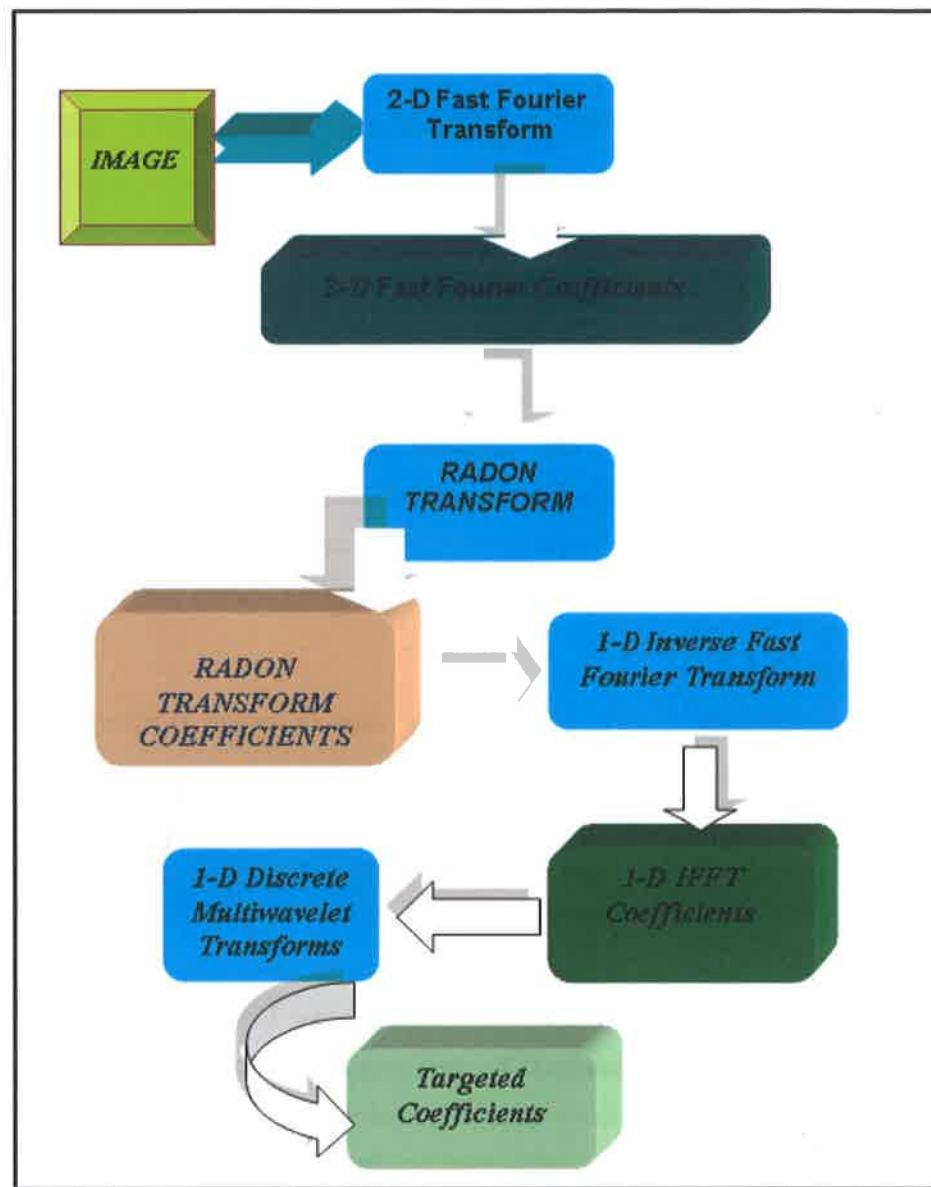


Figure (1) Flow diagram of Hybrid Transform

2.1 Fast Fourier Transform (2-D FFT)

The multidimensional DFT:

$$X_k = \sum_{n=0}^{N-1} e^{-2\pi i k \cdot (n/N)} x_n$$

transforms an array x_n with a d-dimensional vector of indices $n = (n_1, n_2, \dots, n_d)$ by a set of d nested summations (over $n_j = 0 \dots N_j - 1$ for each j), where the division n/N , defined as $n/N = (n_1/N_1, \dots, n_d/N_d)$, is performed element-wise. Equivalently, it is simply the composition of a sequence of d sets of one-dimensional DFTs, performed along one dimension at a time (in any order).

This compositional viewpoint immediately provides the simplest and most common multidimensional DFT algorithm, known as the row-column algorithm (after the two-dimensional case, below). That is, one simply performs a sequence of d one-dimensional FFTs (by any of the above algorithms): first you transform along the n_1 dimension, then along the n_2 dimension, and so on (or actually, any ordering will work) [7].

2.2 Radon transform [8]

Let $f(x) = f(x, y)$ be a continuous function vanishing outside some large disc in the Euclidean plane R_2 . The Radon transform, Rf , is a function defined on the space of lines L in R_2 by the line integral along each line:

$$Rf(L) = \int_L f(x) |dx|.$$

Concretely, any straight line L can be parametrized by

$$(x(t), y(t)) = t(\sin \alpha, -\cos \alpha) + s(\cos \alpha, \sin \alpha)$$

Where s is the distance of L from the origin and α is the angle the normal vector to L makes with the x axis. It follows that the quantities (α, s) can be considered as coordinates on the space of all lines in R_2 , and the Radon transform can be expressed in these coordinates by

$$\begin{aligned} Rf(\alpha, s) &= \int_{-\infty}^{\infty} f(x(t), y(t)) dt \\ &= \int_{-\infty}^{\infty} f(t(\sin \alpha, -\cos \alpha) + s(\cos \alpha, \sin \alpha)) dt \end{aligned}$$

More generally, in the n -dimensional Euclidean space R_n , the Radon transform of a compactly supported continuous function f is a function Rf on the space Σ_n of all hyperplanes in R_n . It is defined by

$$Rf(\xi) = \int_{\xi} f(x) d\sigma(x)$$

for $\xi \in \Sigma_n$, where the integral is taken with respect to the natural hypersurface measure, $d\sigma$ (generalizing the $|dx|$ term from the 2-dimensional case). Observe that any element of Σ_n is characterized as the solution locus of an equation:

$$x \cdot \alpha = s$$

where $\alpha \in S_{n-1}$ is a unit vector and $s \in R$. Thus the n -dimensional Radon transform may be rewritten as a function on $S_{n-1} \times R$ via

$$Rf(\alpha, s) = \int_{x \cdot \alpha = s} f(x) d\sigma(x).$$

It is also possible to generalize the Radon transform still further by integrating instead over k -dimensional affine subspaces of R_n . The X-ray transform is the most widely used special case of this construction, and is obtained by integrating over straight lines [9].

2.3 Inverse Fast Fourier Transform (1-D IFFT)

The inverse discrete Fourier transform (IDFT) is given by

$$x_n = \frac{1}{N} \sum_{k=0}^{N-1} X_k e^{\frac{2\pi i}{N} kn} \quad n = 0, \dots, N-1.$$

A simple description of these equations is that the complex numbers X_k represent the amplitude and phase of the different sinusoidal components of the input "signal" x_n .

The DFT computes the X_k from the x_n , while the IDFT shows how to compute the x_n as a sum of sinusoidal components $(1/N)X_k e^{\frac{2\pi i}{N} kn}$ with frequency k/N cycles per sample.

2.4 Multiwavelet transform

Multiwavelets provide one alternative to the wavelet transform. Despite its general success, the wavelet transform often fails to accurately capture high frequency information, especially at low bit rates where such information is lost in quantization noise.

2.5 Algorithm for Computation of the Hybrid

Step (1): Resizing. To implement the Radon Transform and the best sequence of directions in the following steps, a two dimensional signal must be first converted into a prime square form.

Step (2): Computation of two Dimensional Fast Fourier Transform (2D FFT).

Step (3): Obtaining the Direct Current (DC) component.

The first element of the produced transformed matrix represents the DC value of the

two dimensional signals. This value will be replaced with a zero value.

Step (4): Computation of the Radon Transform, to compute the Radon Transform.

Step (5): Compute the One Dimension Inverse Fast Fourier Transform (1-D IFFT); here the one dimension inverse Fast Fourier transform should be applied to each column of the produced matrix from step (5).

Step (6): Resizing of The 1-D IFFT Resultant Matrix, it is necessary here to resize the vectors into a power of two elements in order to enable the computation of the last step.

Step (7) : Compute The One Dimension Multiwavelet Transform (1-D DMWT), The final step of the hybrid transform is to apply the one dimension Multiwavelet transform to each row of the produced matrix from step (6) using critically sampling scheme .

2.6 A General Algorithm for Computing 1-D DMWT Using a Critically-Sampled Scheme of Preprocessing. [6]

The following algorithm for computing single-level 1-D Discrete Multiwavelet Transforms using GHM four multifilters and using a critically-sampled scheme of preprocessing 1st order:

1. **Checking Image Dimensions:** Image matrix should be a square matrix, $N \times N$ matrix, where N must be power of 2. So checking input image dimensions is the first step of the transform procedure. If the image is not a square matrix some operation must be done to the image like resizing the image or adding rows or column of zeros to get a square matrix.
2. **Constructing A Transformation Matrix :** Use the transformation matrix, such as given in the following matrix format:

$$\begin{bmatrix} H_0 & H_1 & H_2 & H_3 & 0 & 0 & \dots \\ G_0 & G_1 & G_2 & G_3 & 0 & 0 & \dots \\ 0 & 0 & H_0 & H_1 & H_2 & H_3 & \dots \\ 0 & 0 & G_0 & G_1 & G_2 & G_3 & \dots \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots \end{bmatrix}$$

An $N/2 \times N/2$ transformation matrix should be constructed using GHM low- and high-pass filters matrices. After substituting GHM matrix filter coefficients

values, an $N \times N$ transformation matrix results with the same dimensions of input image dimensions after preprocessing.

3. **Preprocessing Rows:** Approximation-based row preprocessing can be computed by applying equations (1) and (2) to the odd- and even-rows of the input $N \times N$ matrix respectively for the 1st order approximation preprocessing.

Term1= (0.3736150) [same odd-row]

Term2= (0.11086198) [next even row]

Term3= (0.11086198) [previous even -row]

$$\text{new even-row} = (\sqrt{2}-1) [\text{same even -row}] \quad \dots \dots \dots \quad (2)$$

4. Transformation of Image Rows:

- I. Apply matrix multiplication to the $N \times N$ constructed transformation matrix by
 - a. The $N \times N$ row preprocessed input image matrix.
- II. Permute the resulting $N \times N$ matrix rows by arranging the row pairs 1,2 and 5,6,
..., $N-3$, $N-2$ after each other at the upper half of the resulting matrix rows, then
the row pairs 3, 4 and 7, 8, ..., $N-1$, N below them at the next lower half.

5. The Final Transformed Matrix: to get the final transformed matrix:
Apply coefficients permutation to the resulting transpose matrix. The final DMWT matrix using approximation-based preprocessing has the same dimensions, $N \times N$, of the original image matrix.

3. Feature Vectors Extraction

Before proceeding with feature extraction, no normalization step is performed on the Arabic handwriting images selected, regarding the translation, rotation and scaling and no lighting equalization is performed. Hybrid transform is performed on the whole image but the hybrid coefficients will be considered only in the column wise aspect. The first four columns of the coefficients matrix are considered (Low resolution information). There is no need to consider more columns because, the size of the coefficients is becoming too large and no more valuable information is obtained. Therefore, the Arabic handwriting images are described by 256 hybrid coefficients vector which represents reasonable amount of information. It is important to take decisions only on the most essential, so-called discriminatory information, which is conveyed by the extracted features. Each of the 256 coefficient vectors contains information about the Arabic handwriting images.

3.1 Feature Vectors Clustering

When solving a pattern recognition problem, the ultimate objective is to design a recognition system which will classify unknown patterns with the lowest possible probability of misrecognition. It is well known that, the complexity of a classifier grows rapidly with the number of dimensions of the pattern space. Thus, the problem is how to face the dimensionality reduction? An efficient way of reducing dimensionality and clustering expression information is to compute the topological relationships between these feature vectors. Thus, Self Organizing Map neural network is used, the dimensionality is reduced to two dimensions instead of 256. Also the discriminating information regarding the face expressions is presented on the Self-Organizing feature Map in terms of the topological relationships.

3.2 Discriminate Feature's Vectors of hybrid transform.

Several experiments have been performed to decide which columns among the first four coefficients columns of the hybrid transform may hold the most discriminating information. Self-Organizing Map neural network is used to show the topological relationships between these features vectors regarding all the facial image data collected from the set of the facial images.

3.3 The experiments

The experiment is performed as follows:

- 1) Several different Arabic handwriting words images selected from the Comprehensive Arabic Database [1]; each one is for different person. (Figures 2).
- 2) The Hybrid Transform is applied to these collected images individually to obtain the coefficient matrix for each single image figure (3). (Matlab is used to implement the hybrid transforms individually on each transform).
- 3) The first four columns coefficients of the hybrid transform final matrix are extracted for each image of the total 63 images.
- 4) Each word then has its own file that contains four sub-files one for each different column (column 1, column 2 column 3 and column 4) a total of 63×4 sets of data.
- 5) All the coefficients for all Arabic handwriting words images for each column are appended to construct a source file, a total of 63 sets of data for each column.
- 6) In order to cluster the coefficient vectors for all Arabic handwriting words images, the Self-Organizing Feature Map neural network is used which is considered an unsupervised classification method.
- 7) The SOFM is applied to the coefficients columns constructed in step 6. The clustering of the image coefficients, such that similar coefficient vectors are grouped together and dissimilar coefficient vectors are grouped into other clusters. From the visual inspection, column one is

found to be the most significant feature vector (see SOFM node clusters maps Figures 8,10) , in which capture the features irrelevant to the writer. Therefore only column one is considered in this study and neglecting columns (2,3 and 4).

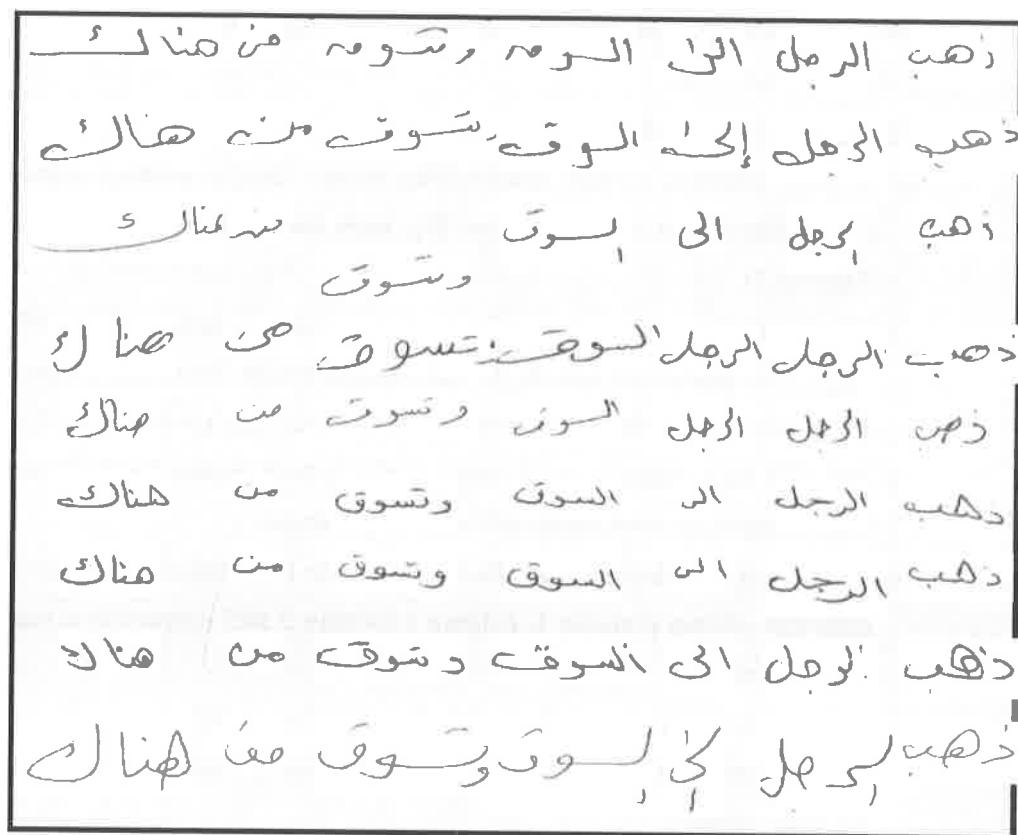


Figure (2) a sample of Arabic handwriting images
for seven different words (7 columns) written by nine different writers (9
rows)

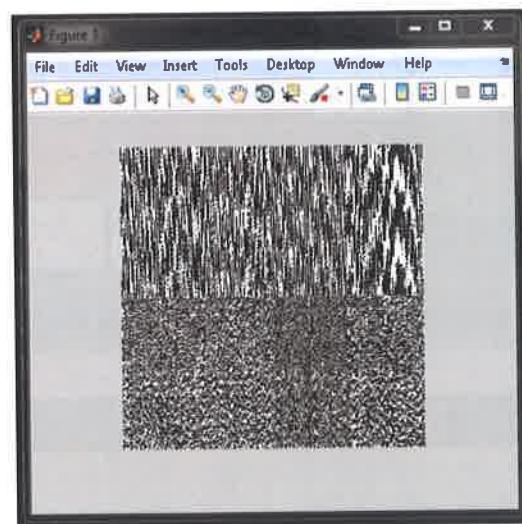


Figure (3) Hybrid transforms coefficients' representation.

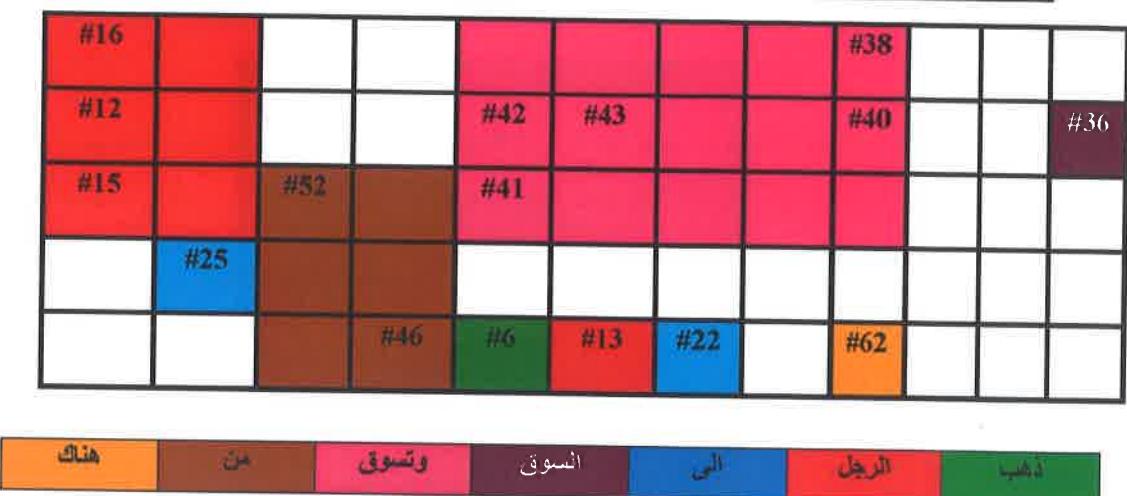


Figure (4) SOFM's nodes clusters for 63 patterns, using hybrid Transform coefficients for seven words written by nine writers labeled from 1-63, presented on their relative feature's position on the map of the self-organized feature map neural network.

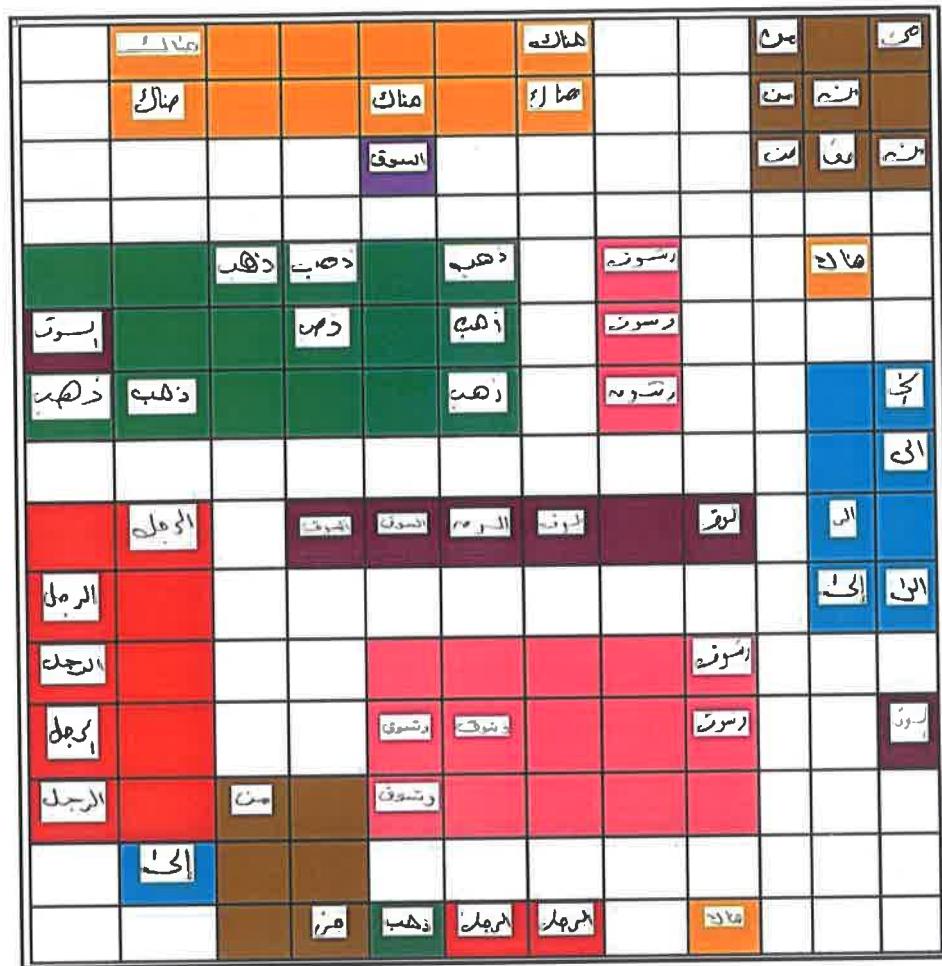


Figure (5) SOFM's nodes clusters for 63 words images, using hybrid Transform for seven words written by nine writers, presented on their relative feature's position on the map of the self-organized feature map.

Arabic Handwriting Words	Labels ranges
ذهب	1-9
الرجل	10-18
الى	19-27
السوق	28-36
ونسق	37-45
من	46-54
هناك	55-63

Table 1 Labels ranges for the Arabic Handwriting Words

The summary of SOFM parameters used is shown in (Table 2) and the topological relationships between the Arabic handwriting images coefficient vectors are shown in (Figure 4 & 5).

Pattern Vector Size	No. Of Patterns	Map size	Accepted Error
256	63	12x15	1.E-7

Table 2 Parameters used in SOFM for the classification task.

In order to determine the ability of the Hybrid Transform to capture the Arabic handwriting words images identity's discriminatory information, the code-book vectors (every node in the map may be thought of as code-book vector) are compared with each input vector (63 image vectors). Since the labels of the input vectors are arranged in ranges according to the two different Arabic handwriting words images, then every node on the map is labeled by assigning it the label of the input vector most similar to it (difference value between the two vectors is less than 1.E-7). This allows the topological structure of the map to be examined as shown in Figure 4). The detail clustering information is presented in (Table 2). It can be observed from the results presented in the first

column map and as shown in (Figure 4) that the hybrid transform possesses the following properties:-

The pattern identity is not filtered out by the transform, so that the different Arabic handwriting words patterns for the same object are well clustered within their similar pattern group. This property is considered most important in pattern recognition, since the identity features could not overcomes by the identity of different pattern. This may lead to cluster all the words of the same identity in one group regardless of who the writer was.

The clustering efficiency is described using equation (3) in which calculate the overlapping of handwriting words patterns.

The Clustering Efficiency:

$$\frac{(\text{Number Of ClusterNodes}) - (\text{Number Of NonClusterNodes})}{(\text{Number Of PatternNodes})} \dots \quad (3)$$

Cluster No.	Arabic word	Nodes in Cluster	No. of Overlapping	Overlapping Percentage
1	ذهب	8	1	13%
2-1	الرجل	5	0	0%
2-2	الرجل	2	0	0%
3	الى	5	0	0%
4	السوق	5	0	0%
5	ونسوق	5	0	0%
6-1	من	7	0	0%
6-2	من	2	0	0%
7	هناك	5	0	0%

Table 3 Clustering Efficiency with overlapping percentage

The results of (Table 3) show that the clustering efficiency for each Arabic word, no overlapping had occurred for all the words except for the cluster ذهب in which the word السوق had overlapped. The test showed that the clustering efficiency with overlapping percentage was 99%.

Feature similarity is also calculated for each group to measure the overall similarity efficiency for the hybrid approach suggested. The similarity measure represents the missing words in which they do not classified in their own classes (in our case 9 classes or groups). The Feature similarity measure is calculated according to the following rule:

$$\text{Feature similarity measure} = 1 - \frac{\text{Missed words of a class}}{\text{No.of words in the same class on the Map}} \times 100$$

Word	Missed words of a class	No.of words in the same class on the Map	similarity measure percentage
ذهب	1	8	88%
الرجل	0	7	100%
الى	1	6	84%
السوق	2	7	71.4%
وتسوق	0	5	100%
من	0	9	100%
هناك	2	7	71.4%
Total	6	49	88%

Table 5 Feature similarity measure

Table 5 above shows that the overall Feature similarity measure was 88%, in which considered a very good result regarding feature extracting problem for a difficult task of Arabic handwriting recognition problem.

5- Conclusions

The hybrid Transform is used (Matlab is used to implement the hybrid transforms individually on each transform), in order to capture the discriminating information for each basic handwriting Arabic word image. The SOFM neural network is used to cluster the feature vectors extracted from hybrid coefficients, using a feature map of (n X m) pattern nodes. Only low resolution information is extracted from its coefficient matrix, this information lie in first coefficient vectors. Research shows that the first column is the most significant in its discriminating power. The test showed that the clustering efficiency with overlapping percentage was 99%. The overall Feature similarity measure was 88%, It can be concluded from these results that, the hybrid coefficients of low resolution can be used as a discriminating features for the task of Arabic words images images recognition invariant to rotation scaling and translation with a very high percent of accuracy. The interesting features of this Hybrid transform are as follow:

- 1- Its ability to capture the Arabic words images without normalization which is considered an excellent achievement since the process of precise normalization is a quite expensive task in terms of the time and efforts. Even with an automatic normalization a human subject must interfere at one or more stages of the process.
- 2- The word's identities not filtered out by the transform, so that different pattern images for different words are well clustered within their similar pattern groups (same word group). This property is considered most important in pattern recognition process, since the identity features could overcome the texture features for the same object. This leads to clustering all the pattern images of the same object in one group regardless of the texture is involved.

3- The suggested method eliminate the process of locating the lines and words then segmenting characters of each word to the process of only segmenting the words.

References

[1] Arabic, "A New Comprehensive Database of Hadritten Arabic Words, Numbers, and Signatures used for OCR Testing", by Nawwaf Kharma. Maher Ahmed, and RababWard,1999,<http://users.encs.concordia.ca/~kharma/ExchangeWeb/Databases/ArabicDBases/>

[2] https://en.wikipedia.org/wiki/Handwriting_recognition

[3] A. Amin, "offline Arabic character recognition," in pattern recognition, 1998.

[4] Mahmoud, W.A. and Ramahi, N., "The Walidlet Transform", Proceeding of 1st International Conference on Digital Communication and Computer Applications, PP.738-741, Jordan, 2007.

[5] [10] Kohonen, T. "Self-Organizing Formation of topologically correct feature maps". Biol Cyb., 43(1):59-69, (1989).

[6] Mahmoud, W.A. and Ramahi, N., "The Walidlet Transform", Proceeding of 1st International Conference on Digital Communication and Computer Applications, PP.738-741, Jordan, 2007.

[7] http://en.wikipedia.org/wiki/Fast_Fourier_transform

[8] http://wapedia.mobi/en/Discrete_Fourier_transform

[9] http://en.wikipedia.org/wiki/Radon_transfo