

Research Article

A Novel Criterion for Writer Enrolment Based on a Time-Normalized Signature Sample Entropy Measure

Sonia Garcia-Salicetti, Nesma Houmani, and Bernadette Dorizzi

Department of EPH, Institut TELECOM, TELECOM & Management SudParis, 91011 Evry, France

Correspondence should be addressed to Nesma Houmani, nesma.houmani@it-sudparis.eu

Received 15 October 2008; Revised 8 March 2009; Accepted 9 June 2009

Recommended by Natalia A. Schmid

This paper proposes a novel criterion for an improved writer enrolment based on an entropy measure for online genuine signatures. As online signature is a temporal signal, we measure the time-normalized entropy of each genuine signature, namely, its average entropy per second. Entropy is computed locally, on portions of a genuine signature, based on local density estimation by a Client-Hidden Markov Model. The average time-normalized entropy computed on a set of genuine signatures allows then categorizing writers in an unsupervised way, using a K-Means algorithm. Linearly separable and visually coherent classes of writers are obtained on MCYT-100 database and on a subset of BioSecure DS2 containing 104 persons (DS2-104). These categories can be analyzed in terms of variability and complexity measures that we have defined in this work. Moreover, as each category can be associated with a signature prototype inherited from the K-Means procedure, we can generalize the writer categorization process on the large subset DS2-382 from the same DS2 database, containing 382 persons. Performance assessment shows that one category of signatures is significantly more reliable in the recognition phase, and given the fact that our categorization can be used online, we propose a novel criterion for enhanced writer enrolment.

Copyright © 2009 Sonia Garcia-Salicetti et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

1. Introduction

Handwritten signature is a behavioural biometric modality showing high variability from one instance to another of a same writer. This high variability explains indeed that the best verification approaches, as particularly reflected for Online Signature in the results of the First International Signature Verification Competition SVC2004 [1] and the Signature Evaluation carried out in the framework of BioSecure Multimodal Evaluation Campaign BMEC2007 [2], are those tolerating random local variations of the signature, as elastic matching techniques (Dynamic Time Warping [3–5] or statistical models, as Hidden Markov Models (HMM) [3, 6–13] and Gaussian Mixture Models (GMMs) [14, 15]. Nevertheless, the amount of this variability is writer dependent, in the sense that some writers have a signature by far more variable from one instance to the next than other writers.

An automatic signature verification system involves two steps: the enrolment step and the verification step. In order

to provide a given level of security to an individual signer, writer enrolment must guarantee that enrolment signatures are stable and complex enough. Indeed, as studied in [16], when enrolling a writer, his/her signature will be acceptable as a reference signature, or as part of a reference set, for any verification system, only if it is complex enough. In [16], a “difficulty coefficient” estimates the difficulty to reproduce a given signature as a function of the rate of geometric modifications (length and direction of strokes) per unit of time, in other words as a function of complexity of the hand draw. Such study concludes that “problematic” signers in terms of performance of Automatic Verification Systems are those with signatures which have a low “difficulty coefficient” (not complex enough signatures).

On the other hand, when enrolling a writer, his/her signature will be suitable as reference or as part of a reference set for any verification system only if it is not too variable; in [16], enrolment signatures are selected by using a comparison algorithm that computes the spatiotemporal difference between two signatures (elastic matching). By this

way, a “dissimilarity index” is proposed to quantify intraclass variability between different signature samples of a same writer. In [17], a procedure relying on a correlation-based criterion detecting local distortions of the hand draw is proposed to select the reference signatures for a signature verification system. Such correlation criterion measures how much local stability values, computed on different signatures when being matched by elastic matching techniques, are correlated. Finally, the subset of signatures with highest correlation is selected as reference set. Alternatively, in [18], the stability criterion is based on the lowest intraclass Euclidean distance between feature vectors representing globally the candidate reference signatures. Finally, in [19], both complexity and variability criteria were proposed for offline signature verification by a human expert. A human operator labels signatures according to both criteria and their impact on performance is studied. Also in [20], signature analysis by means of fractal geometry led to the emergence of a complexity criterion to categorize writers.

All these works suggest the strong impact of complexity and variability criteria on the classifier performing signature verification. Indeed, stability is required in genuine signatures in order to be able to characterize a given writer, since the less stable a signature is, the more likely it is that a forgery gets dangerously close to genuine signatures in terms of the metric of any classifier. Also, complex enough signatures are required at the enrolment step to generate a certain level of security.

In this work, we propose to exploit for writer enrolment a time-normalized entropy measure that allows quantifying both the stability and the complexity of a writer’s genuine signatures. This entropy, measured in bits per second, is computed on portions of the signature, and averaged over such portions. As the entropy of a random variable only depends on its probability density values, a good estimation of this probability density is important [21]. As in online signatures there are local time dependencies in the dynamics of the hand-draw, a local paradigm for density estimation is natural.

In the previous works [22, 23], we proposed to estimate the probability density of a writer’s dynamics locally, by a Hidden Markov Model (HMM) trained on a set of ten genuine signatures, to extract a Personal Entropy measure globally from such set. In this work, we follow the same local paradigm, but we compute the time-normalized entropy of a signature sample “Sample Entropy”, namely, the average entropy per second of such sample, therefore quantified in bits per second. It is worth noticing that the above mentioned Hidden Markov Model, whose complexity (topology) is related to the length of the genuine signatures of a writer, is only used in our work as a local refined density estimator, devoted to compute the time-normalized entropy of a signature sample, and not as a classifier.

Based on the “Sample Entropy”, we then propose to generate for each writer a “Personal Entropy measure” value, by averaging the “Sample Entropy” associated to each of his/her genuine signatures. We show in this work that this measure allows categorizing writers in linearly separable and visually coherent categories, by a K-Means procedure [24]. Moreover,

we related this categorisation to variability and complexity measures, this way showing quantitatively the link between our new Personal Entropy measure and some behavioural characteristics of the signature. Our previous performance assessment study [23], carried out only on random forgeries, with different classifiers, showed that system performance changes in function of the different writer categories. In this work, we first extend our performance assessment study to skilled forgeries and confirm this interesting result: there is one category of users, which can be detected by their Personal Entropy, and are “problematic”, since their signatures are vulnerable because of their strong variability and low complexity. At the opposite, there is a category of “safer” signatures, highly complex and stable, that can also be detected by their associated writer’s Personal Entropy. Our aim in this work is to exploit this entropy measure in order to enhance enrolment in the following ways.

- (i) To inform the user of the intrinsic risk related to his/her signature.
- (ii) In case of a “problematic” signature, to leave the possibility to the signer of choosing between deciding to pursue enrolment knowing the intrinsic risk of his/her signature, or alternatively to change his/her signature for security purposes.
- (iii) To adjust the quality of enrolment data according to the level of security required by the application.

As previously mentioned, writer categories emerge from our entropy measure, by means of a K-means procedure. Given this fact, each writer category is naturally associated to a signature “prototype” or Entropy-Prototype (EP), which corresponds to the mean of the class. We propose in this work to exploit such Entropy-Prototypes to identify beforehand “problematic” signers. We show indeed that after having generated prototypes on a given reduced data set of 104 writers from the complete DS2 database [25], it is possible to generalize the writer categorization process on other writers belonging to the same database. Given the fact that our writer categorization process is totally automatic, independent of any classifier (it only relies on our proposed Personal Entropy measure), and besides can be generalized to new writers acquired in similar conditions (same digitizer, same acquisition protocol, similar sampling frequency, similar resolution), we propose a novel criterion for a better writer enrolment process targeting enhanced signature verification. Indeed, our writer categorization process gives as outputs one Entropy-Prototype per category, which combined to a Nearest Neighbour Rule [24], naturally allows classifying a signature sample during the enrolment step. This classification allows therefore measuring the intrinsic level of security of a user’s signature at the enrolment step.

This paper is organized as follows. The next section describes how the “Sample Entropy” measure associated to each genuine signature sample is computed by means of a Writer-HMM, and the resulting “Personal Entropy” value of each writer. Also, we present the automatically generated categories of writers, obtained when performing a K-Means procedure on such “Personal Entropy measure” of each

writer, on a subset of the BioSecure Data Set 2 (DS2-104) and on MCYT-100 database, both captured on a digitizer. In order to give a quantitative interpretation of these categories, we have defined complexity and variability measures, and we have shown the strong relationship between our Personal-Entropy measure and both the complexity and the variability in signatures. Section 3 presents performance assessment across such writer categories, by means of two statistical classifiers of same complexity (number of parameters), namely, a Hidden Markov Model (HMM) and a Gaussian Mixture Model (GMM), on DS2-104 and MCYT-100 databases. Such statistical approaches gave indeed the best signature verification results in the last Signature Evaluation Campaign in the framework of BioSecure Multimodal Evaluation Campaign BMEC'2007 [2]. Section 4 describes the generalization of the writer categorization process, relying on Entropy-Prototypes built on a subset of Data Set 2 (DS2-104) and evaluated on the large data set DS2-382 of 382 persons; the resulting global performance on DS2-382 are compared with performance on each category. Finally, the proposed enhanced writer enrolment procedure relying on Personal-Entropy is described in detail.

2. Time-Normalized Sample Entropy and Writer Categories

2.1. Measuring Time-Normalized Sample Entropy with a Hidden Markov Model. We consider in this work a signature as a sequence of two time functions, namely, its raw coordinates (x, y) . Indeed, raw coordinates are the only time functions available on all sorts of databases, whether acquired on fixed platforms (as digitizing tablets) or on mobile platforms (as Personal Digital Assistants).

The entropy of a random variable only depends on its probability density values; therefore a good estimation of this probability density must be performed to compute reliably an entropy value. As the online signature is piecewise stationary, it is natural to estimate the probability density locally, namely, on portions of the signature. In this framework, Hidden Markov Models [3] (HMM) appear as a natural tool as they both allow performing a segmentation of the signature into portions and a local estimation of the probability density on each portion.

We thus consider each genuine signature of a given writer as a succession of portions, generated by its segmentation via such writer's Hidden Markov Model (HMM). Therefore, we obtain as many portions in each signature as there are states in the Writer-HMM. Then we consider each point (x, y) in a given portion S_i as the outcome of one random variable Z_i (see the top of Figure 1) that follows a given probability mass function $p_i(z) = \Pr(Z_i = z)$, where z belongs to the Alphabet A of ordered pairs (x, y) . Such random variable associated to a given portion of the signature is discrete since its alphabet A has a finite number of values, thus its entropy in bits is defined as

$$H(Z_i) = - \sum_{z \in S_i} p(z) \cdot \log_2(p(z)). \quad (1)$$

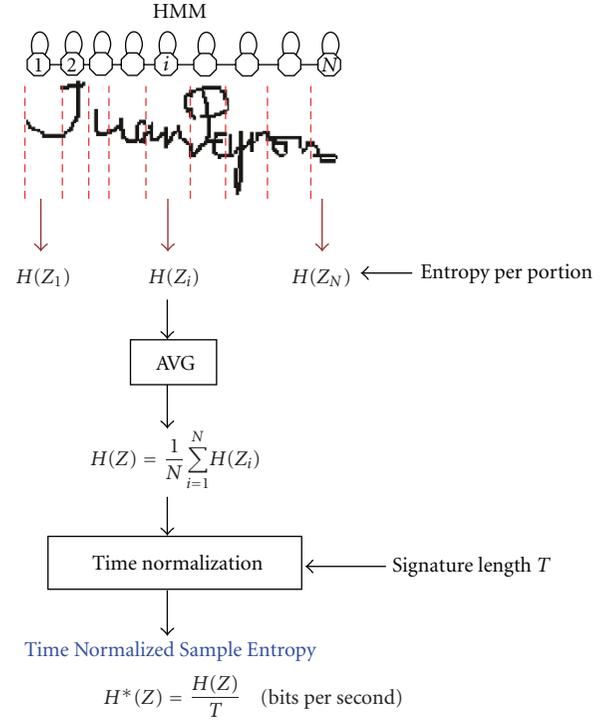


FIGURE 1: The Time-Normalized Sample Entropy computation.

Nevertheless, the hand-drawing as a time function is a continuous process from which we retrieve a sequence of discrete values via a digitizer. For this reason, although $Z = (x, y)$ is discrete, we take advantage of the continuous emission probability law estimated on each portion by the Writer-HMM. Such density function is modelled as a mixture of Gaussian components.

To compute the Time-Normalized Sample Entropy of a signature sample, we first train the Writer-HMM on 10 genuine signatures of such writer, after computing a personalized number of states, as follows:

$$N = \frac{T_{\text{Total}}}{M * 30}, \quad (2)$$

where T_{Total} is the total number of sampled points available in the genuine signatures, and $M = 4$ is the number of Gaussian components per state. We ensure this way that the number of sample points per state is at least 120, in order to obtain a good estimation of the Gaussian Mixture in each state (four Gaussian components).

Then we exploit the Writer-HMM to generate by the Viterbi algorithm [3] the portions on which the entropy is computed for each genuine signature. On each portion, we consider the probability density estimated by the Writer-HMM to compute locally this entropy. We then average the entropy over all the portions of a signature and normalize the result by the signing time of the signature sample (see Figure 1). This measure is a Time-Normalized Sample Entropy, expressed in bits per second. Our experiments show that in order to get a good estimation of Personal Entropy, it is necessary to have at least 10 signatures of each writer.

Averaging this measure across the 10 genuine signatures on which the local probability densities were estimated by the HMM allows generating a measure of Personal Time-Normalized Entropy, denoted “Personal Entropy” in the following of this paper. Time normalization allows comparing users between them in terms of entropy; indeed, without such time normalization, due to the great difference in length between signatures of different persons, entropy tends to be higher on longer signatures.

2.2. Databases Description. We used three databases in this work: the freely available and the widely used MCYT subset of 100 persons [26], and two subsets from the online signature database acquired in the framework of the BioSecure Network of Excellence [25]: DS2 (for Second Data Set of the whole data collection), acquired on a digitizer. The first subset DS2-104 contains data of 104 persons, and the second subset DS2-382 contains data of 382 persons. The whole BioSecure Signature Subcorpus DS2 [25], acquired on several sites in Europe, is the first online signature multisection database acquired in a digitizer.

DS2 contains data from 667 persons acquired in a PC-based offline supervised scenario and the digitizing tablet WACOM INTUOS 3 A6. The pen tablet resolution is 5080 lines per inch, and the precision is 0.25 mm. The maximum detection height is 13 mm, and the capture area is 270 mm (width) \times 216 mm (height). Signatures are captured on paper using an inking pen. At each sampled point of the signature, the digitizer captures at 100 Hz sampling rate the pen coordinates, pen pressure (1024 pressure levels), and pen inclination angles (azimuth and altitude angles of the pen with respect to the tablet). This database contains two sessions, acquired two weeks apart, each containing 15 genuine signatures. The donor was asked to perform, alternatively, three times five genuine signatures and twice five forgeries. Indeed, for skilled forgeries, at each session, a donor is asked to imitate five times the signature of two other persons after several minutes of practice and with the knowledge of the signature dynamics.

2.3. Writer Categories with Personal Entropy Measure. We performed on the two databases described in Section 2.2 (DS2-104 and MCYT-100), containing around 100 persons, a K-Means procedure [24] on Personal Entropy values for different values of K . We reached a good separation of signatures with $K = 3$ on both databases, as shown in Figure 2 for some signatures in DS2, whose owners authorized their publication.

Figure 3 shows that the obtained three categories are actually linearly separable, as represented by indicative lines reporting the automatic classification results given by the K-Means procedure.

As mentioned before, time normalization allows comparing users between them in terms of entropy since there is a great difference in length between signatures of different persons.

We notice that on the two databases, the first category of signatures, those having the highest Personal Entropy



FIGURE 2: Examples of signatures from DS2-104 of (a) high, (b) medium, and (c) low Personal Entropy (with authorization of the writers).

(Figure 2(a)), contains short simply drawn and not legible signatures, often with the shape of a simple flourish. At the opposite, signatures in the third category, those of lowest Personal Entropy (Figure 2(c)), are the longest and their appearance is rather that of handwriting, some being even legible. In between, we notice that signatures with medium Personal Entropy (second category, Figure 2(b)) are longer than those of the first category, often showing the aspect of a complex flourish.

Categories of signatures seem at this step visually related to complexity and variability criteria. We therefore propose quantitative measures of complexity and variability, with which we will analyze the obtained Entropy-based categories.

2.4. Relation between Our Personal Entropy and Complexity and Variability Measures. In order to measure complexity, we consider a vector of seven components related to the shape of handwriting: numbers of local extrema in both x and y directions, changes of pen direction in both x and y directions, cusps points, crossing points, and “star points” [27]. We consider the Euclidean norm of the vector as the indicator of complexity for each signature. We then average such measure on the 10 genuine signatures in order to generate a complexity measure for a given person.

In order to measure the variability of a client’s signature, we use Dynamic Time Warping [3], which relies on a local paradigm to quantify distortions. We compute the distances between all the possible couples of genuine signatures (45 as we consider 10 genuine signatures) and average the obtained distances to get the indicator of signature variability. Four features are extracted locally per point: absolute speed, the angle between the absolute speed vector and the horizontal axis, curvature radius of the signature, and the length to width ratio on a sliding window of size 5.

Figure 4 shows Personal Entropy versus Complexity and Variability indicators, per category on DS2-104 and MCYT-100. We see that signatures of highest Personal Entropy are highly variable and of rather low complexity. At the opposite, signatures of lowest Personal Entropy are by far more complex and more stable (show low variability). We noticed

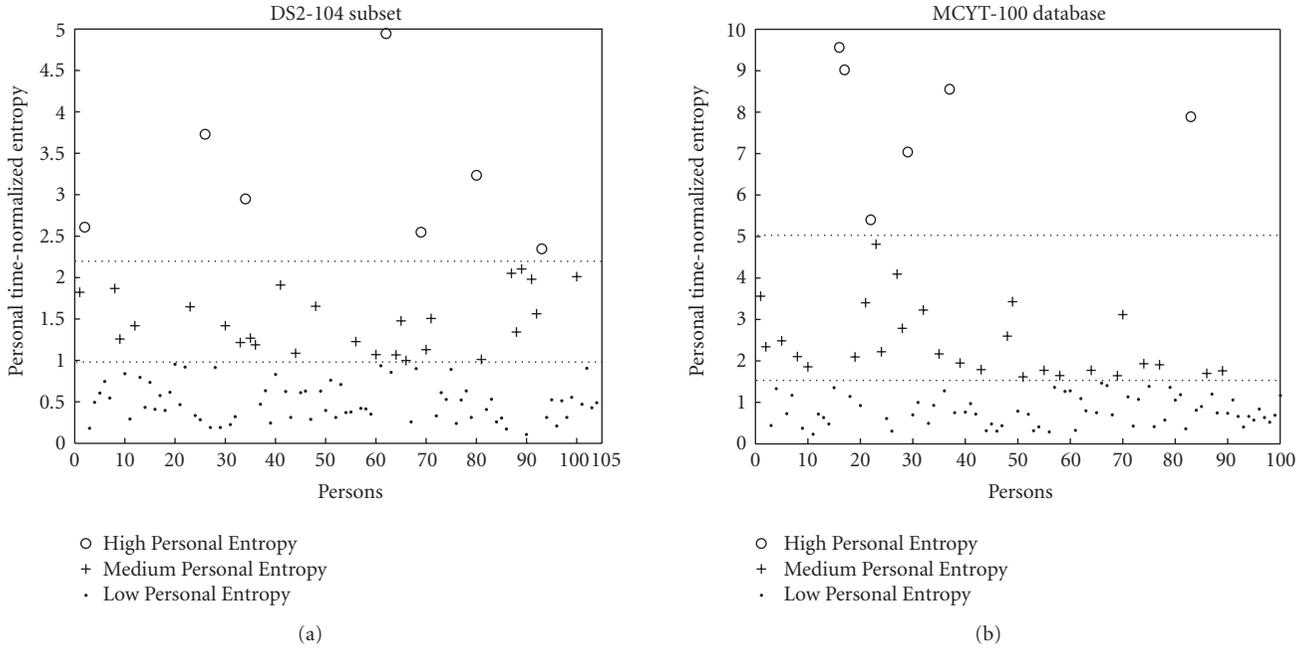


FIGURE 3: Personal Entropy values on data from DS2-104 (a) and from MCYT-100 (b) across the 3 writer categories. Two indicative lines report the separation between categories obtained by the K-Means procedure.

that this behaviour is verified for all the databases considered in this work. We therefore conclude that our Personal Entropy measure allows quantifying both the complexity and variability of a writer’s signatures simultaneously.

3. Verification Performance

In this section, we study the relationship between Personal Entropy-based categories and performance of two different automatic signature verification systems, on two different databases: DS2-104 and MCYT-100.

3.1. Score Computation by the Two Classifiers. Two classifiers are used in this study considering only the raw coordinates description of signatures as input data: a Hidden Markov Model [3] and a Gaussian Mixture Model [14].

For performance assessment, both skilled and random forgeries are considered. Ten random samplings are carried out on genuine and impostor signatures in the following way: each sampling contains five genuine signatures used as the training set for both statistical classifiers. For test purposes, the remaining 25 genuine signatures and 20 skilled forgeries (belonging to two sessions) are used for DS2-104. For MCYT-100, we tested on the remaining 20 genuine signatures and 25 skilled forgeries. Also, 30 impostor signatures randomly sampled in equal number in each Personal Entropy category (10 random forgeries per category) are considered for both databases. The False Acceptance and False Rejection Rates are computed relying on the total number of False Rejections and False Acceptances obtained on the whole ten random samplings.

Concerning the topology of the two statistical models, we used a GMM and a left-to-right HMM of the same complexity in terms of Gaussian components. It is worth noticing that the HMM classifier differs from the HMM used for Personal Entropy computation. Indeed, the former is devoted to classification, while the latter only performs local density estimation. We considered for the HMM classifier a 6 states and 4 Gaussian components per state, as a tradeoff in complexity between the signatures of the two extreme categories. For the GMM, accordingly, we considered 24 Gaussians to model a person’s signatures. The dissimilarity matching score for both statistical models is

$$\text{Score} = |\text{LL} - \text{LL}_{BA}|, \quad (3)$$

where LL is the Log-Likelihood of the test signature (normalized by the length of the test signature), and LL_{BA} is the corresponding average Log-Likelihood of the training signatures.

3.2. Performance Assessment on DS2-104 and MCYT-100 with the Two Classifiers. In our experiments, both HMM and GMM classifiers were intentionally not optimized, since our aim is not to improve absolute system performance but to analyze the relative differences in classifiers’ performance between writer categories.

We notice on Figures 5 and 6 corresponding to DS2-104, and on Figures 7 and 8 corresponding to MCYT-100, that the results lead to different behaviours in terms of performance according to the category of Personal Entropy that we consider.

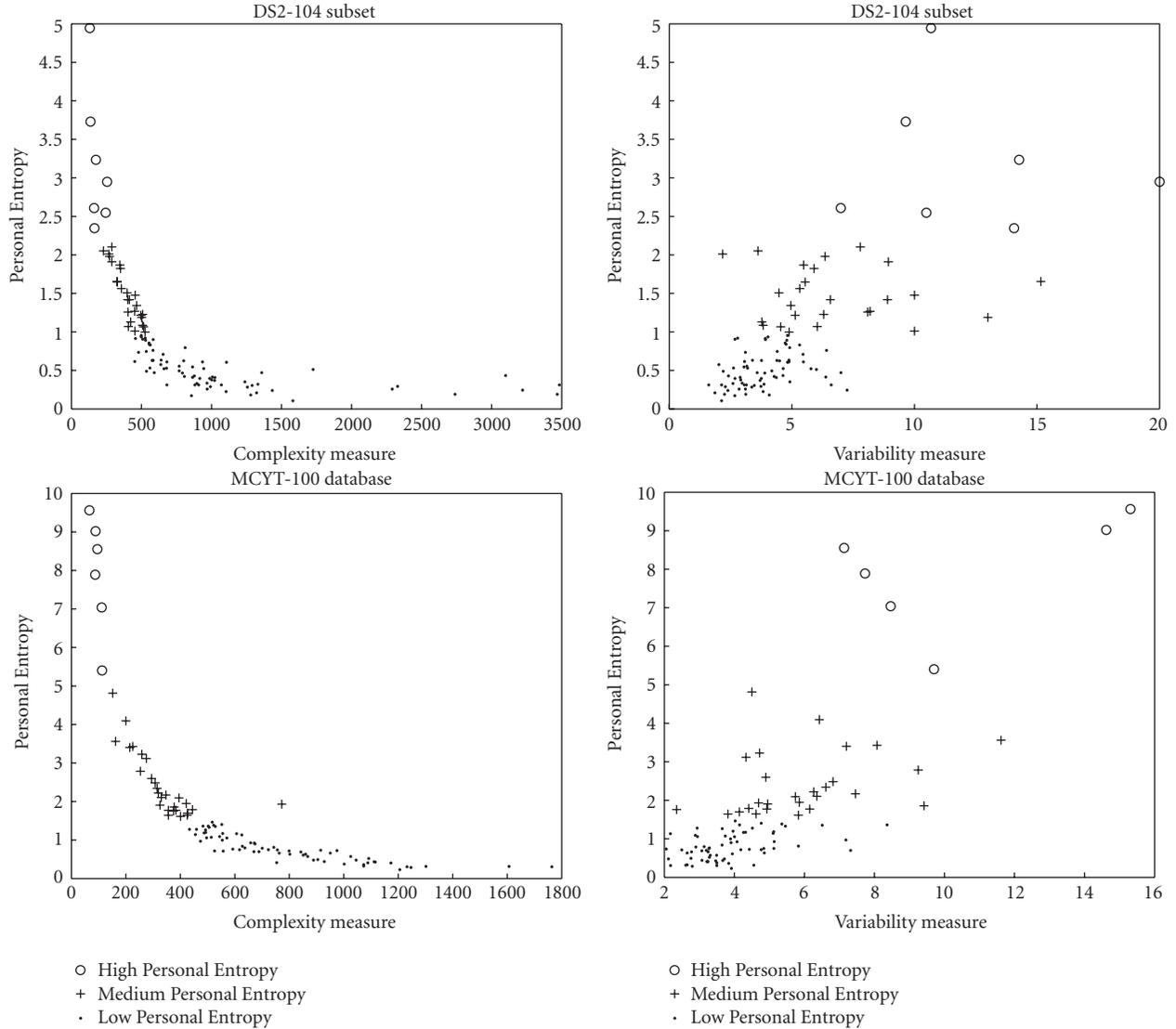


FIGURE 4: Personal Entropy versus complexity (left) and Personal Entropy versus variability (right) on MCYT-100 and DS2-104 databases, for Personal Entropy-based categories.

There is a significant difference in classifiers' performance between the two extreme categories, for both skilled and random forgeries: GMM and HMM classifiers give the best performance on writers belonging to the category of lowest Personal Entropy, that is, those having the longest most complex and most stable signatures, as those shown in Figure 2(c). At the opposite, HMM and GMM classifiers give the worst performance on writers belonging to the highest Personal Entropy, those having the shortest simplest and most unstable signatures, as those shown in Figure 2(a). We also notice that performance values for the category of writers with medium Personal Entropy are in between those of the two extreme writer categories.

As shown in Tables 1 and 2, for the two classifiers, at the Equal Error Rate functioning point, performance is roughly improved by a factor around 2 for skilled and random forgeries when switching from the highest entropy

category to the lowest one, on both DS2-104 and MCYT-100. Confidence Intervals at 95% are given to show the significance of results. At other functioning points, this gap in performance between the two extreme categories is maintained for the two classifiers, as shown in Figures 5, 6, 7, and 8.

For a better insight on the impact of high and medium Personal Entropy categories on system performance, we ordered, in a decreasing way, users from such categories according to their Personal Entropy. Then, we compute when removing the top $x\%$ of such users, the relative improvement $\Delta(x)$ of the Equal Error Rate with regard to the average EER on the whole DS2-104 database (denoted by $\overline{\text{EER}}$) defined as follows:

$$\Delta(x) = \frac{\overline{\text{EER}} - \text{EER}(x)}{\overline{\text{EER}}}, \quad (4)$$

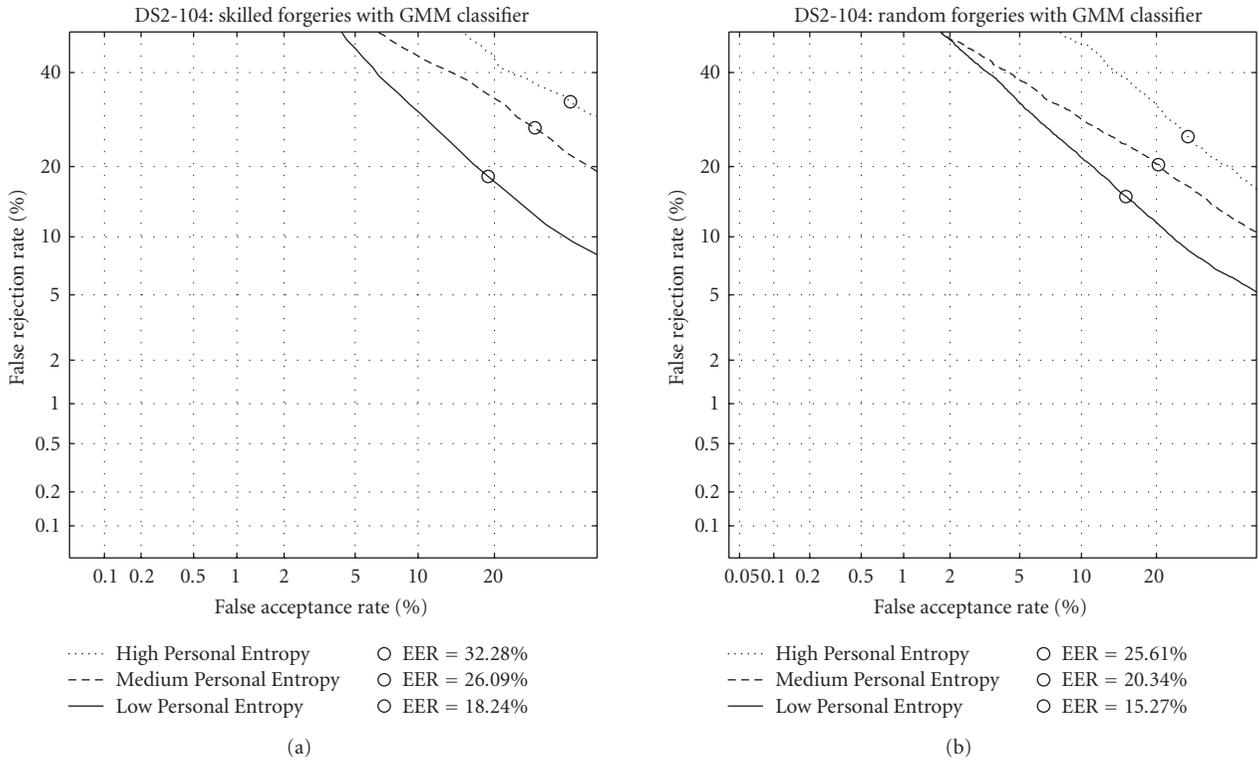


FIGURE 5: DET-curves considering skilled forgeries (a) and random forgeries (b), on each writer category on DS2-104 subset with the GMM classifier.

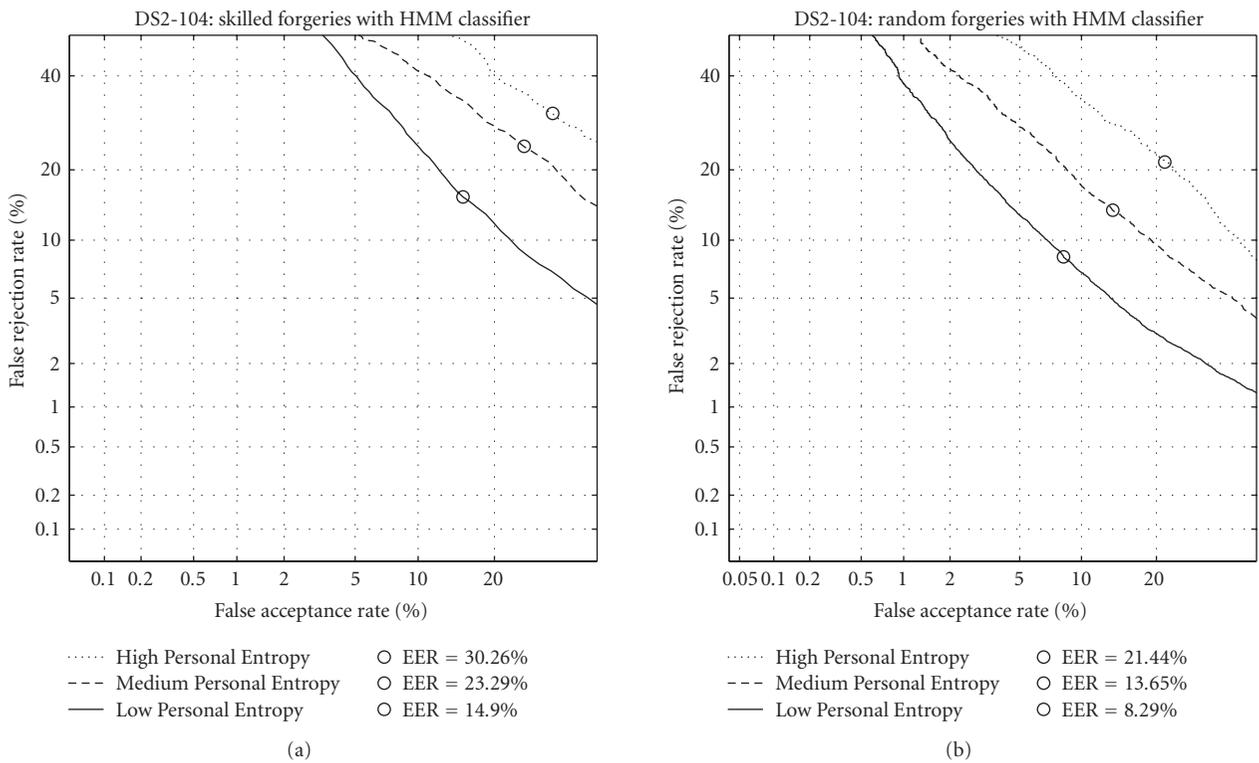


FIGURE 6: DET-curves considering skilled forgeries (a) and random forgeries (b), on each category on DS2-104 subset with the HMM classifier.

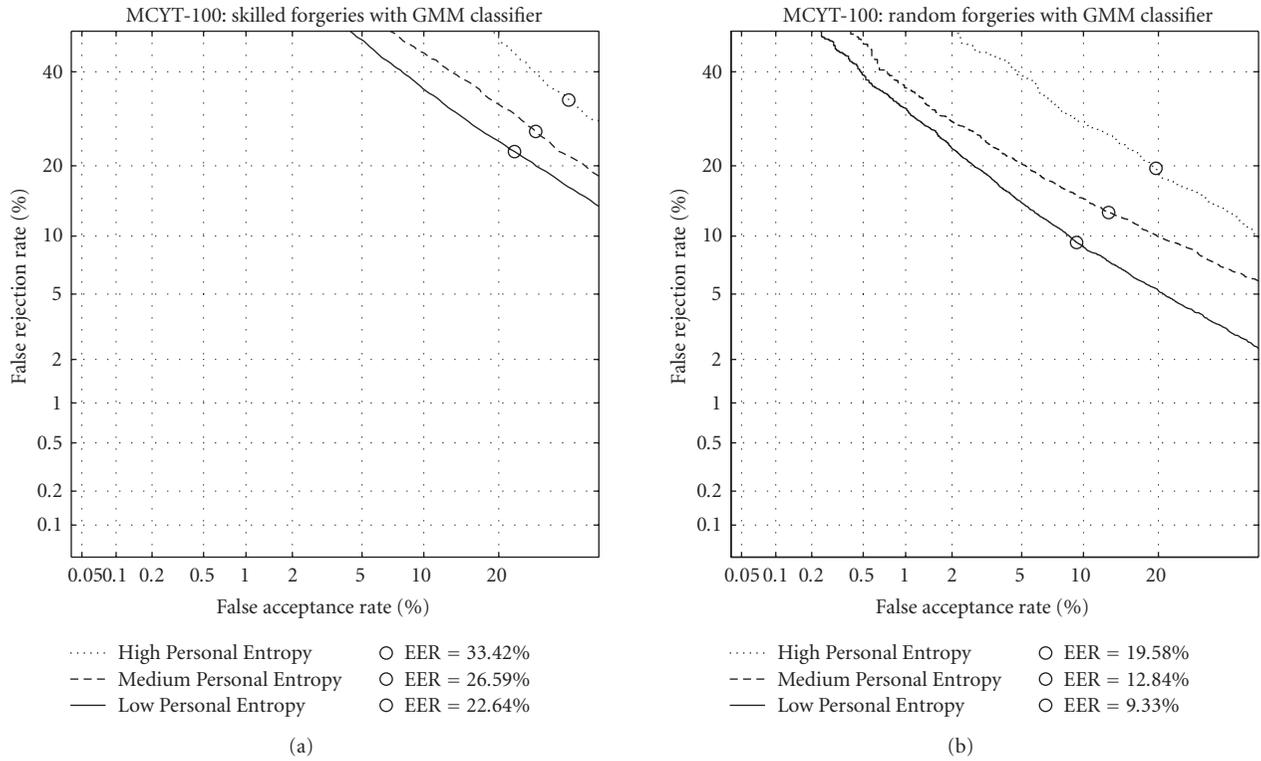


FIGURE 7: DET-curves considering skilled forgeries (a) and random forgeries (b), on each writer category on MCYT-100 database with the GMM classifier.

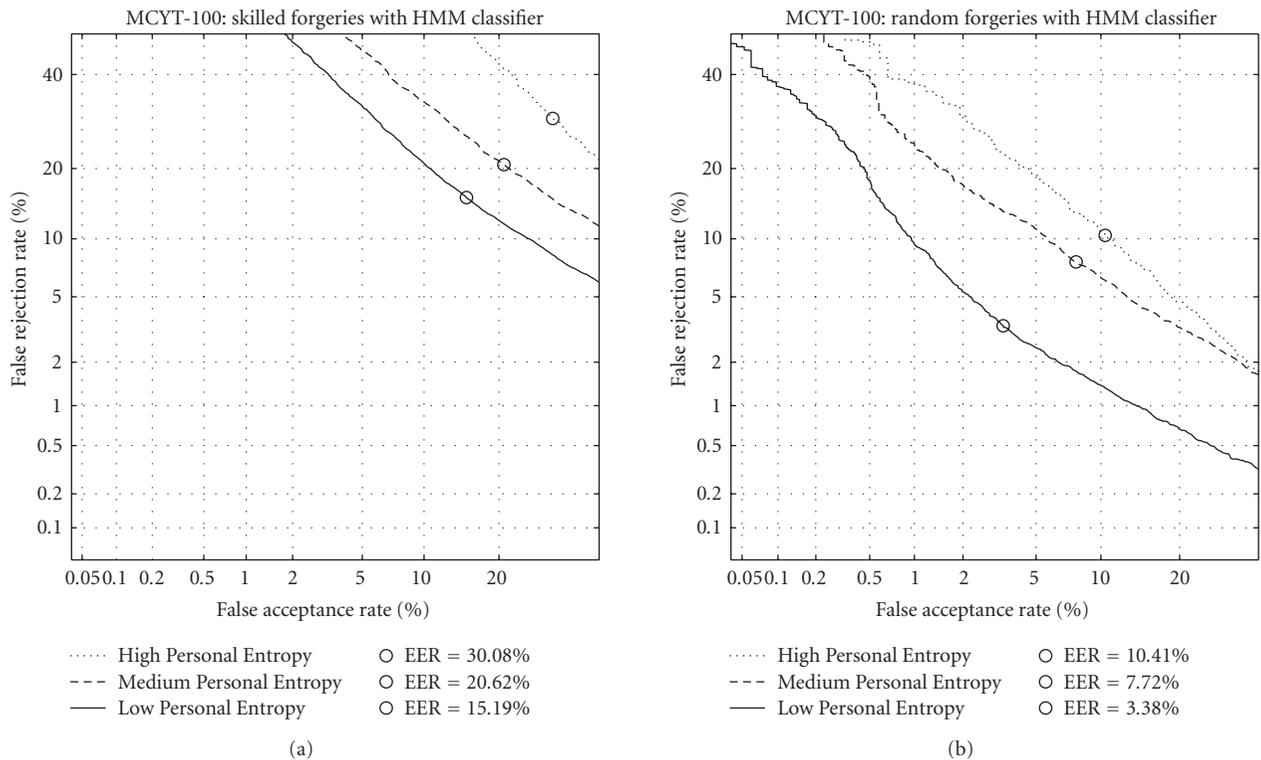


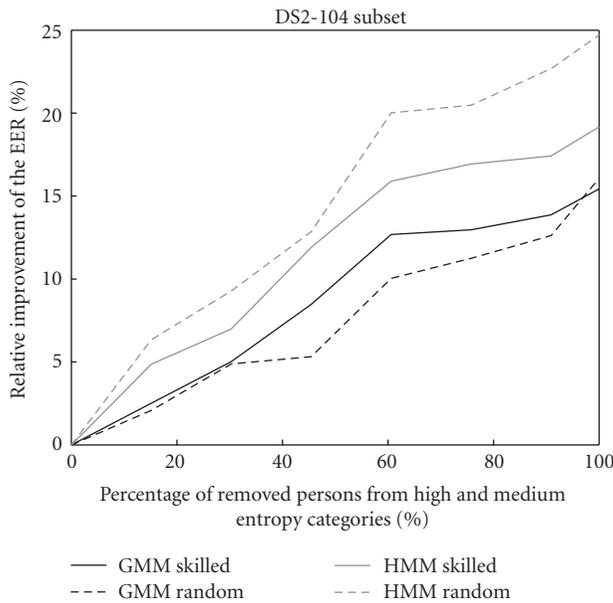
FIGURE 8: DET-curves considering skilled forgeries (a) and random forgeries (b), on each category on MCYT-100 database with the HMM classifier.

TABLE 1: Equal Error Rate and Confidence Interval in each writer category on DS2-104 subset, with HMM and GMM classifiers considering skilled and random forgeries.

	DS2-104 subset							
	GMM classifier				HMM classifier			
	Skilled forgeries		Random forgeries		Skilled forgeries		Random forgeries	
	EER (%)	CI (95%)	EER (%)	CI (95%)	EER (%)	CI (95%)	EER (%)	CI (95%)
High entropy	32.28	± 0.100	25.61	± 0.057	30.26	± 0.100	21.44	± 0.080
Medium entropy	26.09	± 0.040	20.34	± 0.026	23.29	± 0.027	13.65	± 0.018
Low entropy	18.24	± 0.010	15.27	± 0.007	14.90	± 0.009	8.29	± 0.001

TABLE 2: Equal Error Rate and Confidence Interval in each writer category on MCVT-100 database, with HMM and GMM classifiers considering skilled and random forgeries.

	MCYT-100 database							
	GMM classifier				HMM classifier			
	Skilled forgeries		Random forgeries		Skilled forgeries		Random forgeries	
	EER (%)	CI (95%)	EER (%)	CI (95%)	EER (%)	CI (95%)	EER (%)	CI (95%)
High entropy	33.42	± 0.170	19.58	± 0.160	30.08	± 0.200	10.76	± 0.120
Medium entropy	26.59	± 0.050	12.84	± 0.028	20.62	± 0.042	7.56	± 0.023
Low entropy	22.64	± 0.018	9.33	± 0.006	15.74	± 0.010	4.13	± 0.003

FIGURE 9: The relative improvement $\Delta(x)$ of the average EER on DS2-104 subset when removing $x\%$ of users from high and medium Personal Entropy categories.

where $EER(x)$ represents the average Equal Error Rate on the whole DS2-104 database after removing $x\%$ of users from high and medium Personal Entropy categories.

We notice in Figure 9 that for both the GMM and HMM classifiers, and both random and skilled forgeries, when removing progressively an increasing percentage x of users from high and medium Personal Entropy categories (according to their Personal Entropy measure), $\Delta(x)$ increases. When

TABLE 3: The relative improvement $\Delta(x)$ of the average EER on DS2-104 subset when removing all users from high and medium Personal Entropy categories.

Classifier	Type of forgeries	\overline{EER}	Δ (100%)
GMM	Skilled forgeries	21.57%	15.43%
	Random forgeries	18.19%	16.06%
HMM	Skilled forgeries	18.43%	19.16%
	Random forgeries	10.90%	23.94%

all users from high and medium Personal Entropy categories are removed ($x = 100\%$), this relative improvement $\Delta(x)$ reaches in all cases more than 15%, as reported in detail in Table 3. Moreover, given that the first 21% of users belong to the high Personal Entropy category (7 users), and the remaining 79% belong to the medium Personal Entropy category (26 users), we conclude that the main improvement is obtained when the first 60% of users are removed (that is all users from the high Personal Entropy category and 50% of users from the medium Personal Entropy category).

4. Generalizing Writer Categorization

4.1. On Categorizing New Writers Relying on Entropy-Prototypes Obtained Offline. We have this far shown that there is one category of users which are much easier to recognize than others, and much easier to discriminate from skilled and random forgeries, those having a low Personal Entropy value. Alternatively, there is another category of users which are extremely difficult to recognize, those having a high Personal Entropy value.

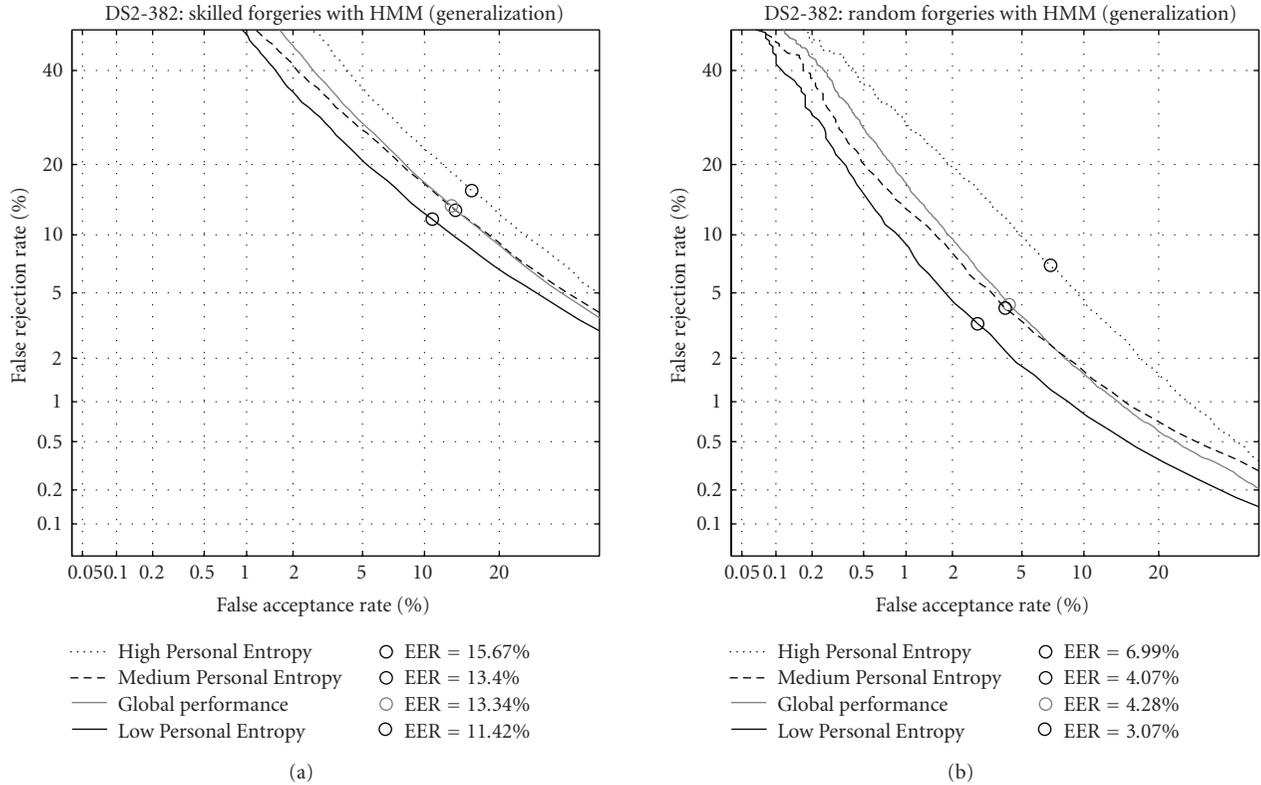


FIGURE 10: DET-curves considering skilled forgeries (a) and random forgeries (b), on each writer category and globally on DS2-382 database with the HMM classifier, after computing entropy-prototypes on DS2-104.

TABLE 4: Equal Error Rate and Confidence Interval in each writer category on DS2-382 database, with the HMM classifier considering skilled and random forgeries.

	DS2-382 with HMM classifier			
	Skilled forgeries		Random forgeries	
	EER (%)	CI (95%)	EER (%)	CI (95%)
High entropy	15.67	± 0.025	6.99	± 0.015
Medium entropy	13.4	± 0.012	4.07	± 0.006
Low entropy	11.42	± 0.014	3.07	± 0.005
Global performance	13.34	± 0.003	4.28	± 0.003

Each writer category is naturally associated to an Entropy-Prototype (EP) inherited from the K-Means procedure used to “cluster” writers. Our aim in this section is to study the possibility of categorizing new writers based on previously generated Entropy-Prototypes (EPs), on a data set of limited size. We carry out this study by generating three Entropy-Prototypes on DS2-104, and using such prototypes to categorize writers from another data set: DS2-382.

Indeed, we categorize a writer belonging to such data set as follows:

- (1) computing the writer’s Personal Entropy with 10 genuine signatures of such writer from DS2-382;
- (2) retrieving the three Entropy-Prototypes (one per category) computed offline on DS2-104 database;
- (3) associating to such writer from DS2-382 the category of closest Entropy-Prototype by the Nearest Neighbor Rule [24].

In order to study the relevance of the previous protocol, we study performance on the obtained categories after generalization. In order to carry out this study, we only consider in the following an HMM classifier, since the same results are obtained with a GMM classifier.

4.2. Generalization on the Same Database from DS2-104 to DS2-382. Figure 10 and Table 4 show the performance obtained on DS2-382 with an HMM classifier on each of the obtained categories after computing Entropy-Prototypes on DS2-104, with skilled and random forgeries respectively. We also compare results per category to global results on the complete DS2-382 database.

As on DS2-104, on which the Entropy-Prototypes have been computed originally, we notice that also on DS2-382 there is a difference in classifiers' performance between the two extreme categories, for both skilled and random forgeries: the HMM classifier gives the best performance on writers belonging to the category of lowest Personal Entropy. At the opposite, the HMM classifier gives the worst performance on writers belonging to the highest Personal Entropy.

As shown in Table 4, at the Equal Error Rate functioning point, performance is roughly improved by a factor 2 for skilled forgeries and 1.4 for random forgeries when switching from the highest entropy category to the lowest one. We also notice that performance values for the category of writers with medium Personal Entropy are in between those of the two extreme writer categories.

Moreover, the global performance on the whole data set DS2-382 is degraded compared to performance on the category of writers with lowest Personal Entropy.

4.3. Our Proposed Criterion for Writer Enrolment. We have shown that Entropy-Prototypes generated offline on a database of limited size (104 persons) can be used to perform writer categorization on new writers from the same database. We thus propose to exploit such Entropy-Prototypes, which are totally independent of the verification system, to identify beforehand signatures that are not secure in terms of performance. The enrolment procedure that we propose has the following steps.

- (1) Ten genuine signatures are requested from the writer to be enrolled.
- (2) A Writer-HMM is built for such writer by training the HMM on such ten genuine signatures.
- (3) The Personal Entropy of such writer is computed.
- (4) The three Entropy-Prototypes computed offline are retrieved.
- (5) The category of the closest Entropy-Prototype by the Nearest Neighbor Rule [24] is associated to the writer.
- (6) When a writer is classified as belonging to the highest Personal Entropy category, he/she should be informed of the intrinsic risk related to his/her signature. Indeed, this category of writers gives unreliable results relatively to other Personal Entropy categories; we thus propose to the user either to pursue enrolment knowing the intrinsic risk of his/her signature, or alternatively to change his/her signature for security purposes.
- (7) When a writer belongs to the category of lowest Personal Entropy, the writer is enrolled.
- (8) When a writer belongs to the category of middle Personal Entropy, we recommend to the writer to do a more complex and less variable signature, but still can retain his/her signature.

Based on our experiments, we can assert that the more Personal Entropy lowers, the more reliable is the signature in

terms of security. This should be to take into account when using online signature in practical applications.

5. Conclusion

We have proposed a novel criterion for writer enrolment that allows guaranteeing a higher level of security to the individual writer, regardless of the verification system that is used. Such criterion relies on an unsupervised automatic writer categorization process, carried out on a Time-Normalized Personal Entropy measure, quantified in bits per second. We first introduce in this work a "Sample Entropy" measure associated to each enrolment signature sample, computed locally by means of a Writer-HMM trained on ten enrolment signatures. Then we explain how the resulting "Time-Normalized Personal Entropy" value of each writer is retrieved.

We show that a writer can be categorized according to this measure and to Entropy-Prototypes computed offline, into one of three categories of writers. This categorization process is crucial because verification systems' performance is significantly different between the extreme categories of highest and lowest Personal Entropy. Indeed, we show across two data sets that our Personal Entropy measure allows classifying writers automatically into three visually coherent and linearly separable categories, opposing long, complex and stable signatures to short, strongly variable and simple signatures. Moreover, we have quantified the behaviour of the signature in terms of complexity and variability, and we have linked these values to our Personal Entropy measure.

We have shown that Entropy-Prototypes, naturally inherited from the K-Means procedure and performed to generate writer categories, can be generated offline on a data set of limited size (around 100 persons) and be used to perform writer categorization on new writers of another data subset, providing the same acquisition conditions. More generally, a database of roughly 100 persons is enough to generate the categories, then allowing the online categorization of any new user whose signatures are acquired in similar conditions (same digitizer, similar tablet resolution, and same acquisition protocol).

Based on this result, we propose an enrolment writer criterion related to such Entropy-Prototypes, totally independent of the verification system, to identify beforehand signatures which are not secure in terms of performance. Indeed, a Nearest Neighbour Rule on Entropy-Prototypes generated offline, on a database of roughly 100 persons, allows categorizing a writer after requesting from him ten instances of his/her signature. A stable and reliable result emerges of our study: the more Personal Entropy lowers, the more reliable is the signature in terms of security. This statement allows adapting the quality of the enrolment data to the level of security requested by the application.

Acknowledgment

The authors thank Javier Ortega-Garcia and his colleagues for putting at disposal the subset of the first 100 users of MCYT Signature Subcorpus.

References

- [1] D.-Y. Yeung, H. Chang, Y. Xiong, et al., "SVC2004: first international signature verification competition," in *Proceedings of International Conference on Biometric Authentication (ICBA '04)*, vol. 3072 of *Lecture Notes in Computer Science*, pp. 16–22, Springer, Hong Kong, July 2004.
- [2] <http://biometrics.it-sudparis.eu/BMEC2007>.
- [3] L. Rabiner and B. H. Juang, *Fundamentals of Speech Recognition*, Signal Processing Series, Prentice-Hall, Englewood Cliffs, NJ, USA, 1993.
- [4] A. K. Jain, F. D. Griess, and S. D. Connell, "On-line signature verification," *Pattern Recognition*, vol. 35, no. 12, pp. 2963–2972, 2002.
- [5] A. Kholmatov and B. Yanikoglu, "Identity authentication using improved online signature verification method," *Pattern Recognition Letters*, vol. 26, no. 15, pp. 2400–2408, 2005.
- [6] J. G. A. Dolfig, *Handwriting recognition and verification, a Hidden Markov approach*, Ph.D. thesis, Philips Electronics N. V., 1998.
- [7] R. Kashi, J. Hu, W. L. Nelson, and W. Turin, "A Hidden Markov Model approach to online handwritten signature verification," *International Journal on Document Analysis and Recognition*, vol. 1, pp. 102–109, 1998.
- [8] G. Rigoll and A. Kosmala, "A systematic comparison of on-line and off-line methods for signature verification with Hidden Markov Models," in *Proceedings of the 14th International Conference on Pattern Recognition (ICPR '98)*, pp. 1755–1757, Brisbane, Australia, August 1998.
- [9] J. Ortega-Garcia, J. Gonzalez-Rodriguez, D. Simon-Zorita, and S. Cruz-Llanas, "From biometrics technology to applications regarding face, voice, signature and fingerprint recognition systems," in *Biometrics Solutions for Authentication in an E-World*, D. Zhang, Ed., pp. 289–337, Kluwer Academic Publishers, Dordrecht, The Netherlands, July 2002.
- [10] J. G. A. Dolfig, E. H. L. Aarts, and J. J. G. M. Van Oosterhout, "On-line signature verification with hidden Markov models," in *Proceedings of the International Conference on Pattern Recognition*, pp. 1309–1312, Brisbane Australia, 1998.
- [11] J. Fierrez-Aguilar, L. Nanni, J. Lopez-Peñalba, J. Ortega-Garcia, and D. Maltoni, "An on-line signature verification system based on fusion of local and global information," in *Proceedings of the 5th IAPR International Conference on Audio- and Video-Based Biometric Person Authentication (AVBPA '05)*, Lecture Notes in Computer Science, Springer, 2005.
- [12] D. Muramatsu and T. Matsumoto, "An HMM on-line signature verifier incorporating signature trajectories," in *Proceeding of the 7th International Conference on Document Analysis and Recognition (ICDAR '03)*, IEEE, Edinburgh, Scotland, August 2003.
- [13] B. Ly Van, S. Garcia-Salicetti, and B. Dorizzi, "On using the Viterbi path along with HMM likelihood information for online signature verification," *IEEE Transactions on Systems, Man, and Cybernetics, Part B*, vol. 37, no. 5, pp. 1237–1247, 2007.
- [14] D. A. Reynolds and R. C. Rose, "Robust text-independent speaker identification using Gaussian mixture speaker models," *IEEE Transactions on Speech and Audio Processing*, vol. 3, no. 1, pp. 72–83, 1995.
- [15] J. Richiardi and A. Drygajlo, "Gaussian mixture models for on-line signature verification," in *Proceedings of the ACM SIGMM Workshop on Multimedia Biometrics Methods and Applications (WBMA '03)*, pp. 115–122, Berkley, Calif, USA, November 2003.
- [16] J.-J. Brault and R. Plamondon, "How to detect problematic signers for automatic signature verification," in *Proceedings of the International Canadian Conference on Security Technology (ICCST '89)*, pp. 127–132, Zurich, Switzerland, 1989.
- [17] V. Di Lecce, G. Di Mauro, A. Guerriero, et al., "Selection of reference signatures for automatic signature verification," in *Proceedings of International Conference on Document Analysis and Recognition (ICDAR '99)*, pp. 597–600, Bangalore, India, 1999.
- [18] C. Allgrove and M. C. Fairhurst, "Enrolment model stability in static signature verification," in *Proceedings of International Workshop on Frontiers in Handwriting Recognition (IWFHR '00)*, pp. 565–570, Amsterdam, The Netherlands, 2000.
- [19] F. Alonso-Fernandez, M. C. Fairhurst, J. Fierrez, and J. Ortega-Garcia, "Impact of signature legibility and signature type in off-line signature verification," in *Proceedings of the IEEE Biometrics Symposium (BSYM '07)*, Baltimore, Md, USA, September 2007.
- [20] V. Boulétreau, *Towards a handwriting classification by fractal methods*, Ph.D. dissertation, Institut National des Sciences Appliquées de Lyon, Lyon, France, 1997.
- [21] T. M. Cover and J. A. Thomas, *Elements of Information Theory*, John Wiley & Sons, New York, NY, USA, 2nd edition, 2006.
- [22] S. Garcia-Salicetti, N. Houmani, and B. Dorizzi, "A client-entropy measure for on-line signatures," in *Proceedings of the IEEE Biometrics Symposium (BSYM '08)*, pp. 83–88, Tampa, Fla, USA, September 2008.
- [23] N. Houmani, S. Garcia-Salicetti, and B. Dorizzi, "A novel personal entropy measure confronted with online signature verification systems' performance," in *Proceedings of the 2nd IEEE International Conference on Biometrics: Theory, Applications and Systems (BTAS '08)*, Washington, DC, USA, September 2008.
- [24] R. O. Duda and P. E. Hart, *Pattern Classification*, Wiley-Interscience, New York, NY, USA, 2nd edition, 2000.
- [25] <http://www.biosecure.info>.
- [26] J. Ortega-Garcia, J. Fierrez-Aguilar, D. Simon, et al., "MCYT baseline corpus: a bimodal biometric database," *IEE Proceedings: Vision, Image and Signal Processing*, vol. 150, no. 6, pp. 395–401, 2003.
- [27] V. S. Chakravarthy and B. Kompella, "The shape of handwritten characters," *Pattern Recognition Letters*, vol. 24, no. 12, pp. 1901–1913, 2003.