IEEE Access
practical innovations : open solutions

Received August 9, 2013, accepted August 21, 2013, date of publication September 16, 2013, date of current version September 25, 2013.

Digital Object Identifier 10.1109/ACCESS.2013.2281773

The DooDB Graphical Password Database: Data Analysis and Benchmark Results

MARCOS MARTINEZ-DIAZ, JULIAN FIERREZ (Member, IEEE), AND JAVIER GALBALLY
Biometric Recognition Group, ATVS, Universidad Autonoma de Madrid, Madrid 28049, Spain

Corresponding author: M. Martinez-Diaz (marcos.martinez@uam.es)

This work was supported by projects Contexts under Grant S2009/TIC-1485 from CAM, Bio-Challenge under Grant TEC2009-11186 and Bio-Shield under Grant TEC2012-34881 from Spanish MECD, and BEAT under Grant FP7-SEC-284989 from EU.

ABSTRACT We present DooDB, a doodle database containing data from 100 users captured with a touch screen-enabled mobile device under realistic conditions following a systematic protocol. The database contains two corpora: 1) doodles and 2) pseudo-signatures, which are simplified finger-drawn versions of the handwritten signature. The dataset includes genuine samples and forgeries, produced under worst-case conditions, where attackers have visual access to the drawing process. Statistical and qualitative analyses of the data are presented, comparing doodles and pseudo-signatures to handwritten signatures. Time variability, learning curves, and discriminative power of different features are also studied. Verification performance against forgeries is analyzed using state-of-the-art algorithms and benchmark results are provided.

INDEX TERMS Graphical password, doodle verification, pseudo-signature.

I. INTRODUCTION

The pervasive presence of touchscreens in entertainment devices, tablets, and mobile phones has led to new interaction capabilities. One of them is user-authentication based on graphical passwords, a topic that has been the subject of active research as a replacement of alphanumerical passwords [1], [2]. The term ''graphical password'' refers to many different graphical authentication methods, which can be broadly classified in three categories: 1) recall, 2) recognition, and 3) cued-recall passwords. Recall-based systems assume that users remember a graphical password during authentication. Recognition systems present graphical information to the user during authentication, from which the user has to perform a selection matching a set of information previously memorized. Cued-recall systems are a hybrid between the two aforementioned, providing graphical cues that help users recall the previously learned password. An extensive survey of graphical password algorithms has been compiled in [1].

In the present work we focus in doodle-based passwords, which are a subset of recall graphical passwords. Individuals are authenticated by using a drawing or sketch, that is captured on a touchscreen during enrollment and is used afterwards for verification. Due to their graphical nature, they are in general easier to remember than classical alphanumerical passwords or PIN codes composed of strings of characters and numbers [3].

It is well known that publicly available databases together with their associated evaluation protocols make possible that researchers develop and objectively compare pattern recognition algorithms on the same benchmark. Experiments carried out using private databases are usually hard to replicate since database-specific effects, which cannot be reproduced by a third party, may take place. Unfortunately, there is no such a public database in the field of doodle-based graphical passwords, to the extent of our knowledge. Experiments carried out in the last years related to doodle verification have used proprietary databases [4]–[7]. Moreover, in these works there is no reference to forgeries, since only genuine doodles are considered.

The main contribution of this work is the presentation and analysis of DooDB, a doodle and pseudo-signature database containing data from 100 users. Pseudo-signatures are doodles based on a simplified version of the user signature, being thus composed of learned and natural movements. The database has been captured on a handheld device under realistic conditions. It has two main advantages compared to other databases used in the literature: two acquisition sessions were performed, so inter-session variability effects can be analyzed, and skilled forgeries are provided for each user. The DooDB database is publicly available from the ATVS - Biometric Recognition Group website (http://atvs.ii.uam.es).

Another objective of this work is to obtain a baseline doodle verification performance that can be used to compare this method with current well known authentication alternatives such as signatures or with future doodle-based recognition algorithms. We also analyze the differences in the verification performance between doodles and pseudo-signatures. Since pseudo-signatures are simplified versions of real signatures, and thus composed of learned movements, it can be hypothesized that they present a lower variability and a better verification performance. The effects of inter-session variability are also studied.

The paper is structured as follows. In Sect. II related works are summarized. In Sect. III the database is described. Quantitative and qualitative properties of the database are analyzed in Sect. IV. Preliminary verification experiments using the data from DooDB are reported in Sect. V and conclusions are finally drawn in Sect. VI.

II. RELATED WORK

Several approaches have been proposed in the literature related to doodle-based user authentication. A survey of the techniques presented over the last years has been reported recently in [1].

One of the first contributions in the field is the Draw-A-Secret system (DAS) [4]. The DAS system implements a grid where users trace their graphical password. The sequence of grid cells (5×5 in that work) that the users follow is then stored and used for validation. Users are accepted by the system only if they follow the same sequence of cells.

The term ''passdoodle'' was later introduced in [5]. In that work, the memorability (i.e. the easiness to remember) of doodles for user authentication is studied, as well as the user preference towards alphanumeric passwords or doodles. However, it is a preliminary study carried out with doodles traced on a sheet of paper by a population of 39 users.

A passdoodle verification system is also proposed in [8]. The stroke spatial distribution and the speed are used for verification. Experiments are performed with a database containing doodles from 10 users, providing thus a limited statistical significance.

A doodle authentication system which uses Dynamic Time Warping (DTW) for matching is described in [7]. Recognition performance results are provided using Tamil characters, instead of doodles.

The Scribble-A-Secret (SAS) scheme for doodle verification was later proposed [6]. This system uses edge orientation patterns as features, extracted from the doodle image (no dynamic information is used). Experiments are carried out with doodles from 87 individuals, traced on a Tablet-PC touchscreen. Forgeries are simulated by comparing doodles from different individuals.

In [9], a verification scheme based on predefined visual cues is presented. The cues are chosen by the user with the aid of a graphical interface. With these cues (which are in general common shapes), each user creates what is called a pseudo-signature. Cryptographic keys are then generated from the pseudo-signatures. Experimental results are reported using a database of 37 subjects, which includes forgeries.

A graphical password verification system based on a set of predefined symbols is proposed in [10]. During enrollment, the user first selects a set of predefined symbols (at least 3) and then draws them. The set of symbols is considered the user password. During authentication, the symbols must be drawn in the same order and are then matched to the predefined templates. If the drawn set is the same as the registered set, the user is validated. No experimental results are provided.

A multi-touch authentication approach has been presented in [11]. In that work, users perform gestures with several fingers at the same time. Since the gesture used for authentication is produced with all the fingers, the authors claim that information from the hand geometry is also captured. Experimental results are reported on a database of 34 users.

The Pattern Lock found in Android OS portable devices is a widespread graphical password scheme. This method presents a square grid of 3×3 points on the screen. The user must trace a pattern between the points without repeating any of them. This resembles a simplified version of the Draw-a-Secret scheme, mitigating the effect of strokes near cell edges. Only the sequence of points is stored as a password, so no dynamic information such as speed or duration is used for verification. Other approaches that also use dynamic information from the Pattern Lock drawing process have been proposed [12], [13].

In [14], an authentication scheme based on continuous touchscreen input, instead of specific gestures, is presented. Results are reported using a set of 41 users.

A. ATTACKS TO GRAPHICAL PASSWORDS

Doodle-based graphical passwords are vulnerable to two main types of attacks. Smudge attacks are those produced when the attacker follows the finger grease path left by the user on the screen [15]. Shoulder-surfing attacks represent the case when the attacker gains visual access to the password drawing process.

In all the aforementioned works, except [9], shoulder-surfing forgeries (i.e. attacks) are not taken into account systematically in the experiments, to the extent of our knowledge. Only random forgeries are considered, which is the case where the attacker claims to be a different user but produces his or her own doodle to try to access the system. Thus, doodles from different users are considered as forgeries of each other, which is a simplified scenario compared to the case of intentional forgeries.

B. GRAPHICAL PASSWORDS VS. HANDWRITTEN SIGNATURES

Doodle-based verification can also be seen as a variant of dynamic handwritten signature verification [16], since signatures can be acquired using touchscreens. Signature verification on handheld devices has also been studied by the authors and is a field which has recently been given

large attention within biometrics [17]. This wide interest is shown by the acquisition of the BioSecure Multimodal Database [18], in which a corpus of signatures was acquired on a PDA. An international evaluation campaign with signatures captured on that database was organized in 2009, with the participation of several research institutions [19].

Doodle verification and signature verification share that in both cases behavioral information (e.g. gesture dynamics) is used for matching. However, compared to signatures, doodles are commonly not composed of natural and trained movements that users have performed for several years. While this may be the source of a larger variability, it is also an advantage for doodles since, unlike signatures, they can be easily replaced if necessary (which is known as revocability in the biometric field). Doodles may also have a higher acceptability, since a number of users are reluctant to provide their digitized signatures in some environments.

III. THE DooDB DATABASE

The DooDB database comprises two subcorpora, each one containing a different modality:

- **Subcorpus 1: Doodles**. Participants were asked to draw with their fingertip a doodle on a handheld device touchscreen that they would use as a graphical password on a regular basis for authentication (e.g. instead of the PIN code). There were no restrictions regarding duration or shape. In most cases, users invented their own doodle at the time of acquisition.
- **Subcorpus 2: Pseudo-signatures**. Participants were also asked to draw with their fingertip a simplified version of their signature, which they would also use as a graphical password on a regular basis. This could be, for example, their initials or part of their signature flourish. The main difference between doodles and this modality is that in this case, the dynamic process to produce the drawing is in general composed of natural and well trained movements.

A. ACQUISITION PROTOCOL

Acquisition was performed using an HTC Touch HD mobile phone (see Fig. 1). The device has a resistive touchscreen of 2×3.5 in (ca. 5×8.5 cm). The x and y coordinates of the fingertip position are sampled at discrete time values t at 100Hz when the user presses the screen. The coordinate values represent milli-inches, so x_t values range between $[0, 2000]$ (width) and y_t values between $[0, 3500]$ (height). The time interval Δ_t between consecutive samples is also stored. However, the device has some sampling errors, such as lost samples or samples that are not captured due to insufficient pressure. The device assigns [0, 0] coordinate values to the erroneous samples. To summarize, each drawing is stored as a sequence of discrete values $[x_t, y_t, \Delta_t]$. Some examples of doodles and pseudo-signatures are shown in Fig. 2.

The acquisition process was divided in two sessions, separated by an average period of two weeks. This period was chosen in order to allow enough inter-session variability while trying to avoid that users forgot their doodles. Participants were briefed in the first session about the purpose of the acquisition. Each modality (doodles and pseudo-signatures) was explained to them following the same instructions so that each user received the same information. The donors were asked to draw with their fingertip on the handset touch screen holding it in their own hand, simulating thus real operating conditions. They were allowed to practice their drawings until they felt comfortable with them.

FIGURE 1. Doodle acquisition setup.

Forgeries have also been captured in this database. To perform forgeries, users had visual access to the doodle or pseudo-signature they had to imitate. The acquisition software replayed the strokes on the screen showing their dynamic properties (e.g. speed). This animation was shown to users up to three times, and then they were allowed to train until they felt confident with their forgery. The usage of the replay software makes possible to produce forgeries with a notable degree of accuracy, as can be observed in Fig. 2.

During the two sessions, the same protocol was followed for each user and modality: 5 genuine samples, then 5 forgeries, 5 genuine samples, followed by 5 forgeries and finally 5 genuine samples. This separation in blocks of 5 signatures allows analyzing intra-session variability. Consequently, at the end of the two sessions, each user had produced 30 genuine drawings (15 per session) and 20 forgeries. In the first session, user n produced forgeries for users $n-1$ and $n-2$, while in the second, forgeries for users $n-3$ and $n-4$ were produced.

B. DEMOGRAPHICS AND MEMORABILITY

The 100 participants in the database present the following age distribution: 75 are less than 25 years old, 14 are between 25 and 40 years old, and 11 are older. The gender distribution

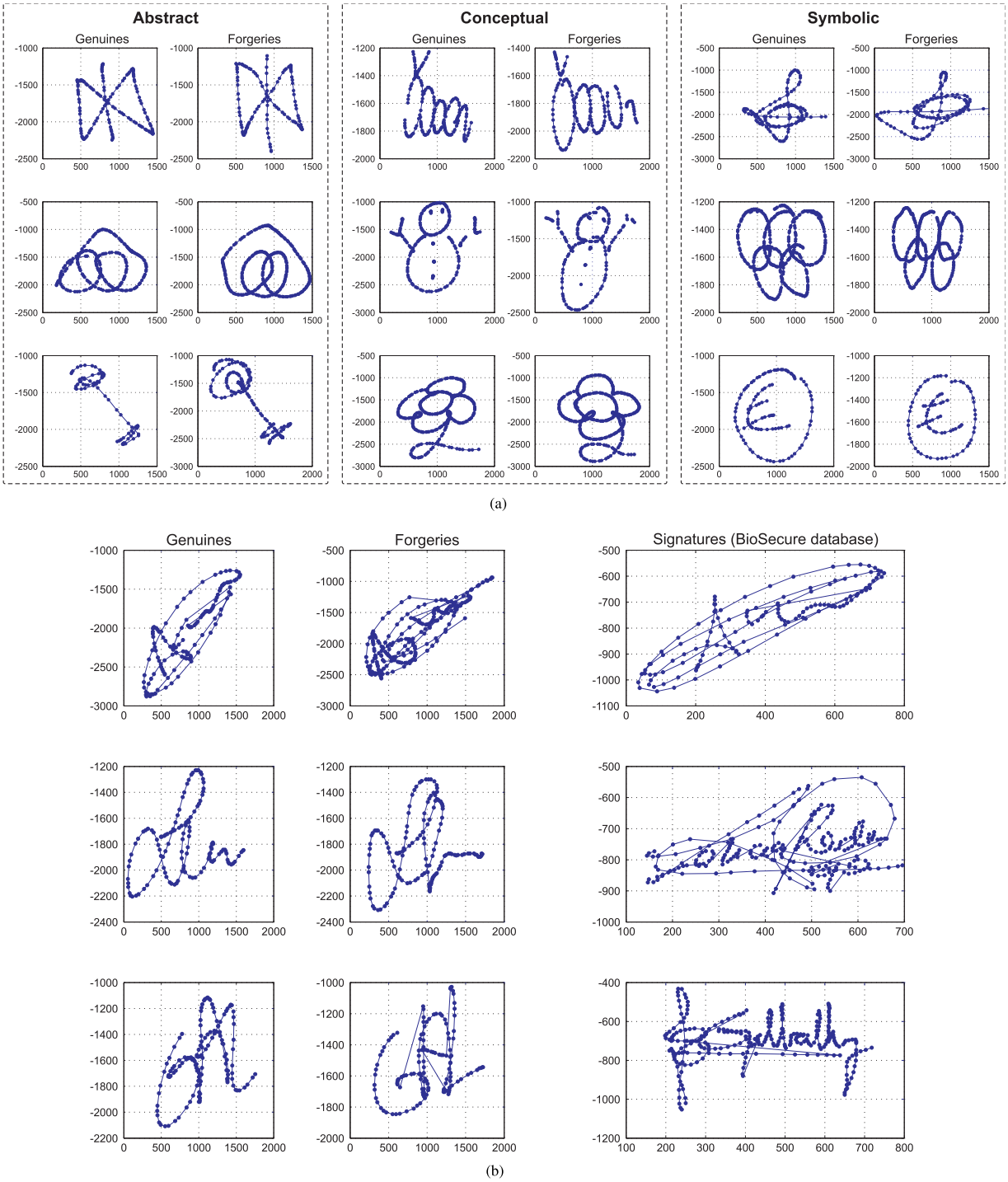


FIGURE 2. **(a) Example of doodles from the database, classified following the criteria explained in Sect. IV. The doodle on the right is a forgery of the one on the left. (b) Example of pseudo-signatures from the database. Genuine pseudo-signatures (left), forgeries (middle) and the corresponding handwritten signature (right) from the BioSecure database [18].**

is 44 women and 56 men. It was observed during the capturing process that participants not familiar with touchscreen devices required a significant longer training time than the rest. This case was more common in older participants.

A subset of 13 participants of this database have also participated in the BioSecure Multimodal Database (BMDB) [18]. In that database, on-line signatures were captured using both a pen-tablet and a PDA with a stylus. This overlap makes possible to observe the evolution of signatures from a controlled scenario (signature with ink pen and paper placed on a pen-tablet), towards more degraded conditions (signature on a PDA with a stylus) and, finally, the most challenging

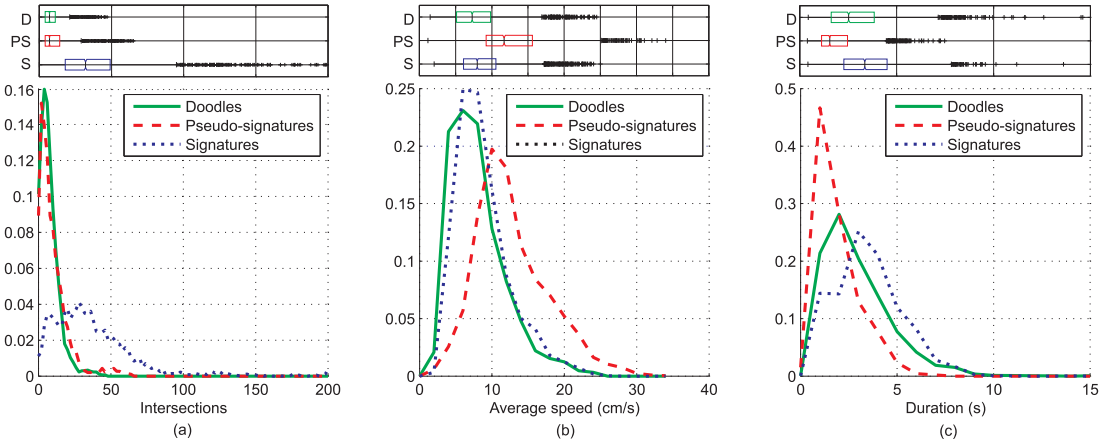


FIGURE 3. Histograms normalized to [0, 1] and box plots of (a) number of intersections, (b) average drawing speed and (c) duration.

case of pseudo-signature (simplified signature traced with the fingertip). Some examples of genuine signatures and their corresponding pseudo-signatures from the same user are shown in Fig. 2(b).

One of the critical issues in graphical passwords is memorability. During the second acquisition session, it was observed that approximately 90% of the participants remembered correctly their pseudo-signature. On the other hand, nearly 40% of the participants had difficulties to recall their doodle from the first session. Users could request to see the tracing process of their own drawings from the first session. This was done by using the aforementioned functionality designed to train forgers. The high percentage of users that requested help to recall their doodles is related to the fact that they did not use them between sessions on a regular basis. In a real scenario with more frequent use, memorability may certainly improve.

IV. DATABASE ANALYSIS

A. STATISTICAL PROPERTIES

Given the different nature of doodles and signatures it is expected that they present differences in their properties such as their length or graphical complexity. A statistical analysis of the properties from the two captured subcorpora has been performed. They have also been compared with the ones from a BioSecure PDA Signature subcorpus of 120 users (also captured by the ATVS - Biometric Recognition Group), allowing thus a comparison between handwritten signatures, finger-traced pseudo-signatures and doodles. The following properties have been analyzed: graphical complexity (as the number of trajectory intersections), average speed and duration.

In Fig. 3(a), the distribution of the number of intersections in the drawings is represented. We observe that signatures present a considerably higher number of intersections, as expected. The difference between doodles and pseudo-signatures is small in this case. A low amount of intersections can be associated to low graphical complexity. This lower complexity indicates that doodles and pseudo-signatures may be easier to forge.

The stroke average speed distributions are compared in Fig. 3(b).[1] As can be seen, doodles are the ''slowest'' from the three datasets. The main cause for this may be that doodles are in general newly invented drawings for the participants, while pseudo-signatures are (or at least contain) previously learned movements. It can also be observed that pseudo-signatures are on average also produced faster than signatures. This is a reasonable result, since the motor process is different for the production of doodles and signatures. When producing a signature, the writer moves the stylus with a combination of his fingers and wrist movements (i.e. the natural writing process), while in the case of finger-drawn sketches, the wrist is the main motor element, as the finger used for drawing is kept almost fixed. This way, signatures are based on more precise movements than doodles, and composed of small graphical elements compared to pseudo-signatures, which are produced by faster movements and larger shapes.

In Fig. 3(c), the statistical distribution of the three sets in terms of their total duration is represented. As can be seen, handwritten signatures tend to have a higher duration than the finger-traced drawings. Moreover, signatures present a higher variability in terms of duration. Doodles also tend to require more time than pseudo-signatures, which are in general composed of initials or simplified signature flourish.

B. VARIABILITY ANALYSIS

Three types of variability may increase the error rate of a verification system. Intra-user variability reflects the difference between genuine samples of the same user. Inter-user variability represents the variance between samples of different users. Last, inter-session variability is related to the difference between samples of the same user over time. In general,

[1]This graph is a corrected version from the one presented in [20], which had an erroneous scaling for the signature duration histogram.

TABLE 1. Verification performance in terms of EER (%) using samples from different sessions for authentication. EER_{sk} refers to the EER for skilled forgeries and EER_{rd} for random forgeries.

Features	Session	Doodles		Pseudo-signatures		Signatures	
		EER_{rd}	EER_{sk}	EER_{rd}	EER_{sk}	EER_{rd}	EER_{sk}
$[x, y]$	1	2.7	28.0	3.5	28.6	3.2	23.9
$[x', y']$	1	3.4	26.7	1.6	23.9	2.1	18.0
$[x'', y'']$	1	4.5	28.1	2.2	19.8	2.8	13.8
$[x, y]$	2	7.6	36.4	5.0	34.5	4.6	27.0
$[x', y']$	2	6.3	33.9	3.8	29.7	3.2	21.5
$[x'', y'']$	2	7.3	34.1	4.3	25.0	4.0	17.8

verification performance will be best if intra-user and inter-session variability are low and inter-user variability is high.

An analysis of the three variability classes in DooDB is carried out in this section. A simple DTW-based verification system trained with the 5 first samples from session 1 is implemented [21], using three pairs of features: the coordinate sequence $[x, y]$, the speed sequence, $[x', y']$ and the acceleration sequence $[x'', y'']$.

Skilled and random forgeries are considered. To compute skilled forgery scores, the 20 available forgeries per user are employed. Random forgeries represent the case where a user claims to be a different one while providing his or her own doodle or pseudo-signature to the system. Random forgery scores are obtained by comparing the user reference set to the first genuine signature sample from each of the remaining users.

The verification performance for the three feature pairs is shown in Table 1 using separately genuine samples from session 1 and from session 2 as test samples. In the case of Session 1, the 10 remaining samples are used for verification (since the first 5 are used for training), while for session 2, all 15 samples are used for verification.

The score distributions of genuine samples from session 2, random forgeries and skilled forgeries are represented for each modality and for each feature pair in Fig. 4. The Equal Error Rates (EERs) of these systems is also shown in Table 1.

Several observations can be made from Fig. 4 and Table 1:

1) INTRA-USER VARIABILITY

In Fig. 4, we observe that the genuine score distribution for doodles presents a long tail towards low scores. This effect reflects the presence of users who vary significantly the aspect or the dynamics (including stroke order) of their doodles. The highest intra-user variability (i.e. the most spread genuine score distribution) is observed for the acceleration features on doodles, which reflects the variation not only in the doodle aspect but also in the dynamics between different sessions. This indicates that in general users concentrate in reproducing the shape of their own doodles, but tend to vary the speed and acceleration of their strokes. The effect is reduced with pseudo-signatures, since generally these are based on better learned movements, and is clearly minimized for signatures, which are the best trained passwords of the three categories.

2) INTER-USER VARIABILITY

Regarding random forgeries, it can be observed in Fig. 4 that random forgery score distributions for doodles are shifted significantly towards lower scores, compared to pseudo-signatures and signatures. This is especially visible for the $[x, y]$ feature pair, revealing a higher inter-user variability, at least in shape, for doodles. This is not reflected in a lower EER in Table 1, since the tail towards lower scores for the genuine score distribution overlaps with forgery scores. When skilled forgeries are considered, inter-user variability is inversely related to the easiness of forging samples from another user. As can be seen in Fig. 4, there is a high overlap between skilled forgeries scores and genuine user scores for doodles. Skilled forgery scores decrease when dynamic features (speed and acceleration) are selected. However, since genuine user scores also decrease for these features on doodles, the overlap does not decrease significantly nor does the EER (for doodles). A predictable effect is that dynamic features such as speed and acceleration provide a higher separation between genuine and skilled forgery scores for signatures since they are harder to imitate, leading to lower EERs.

3) INTER-SESSION VARIABILITY

As expected, the error rates are higher in every case when genuine samples from session 2 are used (see Table 1). We observe that the performance degradation between sessions for doodles and pseudo-signatures is significantly higher than for signatures both in relative and absolute terms. It is also worth noting that the verification performance against random forgeries is in some cases better for doodles and pseudo-signatures than for signatures. This suggests a higher variability in size and shape between users, compared to signatures. However, the higher error rates against skilled forgeries also reflects that pseudo-signatures, and especially doodles are significantly easier to forge.

C. LEARNING CURVE

The learning curve for the three modalities (doodles, pseudo-signatures and signatures) is studied by analyzing the average genuine sample duration for each capture block during the database acquisition. As described in Sect. III-A, during the database acquisition process, users were asked to draw genuine samples in blocks of 5, separated by the production of forgeries.

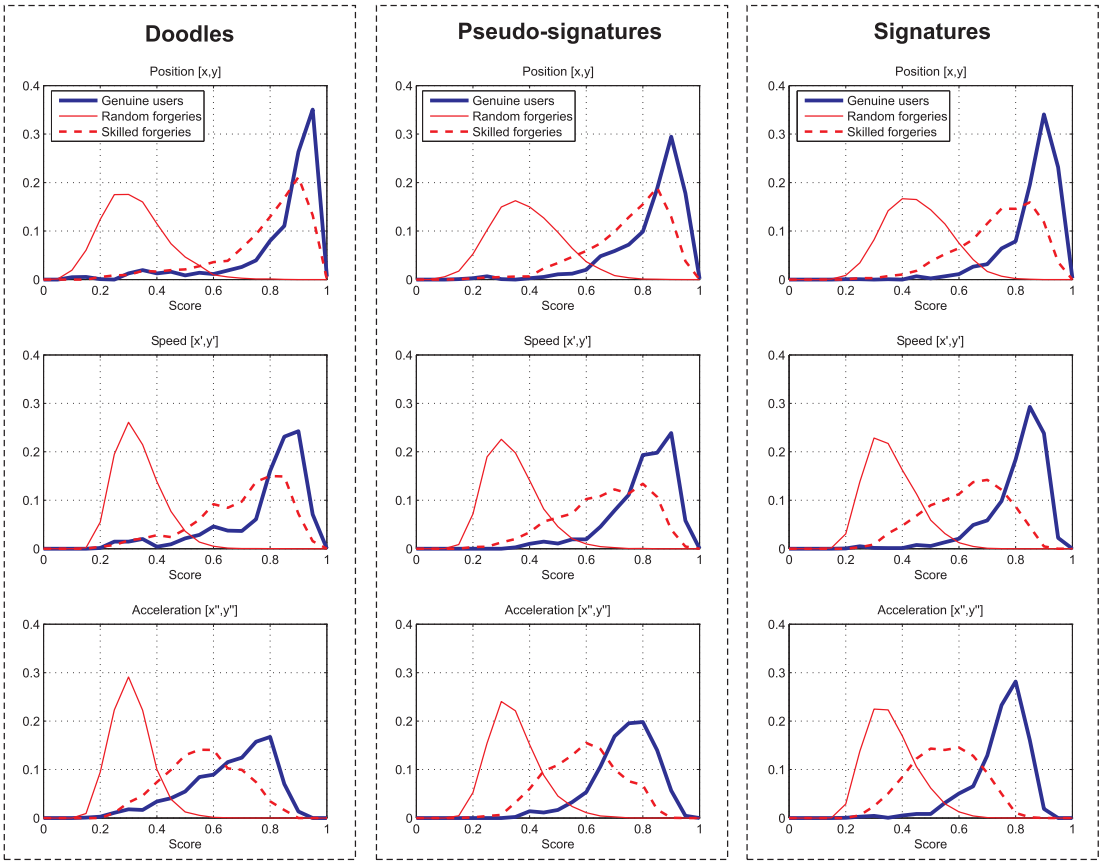


FIGURE 4. Score distributions for doodles (left), pseudo-signatures (middle) and signatures (right) using different feature pairs.

It can be hypothesized that if the average duration significantly decreases between different blocks, the users are still not used to the acquisition method or they are still learning how to produce their graphical password. The average duration for each modality among consecutive blocks is represented in Fig. 5. The average duration between the first block and the last block for the case of doodles has a 20% difference, while for pseudo-signatures and signatures there is a 10% difference.

These observations corroborate the fact that doodles were in general specifically created for the experiments while pseudo-signatures are composed of well-learned movements.

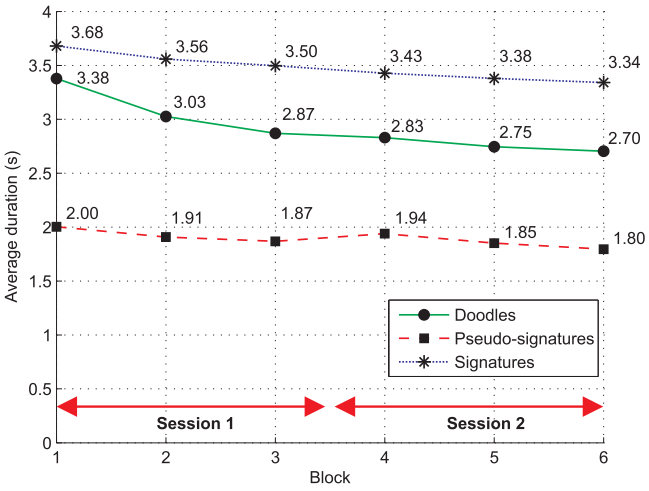


FIGURE 5. Average genuine sample duration for each capture block during database acquisition (3 blocks of 5 samples per session).

D. GRAPHICAL AND QUALITATIVE PROPERTIES

When the whole doodle dataset is visually inspected, it can be seen that there are three main types of doodles:

- **Abstract** doodles, which cannot be directly interpreted as representing an object or idea.
- **Conceptual** doodles, which represent an object or idea (e.g. a flower).
- **Symbolic** doodles, which are known and recognizable symbols, like currency or musical notation.

Doodles that are abstract for an observer may be conceptual to another that is able to interpret them. However, it seems reasonable to assume that abstract doodles may be more resilient to forgers with visual access to them, since they are harder to remember [3]. The proportion of these three doodle types in the DooDB database is: 43 abstract, 37 conceptual, and

TABLE 2. Verification performance in terms of EER (%) using samples from session 2 for authentication. EER_{sk} refers to the EER for skilled forgeries and EER_{rd} for random forgeries.

Features	Doodles		Pseudo-signatures		Signatures	
	EER_{rd}	EER_{sk}	EER_{rd}	EER_{sk}	EER_{rd}	EER_{sk}
HP-LOCAL	5.4	33.8	3.1	28.4	2.1	17.8
ATVS-BSEC	3.4	34.4	3.1	26.9	2.5	15.8

20 symbolic doodles, although this is based on a subjective evaluation. It has also been observed some repetitions among the doodles provided by participants, specially for common drawings. Some examples of repeated doodles are a flower symbol and a smiling face. Examples of each type of doodle are shown in Fig. 2.

Regarding pseudo-signatures, a clear classification between different types cannot be established. It is observed that most participants tend to produce a simplified version of the signature, including flourish. However, approximately 20% of the participants have written their initials, their name or a shortened version of their name without flourish.

V. BENCHMARK RESULTS

In order to assess the authentication performance based on doodles and pseudo-signatures, preliminary experiments have been carried out. A simple verification system, based on Dynamic Time Warping (DTW) to compare the captured time sequences has been used, following the algorithm as described in [21].

Two representative local feature sets from the state of the art are studied in this benchmark. First, the one from the doodle authentication system proposed in [7]. In that system, 6 local features are extracted from the doodle trajectory. These are the coordinate sequence $[x, y]$, and its first and second derivatives (speed and acceleration). Thus, each doodle is described by the 6-dimensional sequence $[x, y, x', y', x'', y'']$. Matching is performed using the DTW algorithm. We refer to this feature set as HP-LOCAL.

The other system is based on the one presented by the Biometric Recognition Group - ATVS to the BioSecure Signature Evaluation Campaign BSEC 2009 [19]. In particular, the system is the one based on DTW that was tuned to maximize its performance against skilled forgeries, identified as system "DTWs" in [19]. It was one of the best performing systems in most evaluation scenarios against skilled forgeries. This feature set is referred to as ATVS-BSEC. The system extracts the following 7 local features:

- x-coordinate, x
- Second-order derivative of x-coordinate, x''
- First-order derivative of y-coordinate, y'
- Second-order derivative of y-coordinate, y''
- Path velocity, $\upsilon = \sqrt{(y')^2 + (x')^2}$
- First-order derivative of path velocity, υ'
- First-order derivative of the *log* curvature radius, ρ', where $\rho = \log(\upsilon/\theta')$ and $\theta = \arctan(y'/x')$ is the curvature of the position trajectory.

A. EXPERIMENTAL PROTOCOL

The experimental protocol follows the one described in Sect. IV-B, but only genuine signatures from session 2 are used for authentication.

The whole sets of doodles and pseudo-signatures from the DooDB database are used for the experiments. The first 5 genuine samples from the first session of each user are used for enrollment as reference templates. The 15 genuine signatures of the second session are used to compute genuine user scores, simulating thus real operating conditions, in which inter-session variability affects the verification performance.

Random and skilled forgery scores are obtained following the same protocol described in Sect. IV-B.

For each comparison against the 5 reference templates, an output score is generated by averaging the inverse of the 5 DTW distances obtained.

B. RESULTS

The verification performance in terms of Equal Error Rate (EER) is shown in Table 2 and DET (Detection Error Trade-off) curves for each dataset are represented in Figure 6. As can be seen, the performance is higher (i.e. lower error) for pseudo-signatures compared to doodles both for random and skilled forgeries.

Comparing Table 2 (which considers state-of-the-art feature sets) to the results shown in Table 1 using only samples from session 2 (with simple feature pairs) for verification, we can see that the performance is similar. This is an indication that the selected state-of-the-art feature sets may not be totally adequate for doodles, and better performance may be achieved by considering feature extraction adjusted to the doodle recognition problem. This is subject to future work.

In Table 1 we also saw that the performance against skilled forgeries improved for pseudo-signatures when dynamic properties (i.e. speed or acceleration) were used. This effect may be due to the higher consistency in the drawing process of pseudo-signatures, since they are composed in general of natural or learned movements. On the other hand, when doodles are considered, the usage of speed or acceleration properties does not increase the performance in the same proportion. This may be due to an increased variability in the drawing process. In fact, it was observed during the doodle subset acquisition, that some users varied the stroke order of their doodles even in the same session. This was not the case for pseudo-signatures.

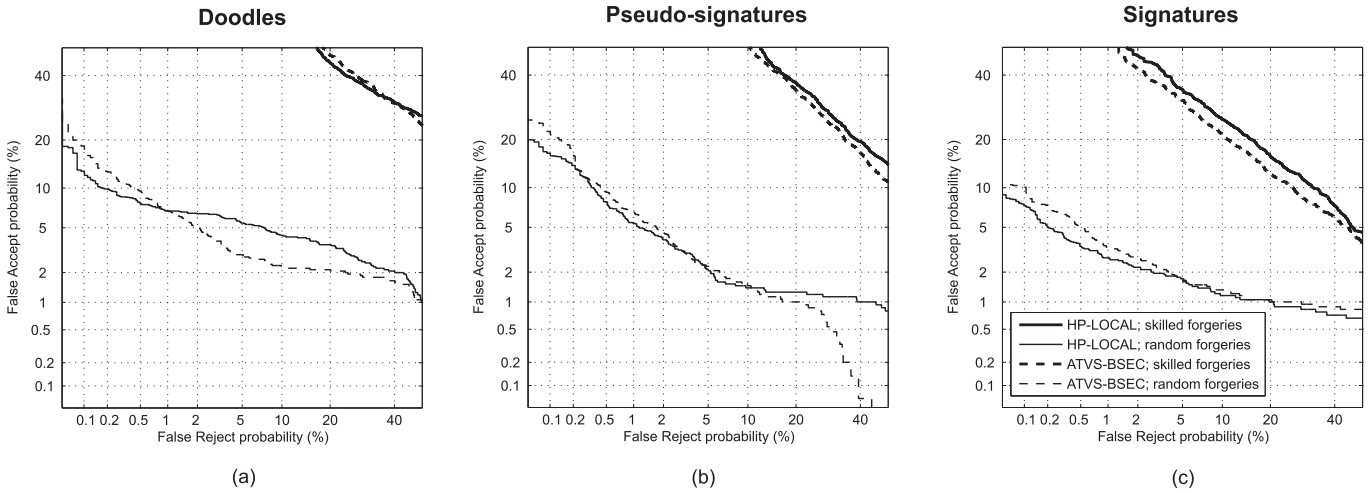


FIGURE 6. DET plots for (a) doodles, (b) pseudo-signatures and (c) signatures.

VI. CONCLUSIONS AND FUTURE WORK

The DooDB database has been presented. This database comprises doodles and pseudo-signatures from 100 users and skilled forgeries for all of them. The acquisition protocol has been described and various data analyses have been performed. Benchmark verification experiments have been carried out, revealing that one of the main challenges of doodle and pseudo-signature verification may be the protection against forgeries.

We have also observed that there is a high intra-user variability in the production of doodles, which negatively affects the verification performance. Unlike the case of signature verification, where dynamic features such as acceleration of velocity clearly increase the verification accuracy [16], the variability found in doodles defies the utility of dynamic features for doodle-based authentication. On the other hand, pseudo-signatures are more stable and thus provide promising results. Users may produce doodles more naturally over time, assuming a frequent usage, leading to an improvement in their verification performance which would become closer to pseudo-signatures in the long term.

Based on the results, doodles and pseudo-signatures are seen as a potential lightweight authentication method oriented to mobile devices. One of the main advantages of this kind of graphical password is its convenience and the possibility of performing user authentication without extra hardware unlike, for example, fingerprint authentication. As previously stated, revocability is an advantage of doodles with respect to other biometric traits.

Future work includes the analysis of the impact of doodle complexity in the performance against skilled forgeries, and the identification of feature sets specifically tuned for doodle verification. Additionally, the impact of each kind of doodle (symbolic, abstract or conceptual) in the quality of forgeries is also source for future research.

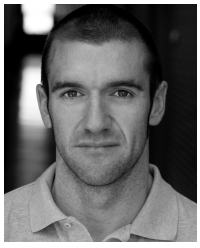
REFERENCES

[1] R. Biddle, S. Chiasson, and P. Van Oorschot, "Graphical passwords: Learning from the first twelve years," *ACM Comput. Surveys*, vol. 44, no. 4, pp. 19:1–19:41, 2012.
[2] X. Suo, Y. Zhu, and G. Owen, "Graphical passwords: A survey," in *Proc. 21st ACSAC*, 2005, pp. 463–472.
[3] K. Renaud, "On user involvement in production of images used in visual authentication," *J. Vis. Lang. Comput.*, vol. 20, no. 1, pp. 1–15, 2009.
[4] I. Jermyn, A. Mayer, F. Monrose, M. K. Reiter, and A. D. Rubin, "The design and analysis of graphical passwords," in *Proc. 8th USENIX Security Symp.*, 1999, pp. 1–14.
[5] J. Goldberg, J. Hagman, and V. Sazawal, "Doodling our way to better authentication," in *Proc. CHI, Extended Abstracts Human Factors Comput. Syst.*, 2002, pp. 868–869.
[6] M. Oka, K. Kato, X. Yingqing, L. Liang, and F. Wen, "Scribble-a-secret: Similarity-based password authentication using sketches," in *Proc. ICPR*, 2008, pp. 1–4.
[7] N. S. Govindarajulu and S. Madhvanath, "Password management using doodles," in *Proc. 9th ICMI*, 2007, pp. 236–239.
[8] C. Varenhorst, "Passdoodles: A lightweight authentication method," (2004) Res. Sci. Inst., Massachusetts Inst. Technol., Cambridge, MA, USA, Tech. Rep. [Online]. Available: http://people.csail.mit.edu/emax/public_html/papers/varenhorst.pdf.
[9] J. Chen, D. Lopresti, and F. Monrose, "Toward resisting forgery attacks via pseudo-signatures," in *Proc. 10th ICDAR*, 2009, pp. 51–55.
[10] W. Zada Khan, M. Y. Aalsalem, and Y. Xiang, "A graphical password based system for small mobile devices," *Int. J. Comput. Sci. Issues*, vol. 8, no. 5, pp. 145–154, Sep. 2011.
[11] N. Sae-Bae, N. Memon, and K. Isbister, "Investigating multi-touch gestures as a novel biometric modality," in *Proc. 5th IEEE Int. Conf. BTAS*, Sep. 2012, pp. 156–161.
[12] A. De Luca, A. Hang, F. Brudy, C. Lindner, and H. Hussmann, "Touch me once and i know it's you!: Implicit authentication based on touch screen patterns," in *Proc. ACM Annu. Conf. Human Factors Comput. Syst.*, 2012, pp. 987–996.
[13] J. Angulo and E. Waestlund, "Exploring touch-screen biometrics for user identification on smart phones," in *Privacy and Identity Management for Life* (IFIP Advances in Information and Communication Technology), vol. 375. New York, NY, USA: Springer-Verlag, 2012, pp. 130–143.
[14] M. Frank, R. Biedert, E. Ma, I. Martinovic, and D. Song, "Touchalytics: On the applicability of touchscreen input as a behavioral biometric for continuous authentication," *IEEE Trans. Inf. Forensics Security*, vol. 8, no. 1, pp. 136–148, Jan. 2013.
[15] A. J. Aviv, K. Gibson, E. Mossop, M. Blaze, and J. M. Smith, "Smudge attacks on smartphone touch screens," in *Proc. 4th USENIX Conf. Offens. Technol.*, 2010, pp. 1–7.

[16] J. Fierrez and J. Ortega-Garcia, "On-line signature verification," in *Handbook of Biometrics*. A. K. Jain, A. Ross, and P. Flynn, Eds. New York, NY, USA: Springer-Verlag, 2008, pp. 189–209.
[17] M. Martinez-Diaz, J. Fierrez, J. Galbally, and J. Ortega-Garcia, "Towards mobile authentication using dynamic signature verification: Useful features and performance evaluation," in *Proc. ICPR*, 2008, pp. 1–6.
[18] J. Ortega-Garcia, J. Fierrez, F. Alonso-Fernandez, J. Galbally, M. R. Freire, J. Gonzalez-Rodriguez, C. Garcia-Mateo, J.-L. Alba-Castro, E. Gonzalez-Agulla, E. Otero-Muras, S. Garcia-Salicetti, L. Allano, B. Ly-Van, B. Dorizzi, J. Kittler, T. Bourlai, N. Poh, F. Deravi, M. Ng, M. Fairhurst, J. Hennebert, A. Humm, M. Tistarelli, L. Brodo, J. Richiardi, A. Drygajlo, H. Ganster, F. M. Sukno, S.-K. Pavani, A. Frangi, L. Akarun, and A. Savran, "The multi-scenario multi-environment biosecure multimodal database (BMDB)," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 32, no. 6, pp. 1097–1111, Jun. 2010.
[19] N. Houmani, A. Mayoue, S. Garcia-Salicetti, B. Dorizzi, M. I. Khalil, M. N. Moustafa, H. Abbas, D. Muramatsu, B. Yanikoglu, A. Kholmatov, M. Martinez-Diaz, J. Fierrez, J. Ortega-Garcia, J. Roure Alcobé, J. Fabregas, M. Faundez-Zanuy, J. M. Pascual-Gaspar, V. Cardeñoso-Payo, and C. Vivaracho-Pascual, "BioSecure signature evaluation campaign (BSEC'2009): Evaluating online signature algorithms depending on the quality of signatures," *Pattern Recognit.*, vol. 45, no. 3, pp. 993–1003, 2012.
[20] M. Martinez-Diaz, J. Fierrez, C. Martin-Diaz, and J. Ortega-Garcia, "DooDB: A graphical password database containing doodles and pseudo-signatures," in *Proc. ICFHR*, 2010, pp. 339–344.
[21] M. Martinez-Diaz, J. Fierrez, and S. Hangai, "Signature matching," in *Encyclopedia of Biometrics*. New York, NY, USA: Springer-Verlag, 2009, pp. 1192–1196.

MARCOS MARTINEZ-DIAZ received the M.Sc. degree in telecommunication engineering from the Universidad Autonoma de Madrid, Madrid, Spain, in 2006. He was an IT Strategy Consultant in Deloitte, Madrid, Spain until 2008. He is currently working in the Technology area with a telecom company in Spain, where he has been involved in different activities such as IT project management, Data Warehouse operations, and IT Architecture. Since 2005, he has been with the Biometric Recognition Group—ATVS with the Universidad Autonoma de Madrid, Madrid, where he is collaborating as Student Researcher pursuing the Ph.D. degree. His current research interests include biometrics, pattern recognition, and signal processing primarily focused on signature verification and graphical passwords. He is the recipient of a number of awards such as the Honeywell Honorable Mention at the Best Student Paper Award at BTAS 2007 and a Special Mention at the National Awards for Telecommunication Engineering studies in Spain.

JULIAN FIERREZ received the M.Sc. and the Ph.D. degrees in telecommunications engineering from the Universidad Politecnica de Madrid, Madrid, Spain, in 2001 and 2006, respectively. Since 2002, he has been with the Biometric Recognition Group, the Universidad Politecnica de Madrid, Madrid. Since 2004, he has been with the Universidad Autonoma de Madrid, Madrid, where he is currently an Associate Professor. From 2007 to 2009, he was a Visiting Researcher with Michigan State University, East Lansing, MI, USA under a Marie Curie Fellowship. His current research interests include signal and image processing, pattern recognition, and biometrics, with emphasis on signature and fingerprint verification, multi-biometrics, biometric databases, and system security. He has been and is actively involved in European projects focused on biometrics, and is the recipient of a number of distinctions for his research, including best Ph.D. in computer vision and pattern recognition from 2005 to 2007 by the IAPR Spanish liaison, Motorola Best Student Paper at ICB 2006, EBF European Biometric Industry Award 2006, IBM Best Student Paper at ICPR 2008, and EURASIP Best Ph.D. Award 2012.



JAVIER GALBALLY received the M.Sc. degree in electrical engineering from the Universidad de Cantabria, Cantabria, Spain, in 2005 and the Ph.D. degree in electrical engineering from the Universidad Autonoma de Madrid, Madrid, Spain, in 2009, where he is currently an Assistant Professor. He has carried out different research internships at worldwide leading groups in biometric recognition such as BioLab from the Universita di Bologna, Bologna, Italy, IDIAP Research Institute in Switzerland, the Scribens Laboratory at the Ecole Polytechnique de Montreal, Montreal, QC, Canada, or the Integrated Pattern Recognition and Biometrics Laboratory (i-PRoBe) at the West Virginia University, Morgantown, WV, USA. His current research interests include the security evaluation of biometric systems, pattern and biometric recognition, synthetic generation of biometric traits, and inverse biometrics. He is actively involved in European projects focused on vulnerability assessment of biometrics (e.g., STREP Tabula Rasa, STREP BEAT) and is the recipient of a number of distinctions including IBM Best Student Paper Award at ICPR 2008, finalist of the EBF European Biometric Research Award in 2009, and Best Ph.D. Thesis Award by the Universidad Autonoma de Madrid in 2010.