

Handwritten Signature Verification by Independent Component Analysis

Marco Desira

*Department of Computer Science
St Martins Institute of IT*

Email: marco.desira@marcodesira.com

Kenneth P. Camilleri

*Department of Systems and Control Engineering
University of Malta*

And

*Department of Computer Science
St Martins Institute of IT*

Email: kpcami@eng.um.edu.mt

Abstract

Abstract This study explores a method that learns about the image structure directly from the image ensemble in contrast to other methods where the relevant structure is determined in advance and extracted using hand-engineered techniques. In tasks involving the analysis of image ensembles, important information is often found in the higher-order relationships among the image pixels. Independent Component Analysis (ICA) is a method that learns high-order dependencies found in the input. ICA has been extensively used in several applications but its potential for the unsupervised extraction of features for handwritten signature verification has not been explored. This study investigates the suitability of features extracted from images of handwritten signatures using the unsupervised method of ICA to successfully discriminate between different classes of signatures.

Index Terms

handwritten signature verification, independent component analysis, unsupervised learning

1. Introduction

Handwritten Signature Verification is the process of verifying whether the features of a handwritten signature can be classified as pertaining to the class of features of handwritten signatures of an individual.

The analysis of handwritten signatures is divided into two main categories, based on: the method used to capture the signatures; and subsequently the features which can be extracted for analysis. The analysis of features extracted from scanned images of handwritten signatures is referred to as off-line or static,

whereas the analysis of handwritten signatures which are captured via digitizing tablets, tablet PCs or other electronic devices which can capture the trajectory of handwriting is referred to as on-line or dynamic [7].

Off-line methods do not require the signer to be present during the verification process and rely only on features which can be extracted from the scanned signature image, and thus require more complex pre-processing steps [13]. Off-line methods do not capture dynamics of the signature such as pressure and velocity the lack of signer-sensitive characteristics which are captured in dynamic features makes off-line signature verification a more difficult challenge [9].

On-line signature verification methods have proved to be more accurate than off-line methods [11], [13] yet off-line signature verification systems are required when the signatory is not present at the verification stage as may be required for verifying signatures on bank cheques.

The process of signature verification should be able to detect forgeries. Forgeries can be classified into two main categories: casual or random forgeries; and skilled and traced forgeries [11].

Casual or random forgeries are created without prior knowledge about the appearance of the signature which is being forged. In this case all that the forger knows is the name of the person, whose signature is being forged. Substitution forgeries where the forger provides his own signature as a forgery are also classified as casual forgeries [1], [11]. Although random forgeries are less difficult to reject than skilled forgeries, random forgeries amount to over 90

In the case of skilled forgeries, the forger has access to one or more copies of the signature which is being forged. The forger has had time to practice the process of creating a forgery to best imitate the signature which is being forged [11]. Skilled forgeries are more difficult

to detect than random forgeries as the characteristic features of a skilled and traced forgeries resemble closely those of the imitated signature [10] thus making skilled forgeries more difficult to discriminate from authentic signatures.

In this study we propose the use of the unsupervised learning technique of independent component analysis (ICA) to extract structure directly from the signature image ensemble. Literature has shown that ICA can successfully extract a representation of data from natural mixtures such as those of sound sources and images. The properties of the independent components extracted by ICA of signatures were investigated and the outcome of the classification exercises which test for the acceptance of genuine signatures and for the rejection of random forgeries are discussed. Skilled forgeries are not used in this study.

2. Independent Component Analysis

Independent Component Analysis (ICA) is a method which transforms multivariate (multidimensional) data so as to make its essential structure more visible or more accessible, thus facilitating the analysis of the data. This representation is learnt from the data itself, thus the learning is unsupervised[12].

2.1. ICA and high-order dependencies

Independent Component Analysis learns both the high-order dependencies and the correlations found in the input, and is thus considered to be a generalisation of Principal Component Analysis [4], [5]. Independent Component Analysis attempts to place the ICA basis vectors in the direction of maximum statistical dependencies and does not restrict them to be orthogonal or orthonormal [4].

Higher-order methods such as ICA use information about the distribution of the input which is not found in the covariance matrix, and thus the distribution of the input is not assumed to be Gaussian [12].

One of the measures of nongaussianity is kurtosis or the fourth-order cumulant [12] (Equation 1). In this implementation we apply this measure to the independent components which are generated when ICA is applied to handwritten signature images. Gaussian random variables have a kurtosis of zero. We expect the kurtosis of independent components to be nonzero.

$$kurt(y) = E\{y\}^4 - 3(E\{y^2\})^2 \quad (1)$$

Barlow [3] argues that it is important for our perceptual system to detect suspicious coincidences which

are new statistical regularities gathered as inputs from the environment. Barlow proceeds to argue that in the visual input, edges constitute a suspicious coincidence. Barlow states that the detection of features by our visual cortex could be based on a redundancy reduction process which is essential to detect suspicious coincidences, such as edges which are localized and oriented filters.

Bell and Sejnowski [6] equate Barlows redundancy reduction problem to that of finding independent components by using a sigmoidal network of non-linear units to maximize information transfer [5]. In their experiments Bell and Sejnowski [6] show that the independent components resulting from ICA based on information maximization, are localized and oriented filters which are in fact edges.

In her study on Facial Image Analysis, Bartlett [4] explains that edges and elements of shape and curvature are examples of high-order dependencies, where the relationship amongst three or more pixels must be addressed. Bartlett argues that in tasks such as face recognition, important information may be found in the high-order relationships among the mage pixels.

3. Architecture for the Processing of Signatures

Bartlett [4] proposes two architectures for representing face images for the purpose of face recognition. The representation is based on statistically independent components which are derived from the image set. In this study of handwritten signatures we adopt one of the architectures, which is referred to as Architecture 1 in [4].

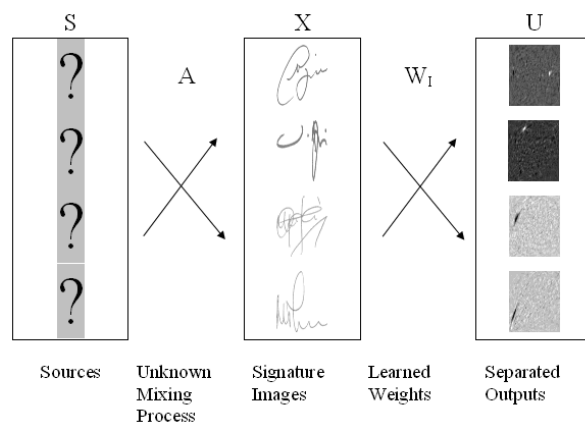


Figure 1. Image synthesis model adopted from [4].

Architecture 1 is characterized by the image synthesis model depicted in Figure 1.

We assume that each signature in \mathbf{X} is a mixture comprised of the independent components in \mathbf{S} . The observed mixtures in \mathbf{X} are a linear combination of the independent sources found in \mathbf{S} and the weights found in \mathbf{A} . By using the unmixing matrix \mathbf{W}_I , the mixtures in \mathbf{X} are transformed into approximations of the independent sources thus generating the rows of \mathbf{U} , where $\mathbf{W}_I\mathbf{X} = \mathbf{U}$.

The independent components in the rows of \mathbf{U} are a set of basis images that are statistically independent.. By using $\mathbf{A} = \mathbf{W}_I^{-1}$ as coefficients, the linear combination of the basis images (independent components) regenerates the observed mixtures i.e. the signature images (Figure 2 below).

$$\text{Signature} = b_1 * \mathbf{u}_1 + b_2 * \mathbf{u}_2 + \dots + b_n * \mathbf{u}_n$$

Figure 2. The linear combination of the basis images i.e. the linear combination of the independent components as adopted from [4].

Each row in \mathbf{A} contains the set of coefficients \mathbf{b} for one image \mathbf{x} . The coefficients in \mathbf{A} constitute the feature space which will be used for signature verification. The classification process adopted for signature verification in this study varies slightly from that used for face recognition in [4].

4. Preprocessing of the Signature Images

The different approaches to offline-signature verification have necessitated different preprocessing steps. These preprocessing steps prepare the signature images for the different type of processing as may be required at the feature extraction phase.

Some of the preprocessing steps encountered in the literature [13] included: removal of blank edges of the signature images (also referred to as data-area cropping); thresholding or binarization, noise reduction, removal of textured backgrounds, size normalization of the signature images, thinning of the signature strokes to a width of one pixel also referred to as skeletonization, close contour tracing; blurring; and rotation of the signature images.

Most of these preprocessing steps are computationally expensive. The preprocessing proposed in this study prior to executing the step of unsupervised learning by ICA entails: converting the signature images to their grey-level equivalent by linearly rescaling the luminance of each image to the interval [0,255]; data-area cropping and scaling By bicubic interpolation.

The signature images were scaled to 60 rows by 50 columns. Such a small size was selected so as to reduce the size of the covariance matrix which has to be created whilst computing PCA and during the whitening phase.

An original signature sample and a preprocessed sample are produced hereunder.



Figure 3. Signature prior to preprocessing on the left. Preprocessed image on the right.

4.1. Reducing the feature space

We perform PCA on the preprocessed images for the purpose of reducing the feature space. The pixel-grey values of the training image set are loaded into vectors \mathbf{x} , by using row-by-row scanning of each of the 60-by-50 pixel window. The rows of the images were concatenated thus producing 13000 dimensional vectors. The vector \mathbf{x} , is centered by subtracting its mean as defined in Equation 2 below.

$$\mathbf{x} \leftarrow \mathbf{x} - E\{\mathbf{x}\} \quad (2)$$

Reduction of the feature space is achieved by selecting m from n principal components, where $m \leq n$. The selection of the vectors is based on the amount of variance that is required to be represented in the application of use. PCA does not throw away high-order relationships, it simply does not separate them [4].

4.2. Whitening

Whitening or sphering is a preprocessing step which simplifies the ICA problem. Prior to executing the learning steps of ICA, the data in \mathbf{X} is passed through a zero-phase whitening filter \mathbf{W}_Z which is defined as twice the inverse square root of the covariance matrix of the data (Equation 3) [4]. Whitening removes the first and second order statistics of the data [4].

$$\mathbf{W}_Z = 2 * \langle \mathbf{X}\mathbf{X}^T \rangle^{-1/2} \quad (3)$$

5. ICA through Information Maximization

Different ICA methods exist. Each ICA method is comprised of an objective function which is implemented by an optimization algorithm [12]. The algorithms for implementing ICA include [12]: non-linear decorrelation algorithms; algorithms for maximum likelihood; algorithms for information maximization; non-linear PCA algorithms; the FASTICA algorithm; and tensor-based algorithms. This study makes use of information maximisation [5] as used in [4] to implement the ICA Architecture 1.

5.1. Training, Feature Extraction and Testing

Let \mathbf{X} be the matrix containing n training signature images in its rows. Let \mathbf{P}_m be the matrix containing the first m principal components of \mathbf{X} which have been selected as mixtures for input to ICA, where $m \ll n$. ICA is performed on $\mathbf{P}_m \mathbf{T}$ so as to produce m independent source images in the rows of the matrix \mathbf{U} . The steps to compute \mathbf{U} follow:

Let \mathbf{W}_X be the whitened version of \mathbf{P}_m^T . \mathbf{W}_X is used as input to the nodes with sigmoidal activation. The matrix of learned weights \mathbf{W} is produced. \mathbf{U} is derived as Equation 4 and 5 [4].

$$\mathbf{W}_I = \mathbf{W} * \mathbf{W}_z \quad (4)$$

$$\mathbf{U} = \mathbf{W}_I * \mathbf{P}_m^T \quad (5)$$

The principal component representation of the set of zero mean images in \mathbf{X} based on \mathbf{P}_m is \mathbf{R}_m defined by Equation 6 [4].

$$\mathbf{R}_m = \mathbf{X} * \mathbf{P}_m \quad (6)$$

The learning rate was initialized at 0.0005 and annealed down to 0.0001 as detailed in Table 1 below.

Table 1. Learning rates and associated iterations.

Learning Rate	Number of Iterations
0.0005	1000
0.0003	200
0.0002	200
0.0001	200

The feature set of the training data are obtained from the coefficients for the linear combination of the statistically independent sources \mathbf{U} as defined in Equation 7 [4], hence, the coefficients held in \mathbf{B} constitute the feature set of the training data.

$$\mathbf{B} = \mathbf{R}_m * \mathbf{W}_I^{-1} \quad (7)$$

The extraction of features from the test data set requires that the test data be centered by deducting the mean of the training data. A PCA representation of the test data is generated by using the principal components \mathbf{P}_m which were used during the training phase as defined in the equation below [4].

$$\mathbf{R}_{Test} = \mathbf{X}_{Test} * \mathbf{P}_m \quad (8)$$

The coefficients required to represent the test data in the ICA feature space are computed by applying Equation 9 [4].

$$\mathbf{B}_{Test} = \mathbf{R}_{Test} * \mathbf{W}_I^{-1} \quad (9)$$

5.2. A Similarity Measure

Evaluation of the performance of signature verification was carried out for the coefficient vectors \mathbf{b} , by the nearest neighbour algorithm based on the cosine of the angle (Equation 10) between the vectors.

$$d = \frac{\mathbf{b}_{test} \cdot \mathbf{b}_{train}}{\|\mathbf{b}_{test}\| \|\mathbf{b}_{train}\|} \quad (10)$$

Testing was carried out using the Leave-One-Out cross validation technique. Features of the test signatures were extracted as defined by Equation 9. Feature extraction from the test signatures was carried out by using the principal component representation of the data with which the testing was performed. Since the objective is to verify the signature, the class with which the test signature is expected to match is known a priori.

6. Implementation

ICA was performed to extract features from three sets of data each containing a different number of signature classes as detailed in Table 2. The details of Test C which yielded the best results are discussed below. The signatures were randomly selected from a database of signatures collated by Azzopardi [2].

Table 2. Training Signature Classes.

Test	Signers	Signatures per Signature Class	Total
A	5	5	25
B	37	6	222
C	26	24	624

6.1. Feature Extraction

The database of signatures contained 40 signatures for each signer. Twenty four signatures of 26 signers were randomly selected for the training phase. PCA of the 624 signatures was carried out. PCA generated 624 eigenvectors (principal components), of which the first 300 principal components captured 90% of the total variation in the data (see Figures 4 and 5 below). ICA was performed on the first 300 principal components.

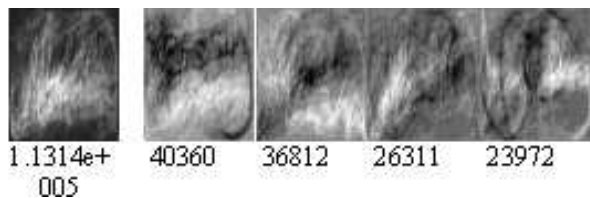


Figure 4. The 5 principal components with largest eigenvalues ordered by the magnitude of the corresponding eigenvalues. Eigenvalue displayed under each principal component.

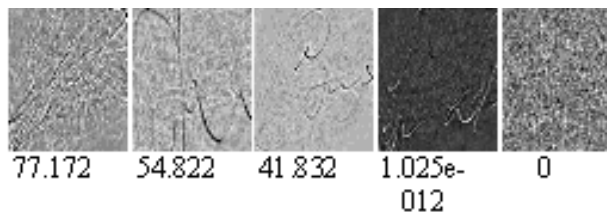


Figure 5. The 5 principal components with lowest eigenvalues ordered by the magnitude of the corresponding eigenvalues. Eigenvalue displayed under each principal component.

On comparing the independent components a sample of which is displayed below to the independent components generated in tests A and B there seems to be a higher degree of separation or there is less apparent mixture in the images of the independent components.

The vast majority of the ICs can be interpreted as bars or edges.

7. Signature Verification and Results

Signature Verification using the remaining 16 signatures of each of the 24 signers, totaling to 416 signatures is carried out against the set of features extracted from the 624 signatures.

Tests using a similarity measure based on the cosine between the coefficient vectors b were carried out. A False Rejection Rate (FRR) of 6.01

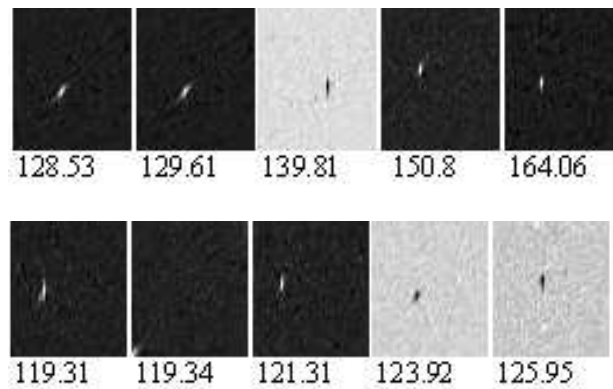


Figure 6. ICs having the highest level of kurtosis

In a study of off-line signature verification systems using HMM, Yacoubi et al [9] reported an FRR of 1.17

8. Conclusion

As we increased the number of signatures at the input a higher level of kurtosis was achieved. From the images of the independent components it was clear that a higher degree of separation was achieved where the kurtosis was further away from 0. Signature verification against features extracted by using ICA with the largest data set at the input yielded the best results. This study has shown that signature verification can be achieved by blindly obtaining the independent components of a set of signatures and classifying the ICA mixing coefficients without the need for detailed pre-defined extraction of features from the signature images, the latter being so far the classical approach towards signature verification.

References

- [1] R. Abbas, "Backpropagation networks prototype for off-line signature verification," *Minor thesis. Master of Applied Science in Information Technology*, 11–15, Jun. 1999.
- [2] G. Azzopardi, "How effective are Radial Basis Function Neural Networks for Offline Handwritten Signature Verification?," *Thesis, B.Sc. Computing and Information Systems, University of London.*, 2006.
- [3] H.B. Barlow, "What is the computational goal of the neocortex?," *Large -Scale Neuronal Theories of the Brain*, 1994.
- [4] M.S. Bartlett, "Face Image Analysis by Unsupervised Learning," *Kluwer Academic Publishers, Boston*, 2001.

- [5] [A.J. Bell,T.J. Sejnowski, "An information-maximization approach to blind separation and blind deconvolution."](#) *Neural Computation*, vol. 7, no. 6, pp. 1129–1159, 1995.
- [6] [A.J. Bell,T.J. Sejnowski, "The "Independent Components" of Natural Scenes are Edge Filters."](#) *Vision Research*, vol. 37, no. 23, pp. 3327–3338, 1997.
- [7] [G. Dimauro, S. Impedovo, S. Modugno,et al., "Recent Advancements in Automatic Signature Verification,"](#) *Proceedings of the 9th Intl' Workshop on Frontiers in Handwriting Recognition*, 2004.
- [8] [J. Drouhard,and R. Sabourin, "A Neural Network Approach to Off-Line Signature Verification using Directional PDF."](#) *Pattern Recognition*, vol. 29, no. -, pp. 415–424, 1996.
- [9] [A. El-Yacoubi,A. Justino,and R. Sabourin, "An off-line signature verification system using hidden Markov model and cross-validation."](#) *Proceedings XIII Brazilian Symposium on Computer Graphics and Image Processing*, 2000.
- [10] [B. Fang,Y.Y. Yang,C.H. Leung, "A smoothness index based approach for off-line signature verification,"](#) *Proceedings of the Fifth International Conference on Document Analysis and Recognition. ICDAR '99*, pp. 785–787, 1999.
- [11] [W. Hou,X. Ye,K. Wang, "A Survey of Off-line Signature Verification,"](#) *Proceedings of the 2004 International Conference on Intelligent Mechatronics and Automation*, pp. 536–541, 2004.
- [12] [A. Hyvrinen,J. Hyvrinen,E. Oja, "Independent Component Analysis,"](#) *Wiley, New York.*, 2001.
- [13] [F. Leclerc,R. Plamondon, "Automatic Signature Verification: The State of the Art -1989 - 1993,"](#) *International Journal of Pattern Recognition and Artificial Intelligence*, vol. 8, no. 3, pp. 643–660, Jun. 1994.
- [14] [K. Ueda, "Investigation of off-line Japanese signature verification using a pattern matching,"](#) *Proceedings. Seventh International Conference on Document Analysis and Recognition*, vol. 2, no. 3, pp. 951–955, 2003.
- [15] [A. Zimmer,L.L. Ling, "A Hybrid on/Off Line Handwritten Signature Verification System,"](#) *ICDAR'03*, vol. 1,pp. 424–428, Jun. 2003.