# Using Complex Networks for Offline Handwritten Signature Characterization

César Armando Beltrán Castañón and Ronald Juárez Chambi

Pontificia Universidad Católica del Perú, Departamento de Ingeniería
Grupo de Reconocimiento de Patrones e Inteligencia Artificial
Av. Universitaria 1801, San Miguel, Lima, Perú
{cbeltran,ronald.juarez}@pucp.pe
http://inform.pucp.edu.pe/~grpiaa/

**Abstract.** This paper develops a novel way for offline handwritten signature characterization using a complex networks approach in order to apply for signature verification and identification process. Complex networks can be considered among the areas of graph theory and statistical mechanics. They are suitable for shape recognition due to their properties as invariance to rotation, scale, thickness and noise. Offline signatures images were pre-processed to obtain a skeletonized version. This is represented as an adjacency matrix where there are applied degree descriptors and dynamic evolution property of complex networks in order to generate the feature vector of offline signatures. We used a database composed of 960 offline signatures groups; every group corresponds to one person with 24 genuine and 30 forged signatures. We obtained a true rate of 85.12% for identification and 76.11% for verification. With our proposal it is demonstrated that complex networks provide a promising methodology for the process of identification and verification of offline handwritten signatures and it could be used in applications like document validation.

**Keywords:** complex networks, pattern recognition, offline handwritten signature verification and identification, shape analysis, image processing.

## 1 Introduction

The properties of complex networks are suitable to solve shape analysis problems, and they are used in other applications like social networks, internet, genetic and so on [12]. Nevertheless, despite the existence of many signature identification and verification techniques, there is no evidence in the literature of application of complex networks for signature characterization [16].

The contribution of this paper is referring to the use of complex network theory for offline handwritten signature characterization. First, we preprocessed signature images using Pavlidis thinning algorithm [1][3], in order to obtain a skeletonized representation of the signature. Second, an adjacency matrix structure is used for complex network representation and is applied the connectivity and evolution properties for features extraction, resulting a 26-dimensional feature vector. Third, these feature vectors constitute the input to the comparison process where we are using a multilayer

E. Bayro-Corrochano and E. Hancock (Eds.): CIARP 2014, LNCS 8827, pp. 580–587, 2014.

perceptron neural network [4][6]. The proposed method makes an efficient offline signature characterization showing a good degree of robustness for information, noise tolerant and invariant to scale and rotation.

## 2 Complex Networks

A complex network refers to a graph with no trivial properties compared with simple graphs, as well as it has a considerable number of nodes and edges.

There are two theoretical models of complex networks such as Erdös-Rényi, scale-free, random models. In this paper we used Watts-Strogatz network model which presents a small-world property. It means the graph nodes can be accessed from other nodes in a short number of edges and it has high number of three-size loops, i.e. if a vertex $i$ is connected to a vertex $j$ and this is connected with a vertex $k$ then is quite likely $i$ and $k$ are connected —high clustering coefficient [12].

The dynamic model network of Watts-Strogatz was simulated using thresholds in order to remake the network connections. The off-line signature is represented by a network structure through different growth stages. The study of its dynamic properties —where measurements were obtained from its evolution based on number of connected components— will produce a set of descriptors which will be used for its analysis and then for the classification process.

### 2.1 Connectivity Measurements on Complex Networks

The degree of a node is the number of edges incident to the node. When we have a directed graph we can refer to in-degree —number of incoming link— and out-degree —number of outgoing links. If the network is undirected then we simply called degree.

The average degree of a network is simply the average of each degree node of the whole network. In terms of the adjacency matrix $A$, the degree $k_i$ of a node $i$ is given by

$$k_I = \sum_{j=1}^{N} A_{ij} \qquad k_\mu = \frac{1}{N}\sum_{j=1}^{N} A_{ij} \qquad k_\kappa = \max\ k_I \tag{1}$$

where $N$ is the number of vertex in the network. Also, the maximum degree $k_\kappa$ and the average $k_\mu$ of the network were defined as

### 2.2 Dynamic Evolution on Complex Networks

Another important property in complex networks is their dynamic evolution which affects its structural properties. As a consequence, the measurements in complex networks are time functions, i.e. two networks obtained at different times from an original network are represented with different characteristics. Although most time the complex network dynamic evolution we can see only in the nature, we can get the patterns of its evolution from its evidence and represent it as a mathematical model. This allows a more suitable characterization to analyze and classification.

# 3 Offline Signature Characterization Using Complex Networks

## 3.1 Off-Line Image Signature Database

A set of image signature have been requested to 4NSigComp2010 commission. This contest is about off-line signature verification and it was organized by *Grupo de Procesado Digital de Señales* (GDPS) of *Universidad de las Palmas de Gran Canaria* [13]. This public database termed GPDS960Signature contains 960 groups of signatures, where each group belongs to particular person. Each person has 24 genuine signatures and 30 forged signatures.

## 3.2 Signature Skeletonization

As first step, it is necessary to preprocess the signature image in order to obtain a more appropriate representation for feature extraction process. The objective of signature skeletonization should be recovering the movement track of pen tip. To extract the signature skeleton we had used the Pavlidis thinning algorithm, where Skeletal pixels are computed by thinning binary images obtained by simple threshold logic and iterations [1][2][14].

## 3.3 Degree Descriptors of Complex Network Signature

In this section, we describe the process of signature characterization focused on complex network [15]. Let $S$ the signature trace represented as a set of points, where $S = [s_1, s_2, \dots, s_N]$ and $s_i = [x_i, y_i]$ whose components are numeric values that represent coordinates at point $i$ from contour. Now, in order to apply complex network theory to this problem, we create the equivalence of $S$ as a representation of graph $G = \langle V, E \rangle$. Each pixel of the skeleton is represented as a node of the network, i.e. $S = V$. A set of edges $E$ connect each pair of vertex establishing, in this way, the network $E$ is calculated using Euclidean distance $d(s_i, s_j) = \sqrt{(x_i + y_i)^2 + (x_i - y_i)^2}$ . Therefore, the matrix is represented by adjacency matrix with weight $W$ and $N \times N$ dimension, where: $w_{ij} = W([w, w_j]) = d(s_i, s_j)$ .

An advantage of this method is its tolerant properties to scale and rotation of the signature image. Fig. 1 shows the process for signature image normalization, in which it show the properties scale and rotation invariant. For that reason, the $W$ matrix is normalized into $[0,1]$

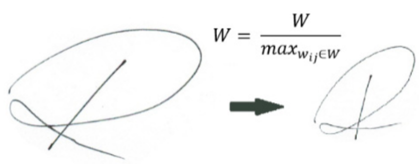


**Fig. 1.** Scale and rotation invariance representation, as a property of complex networks

### 3.4 Dynamic Evolution on Complex Networks

The use of concepts and underlying research tools in complex networks provide relevant information for the characterization of signatures. In this sense, given a specific transformation, the characterization of $S$ can be represented as a feature vector obtained from different values of $T_i$ with a transformation operation $\delta$, that redefines the number of connections in the graph. This operation is represented as $A = \delta_{T_i}(W_i)$, and it is applied to each element of the unweighted matrix W.

$$A_{T_i} = \delta_{T_i}(W_i) = \forall \mathrm{w} \in \mathrm{W} \begin{cases} a_{ij} = 1 \; si \; w_{ij} \geq T_i \\ a_{ij} = 0 \; si \; w_{ij} < T_i \end{cases} \tag{2}$$

In other words, the characterization is performed using several transformations $\delta$ where the threshold $T_i$ is regularly increased at a rate of $T_{inc}$. Hence, given a set $T$, an element $T_i \in T$ is defined by a function $f: T \to T$

This approach allows a characterization that describes a list of transient characteristics of the dynamic evolution of the network, as shown in Fig. 2.

(a) (b) (c)

**Fig. 2.** Signature dynamic evolution represented on a network for the threshold $\boldsymbol{T_l}$: (a) $\boldsymbol{T_l = 0.1}$; (b) $\boldsymbol{T_l = 0.15}$ and (c) $\boldsymbol{T_l = 0.2}$

Therefore, once we obtained the networks from dynamic evolution, the proposed feature vector is the concatenation of all degree descriptors from each network. This descriptors were calculated using adjacency matrix $A$ which is modified by $\delta$ operator, for instance, for $T_l$ threshold we extracted the average $(k_\mu)$ and maximum degree $(k_\kappa)$, presented in previous section. Finally, degree normalization is required, which is computed using the following equation.

$$\forall k_i = \frac{k_i}{N} \tag{3}$$

This normalization was developed in order to reduce the influence of the quantity of the network nodes in descriptors computation. Therefore, after consider the network transformation for a $T_l$ threshold, the feature vector denoted by $\varphi$ is calculated as the concatenation of average $(k_\mu)$ and maximum degree $(k_\kappa)$ for each stage of the network evolution and thus we get the characterization proposed

$$\varphi = [\, k_\mu(T_0), k_\kappa(T_0), k_\mu(T_1), k_\kappa(T_1), \ldots, k_\mu(T_Q), k_\kappa(T_Q)] \tag{4}$$

The off-line signature characterization process is summarized in the following algorithm:

```
Notation:
            S   is the initial image to the signature
            V   is the set of nodes, each node is a pixel represented
                by its coordinates
            W   is the weighted adjacency matrix that represent the
                initial regular network generated from skeleton signa
                ture
            φ   is the feature vector that it's obtained by applying
                the algorithm to S, initially empty.
            δ   is the transform operation to the complex network

CHARACTERIZATION (S, φ)

1. Skeletonization of signature from S
2. Get the set of vertex V from skeleton obtained in S
3. Get the weighted adjacency matrix W from V
4. Normalize the weight in W in the interval [0 1]
5. For each threshold T_l from  T_0 to T_Q with  T_inc do steps 6 to 9
            Begin
6           Apply transform operation  δ_{T_i}(W) to get unweighted
            matrix A_{T_i}
7           Get avg. k_μ(T_l) and max. k_κ(T_l) degree
8           Append k_μ(T_l) and  k_κ(T_l) to φ
9           Normalize φ according the number of nodes N in [0,1]
            end
   end
10 End of algorithm
```

# 4 Results

To evaluate the performance of the feature vector proposed, we used a classification model in order to measure the similarity of features in the GDPS960 signature database. Experiments were performed using scale, rotation, noise and thickness variation. For that, we used a Multilayer Perceptron Neural Network —MLP— as machine learning and leave-one-out cross-validation to validate the model.

We conducted two types of experiments. The first was to measure the verification effectiveness of the algorithm and the second for identification. For verification, we defined two classes: one for genuine signatures and the other for forged signatures. For identification, we defined 960 classes, each class for a person, and each of them has only genuine signature of a person.

## 4.1 Verification Results

The efficiency histogram is shown on Fig. 3 which has high capability for discrimination between two persons. We had obtained rates up 98% and an average of 76% for verification.

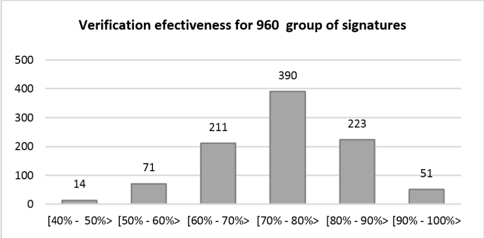


**Fig. 3.** Histogram of effectiveness for verification

### 4.2 Identification Results

We used whole genuine signature from each person for identification experiments, each group of genuine signatures represented a class, doing more complex the computation due to the 960 classes and 23040 instances. We obtained 85.12% of correct classifications for identification, with 14.88% of overall error, see Table 1. This high rate shows the effectiveness of the algorithm to identify signatures. The final confusion matrix is of 960-dimensional.

**Table 1.** Summary of success rate for identification of 960 persons

| Success rate | Percentage |
|---|---|
| Correct classification | 85.12% |
| Incorrect classification | 14.88% |
| Total Instances | 960 x 24 = 23040 |

## 5 Validation of Classification Model

We used as a measure to analyze the performance in pattern recognition the precision and recall. Figure 4 shows the precision-recall curve for identification which trend to be an horizontal line what demonstrate the classifier high performance. We computed 54 pair of this measures —a pair for each instance— for verification and 960 pairs for identification.

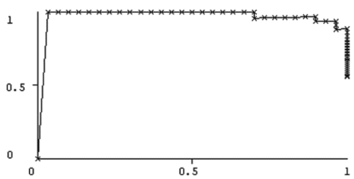


**Fig. 4.** Experiment with 85.12 % true classification to identification process with precision (x-axis) recall (y-axis) graphic

Figure 5 shows the 54 pairs average for the 960 experiments to verification signature. Each pair —precision and recall— for the 2 group of experiments — verification and identification— we obtained the average precision-recall curve showed in Table 2 and 5, respectively.

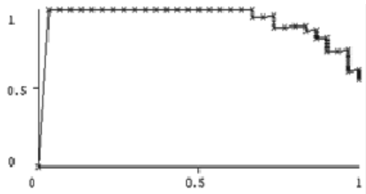


**Fig. 5.** Experiment with 76.11 % of true classification to verification process with precision (x-axis) recall (y-axis) graphic

**Table 2.** Details of precision and recall for class 0 (genuine signatures) and class 1 (forged signatures) of Fig. 4

| Class | TP | FP | Precision | Recall |
|---|---|---|---|---|
| 0 | 0.833 | 0.125 | 0.893 | 0.833 |
| 1 | 0.875 | 0.167 | 0.808 | 0.875 |
| **Avg.** | 0.852 | 0.144 | 0.855 | 0.852 |

**Table 3.** Details of precision and recall for class 0 (genuine signatures) and class 1 (forged signatures) of Fig. 5

| Class | TP | FP | Precision | Recall |
|---|---|---|---|---|
| 0 | 0.733 | 0.242 | 0.797 | 0.733 |
| 1 | 0.785 | 0.267 | 0.768 | 0.785 |
| Avg | 0.752 | 0.244 | 0.783 | 0.752 |

## 6 Discussion and Conclusions

Results were obtained using images with variation in scale, rotation, thickness and noise, which confirm invariance properties of complex network approach making it robust for shape representation. This is due to signature representation as a network or graph in which position of nodes is not important while connections and weights are well defined.

From the signature representation as a dynamic evolution of complex network, experiments were conducted for verification and identification obtaining an average effectiveness of 76% and 85%, respectively. Although there were cases with a low rate; this is a common problem due to limitations in intra-class variability.

To verify the performance of the classifier it was used the precision-recall metrics. Which confirm that features extracted are suitable for offline handwritten signature recognition. The area under the curve, showed in Fig. 4 and Fig. 5, tends to 1, which demonstrates its effectiveness. For identification process, we can see in Table 2 an average accuracy of 85.5% as overall result.

Finally, in this paper we presented an effective approach for offline handwritten signature characterization, for verification and identification processes, for which it has been applied the novelty complex networks approach.

## 7 Future Research

Complex networks have many properties that can be used as feature vector. This work allows to looking for new properties to represents signature shape. Also, we can extends the scope of research to mix complex networks approach with some *online handwritten verification* techniques in order to improve the accuracy of signature recognition.

## References


[1] Pavlidis, T.: A thinning algorithms for discrete binary images. Computer Graphics and Image Processing, 142–157 (1980)

[2] Dimauro, G., Impedovo, S., Pirlo, G.: A stroke-oriented approach to signature verification. In: From Pixels to Features III—Frontiers in Handwriting Recognition, pp. 371–384 (1992)

[3] Dimauro, G., Impedovo, S., Pirlo, G.: Component-oriented algorithms for signature verification. Pattern Recognition 8(3), 771–794 (1994)

[4] Herbst, N.M., Liu, C.N.: Automatic signature verification based on accelerometry. IBM J. Res. Dev. 21, 245–253 (1977)

[5] Brault, J.J., Plamondon, R.: Segmenting handwritten signatures at their perceptually important points. IEEE Trans. Pattern Anal. 15(9), 953–957 (1993)

[6] Shafiei, M.M., Rabiee, H.R.: New on-line signature verification algorithm using variable length segmentation and hidden Markov models. In: Proc. 7th Int. Conf. Doc. Anal. Recognit, vol. 1, pp. 443–446 (2003)

[7] Sabourin., R., Drouhard, J.P.: Shape matrices as a mixed shape factor for offline signature verification. In: Proc. 4th Int. Conf. Doc. Anal. Recognit, vol. 2, pp. 661–665 (1997)

[8] Al-Shoshan, A.I.: Handwritten signature verification using image invariant and dynamic features. In: Proc. Int. Conf. Comput. Graphics, pp. 173–176 (2008)

[9] Fierrez-Aguilar, J., Ortega-Garcia, J., Ramos, D.D., Gonzalez-Rodriguez, J.: HMM-based on-line signature verification: Feature extraction and signature modelling. Pattern Recognition 28(16), 2325–2334 (2007)

[10] Huang, K., Yan, H.: Offline signature verification using structural feature correspondence. Pattern Recognition 11, 2467–2477 (2002)

[11] Impedovo, D., Giussepe, P.: Automatic Signature Verification: The State of Art. IEEE Transactions on System, Man and Cybernetics 38(5), 609–635 (2008)

[12] Barabási, A.-L.: Linked: The New Science of Networks, USA: Perseus Books Group (2002)

[13] Blumenstein, M., Ferrer, M.A., Vargas, J.F.: The 4NSigComp2010 off-line signature verification competition. In: Proceedings of 12th International Conference on Frontiers Handwriting Recognition, Kolkata, India, pp. 16–18 (November 2010)

[14] Pavlidis, T.: Algorithms for Graphics and Image Processing, Germany. Springer (1982)

[15] Backes, A.R., Casanova, D., Bruno, O.M.: A complex network-based approach for boundary shape analysis. Pattern Recognition 42, 54–67 (2009)

[16] Rivard, D., Granger, E., Sabourin, R.: Multi-feature extraction and selection in writer independent off-line signature verification. IJDAR 16, 83–103 (2013)