Writer-specific Dissimilarity Normalisation for Improved Writer-independent Off-line Signature Verification

Jacques Swanepoel and Johannes Coetzer
Department of Mathematical Sciences, Stellenbosch University, South Africa
{jpswanepoel,jcoetzer}@sun.ac.za

Abstract

In this paper we present a novel writer-independent off-line signature verification system. This system utilises the discrete Radon transform and a dynamic time warping algorithm for writer-independent signature representation in dissimilarity space. The system also considers writer-specific statistics for dissimilarity normalisation. A discriminant function, either linear or quadratic, is utilised for signature modelling and verification.

We show that the feature extraction and dissimilarity representation framework proposed in this paper provides a successful platform for signature modelling and verification. We also show that the inclusion of writer-specific statistics during dissimilarity normalisation improves the proficiency of the proposed writer-independent verification system.

When evaluated on Dolfing's data set, a signature database containing 1530 genuine signatures and 3000 amateur skilled forgeries, we show that the system presented in this paper outperforms existing systems also evaluated on this data set.

1 Introduction

In the traditional *writer-dependent* approach to off-line signature verification, each writer submits a set of genuine signature samples, in order to train a model specific to said writer. This popular approach, however, has two notable disadvantages: (1) A relatively large training set, not realistically obtainable in practice, is required to produce a sufficiently representative writer model; (2) When utilising a discriminative model for classification, only random forgeries may be used for model training, since it is not reasonable to assume the availability of skilled forgeries for every writer enrolled during deployment.

In contrast, the *writer-independent* approach constructs a signature model that discriminates between two classes only, namely genuine and forged signatures (also referred to as positive and negative instances) belonging to *any* writer. This is achieved by utilising a dissimilarity representation, obtained by comparing writer-specific positive and negative training samples to a writer-specific set of positive reference samples, for model training. During system deployment, any newly enrolled writer only needs to submit a positive reference set, in order to obtain a dissimilarity representation suitable for consideration by the trained model. The writer-independent approach therefore provides effective solutions to the problems of data scarcity and training with skilled forgeries.

As a result, the use of writer-independent verification frameworks has gained notable popularity in the literature [9, 1, 2]. In this paper, we propose a novel writer-independent verification framework for skilled forgery detection.

2 System overview

The system presented in this paper is trained and evaluated using signatures from different writers. Every writer considered provides a collection of labelled positive samples that constitutes a writer-specific reference set. The training set is composed of labelled positive and negative samples from a collection of *guinea-pig* writers (e.g. banking staff or a control group). The evaluation set comprises unlabelled positive and negative samples from a collection of writers different than those in the guinea-pig set (e.g. banking clients). System performance is gauged using the *equal error rate* (EER).

During signature modelling, the required dissimilarity representation is achieved by employing a two-stage process. Binary signature images are first converted into feature sets using the *discrete Radon*

978-0-7695-4774-9/12 $26.00 © 2012 IEEE
DOI 10.1109/ICFHR.2012.290

transform (DRT) [4]. Using a *dynamic time warping* (DTW) algorithm [3], the above-mentioned feature sets are subsequently matched to those extracted from writer-specific reference signatures, so that a set of dissimilarity vectors is obtained.

The set of dissimilarity vectors obtained from signatures in the training set is used to train a *discriminant function* (DF) [6]. A *linear discriminant function* (LDF) and *quadratic discriminant function* (QDF) are considered. During evaluation, questioned signatures from the evaluation set are encoded into dissimilarity vectors, by comparing said signatures to the appropriate writer-specific reference signatures. The trained DF is subsequently used to predict class membership.

3 Signature representation

Feature extraction. The use of projection profiles for feature extraction is popular, since it effectively captures signature shape information. Many systems in the literature, however, rely solely on horizontal and vertical projection profiles.

The system presented in this paper utilises the discrete Radon transform, since it enables the use of an expanded projection angle set and therefore constitutes a natural progression of the projection-based method. The DRT has been shown to be well suited for signature representation [4].

Any binary signature image I presented to the system proposed in this paper is converted into a set of T, d-dimensional feature vectors. This is achieved by calculating a projection profile for each angle θ_i in the set $\Theta = \{\theta_1, \theta_2, \ldots, \theta_T\}$. This set contains T equally distributed angles in the range $[0°, 180°)$[1]. Each of the T feature vectors obtained by employing the DRT is then interpolated to have a length of d and normalised to have unit variance. This process produces a scale invariant feature set $\mathbf{X} = \{\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_T\}$.

Dissimilarity representation. The DRT-based feature extraction technique described above produces a collection of feature sets for each writer enrolled into the system. In a writer-dependent verification scenario, these feature sets may be used to train a writer-specific signature model. In order to construct a writer-independent system, however, each writer-dependent collection of feature sets is converted into a set of dissimilarity vectors.

[1]Initially, $T+1$ equally distributed angles in the range $[0°, 180°]$ are obtained. The angle $180°$ is then discarded, since its projection is equivalent to that of $0°$.

Given a feature set $\mathbf{X}_k^{(\omega)}$, extracted from a positive reference signature belonging to writer ω, any other feature set $\mathbf{X}_q^{(\omega)}$ that is claimed to belong to this writer can be converted into a dissimilarity vector $\chi_{(q,k)}^{(\omega)}$ by calculating the dissimilarity between $\mathbf{X}_k^{(\omega)}$ and $\mathbf{X}_q^{(\omega)}$. It is proposed in [9] that the dissimilarity between two d-dimensional *feature vectors* $\mathbf{x}_k$ and $\mathbf{x}_q$ is obtained using the Euclidean distance, such that

$$\chi_{(q,k)}^{(\omega)} = \bigcup_{i=1}^{d} \sqrt{(x_i^{(k)} - x_i^{(q)})^2}, \tag{1}$$

where $\bigcup$ denotes vector concatenation, whilst $x_i^{(k)}$ and $x_i^{(q)}$ denote the i^{th} pair of elements belonging to feature vectors $\mathbf{x}_k$ and $\mathbf{x}_q$, respectively.

In contrast, the system presented in this paper obtains the dissimilarity between two *feature sets* $\mathbf{X}_k$ and $\mathbf{X}_q$ by using a DTW-algorithm, such that

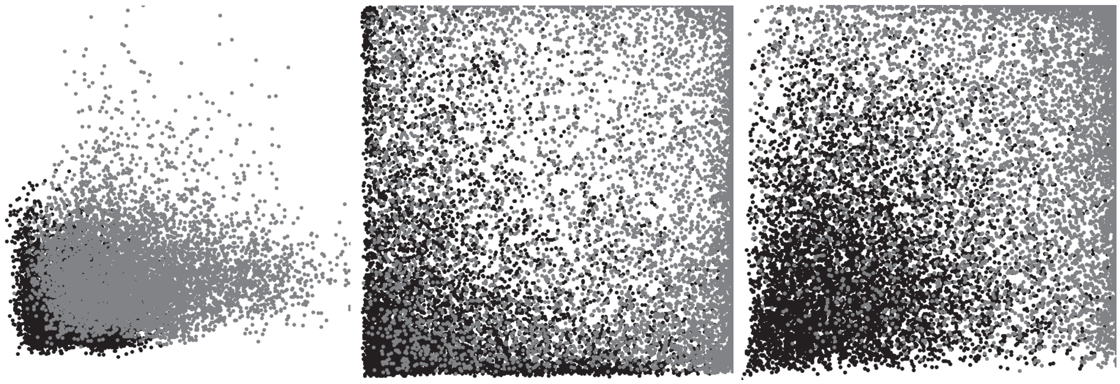
$$\chi_{(q,k)}^{(\omega)} = \bigcup_{i=1}^{T} D(\mathbf{x}_i^{(k)}, \mathbf{x}_i^{(q)}), \tag{2}$$

where $D(\mathbf{x}_i^{(k)}, \mathbf{x}_i^{(q)})$ denotes the DTW-based distance between the i^{th} pair of d-dimensional feature vectors $\mathbf{x}_i^{(k)} \in \mathbf{X}_k$ and $\mathbf{x}_i^{(q)} \in \mathbf{X}_q$. This process is therefore able to convert any set of T, d-dimensional feature vectors into a *single* T-dimensional dissimilarity vector, which is well-suited for writer-independent signature modelling.

4 Signature modelling

The system presented in this paper uses a set of dissimilarity vectors, obtained from several different writers, for signature modelling. In order to produce these dissimilarity vectors, each writer submits K genuine signatures during enrolment, that serve as a writer-specific reference set.

A writer-independent signature model is trained using samples of genuine signatures *and* skilled forgeries, obtained from a set of guinea pig writers. These writers are considered representative of the general public, and their signatures are used for training purposes only. Given a set of K reference signatures and N labelled training signatures (that include both positive and negative samples) for each of the Ω guinea-pig writers, the system generates a set of $K N \Omega$ dissimilarity vectors by computing $\chi_{(n,k)}^{(\omega)}$ for $k = \{1, 2, \ldots, K\}$, $n = \{1, 2, \ldots, N\}$,



(a) (b) (c)

Figure 1. Typical separation of negative and positive dissimilarity vectors, denoted by grey markers superimposed onto black markers, respectively, when $\mathbf{T} = 2$, $\mathbf{K} = 10$ and (a) no normalisation, (b) global normalisation or (c) writer-specific normalisation is employed.

and $\omega = \{1, 2, \ldots, \Omega\}$. We henceforth use the simplified notation χ to denote an arbitrary dissimilarity vector and refer to dissimilarity vectors representative of genuine signatures and forgeries as being positive and negative, respectively.

Dissimilarity normalisation. Let χ^+ and χ^- denote the respective sets containing positive and negative dissimilarity vectors, obtained from all the guinea-pig writers. These sets provide a sufficient platform for obtaining a DF for verification purposes. However, since the data contained in χ^- represent *skilled* forgeries, a significant degree of overlap is observed between χ^+ and χ^- in dissimilarity space (see Figure 1 (a)). In order to maximally separate these sets, thereby providing an optimal platform for model training, each vector in χ^+ and χ^- has to be appropriately normalised.

A wide variety of suitable normalisation techniques are documented in the literature [7]. The system presented in this paper performs dissimilarity normalisation using a rescaled version of the well-known *logistic function*, such that any given dissimilarity vector χ can be converted into a normalised dissimilarity vector $\bar{\chi}$ as follows

$$\begin{aligned}\bar{\chi} &= \eta(\chi, \mu^+, \sigma^+) \\ &= \left[1 + \exp\left(\frac{-6\chi}{\mu^+ + \sigma^+} + \mu^+ + \sigma^+\right)\right]^{-1}. \quad (3)\end{aligned}$$

We therefore shift and rescale the numerically significant domain of the logistic function from $[-6, 6]$ to $[0, 2(\mu^+ + \sigma^+)]$, where the parameters μ^+ and σ^+ contain the dimension-specific means and standard deviations of all the vectors in χ^+, as proposed in [7]. The effect of this *global normalisation* strategy is illustrated in Figure 1 (b).

In this paper we propose the incorporation of dissimilarity statistics on a writer-specific level, thereby yielding a more appropriate overall normalisation strategy. First, the writer-specific statistics μ^+_ω and σ^+_ω are determined, using the dissimilarity vectors generated by comparing every reference signature belonging to writer ω to every *other* reference signature belonging to this writer. Writer-specific normalised dissimilarity vectors are then computed as $\bar{\chi} = \eta(\chi, \mu^+_\omega, \sigma^+_\omega)$. This separation of positive and negative dissimilarity vectors of each *individual* writer leads to improved separation of positive and negative dissimilarity vectors across the *entire* set of writers, since only strictly relevant information is used in the normalisation process. The impact of this *writer-specific normalisation* strategy is illustrated in Figure 1 (c).

It should be noted that the proposed writer-specific approach to dissimilarity normalisation is only possible if $K > 1$. For $K = 1$, no writer-specific statistics can be obtained and the global normalisation strategy has to be used.

Discriminant analysis. Once χ^+ and χ^- have been maximally separated using the proposed normalisation technique, a discriminant function

$$f(\bar{\chi}) = C + L\bar{\chi} + Q\bar{\chi}^T\bar{\chi} \quad (4)$$

is used to obtain the decision boundary which maximises class discrimination. The values C, L and

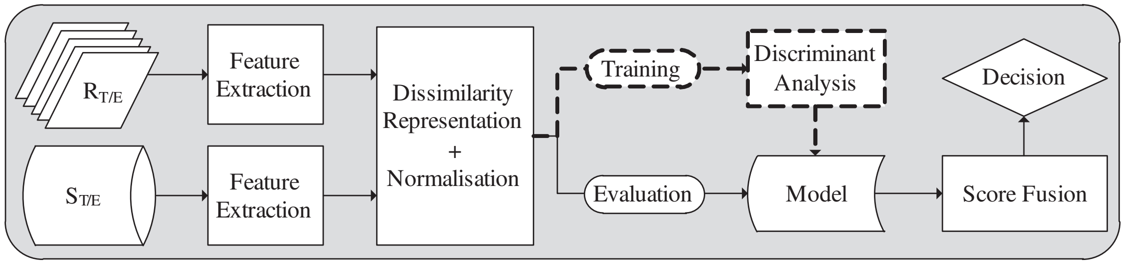


Figure 2. Schematic representation of the experimental protocol utilised in this paper.

Q denote the constant, linear and quadratic coefficients of said decision boundary, respectively. In this paper we consider both linear ($Q = 0$) and quadratic ($Q \neq 0$) discriminant functions. LDF coefficients are computed using a pooled covariance estimate, whilst QDF coefficients are computed using class-specific covariance estimates.

5 Verification

The verification protocol utilised in this paper is based on the questioned document expert's approach [9]. Any questioned signature $I_q^{(\omega)}$, claimed to belong to writer ω, is first converted into a DRT-based feature set, which is subsequently compared to that of each of the K reference signatures belonging to writer ω, thereby yielding a set of K normalised DTW-based dissimilarity vectors.

Each questioned dissimilarity vector is then presented to the trained DF, in order to obtain a signed distance measure relative to the corresponding decision boundary. Each distance measure is then converted, using the conventional logistic function, into a *partial confidence score* $s \in [0, 1]$. Finally, the set of K partial confidence scores are averaged, yielding the *final confidence score* s^* as follows

$$s^* = \frac{1}{K} \sum_{k=1}^{K} \left[1 + \exp\left(-f(\bar{X}_{(q,k)}^{(\omega)}) \right) \right]^{-1}. \quad (5)$$

This final confidence score is used to predict class membership, by imposing a sliding threshold $\tau \in [0, 1]$, such that the questioned signature $I_q^{(\omega)}$ is accepted as genuine if and only if $s^* \geq \tau$.

6 Experiments

Data. The signature database considered in this paper contains 4530 signatures (1530 genuine signatures and 3000 amateur skilled forgeries) obtained from 51 writers. For each writer, 30 genuine signatures and 60 forgeries are available (except for two writers, for whom only 30 forgeries are available). This database, known as *Dolfing's data set*, was originally captured on-line [5], but has since been converted into an off-line representation [3], thereby rendering it suitable for the evaluation of the system presented in this paper.

Protocol. The experimental protocol utilised in this paper is illustrated in Figure 2. Prior to experimentation, Dolfing's data set is partitioned into two disjoint subsets. This partitioning ensures that data from *different* writers are used for model training and evaluation, thereby ensuring that the results reported represent an unbiased estimation of system performance. These partitions, referred to as the *training set* and *evaluation set*, contain the signatures of 34 writers and 17 writers, respectively.

During model training, only the training set is used. For every writer, K genuine signatures are reserved for the reference set R_T, whilst $30 - K$ genuine signatures and $30 - K$ forgeries constitute the set S_T considered for training. As a result, $30 + K$ forgeries therefore remain unused, in order to ensure an unbiased training set. Each of the K reference signatures is used to obtain $K(30 - K)$ positive dissimilarity vectors and $K(30 - K)$ negative dissimilarity vectors. The entire set of normalised positive and negative dissimilarity vectors, obtained from all 34 writers in the training set, is used to determine the optimal DF, which is retained for subsequent verification.

During model evaluation, only the evaluation set is used. For every writer, K genuine signatures are reserved for the reference set R_E, whilst $30 - K$ genuine signatures and 60 forgeries constitute the set S_E considered for verification. The entire set of genuine signatures and forgeries, obtained from all 17 writers in the evaluation set, is used to gauge system performance.

Table 1. Average EERs obtained when considering Dolfing's evaluation set.

LDF		K							
$\mu_{\mathrm{EER}}(\%)$		1	3	5	7	9	11	13	15
T	2	27.28	23.08	17.08	13.20	10.73	9.24	8.31	7.92
	3	28.30	21.95	17.74	14.34	11.79	10.40	9.96	9.11
	30	29.29	14.95	12.44	9.03	6.90	5.86	5.09	**4.54**
	60	30.52	14.65	12.75	9.07	6.92	6.14	5.34	4.59
	90	30.92	16.71	12.63	9.09	7.32	6.21	5.42	4.84
	180	30.38	17.56	13.84	9.60	7.46	6.53	5.83	5.23
QDF		K							
$\mu_{\mathrm{EER}}(\%)$		1	3	5	7	9	11	13	15
T	2	28.75	22.68	17.64	13.95	11.31	10.38	9.92	9.25
	3	29.36	21.99	18.35	15.36	13.64	11.63	10.90	10.38
	30	28.81	22.23	16.98	11.92	7.48	6.09	5.33	**4.93**
	60	31.41	26.99	21.52	14.90	9.58	7.80	6.44	6.61
	90	32.56	30.30	24.80	18.19	12.03	10.33	8.08	8.54
	180	35.98	33.66	29.71	24.69	18.29	16.91	13.32	15.46

Since only 17 writers are considered for evaluation, the protocol utilised in this paper employs 3-fold *cross validation* and repetitive *data randomisation*, which proceeds as follows: (1) The data set is split into three equal subsets, each containing signatures from 17 writers; (2) Each subset, in turn, is used as an evaluation set, whilst signatures from the remaining 34 writers constitute the training set; (3) The order of the writers is randomised and the process is repeated. We consider 10 repetitions and therefore report the results from 30 trials.

Results. We henceforth refer to the LDF-based and QDF-based signature modelling techniques considered by the system presented in this paper, as the LDF system and QDF system, respectively. The average equal error rate yielded by these systems is presented in Table 1. The average EER, denoted by μ_{EER}, is obtained by considering the evaluation set for all 30 trials. Table 1 illustrates the influence of the parameters T and K, whilst $d = 128$ remains fixed.

It is clear from Table 1 that increasing the reference set size K invariably (except for the two underlined cases) leads to an improvement in verification proficiency. This is understandable, since this parameter plays a central role in several key phases regarding signature modelling and verification – it determines the training set size, the representation potential of the writer-specific normalisation statistics, as well as the size of the partial scoring pool. Note in particular the generally significant decrease in system performance when $K = 1$. When using a single reference signature, the writer-specific normalisation technique proposed in this paper can not be utilised, which leads to sub-optimal class separation in dissimilarity space.

It is also clear from Table 1 that an increase in the projection angle set size T does not necessarily improve system performance. In fact, we find that an expanded projection angle set may in some cases severely impede verification proficiency. This suggests that it is not the number of projection profiles included in the feature set, but rather *which* projection profiles are included, that determines class separation in dissimilarity space. This is an understandable observation, since it is expected that projections generated from a small set of significant directions (e.g. horizontal, vertical, baseline direction etc.) will capture the majority of signature shape variation. The inclusion of non-essential projection profiles therefore not only leads to information redundancy in the feature set, but also minimises the role of more significant projection profiles in dissimilarity space, thereby adversely affecting the system's discrimination potential.

Interestingly, the LDF consistently outperforms the QDF. Given the dissimilarity representation of the positive and negative classes, one may expect a non-linear DF to model class separation more effectively. We find the inferior verification proficiency of the QDF is due to *over-training*, resulting in inadequate generalisation. This assertion is substantiated by noting that when the *training set* is also considered for evaluation, the QDF outperforms the LDF. These results are not tabulated in this paper.

To our knowledge, the system presented in this paper constitutes the first writer-independent verification system evaluated on Dolfing's data set. We therefore compare the results reported in this paper to results reported for existing *writer-dependent* systems [4, 10, 8], also evaluated on this data set (see Table 2). Each of these existing systems requires 15 genuine training signatures per writer, prior to model optimisation. When employing a writer-independent approach, this is analogous to requiring $K = 15$ reference signatures per writer, prior to model training.

Table 2. EERs obtained for existing systems evaluated on Dolfing's data set.

System	EER (%)
[4] (2004)	12.2
[10] (2010)	10.23
[8] (2010)	8.89
QDF (This paper)	**4.93**
LDF (This paper)	**4.54**

It is clear from Table 2 that, under similar operating conditions, the LDF and QDF systems significantly outperform the systems proposed in [4, 10, 8]. In fact, from Table 1 we note that even when reducing the reference set size to $K = 9$, the systems proposed in [4, 10, 8] are still outperformed by the LDF *and* QDF systems presented in this paper. The most promising result, however, is that the LDF achieves an EER of 12.44% for $K = 5$. This system compares well with the systems proposed in [4, 10, 8], despite having a drastically reduced reference set size i.e. a reference set size that one is likely to encounter in practical scenarios.

7 Conclusion

In this paper we demonstrated that: (1) Dissimilarity vectors obtained by employing the DRT and DTW provide a successful framework for writer-independent signature representation; (2) The inclusion of writer-specific statistics during dissimilarity normalisation improves overall class separation when training with skilled forgeries; (3) The utilisation of these proposed techniques yields a novel verification system that outperforms existing systems evaluated on the same data set.

The results reported in this paper also provide considerable insight into potential future work regarding system improvement. We found that the inclusion of non-essential projection angles in the DRT impedes system performance. Given a full projection angle set, one may for instance utilise a technique such as principal component analysis (PCA) for dynamic feature selection and consequent dimension reduction. Also, in order to obtain a more robust non-linear decision boundary in dissimilarity space, the use of more advanced discriminative classifiers such as support vector machines (SVMs) may potentially be advantageous.

The incorporation of PCA and SVMs into the signature representation framework presented in this paper is currently under investigation.

References

[1] L. Batista, E. Granger, and R. Sabourin. Applying dissimilarity representation to off-line signature verification. *International Conference on Pattern Recognition*, pages 1293–1297, 2010.

[2] D. Bertolini, L. Oliveira, E. Justino, and R. Sabourin. Reducing forgeries in writer-independent off-line signature verification through ensemble of classifiers. *Pattern Recognition*, 43:387–396, 2010.

[3] J. Coetzer. *Off-line Signature Verification*. PhD thesis, Stellenbosch University, 2005.

[4] J. Coetzer, B. Herbst, and J. du Preez. Offline signature verification using the discrete radon transform and a hidden markov model. *European Association for Signal Processing Journal on Applied Signal Processing*, 4:559–571, 2004.

[5] J. Dolfing, E. Aarts, and J. van Oosterhout. On-line signature verification with hidden markov models. *International Conference on Pattern Recognition*, 2:1309–1312, 1998.

[6] J. Friedman. Regularized discriminant analysis. *Journal of the American Statistical Association*, 84(405):165–175, 1989.

[7] A. Jain, K. Nandakumar, and A. Ross. Score normalization in multimodal biometric systems. *Pattern Recognition*, 38:2270–2285, 2005.

[8] M. Panton and J. Coetzer. Off-line signature verification using ensembles of local radon transform-based hmms. *Annual Symposium of the Pattern Recognition Association of South Africa*, pages 201–206, 2010.

[9] C. Santos, E. Justino, F. Bortolozzi, and R. Sabourin. An off-line signature verification system based on the questioned document expert's approach and a neural network classifier. *International Workshop on Frontiers in Handwriting Recognition*, pages 498–502, 2004.

[10] J. Swanepoel and J. Coetzer. Off-line signature verification using flexible grid features and classifier fusion. *International Conference on Frontiers in Handwriting Recognition*, pages 297–302, 2010.