University of Thi-Qar Journal Of Science (UTsci)
Website: http://jsci.utq.edu.iq **Email: utjsci@utq.edu.iq**
Volume 6, Number 1, December 2016

A Hybrid Features for Signature Recognition Using Neural Network

Zamen F. Jabr(1) Shaymaa R. Saleh(2) Abeer N. Fasial(3)

(1)College of Education for Girls- University of Thi-Qar
za.fa.ja@gmail.com

(2)College of Computers and Mathematics Sciences- University of Thi-Qar
shaymaarashid30@gmail.com

(3)College of Computer Sciences and Information Technology- University of Sumer
abeernaser13@gmail.com

Abstract

In automatic personal recognition systems, biometric features is used as recognition measure based on biological traits such as face, iris, fingerprint, etc…or gait, signature which is considered behavioral characteristics. Signature verification is one of the authentication methods which can provide security at maintenance and low cost. The most essential and challenging stage of any off-line signature system is feature extraction stage. The accuracy and robust of the recognition system depends basically on the usefulness of the signature features extracted by this system. If the extracted features from a signature's image doesn't robust this will cause to higher verification error-rates especially for skilled forgeries in hacker the system. In this paper, we present a new offline handwritten signature recognition system based on combination of global with Statistical and GLCM (Grey Level Co-occurrence Matrix) features using neural network as classifier tool. The global, Statistical and GLCM features are combined to consist a vector of 14 features for the authentication of the signature. Verification of signatures is decided using neural network. The experimental results obtained by using a database of 7 individuals' signatures. A total number of 70 images are collected with 10 signatures for each person, 5 of the signatures are used in training phase, and the remaining 5 signatures are used in testing phase. In this proposed method the results show 100% recognition accuracy for training and 97.1% recognition accuracy for the testing.

Keywords: signature recognition, features extraction, neural network.

الخصائص الهجينة لتمييز التواقيع الشخصية باستخدام الشبكات العصبية

شيماء راشد صالح عبير ناصر فيصل زمن فاضل جبر

المستخلص:

يستخدم القياس الحيوي للتميز التلقائي للشخص بالاعتماد على الصفات البيولوجية مثل بصمات الأصابع وقزحية العين والوجه، الخ ... أو الخصائص السلوكية مثل مشية والتوقيع. التحقق من صحة التوقيع هي واحدة من أساليب المصادقة التي يمكن أن توفر الأمن في الصيانة والتكلفة المنخفضة.تعد مرحلة استخلاص الخصائص المرحلة الاكثر اهمية والاكثر صعوبة في أي نظام تحقق عن طريق التواقيع الشخصية. دقة النظام تعتمد بصورة اساسية على فعالية ودقة خصائص التوقيع المستخدمة في نظام التميز. كما ان عدم القدرة على استخلاص خصائص قوية من صورة التوقيع تسبب نسبة خطا كبيرة في نظام التحقق وهذا يفيد المزورين الماهرين في اختراق النظام. في هذه الورقة البحثية نحن قدمنا نظام تميز جديد للتواقيع الشخصية بالاعتماد على خليط من الخصائص العامة والخصائص الاحصائية وخصائص ناتجة من تطبيق مصفوفة تكرار حدوث المستويات الرماديةGLCM على صورة التوقيع وباستعمال الشبكات العصبية كأداة تصنيف. الخصائص المذكورة اشتركت لتكون متجه خصائص يضم 14 خاصية تمييز كل توقيع لشخص معين عن توقيع الشخص الاخر لتوثيق التوقيع رقميا. النتائج التجريبية المستحصلة باستخدام قاعدة بيانات تضم تواقيع 7 اشخاص . العدد الكامل لصور التواقيع كان 70 صورة جمعت بواقع 10 صور توقيع لكل شخص خمس منها استخدمت في مرحلة التدريب والخمس المتبقية في مرحلة الاختبار . النتائج اظهرت بان الطريقة المقترحة أنجزت دقة مقدارها 100% في مرحلة التدريب و 97.1% في مرحلة الاختبار .

1. Introduction

Biometric systems are mainly based on fingerprint, iris scanning, face recognition, ECG and DNA analysis etc. Although attributes associated with fingerprint, iris, and retina do not change overtime, but they need special and relatively expensive device to acquire the image. Thus, all these systems provide better security than traditional systems but with expensive device and processing systems. Signature verification is one of the authentication methods which can provide security at maintenance and low cost because the instrument for acquisition of signature is much cheaper. Signature verification systems are method of authentication which decides whether a particular signature is authenticated or faked.[Madhuri, Alok,2015]

Unlike passwords and PIN (Personal Identification Number) codes that can be forgotten, lost, robbed or shared, signatures are special to an individual and difficult to be forgotten or even reproduce by others; therefore, authentication using signature has been widely used as a secure way of identification by people. During information technology time, where fast recovering information is acquired using communication technologies and computer networking, signatures is used in many information recovery applications, like border security and banking systems [H. Hiary,etc,2013].

There are two type of signature verification systems the first is offline (static) and the second is online (dynamic) [Kholmatov and B. Yanikoglu,2005]. The dynamic features of the signature (signature trajectory, pen pressure, pen downs and pen ups, time stamping, etc) are used in online signature verification this features are hold by a pen based tablet, while scanner is used to capture an input image in offline signature verification to acquire a sample of signature image on writing paper this type of recognition system generally observe on documents signature and bank checks. The operation of fabricating an online signature more complex because its dynamic feature of the online signature verification is more robust, reliable, and accurate than offline signature verification [T.S. Ong,etc,2009]. Online signature verification has a main drawback: it is on-line so no observation and registration of the signing process. Therefore, it cannot be used in some important applications that the signer could not exist in signing position [M. R. Pourshahabi,etc,2009]. For this reason, static signature analysis is still in attention of many researchers because offline methods do not need special acquisition hardware, just a paper and a pen, so they are less invasive and more user friendly [B. Kovari,etc,2007]. This paper is ordered as follows: In section 2, Feature Extraction is introduced, in Section 3 neural network is briefly introduced, in section 4 our proposed method is presented it consists of read signature's image, preprocessing, feature extraction, normalization, neural network training and testing. The experimental results are given in section 5 and finally conclusions are presented in section 6.

2. Features Extraction

Prior to application of any technique it should be kept in mind that we must to choose features that provide variance which is large enough to identify the forgery and small enough to recognize different signatures of same person. By feature extraction of signatures we mean extraction of the properties of the signatures such as shape, texture, color or contents of the signature [Madhuri, Alok,2015]. In this work different types of features are extracted such as:

2.1 **Statistical features**: The distribution of pixels of a signature's image consist the statistical features, e.g. statistics of high gray-level pixels to describe pseudo-dynamic characteristics of signatures. Three Statistical features are used in this work which is [H. Anand and. D.L Bhombe,2014]

a- **Standard Deviation**: is the square root of an unbiased estimator of the variance of the population.

b- **Kurtosis**

Kurtosis is any measurement of the "peakedness" of the probability distribution of a real random variable and it is can be consider as a descriptor of the shape of a probability distribution. There is various explanation of kurtosis these are mainly peakedness (width of peak) tail weight etc.

c- **Entropy**

Entropy is a statistical measure of randomness that can be used to represent the texture of the input image

2.2 **Grey Level Co-occurrence Matrix (GLCM)**

The Gray Level Coocurrence Matrix (GLCM) is a method of find second order statistical texture features. This method can be defined as a matrix with the of columns and rows equal to the number of gray levels, G, in the image. In any space of image two pixels, separated by distance of a pixel (Δx, Δy), follow within a given neighborhood, one with intensity ‘i’ and the other with intensity ‘j’ the relative frequency can be defined with matrix element P (i, j | Δx, Δy) . At a specific displacement distance d and at a specific angle (ɵ) the matrix element P (i, j | d, ɵ) contain the second order statistical probability values for variation between gray levels ‘i’ and ‘j’. Using a large number of intensity levels G involves saving a lot of temporary data, i.e. a G × G matrix for each mixture of (Δx, Δy) or (d, ɵ). The GLCM’s are very sensitive to the size of the texture samples image on which they are evaluated because of their large dimensionality. For this reason the reduced to the , gray levels number is necessary step in this statistical method .In fig1 explain a simple example for the formulation of GLCM matrix for four different gray levels. In this example an offset of one pixel is used (a mention pixel and its close neighbor). Using a larger offset is possible, If the window is big enough. The number of times the combination i,j occurs will be filled in the cell of this matrix. Based on the concept of co-occurrence matrix, Haralick take the probability matrix to defines fourteen textural features from it in order to extract the feature of texture statistics of images [P. Mohanaiah,etc,2013]. In this paper three important features, energy, Correlation, and Contrast are selected which is [P. Mohanaiah,etc,2013],[Tuan Anh Pham,2010] :

a- Energy: It is the sum of squares of entries in the GLCM which use to measure the image homogeneity. Energy is high when image has very good homogeneity (Uniformity) or when pixels in the image are very similar.

b- Correlation: is use to measure the linear dependency of grey levels of neighboring pixels.

c- Contrast: is a measure of intensity or gray-level changes between the reference pixel and its neighbor.

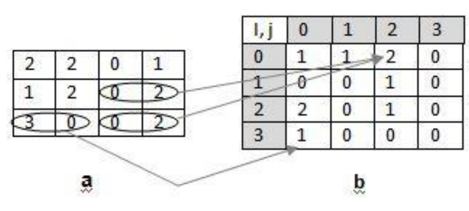

Fig.1:a is an 3x4 image and b is GLCM for image a

2.3 **Global Features**

Global features Supply information about specific cases related to the structure of the signature. Eight global features are used which is [E. Ozgunduz and M. Elif Karslıgil,2003]:

1-Signature Height
2- Image Area
3- Maximum Horizontal Projection
4- Maximum Vertical Projection
5- Horizontal Center of the Signature
6- Vertical Center of the Signature
7- Number of Edge Point (En): is defined as a point that has only one 8-neighbor as shown in fig.2 where it is defined as (E1, E2, E3, E4).
8- Number of Cross Point (Cn): is a point that has at least three 8-neighbors as shown in fig.2 where it is defined as (C1, C2, C3, C4).

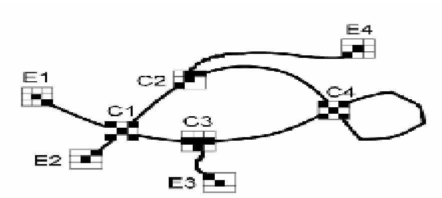

Fig.2: Edge point (E1,E2,E3,E4) and cross point (C1,C2,C3,C4)

3. Artificial neural networks

Artificial neural networks (ANNs) are a kind of artificial intelligence which tries to mimic the actions of the human brain and nervous system. Neural networks (NN) composed from a number of neuron or processing elements (nodes) that are organized in layers: one input layer, one or more hidden layers and one output layer. Each neuron is linked to other neuron in the next layers. In hidden and output layers neurons are process its input by multiplying each input by its connection weights and then this will be passed to a transfer function to obtain the output [M. A. Shahin,etc,2001].

The most popular algorithms for training NN is Back propagation (BP) algorithm because of its simplicity, generalization and applicability viewpoints. The algorithm composed from two stages: Training and Testing stage. In the stage of training, firstly the connection weights of the network are initialized randomly. Secondly, neurons are process its input, then the network output is computed using transfer function and this output will be compared to a target value. Then, the network error is computed and utilized to modify the output layer connection weights. In a similar way, the network error is propagated inversely and utilized to adjust the prior layers weights. In the phase of testing, only the feedforward calculation using final training weights and input sample will be occur [A. Zilouchian and M. Jamshidi,2001].

4. Proposed method:

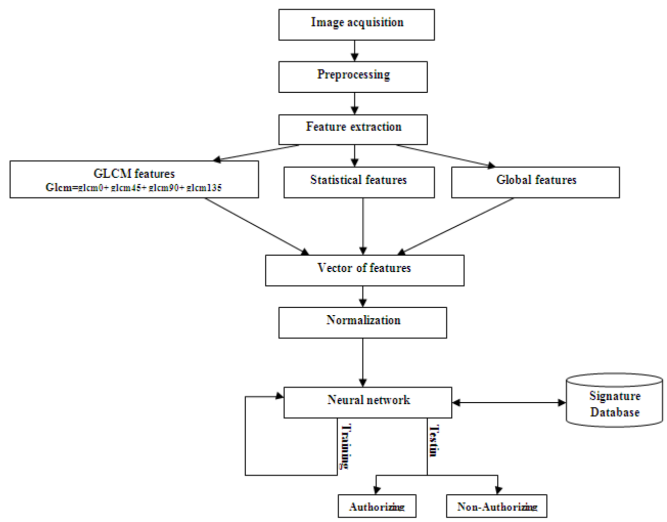

Fig.3: proposed method

Each step in our proposed method is explaining bellow
1- **Image acquisition**: signature's image for each person entered to our system using scanner with jpg colored format in 3 levels (red, green, blue).
2-Pre-processing: contain the following steps:
a) Convert color image to gray scale image
b) Convert gray scale image to binary image using automatic thresholding.
c) Image resizing: image size is adjusted to 250 ×250.
d) Image Closing: apply morphological closing in order to smooth the boundary of signature image and filling the holes in it.
e) Image Thinning: The goal of this is to remove the width of signature and convert from several pixels to single pixel.

3-Features extraction

In this stage three groups of features are used in order to construct vector of 14 features, for each person's signature image. These features are:

a. **Statistical features**: which is (Standard Deviation, Kurtosis, Entropy).
b. **GLCM features**: We extract four GLCM In degrees 0°, 45°, 90°, 135° and then add all to construct one matrix; from this matrix we extract three features which is (Energy, Contrast, correlation).
c. **Global features**: which is (Signature Height, Image Area, Maximum Horizontal Projection, Vertical Center of the Signature, Number of Edge Point, Number of Cross Point, Maximum Vertical Projection, Horizontal Center of the Signature)

So we have a vector of 14 features which consist of 8 features from global features, 3 from statistical and 3 from GLCM table.1 show this feature for one person. These features are dividing to 5 as training group and 5 as testing group.

Table 1: Feature vector for all 10 image that belong to a person No.1

Image No.	Feature vector													
	Statistical			GLCM			Global							
	Std	Kur	Entropy	cont	corr	enrg	area	H_to_w	H_histo	V_histo	CX	CY	edge	cross
1	0.5	32.22	42.43	649.05	0.91	0.96	1171	0.68	15	36	150	124	626	434
2	0.41	30.2	47.28	533.24	0.93	0.97	708	0.67	14	16	156	110	550	125
3	0.43	30.16	47.93	532.74	0.92	0.96	730	0.81	22	16	160	114	575	128
4	0.51	30.99	42.61	605.06	0.91	0.96	683	0.85	14	22	145	107	575	86
5	0.48	30.21	46.19	530.58	0.91	0.96	698	0.82	20	27	154	103	547	121
6	0.59	36.51	30.45	672.39	0.9	0.95	930	0.9	15	26	153	118	666	214
7	0.54	32.77	36.46	573.55	0.91	0.96	774	0.7	25	23	148	117	649	100
8	0.53	33.36	36.28	615.6	0.91	0.95	773	0.88	13	27	160	109	575	164
9	0.63	37.04	29	740.15	0.89	0.95	992	0.89	14	20	149	128	752	197
10	0.67	38.34	26.86	795.69	0.88	0.94	992	0.83	14	20	149	128	752	197

Table2: Target vector of person No.1

Target vector
0000001
0000001
0000001
0000001
0000001
0000001
0000001
0000001
0000001
0000001

4.Normalazition

It is an important stage which used to increase the training capabilities without normalization stage; network learning will be done slowly. Vector of features are normalized using Min-Max Normalization as the following equation:

y = (xmax-xmin)*(x1-xmin)/(xmax-xmin) + xmin…..(1)

This Normalization performs rescaling for input feature where it is converting from a range of values to another one. Often the ranges of values are mostly between [0, 1] or [-1, 1] [T. Jayalakshmi, A.Santhakumaran,2011].

5. Training & Testing of NN

The NN architecture that used in this work contains one input layer, one hidden layer and one output layer. Neurons Number of input layer is 14 neurons one for each feature, 20 neurons in the layer of hidden and 7 neuron in the layer of output which is equal to number of persons that the signature are taken from them. Activation function that is used in layer of hidden is "Tansig" and" purelin" is used for layer of output. The network will be training with target matrix has Rows equivalent to image numbers and columns equivalent to persons number. Each rows in this matrix consists of zeros value with a 1 in element i, where i is the person number that we want to process. For example if we have 7 person the coding will be as in fig.4

Person no.1	Person no.2	Person no.3	Person no.4	Person no.5	Person no.6	Person no.7
0000001	0000010	0000100	0001000	0010000	0100000	1000000

Fig.4: target coding for 7 person

If we have more than 7 we will expand the coding table2 contain a target vector for images of person no.1. Table3 gives an algorithm for the verification system of offline signature using NN for checking the authenticity of signatures.

University of Thi-Qar Journal Of Science (UTsci)
Website: http://jsci.utq.edu.iq Email: utjsci@utq.edu.iq
Volume 6, Number 1, December 2016

Table3: Training and testing algorithm for neural network

Training algorithm	Testing algorithm
1- Load training feature matrix 2- load target matrix of training 3- normalize feature matrix 4- design neural network 5- initialize weight of network randomly 6- Training a neural network with normalized feature and target of training 7- Compute performance of training neural network using MSE 8- Save training neural network	1- Loadt testing feature matrix 2- load target matrix of testing 3- normalize feature matrix 4- load trained neural network with weights 5- testing a trained neural network with normalized feature and target of testing 6- extract the output of network 7- find max value in each column of output 8- find indices of max value 9- process the output by putting 1 in location of max value and 0 for the other 10- compare target with the output after processing 11- compute classification rate

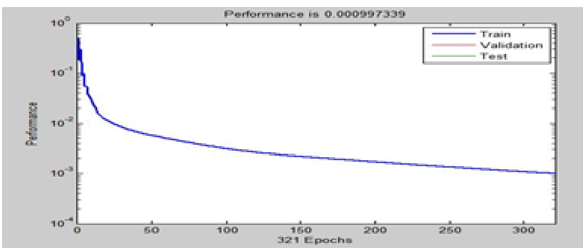

Fig5: performance graph of the training feed forward neural network

Table4: Classification Rate of Testing Image.

Person NO.	1	2	3	4	5	6	7
1	5	0	0	0	0	0	0
2	0	5	1	0	0	0	0
3	0	0	4	0	0	0	0
4	0	0	0	5	0	0	0
5	0	0	0	0	5	0	0
6	0	0	0	0	0	5	0
7	0	0	0	0	0	0	5
correct classification rate	100%	100%	80%	100%	100%	100%	100%
The average correct rate	97.1%						

5.Conclusion

In this paper a method for verification of off line signature using NN are present where the verification is based on features that obtained from the images of sign ature which is global, Statistical and GLCM. These features are utilized for training the NN by using BP algorithm. The feature set is showed to be efficient in capture finer variations in the signature. Our system obtain 100% rate of accuracy for training where all sign atures that it was fed to NN during training are organized exactly and 97.1% for the sign atures that it was fed to NN during testing this mean high performance when it was presented with signatures that it was not seen in training which mean generalization ability of the network to non-seen data and also the efficient of features set that extracted from signature.

References

A. Zilouchian and M. Jamshidi, "Intelligent Control Systems Using Soft Computing Methodologies", CRC Press, 2001.

B. Kovari, Z. Kertesz, and A. Major. Off-line signature verification based on feature matching. 11th IEEE Int Conf on Intelligent Engineering Systems, Budapest, Hungary, 29 June - 1 July 2007.

E. Ozgunduz and M. Elif Karslıgil, "Off-line Signature Verification and Recognition by Support Vector Matching", Computer Engineering Department, Yıldız Technical University, Yıldız, Istanbul, Turkey, 2003.

H. Anand and. D.L Bhombe," Enhanced Signature Verification and Recognition Using Matlab", International Journal of Innovative Research in Advanced Engineering (IJIRAE) ,Volume 1 Issue 4, May 2014.

H. Hiary,R. Alomari, TH. Kobbaey, R. Z. AL-Khatib, M. A. AL-ZU'BI and H. Hasan, "Off-Line Signature Verification System Based On Dwt and Common Features Extraction", Journal of Theoretical and Applied Information Technology, Vol. 51 No.2, 20th May 2013.

Kholmatov and B. Yanikoglu. Identity authentication using improved on-line signature verification method. Pattern Recognition Letters, Volume 26, Issue 15, November 2005, pp. 2400-2408.

M. A. Shahin, M. B. Jaksa and H. R. Maier ," Artificial Neural Network Applications In Geotechnical Engineering", Australian Geomechanics Journal 36(1):49–62 · March 2001

M. R. Pourshahabi, M. H. Sigari and H. R. Pourreza," Offline Handwritten Signature Identification and Verification Using Contourlet Transform ", International Conference of Soft Computing and Pattern Recognition, 2009.

Madhuri, Alok," Offline Signature Verification Using Contour Tracing ",International Journal of Science and Research (IJSR), Volume 4 Issue 8, August 2015.

P. Mohanaiah, P. Sathyanarayana and L. GuruKumar, " Image Texture Feature Extraction Using GLCM Approach", International Journal of Scientific and Research Publications, Volume 3, Issue 5, May 2013

T. Jayalakshmi, A.Santhakumaran, “Statistical Normalization and Back Propagation for Classification”, International Journal of Computer Theory and Engineering, Vol.3, No.1, PP. 89-93,IACSIT, 2011.

T.S. Ong, W.H. Khoh, A. Teoh. Dynamic handwritten signature verification based on statistical quantization Mechanism. IEEE Int Confon Computer Engineering and Technology, 2009.

Tuan Anh Pham “Optimization of Texture Feature Extraction Algorithm” M.Sc thesis, Delft university, Netherlands,2010.