Offline Handwritten Signature Modeling and Verification Based on Archetypal Analysis

Elias N. Zois[1], Ilias Theodorakopoulos[2] and George Economou[2]

[1]Athens Technological & Educational Inst.
Agiou Spiridonos Str., 12243 Egaleo, Greece
ezois@teiath.gr

[2]University of Patras
Rio, 24504, Greece.
{iltheodorako, economou}@upatras.gr

Abstract

The handwritten signature is perhaps the most accustomed way for the acknowledgement of the consent of an individual or the authentication of the identity of a person in numerous transactions. In addition, the authenticity of a questioned offline or static handwritten signature still poses a case of interest, especially in forensic related applications. A common approach in offline signature verification system is to apply several predetermined image analysis models. Consequently, any offline signature sample which originates from either authentic persons or forgers, utilizes a fixed feature extraction base. In this proposed study, the feature space and the corresponding projection values depend on the training samples only; thus the proposed method can be found useful in forensic cases. In order to do so, we reenter a groundbreaking unsupervised learning method named archetypal analysis, which is connected to effective data analysis approaches such as sparse coding. Due to the fact that until recently there was no efficient implementation publicly available, archetypal analysis had only few cases of use. However, a fast optimization scheme using an active set strategy is now available. The main goal of this work is to introduce archetypal analysis for offline signature verification. The output of the archetypal analysis of few reference samples is a set of archetypes which are used to form the base of the feature space. Then, given a set of archetypes and a signature sample under examination archetypal analysis and average pooling provides the corresponding features. The promising performance of the proposed approach is demonstrated with the use of an evaluation method which employs the popular CEDAR and MCYT75 signature datasets.

1. Introduction

Biometric characteristics have been found to be utilizable in numerous conditions which eventually require the authentication of a person or evidence of his/hers deliberate consent [1]. In general, they can be partitioned into two main groups. The first addresses the verification of one person's physical attendance; physiological biometric traits like face [2], fingerprint [3] and iris [4] are typically utilized. The second group usually addresses the confirmation of either the attendance of a person or his/hers deliberate consent on a transaction; behavioral traits are usually employed like handwriting [5, 6]. The most acknowledged mean for validating the identity of a person is the handwritten signature. It is a private motoric pattern shaped from a potential mixture of letters and/or an intricate flourish [7]. The production of the signature is the combined result of a person's individual motoric process as it is portrayed by a trace onto a sheet of paper or an electronic device.

Signatures are a part of behavioral traits; hence they are carriers of intrinsic ambiguity. They are different even when they are created from the same person [6]. Upon request, forensic document examiners (FDEs) provide a justified estimation concerning the genuineness or not of a sample under question [8]. Automated signature verification (ASV/SV) systems [6] have been also employed in order to facilitate the verification of an individual with the use of machine vision and pattern recognition (PR) techniques. Bridging the gap between FDEs and ASVs presents several obstacles to overcome, as there are numerous terminologies and modalities which the PR and FDEs communities are treating in a dissimilar way [9, 10, 11, 12, 13, 14].

A key issue of an SV system is the feature extraction procedure, which converts the input signature image into a set of numbers-features that eventually form now a new feature space. This conversion preferably must preserve all the essential intrapersonal information, which is vital for the subsequent verification step. A substantial amount regarding feature extraction methods depends on the utilization of global and/or local signature descriptors [15]. In particular, local features tend to exploit statistics between the pixels-members of the signature image. During the preceding decade, a variety of methods which employ local feature analysis for offline SV has been presented with notable results for verification tasks [16], [17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28]. An

important issue that is encountered often in offline signature feature extraction is that the conversion of a signature sample, emanating from genuine persons or imitators, involves a fixed image analysis model. Following are the novelty characteristics of the proposed approach:

1. The feature extraction procedure adapts to a narrow set of randomly selected genuine reference samples, making it attractable to forensic cases. Thus, for any questioned signature sample, the extracted features exploit the relation between it and a representation of genuine samples.

2. In order to provide a mean for the representation of the reference signatures we reenter a groundbreaking unsupervised learning method named archetypal analysis [29], which is connected to effective data analysis approaches such as sparse coding [30]. Due to the fact that until recently there was no efficient implementation publicly available, archetypal analysis had only a few cases of use [31, 32, 33]. Recently, a fast optimization scheme using an active set strategy was presented [34] and an efficient open-source implementation interface has been also provided to facilitate computations [35].

The main goal of this work is to introduce archetypal analysis for offline signature verification. The output of archetypal analysis performed on a collection of a user's signature samples is a set of learned archetypes, constituting the model of the particular individual's signature. Archetypes are a special case of dictionary elements, which in the context of this work are used in order to form a feature space for the representation of test samples. Thus, given a set of archetypes and a signature sample under examination, archetypal analysis is used again, only in that stage it provides the corresponding representation coefficients, in a manner similar to other popular coding schemes (i.e. LLC, sparse coding etc.). In the presented approach, archetypal analysis is used as a means to learn and compute efficient local features, operating on a patch-level rather that in whole-image level. The global signature descriptor is derived through average pooling of the local features, over a specially designed 2-level spatial pyramid.

The rest of the paper is organized as follows. Section 2 presents the basic postulates of the archetypal analysis. Section 3 provides the systems architecture, establishes the link between archetypal analysis and signature verification and provides the proposed feature extraction method. Section 4 reports experimental methods and results. Finally, section 5 provides the conclusion.

2. Postulates of Archetypal Analysis

Following the seminal work of Chen et al. [34] let us consider a matrix $X=[x_1,x_2,...,x_n]\in R^{m\times n}$ in which each x_i-column is a vector which belongs to R^m. The role of the x_i-vector is to denote any data point in an m-dimensional data space. Archetypal analysis learns a factorial representation of X by solving a corresponding archetypal problem. Specifically, it looks for a set $\mathbf{Z}$ of p-archetypes in which each z_j-column is a vector which also belongs to R^m. These $z_j,\ j=1:p$ column vectors are denoted as the archetypes; they are elements which are formed in order to comply with two geometrical constraints. The first one, points out that, each x_i-vector must be approximated adequately by a convex combination of some archetypes z_j, while the second one, addresses a dual constraint; each z_j-archetype must also be approximated by a convex combination of the x_i-vector. Consequently, given a set of archetypes $\mathbf{Z}$ each x_i-vector should be approximated by the product Za_i, where $a_i, i=1:n$, $a_i\in R^p$ is a coefficient column vector which resides in the simplex Δ_p:

$$\Delta_p \triangleq \left\{a_i\in R^p \ \text{ s.t } a_i\geq 0 \text{ and } \textstyle\sum_{j=1}^{p} a_i(\mathrm{j})=1\right\}, \forall i\in[1:n] \quad (1)$$

Likewise, each archetype z_j should be approximated by the product $X\beta_j$, where $\beta_j, j=1:p$, $\beta_j\in R^n$ is another coefficient column vector which resides in the simplex Δ_n:

$$\Delta_n \triangleq \left\{\beta_j\in R^n \ \text{ s.t } \beta_j\geq 0 \text{ and } \textstyle\sum_{i=1}^{n} \beta_j(\mathrm{i})=1\right\}, \forall j\in[1:\mathrm{p}] \quad (2)$$

The resulting formulation is the following problem of minimizing the residual sum of squares (RSS):

$$\min(\mathrm{RSS}) \triangleq \min_{a_i\in\Delta_p \text{ for } i=1:\mathrm{n}} \sum_{i=1}^{n} \|x_i - Za_i\|_2^2 \quad (3)$$

where $\forall j,\ z_j = X\beta_j$, which in the matrix factorization form is equivalent to:

$$\min_{\substack{a_i\in\Delta_p \text{ for } i=1:\mathrm{n}\\ \beta_j\in\Delta_n \text{ for j=1:p}}} \|X - XBA\|_F^2, \quad (4)$$

with,

$A=[a_1,a_2,...a_n]\in \mathrm{R}^{p\times n}$, $B=[\beta_1,\beta_2,...\beta_p]\in \mathrm{R}^{n\times p}$, $\mathrm{Z}=X\,\mathrm{B}$

The formulation of eq. (3) is a non-convex optimization problem, but it is convex with respect to one of the variables A or B, when the other is kept fixed. Then, a block-coordinate descent scheme guarantees in a asymptotically way a stationary point of the problem [36]. When fixing all variables except one column a_i of A and minimizing with respect to a_i, the problem is solved with

the use of a quadratic program:

$$\min_{a_i \in \Delta_p} \sum_{i=1}^{n} \|x_i - Za_i\|_2^2 \quad (5)$$

In a similar way, when fixing all variables but one column β_j of B, the problem is again solved with the use of a quadratic program:

$$\min_{\beta_j \in \Delta_n} \left\| X - XB_{old}A + X(\beta_{j,old} - \beta_j)a^j \right\|_F^2 \quad (6)$$

where $\beta_{j,old}$ is the current value of β_j, and $a^j \in R^{1\times n}$ corresponds to the j-th row of A. Both situations as described by eqs. (5)-(6) share the same form of least-squares optimization problem equipped with a simplicial constraint. In order to proceed with an intuitional description of the solution we recall some elements from the sparse coding theory [30]. In particular, eq. (1) states that each x_i -vector should be approximated by the product Za_i, under the constraints a) that the sum of any a_i coefficients is equal to one, and b) $a_i \geq 0$. Consequently, this is a direct example of the utilization of the l_1 -norm of the a_i coefficients constrained to be one. Thus, archetypal analysis provides sparse approximation of the data and the z_j -archetypes act similar to the dictionary elements d_j in the following sparse coding formulation [30]:

$$\min_{\substack{D\in R^{m\times p} \\ A\in R^{p\times n}}} \frac{1}{2}\|X - DA\|_F^2 + \lambda\|A\|_1 \text{ s.t. } \|d_j\|_2 \leq 1, \ \forall j \quad (7)$$

Therefore, the main difference between sparse coding and archetypal analysis (aside from the non- negativity of a_i) is that archetypes must be convex mixtures of the data points X. Consequently, the vectors β_j are also constrained to be in the simplex Δ_n, something that provides them with the sparsity attribute. Then, an active set algorithm, [37] can be used to efficiently solve the quadratic programs provided by eqs. (5) - (6). In our experiments, a dedicated fast and accurate implementation of the SPAMS toolbox [35], has been employed in order a) to realize the aforementioned evaluation of the reference-Z^{REF} archetypal matrix given the union of the genuine reference signature set $X^{REF} = \bigcup X^{(1^{st},2^{nd}.3^{rd},\ldots)}$ and b) the coefficients: $A^{(sample)}$, $B^{(sample)}$ given the reference archetypal matrix Z^{REF} along with any other input signature. The number of iterations for archetypal analysis was set to 100, which has been reported to provide a good performance in terms of time and accuracy in several experiments [34].

At this point we present some necessary terminology in order to facilitate the readability of the following paragraphs. The $X \in R^{m\times n_k}$ data matrix contains the union of the n_k -signature pixels $\{x_i \in R^m\}$, $i = 1:n_k$ information, where m represents the raw data dimensionality; the value of m is directly related to the selection of the granularity of the image probing element, defined hereafter as patch size N_{sp}. The value of n_k corresponds to the total number of the signature pixels of the k^{th} signature sample, while the value of p represents the selection of the number of archetypes. Summarizing, each z_j archetype becomes a convex combination of a few data points only, which is useful for interpreting z_j by means of the information stored on the non-zero entries in β_j. Clearly these values specify the analogy in which an input data point $\{x_i \in R^m\}$ is associated to the z_j archetype. On the other, each non-zero entry stored in the a_i coefficients provides information regarding the use of each z_j archetype for efficient reconstruction of the x_i signature input data point.

3. System design

The proposed architecture is summarized in figure 1. During the first phase of the learning stage-enrollment, a population of $Ngen$ -reference genuine signature samples is enrolled. Then, the union of the signature pixels of all reference images forms the matrix $X^{REF} \in R^{[m\times\sum_{k=1}^{Ngen} n_k]}$. With the use of eq. (3) the Z^{REF} reference matrix is evaluated and stored as the representative entity of a set of reference signatures.

In the second phase of the learning stage, which includes training, for each of the genuine reference input

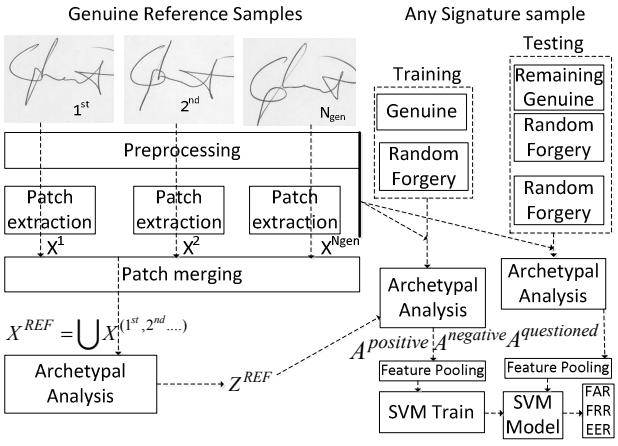

Figure 1: Proposed system architecture.

signatures along with a number of random forgeries i.e. other writers genuine signatures, similar data matrices X^T are formed, and analyzed into the corresponding A^T matrices by keeping Z^{REF} fixed while solving eq. (3). Average pooling across the rows of the A^T matrices creates the signature features with dimensionality proportional to the value of p. Next a binary SVM classifier [10] is built with the genuine reference features as the positive class and the random forgeries features as the negative class.

In the testing stage, any questioned input signature sample and corresponding data form X^Q is analyzed with the use of average pooling on the derived A^Q matrix that provides the associated feature vector. In turn, the classifier provides a score, which will be used in order to assess the verification efficiency by means of receiver operating characteristics (ROC) or detection error tradeoff (DET) analysis. In our study the equal error rate (EER), which is the point in which the false acceptance rate (FAR) and the false rejection rate (FRR) are equal, has been chosen as the performance measure.

3.1. Preprocessing and patch extraction

Typical preprocessing includes thresholding, thinning and equimass segmentation [38, 39] into four equimass segments, denoted hereafter as $EMS_g, g = 1:4$ in order to distinguish them from the entire signature image $IMG^{1\times1}$. Following the description provided in the previous section and prior to the Z^{REF} estimation, the format of the matrix $X \in R^{m\times n}$ of any input signature image is clarified. Following, for every *i*-pixel with *(r,c)*-coordinates being part of entire signature trace $IMG^{1\times1}$ an elementary rectangular window of size $m = N_{sp} \times N_{sp}$ denoted

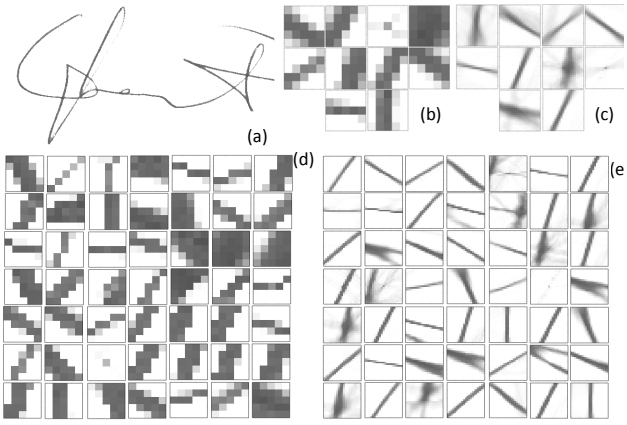

Figure 2: a) A signature image. b-e) Archetypes for parameters (N_{sp}, p): (5,10), (20,10), (5,49), (20,49).

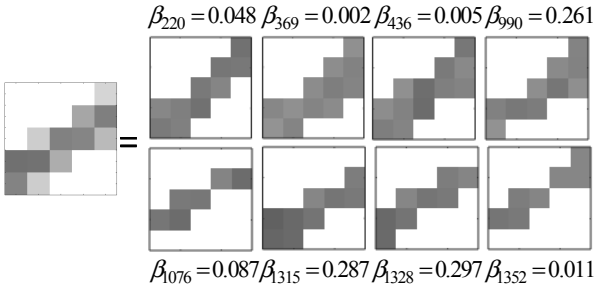

Figure 3: An archetype and its convex decomposition into 8 components and corresponding β_j weights. (N_{sp}, p): (5, 10)

hereafter as the patch $p_{Nsp}(r,c) \triangleq p^i_{Nsp}$, is imposed in order to locate and store the local neighborhood pixel intensities. Next, the patch p^i_{Nsp} is stored with a one-dimensional column vector format $p^{i,1D}_{NSp}$, $i = 1:n_k$. The concatenation of all the image patches of a specific signature image $\bigcup_i p^{i,1D}_{NSp}, i = 1:n_k$, ultimately defines the matrix of the specific signature sample as: $X^k = \bigcup_i p^i_{Nsp}, i = 1:n_k$.

3.2. Application of archetypal analysis

In accordance to the exposed material so far, the genuine reference data matrix is defined by the union of the patches of all reference signature images $X^{REF} = \bigcup_k X^k$, $k = 1:Ngen$. Then, given X^{REF}, along with the number of archetypes p, eq. (3) is solved in order to provide the Z^{REF} matrix of archetypes. In order to illustrate the method, figure 2 presents the derived archetypes for the case of one signature sample for various values of the patch size N_{sp} and the number of archetypes p as a system parameter. Additionally, figure 3 displays an example of the way that one archetype is

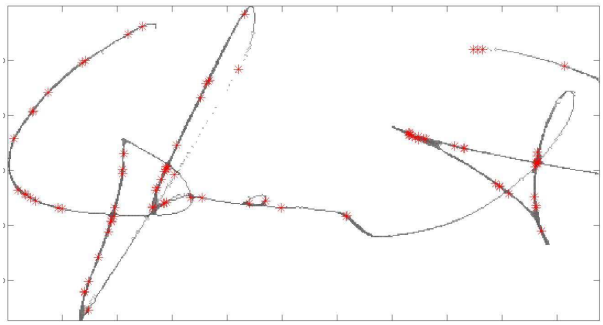

Figure 4: Location of patches (marked with red color) on a signature image which are forming the set of archetypes. (N_{sp}, p) (5, 10)

synthesized from some convex combination of patches and figure 4 provides an illustrative example of the location of the patches, used to form the entire set of archetypes, on the signature. The minimum value of the patch size is usually set such that to confine a minimum amount of the signature stroke, after the thinning operation. For example, the CEDAR signature dataset [40] contains signatures, which after a minimal thinning-trimming result into traces that fit within a patch size $N_{sp}=5$, thus enabling the detection of signature microstructures. Consequently, the minimum patch size along with the corresponding data dimensionality m equals to $N_{sp}\times N_{sp}=25$. The maximum patch size clearly can be chosen according to the amount of the signature trace that one wishes to capture and contains more shaped characteristics.

An interesting analysis has been drawn in order to investigate the effect of the number of patches per archetypes as a function of the patch size – dimensionality of the problem N_{sp} and the number of archetypes p. Figure 5 presents graphically the results of this analysis for one writer although similar patterns apply for any other writer. According to it, the average number of signature patches which participates to the forming of the archetypes decreases as the number of archetypes is increasing while satisfies the sparsity property. Additionally, given that the number of archetypes p is kept fixed, then the number of average number of signature patches that forms the archetypes increases as a function of the patch size N_{sp}. Furthermore, an additional analysis of the magnitude of the β_j coefficients indicates that if the patch size N_{sp} is fixed then the increase of the number p of requested archetypes leads to a lesser number of signature patches that forms the archetypes, but with β_j magnitudes of lower uncertainty. This is demonstrated with the use of the following measure. Since a) $B=[\beta_1,\beta_2,...\beta_p]\in \mathrm{R}^{n\times p}$, b) the sum of each β_j coefficients is one and c) $\beta_j\geq 0$ one can make use of the entropy definition:

$$H(\beta_j)=-\sum_{i=1}^{n_k}\beta_j(i)\times\log_2(\beta_j(i)),\ \forall j, j=1:p \qquad (8)$$

in order to provide a qualitative measure of the uncertainty which accompanies the formation of the archetypes from each participating patch. Figure 6, presents graphically the evolution of the uncertainty of the participating patches, as expressed with the average entropy of the archetypes $\overline{H}=\frac{1}{p}\sum_{j=1}^{p}H(\beta_j)$, as a function of a) the number of archetypes p and b) the patch size N_{sp}. Inspection of both figures 5 and 6 indicates that as the number of archetypes increases, the number of participating patches and their average entropy $\overline{H}$ decreases.

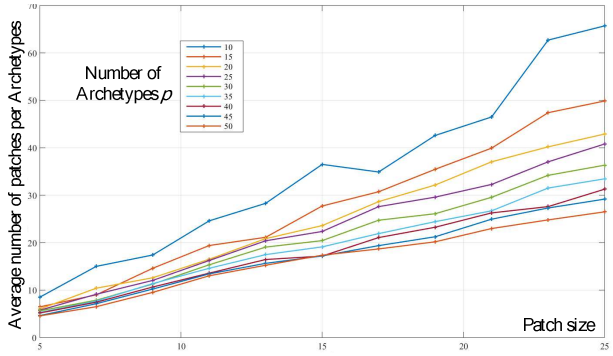

Figure 5: Example of the average number of patches per archetype as a function of the patch size and the number of archetypes for one writer.

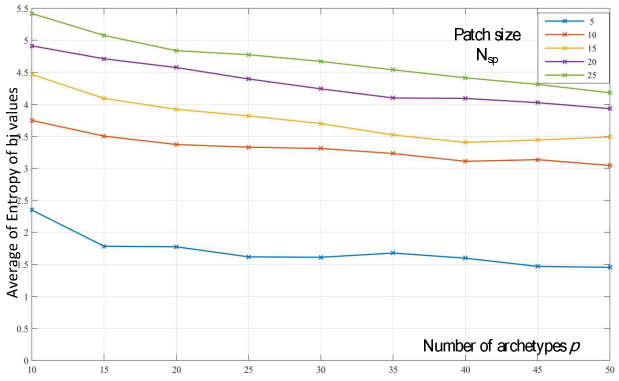

Figure 6: Average of entropy of the β_j values as a function of the number of archetypes and the patch size for one writer.

3.3. Feature extraction

As figure 1 describes, the genuine reference samples for one writer are forming the model of his handwriting style by means of the claimed identity as it is represented by the Z_{writer}^{REF} archetype matrix. For any other signature sample Q claiming this identity, the feature extraction stage utilizes its set of $X^Q\in R^{m\times n_Q}$ patches and the Z_{writer}^{REF}. The solution of eq. (3) provides a) the $A_{1\times 1}^Q\in R^{p\times n_Q}$ coefficient matrix of the entire signature in order to provide its global representation. Additionally, the $X_{EMSg}^Q\in R^{m\times n_{EMSg}}$, $g=1:4$ image patches, which belong to the corresponding EMS_g image segments, are also analyzed and four corresponding matrices $A_{EMSg}^Q\in R^{p\times n_{EMSg}}$ also provide localized information. In this work, inspired by [35], the feature vector $f^Q\in R^{5p\times 1}$ is simply formed by a) average

pooling of the A coefficients of each archetype and b) concatenation of the resulting values for all segments. Let us denote with the notation: $f^{Q}=\left\{f^{Q}_{1\times 1}\in R^{p\times 1},\{f^{Q}_{EMSg}\in R^{4p\times 1}\}\right\}$ the features which represents any Q-signature sample where each j-feature component is evaluated as:

$$f^{Q}_{1\times 1}(\mathrm{j})=\frac{1}{\left|IMG^{1\times 1}\right|}\sum_{c=1}^{\left|IMG^{1\times 1}\right|}\mathbf{A}^{Q}_{1\times 1}(\mathrm{j},\mathrm{c}),j=1:p$$
$$f^{Q}_{EMGg}(\mathrm{j})=\frac{1}{\left|EMS_{g}\right|}\sum_{c=1}^{\left|EMS_{g}\right|}\mathbf{A}^{Q}_{EMSg}(\mathrm{j},\mathrm{c}),j=1:p,g=1:4 \quad (9)$$

It is easily deducted that the dimensionality of the average feature vector is directly related to the number of selected archetypes p.

4. Experimental protocol

In order to demonstrate and evaluate the efficiency of the proposed method, the CEDAR [40] and MCYT75 [41] signature datasets, which have been widely used in signature verification research were adopted. The CEDAR dataset contains offline signatures from 55 volunteers. For each writer, 24 genuine and 24 skillfully forged signatures were provided (i.e., 1320 genuine and 1320 forged signatures in total). The scanned signatures are composed of 8-bit gray-level images at 300 dpi. The second signature database used was the off-line version of the MCYT75 signature database. A whole of thirty (15 genuine and 15 simulated) signature samples were recorded for each one of the 75 enrolled writers at a resolution of 600 dpi.

For each one of the writers a specific model is built by randomly gathering $Ngen$ reference signature samples. The number of $Ngen$ in this work has been set to five ($Ngen$=5) for representation of cases in which only a few samples are available. Then, the Z^{REF}_{writer} archetype matrix is evaluated from all the genuine reference samples and henceforth models the handwriting of the specific writer. In the training procedure, each one of the $Ngen$ genuine samples is once more analyzed with the use of eq. (3) in order to provide the positive class $\omega^{\oplus}\in R^{Ngen\times(5p)}$ with the feature extraction procedure of eq. (9). The negative class $\omega^{-}\in R^{10\times(5p)}$ is composed from 10 out of 54 or 74 samples, by taking one random sample from 10 out of the remaining writers. Thus, a corresponding learning feature population $[\omega^{\oplus}\omega^{-}]\in R^{(\mathrm{N}gen+10)\times(5\mathrm{p})}$ is used as an input to a binary SVM classifier, with a radial basis, while a holdout cross-validation procedure returns the optimal values of the C^{opt} and gamma-γ^{opt} parameters with respect to a maximum cross validation value of the Area Under Curve. In addition, the cross-validation procedure returns the SVM output scores conditioned on the positive $\omega^{\oplus}$ class samples $CVS^{\oplus}$.

The testing stage utilizes primarily the remaining genuine signatures and the skilled forgeries (S). The receiver operating characteristic parameters FAR(S) and 1-FRR are computed as a function of the sliding threshold whose extremes are the minimum and maximum values of the $CVS^{\oplus}$. The EER: FAR(S) = FRR is then computed and reported. Additionally, at the threshold point of the EER the genuine samples of the remaining writers are used for the evaluation of the random forgery-(R) FAR(R) error rate. The experiments were repeated ten times and their average values are reported.

4.1. Overall performance

The influence of two critical parameters, namely the patch size N_{sp} and the number of archetypes p is examined and reported by allowing several values. Specifically, for the patch size N_{sp} two distinct values were selected in order to define two major categories. The first utilizes a patch size of five (5), which corresponds to the category of analyzing the signature trace by microstructural probing; the second category utilizes a patch size of nineteen (19) which corresponds to the case of having archetypes representing elongated shape characteristics. Regarding the value of p, we employed values starting from ten (10) up to fifty (50) with a step of 10 and corresponding feature dimensionality of 50 – 250

TABLE 1: ERROR RATES (%) FOR THE PROPOSED METHOD. NGEN=5. CEDAR DATASET

N_{SP}	Number of archetypes p									
	10		20		30		40		50	
	EER	FAR(R)	EER	FAR(R)	EER	FAR(R)	EER	FAR(R)	EER	FAR(R)
5	3.10	0.93	2.83	0.45	2.07	0.17	2.14	0.19	2.26	0.23
19	5.36	1.22	5.58	0.96	6.14	0.98	6.02	0.88	6.12	1.22

TABLE 2: ERROR RATES (%) FOR THE PROPOSED METHOD. NGEN=5. MCYT75 DATASET

N_{SP}	Number of archetypes p									
	10		20		30		40		50	
	EER	FAR(R)	EER	FAR(R)	EER	FAR(R)	EER	FAR(R)	EER	FAR(R)
5	6.24	0.96	4.72	0.73	3.97	0.23	4.08	0.25	4.31	0.26
19	7.57	0.79	5.59	0.63	4.99	0.37	5.15	0.45	5.46	0.61

($5p$). Table 1 and table 2 depict the error rates of the proposed method. Examination of these results reveals that the analysis of a signature with the use of archetypes which are synthesized from signature patches of small size, provide superior error rates when compared to the ones derived with the use of archetypes created from larger signature patches. This is a potential indication that larger shape archetypes are more sensitive to the intra-class handwriting variability, especially for the unstable CEDAR dataset [22, 27]. Regarding the observed high efficiency of the small patch size, examination of the associated plots of figure 6, shows that the use of a small patch size is related with smaller amounts of mean entropy $\overline{H}$. Consequently, when a small patch size is selected, a small fraction of signature patches is employed, but with lesser uncertainty, in order to assemble the archetypes. Again, one may assume that the archetypes, expressed as a convex combination of signature patches with lower entropy carry an inherent property of the signature generation process. These archetypes are employed to analyze any Q-input signature by means of the coefficient A^Q matrix. Potentially, the $a_i^Q \in R^p$, $i=1{:}n_Q$ coefficients and the corresponding features, map this inherent property regarding the generation of each archetype. Another observation which can be drawn from table 1 is that, for the case of using a small valued patch size, the verification accuracy, tends to increase when the number of archetypes p equals the dimensionality of the patch data m.

4.2. Comparisons with state-of-the-art

The comparison of the derived results with other, state of the art systems, is generally considered a rather hard task given the various degrees of freedom regarding the type or number of signatures utilized during the classifier construction and evaluation [28]. Nevertheless, Table 3 and Table 4 provide evidence that the proposed method achieves a low error of verification when a few genuine samples are available which is at least comparable to the ones derived from state of the art methods. It is also important to mention the fact that the training stage of our method includes only genuine and random forgeries, and not any sort of skilled forgeries while testing in a common context both genuine, random and skilled forgeries.

4.3. Comparison to sparse coding

Finally, it is very interesting to explore other popular variations from the general family of coding problems similar to that of eq. (7), involving different constraints to the dictionary and coefficients. Due to page limitations

TABLE 3: ERROR RATES (%) OF THE PROPOSED METHOD COMPARED TO OTHER SYSTEMS WITH THE CEDAR DATASET.

First Author	Method	#Signatures for training	FRR	FAR (R)	FAR (S)
Kumar R. [22]	Surroundness	24 genuine + 24 forgeries	8.33	-	8.33
Kumar M. [42]	Chord Moments	16 genuine + 16 forgeries	6.36	-	5.68
Guerbai [43]	Curvelet Transform	12 genuine	5.60	-	5.60
Serdouk [24]	Gradient Local Binary Patterns + LRF	16 genuine	2.12	-	4.93
Kalera [40]	Gradient, Structural and Concavity	16 genuine + 16 forgeries	21.9	-	21.9
Srihari [44]	Gradient, structural, and concavity	16 genuine + 5 forgeries	8.5	-	10.1
Serdouk [45]	Gradient Local Binary Patterns	16 genuine + 16 forgeries	4.31	-	6.36
Okawa [28]	Bag of Visual Words with KAZE	16 genuine + 16 forgeries	1.60	-	1.60
Zois [27]	Partially Ordered Sets	5 genuine +10(R) forgeries	4.44	1.61	15.9
Ganapathi [46]	Gradient Direction Histogram	14 genuine + 14 forgeries	6.01	-	6.01
Shekar [47]	Local Morphological Pattern Spectrum	16 genuine + 15 forgeries	9.58	-	9.58
Hamadene [26]	Directional Co-occurrence Matrix	5 genuine	4.21	-	0
Zois [52]	K-SVD dictionary learning – OMP	5 genuine+10(R)forgeries	7.32	0.34	6.83
Proposed	**Archetypes (p=30, N_{sp}=5)**	**5 genuine+10(R)forgeries**	**2.07**	**0.17**	**2.07**

TABLE 4: ERROR (%) RATES OF THE PROPOSED METHOD COMPARED TO OTHER SYSTEMS WITH THE MCYT75 DATASET.

First Author	Method	#Signatures for training	FRR	FAR (R)	FAR (S)
Vargas [20]	Local binary patterns	10 genuine + 74(R)forgeries	8.59 5.07	- 0.15	6.77 -
Zois [27]	Partially Ordered Sets	5 genuine + 10(R)forgeries	4.48	1.50	25.19
Fierrez-Aguilar [19]	Global and local slant	10 genuine	9.28 1.15	- 1.15	9.28 -
Alonso-Fernandez [49]	Slant and envelope	10 genuine	22.9	2.67	22.0
Wen [50]	Invariant ring peripheral	5 genuine	15.0	-	15.0
Yin Ooi [51]	Discrete Radon transform	5 genuine	13.86	-	13.86
Soleimani [48]	Deep multitask metric	10 genuine + 74(R)forgeries	9.86	0.37	9.86
Zois [52]	K-SVD dict. learn. – OMP	5 genuine + 10(R)forgeries	7.32	0.35	10.40
Proposed	**Archetypes (p=30, N_{sp}=5)**	**5 genuine + 10(R)forgeries**	**3.97**	**0.23**	**3.97**

though, we have limited our study to the other extreme of the sparse coding problem defined as:

$$\min_{D,A} \frac{1}{2}\|X - DA\|_F^2 \text{ s.t. } \|a_i\|_0 \le s \le p, \ \forall i \qquad (10)$$

where the only constraint is essentially the sparsity s of the representation coefficients. The solution to the above problem can be approximated using greedy techniques. In reference [52] the K-SVD and OMP [53] algorithms was employed to solve the dictionary construction and sparse coding problems (10) respectively, in a greedy fashion. Results in [52] derived for both CEDAR and MYT75 datasets, are directly comparable to that of archetypal analysis, since the exact same feature pooling scheme, patch size and experimental protocol was utilized, and the dictionary size (60 atoms) is rather similar.

From the corresponding results of Tables 3 and 4 it is evident that the convexity constraints imposed on archetypes and coefficients come to a great benefit to the discrimination capacity of the final descriptor. Although standard sparse coding still exhibits a good ability to express variations of microstructures of different signatures, thus delivering a good FAR on random forgeries, it significantly lacks of discriminative capacity when it comes to the more fine-grained distinction between genuine and skilled forgeries exhibiting larger FAR(S). We believe this behavior has its root to the degenerate structure of the manifold where patches of signatures lie on, in contrast to the corresponding manifold of natural-images, for which standard sparse coding reportedly exhibits [34] better performance on image classification tasks under similar settings.

5. Conclusion

The offline handwritten signature forms a particular gender of image signals which exhibit a degenerate structure and therefore lie on a low dimensional subspace. On the other, parsimony, a biologically inspired notion has been lately exploited in a plethora of pattern recognition, machine learning, and computer vision applications. The proposed method can be summarized as follows:

1) The use of archetypal analysis for establishing primitives, defined as archetypes which can be used as a mean to represent the handwritten strokes.

2) The creation of a reference archetypal matrix which represents a population of some reference genuine samples.

3) The use of average pooling in order to provide features for offline signature verification.

4) Experimental results show that this analysis further improves the verification error rates on two popular signature datasets.

Our future research plan among others includes the use of additional signature datasets, the use of several other pooling feature extraction techniques, and the use of individual reference samples, instead of a reference set as a whole for the construction of the classifier.

References

[1] A. K. Jain, A. A Ross, and K. Nandakumar, Introduction to Biometrics, pp. 1-47, New York: Springer, 2011.

[2] A. Z. Li, and A. K. Jain, Handbook of face recognition, 2nd ed., pp. 1-15, New York: Springer, 2011.

[3] D. Maltoni, D. Maio, A.K. Jain and S. Prabhakar, Handbook of fingerprint recognition, pp. 1-56, London: Springer-Verlag, 2009.

[4] J. Daugman, Iris Recognition, Handbook of Biometrics, A. K. Jain, P. Flynn and A. A. Ross, eds., pp. 71-90, Boston, MA: Springer US, 2008.

[5] R. Plamondon, and S. N. Srihari, Online and off-line handwriting recognition: a comprehensive survey, IEEE Transactions on Pattern Analysis and Machine Intelligence, 22(1):63-84, 2000.

[6] R. A. Huber, and A .M. Headrick, Handwriting identification: facts and fundamentals, pp. 102, Boca Raton: CRC Press, 2010.

[7] H-L. Teulings, Handwriting Movement Control, pp. 561-614, London: Academic Press, 1996.

[8] L. C. Alewijnse, C. E. van den Heuvel, and R. D. Stoel, Analysis of signature complexity, Journal of Forensic Document Examination, 21:37-49, 2011.
[9] J. Coetzer, B. M. Herbst, and J. A. d. Preez, Offline Signature Verification Using the Discrete Radon Transform and a Hidden Markov Model, EURASIP Journal on Advances in Signal Processing, 2004(4):1-13, 2004.
[10] E. J. R. Justino, F. Bortolozzi, and R. Sabourin, A comparison of SVM and HMM classifiers in the off-line signature verification, Pattern Recognition Letters, 26(9):1377-1385, 2005.
[11] M. Hanmandlu, M. H. M. Yusof, and V. K. Madasu, Off-line signature verification and forgery detection using fuzzy modeling, Pattern Recognition, 38(3):341-356, 2005.
[12] S. Pal, S. Chanda, U. Pal, K. Franke, M. Blumenstein, Off-line signature verification using G-SURF ,12th International Conference on Intelligent Systems Design and Applications pp. 586-591, 2012.
[13] M. A. Ferrer, J. B. Alonso, and C. M. Travieso, Offline geometric parameters for automatic signature verification using fixed-point arithmetic, IEEE Transactions on Pattern Analysis and Machine Intelligence, 27(6):993-997, 2005.
[14] M. I. Malik, and M. Liwicki, From Terminology to Evaluation: Performance Assessment of Automatic Signature Verification Systems, International Conference on Frontiers in Handwriting Recognition, pp. 613-618, 2012.
[15] D. Impedovo, and G. Pirlo, Automatic Signature Verification: The State of the Art, IEEE Transactions on Systems, Man and Cybernetics, Part C: Applications and Reviews, 38(5):609-635, 2008.
[16] V. Nguyen, Y. Kawazoe, T. Wakabayashi, U. Pal and M. Blumenstein, Performance Analysis of the Gradient Feature and the Modified Direction Feature for Off-line Signature Verification, International Conference on Frontiers in Handwriting Recognition, pp. 303-307, 2010.
[17] M. B. Yılmaz, and B. Yanıkoğlu, Score level fusion of classifiers in off-line signature verification, Information Fusion, 32(B):109-119, 2016.
[18] R. Sabourin, G. Genest, and F. J. Preteux, Off-line signature verification by local granulometric size distributions, IEEE Transactions on Pattern Analysis and Machine Intelligence, 19(9):976-988, 1997.
[19] J. Fierrez-Aguilar, N. Alonso-Hermira, G. Moreno-Marquez and J. Ortega-Garcia, An Off-line Signature Verification System Based on Fusion of Local and Global Information, Biometric Authentication, Lecture Notes in Computer Science D. Maltoni and A. K. Jain, eds., pp. 295-306: Springer Berlin Heidelberg, 2004.
[20] J. F. Vargas, M. A. Ferrer, C. M. Travieso and J. B. Alonso, Off-line signature verification based on grey level information using texture features, Pattern Recognition, 44(2):375-385, 2011.
[21] L. Batista, E. Granger, and R. Sabourin, Dynamic selection of generative–discriminative ensembles for off-line signature verification, Pattern Recognition, 45(4):1326-1340, 2012.
[22] R. Kumar, J. D. Sharma, and B. Chanda, Writer-independent off-line signature verification using surroundedness feature, Pattern Recognition Letters, 33(3):301-308, 2012.
[23] G. Pirlo, and D. Impedovo, Verification of Static Signatures by Optical Flow Analysis, IEEE Transactions on Human-Machine Systems, vol. 43(5):499-505, 2013.
[24] Y. Serdouk, H. Nemmour, and Y. Chibani, New off-line Handwritten Signature Verification method based on Artificial Immune Recognition System, Expert Systems with Applications, 51:186-194, 2016.
[25] M. A. Ferrer, J. F. Vargas, A. Morales and A. Ordonez, Robustness of Offline Signature Verification Based on Gray Level Features, IEEE Transactions on Information Forensics and Security, 7(3):966-977, 2012.
[26] A. Hamadene, and Y. Chibani, One-Class Writer-Independent Offline Signature Verification Using Feature Dissimilarity Thresholding, IEEE Transactions on Information Forensics and Security, 11(6):1226-1238, 2016.
[27] E. N. Zois, L. Alewijnse, and G. Economou, Offline signature verification and quality characterization using poset-oriented grid features, Pattern Recognition, 54(6):162-177, 2016.
[28] M. Okawa, Offline Signature Verification Based on Bag-Of-Visual Words Model Using KAZE Features and Weighting Schemes, IEEE Conference on Computer Vision and Pattern Recognition (CVPR) Workshops, pp. 184-190, 2016.
[29] A. Cutler, and L. Breiman, Archetypal Analysis, Technometrics, 36(4):338-347, 1994.
[30] B. A. Olshausen, and D. J. Field, Emergence of Simple-Cell Receptive-Field Properties by Learning a Sparse Code for Natural Images, Nature, 381(6583):607-609, 1996.
[31] M. J. A. Eugster, and F. Leisch, Weighted and robust archetypal analysis, Computational Statistics & Data Analysis, 55(3):1215-1225, 2011.
[32] M. Mørup, and L. K. Hansen, Archetypal analysis for machine learning and data mining, Neurocomputing, 80:54-63, 2012.
[33] S. Seth, and M. J. A. Eugster, Probabilistic archetypal analysis, Machine Learning, 102(1):85-113, 2016.
[34] Y. Chen, J. Mairal and Z. Harchaoui, Fast and Robust Archetypal Analysis for Representation Learning, IEEE Conference on Computer Vision and Pattern Recognition, Columbus, pp. 1478-1485, 2014.
[35] J. Mairal, F. Bach, and J. Ponce, Sparse Modeling for Image and Vision Processing, Found. Trends. Comput. Graph. Vis., 8(2-3):85-283, 2014.
[36] D. Bertsekas, and D. Bertsekas, Nonlinear Programming: Athena Scientific, 1999.
[37] J. Nocedal, and S. J. Wright, Numerical optimization 2nd ed., Springer, 2006.
[38] J. T. Favata, and G. Srikantan, A multiple feature resolution approach to handprinted digit and character recognition, International journal of imaging systems and technology, 7(4):304-311, 1996.
[39] G. Pirlo, and D. Impedovo, Cosine similarity for analysis and verification of static signatures, IET Biometrics, 2(4):151-158, 2013.
[40] M. K. Kalera, S. Srihari, and A. Xu, Offline signature verification and identification using distance statistics, International Journal of Pattern Recognition and Artificial Intelligence, 18(7):1339-1360, 2004.

[41] J. Ortega-Garcia, J. Fierrez-Aguilar, D. Simon, J. Gonzalez, M. Faundez-Zanuy, V. Espinosa, A. Satue, I. Hernaez, J.J. Igarza, C. Vivaracho, D. Escudero, and Q.I. Moro, MCYT baseline corpus: a bimodal biometric database, IEE Proceedings Vision, Image and Signal Processing, 150(6):395-401, 2003.

[42] M. M. Kumar, and N. B. Puhan, Off-line signature verification: upper and lower envelope shape analysis using chord moments, IET Biometrics, 3(4):347-354, 2014.

[43] Y. Guerbai, Y. Chibani, and B. Hadjadji, The effective use of the one-class SVM classifier for handwritten signature verification based on writer-independent parameters, Pattern Recognition, 48(1):103-113, 2015.

[44] S. N. Srihari, A. Xu, and M. K. Kalera, Learning strategies and classification methods for off-line signature verification, International Workshop on frontiers in Handwriting Recognition, pp. 161-166, 2004.

[45] Y. Serdouk, H. Nemmour, and Y. Chibani, An improved Artificial Immune Recognition System for off-line handwritten signature verification, 13th International Conference on Document Analysis and Recognition, pp. 196-200, 2015.

[46] G. Ganapathi, and R. Nadarajan, A Fuzzy Hybrid Framework for Offline Signature Verification, Pattern Recognition and Machine Intelligence: Proceedings of 5th International Conference, PReMI 2013, Kolkata, India, P. Maji, A. Ghosh, M. N. Murty et al., eds., pp. 121-127, Berlin, Heidelberg: Springer Berlin Heidelberg, 2013.

[47] B. H. Shekar, R. K. Bharathi, and B. Pilar, Local Morphological Pattern Spectrum Based Approach for Off-line Signature Verification, Pattern Recognition and Machine Intelligence: Proceedings of 5th International Conference, PReMI, P. Maji, A. Ghosh, M. N. Murty et al., eds., pp. 335-342, Berlin, Heidelberg: Springer Berlin Heidelberg, 2013.

[48] A. Soleimani, B. N. Araabi and K. Fouladi, Deep Multitask Metric Learning for Offline Signature Verification. Pattern Recognition Letters, 80:84-90, 2016.

[49] F. Alonso-Fernandez, M. C. Fairhurst, J. Fierrez and J. Ortega-Garcia, Automatic Measures for Predicting Performance in Off-Line Signature, IEEE International Conference on Image Processing, pp. 369-372, 2007.

[50] J. Wen, B. Fang, Y.Y. Tang, and T. Zhang, Model-based signature verification with rotation invariant features. Pattern Recognition, 42:1458-1466, 2009.

[51] S. Y. Ooi, A. B. J. Teoh, Y. H. Pang and B. Y. Hiew, Image-based handwritten signature verification using hybrid methods of discrete Radon transform, principal component analysis and probabilistic neural network. Applied Soft. Computing, 40:274-282, 2016.

[52] E. N. Zois, I. Theodorakopoulos, D. Tsourounis, and G. Economou, Parsimonious Coding and Verification of Offline Handwritten Signatures, IEEE Conference on Computer Vision and Pattern Recognition (CVPR) Workshops, 2017, pp. 134-143. (In Press)

[53] R. Rubinstein, A. M. Bruckstein, and M. Elad, Dictionaries for Sparse Representation Modeling, Proceedings of the IEEE, 98(6):1045-1057, 2010.