A TEST TOOL TO SUPPORT BRUT-FORCE ONLINE AND OFFLINE SIGNATURE FORGERY TESTS ON MOBILE DEVICES

Frank Zoebisch[+], *Claus Vielhauer*[*]

[+]AMSL - Institute ITI, Otto-von-Guericke University Magdeburg, D-39016 Magdeburg, Germany
*Platanista GmbH, D-64289 Darmstadt, Germany

ABSTRACT

Testing of biometric systems requires the consideration of aspects beyond technical and statistical parameters. Especially for testing biometric techniques based on behavior, human factors like intention and forgery strength need to be considered. In this paper, a test tool to support skilled forgeries by test subjects is presented for handwriting verification systems. The software tool has been implemented on two computer platforms and is based on a three level forgery quality model. First experimental results are presented, which indicate that by applying the presented system in attack tests, forgeries of gradual quality can be obtained from test persons.

1. INTRODUCTION

With an increasing spreading of biometric systems, quality measurement by testing becomes essential for a wide user acceptance. In passive biometric systems like fingerprint scanning, false-acceptance-rates are typically determined by verification tests on random sets of samples originating from non-authentic users against authentic ones. Behavioral techniques require a more granular forgery evaluation, for which this paper presents a novel approach for handwriting verification. In the following chapter, we will first discuss human aspects in respect to their intentions, followed by a classification of forgeries. Subsequently, we present design considerations for a handwriting evaluation system and user interface layouts for two different computer platforms. The paper will be concluded with first experimental results, summary and a look at future work.

2. HUMAN AND TECHNICAL ASPECTS OF TESTING BIOMETRIC VERIFICATION SYSTEMS

Testing of biometric systems requires consideration of a number of technical parameters such as device specifics, algorithm properties, environmental conditions, logging and contextual information for behavioral techniques and others, from the variety see for example [1, 2]. Besides those technical aspects, there are aspects related to human factors when operating a biometric system. These need to be evaluated in respect to various application classes, as presented for example in [3], in order to determine application specific test results. For the research presented in this paper, we classify the human aspects as follows. A first distinction is the role of the user during a verification process, i.e. is she or he the authentic subject (AS) exposed to the system or a non-authentic subject (NS). Based on this distinction, we further classify in relation to the users' intentions. For the class of NS users, two possibilities exist: either the intention is an aware attempt to be accepted by the biometric system as authentic users with different identity, in which case we call them Attackers, or the users accidentally try to get accepted by the system as different identities, in this case called Accidental Forgers.

AS can use a biometric verification system with the intention of a conscious and unsolicited, successful verification, which we refer to as Declarer of Intention. In addition to this, the two other sub classes of AS consist of Deniers, who do not want to be identified by the systems and secondly, Compelled Users, i.e. subjects under physical or mental pressure, who are enforced by third parties to attempt a successful verification. Figure 1 illustrates this classification in a tree representation within a two-layer model for the user position (role) and intention aspects.

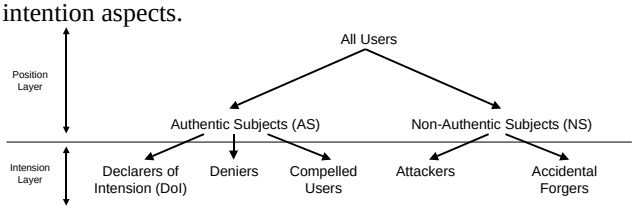


Figure 1. Intention-based user classification

In [3], we have presented a matrix of handwriting application domains for biometric systems and an assignment of handwriting-based biometric feature classes and sampling methods. While in that publication, the aim was to give decision guidelines to find adequate biometric methods for different application domains, the authors

have extended the evaluation matrix by an assignment of user classes based on their intention, which have to be expected in each of the domain. This assignment is shown in Figure 2.

Application	Objective	Domain	Authentic Subjects			Non-Authentic Subjects	
			Declarers of Intention	Deniers	Compelled	Attackers	Accidental Forgers
Document Recognition	Ground Truth Extraction	Automated Content Recognition	√				√
Forensic	Writer Verification	Legal Evidence		√	√	√	
Convenience	Writer Identification	Personal Digital Devices/ Environments	√				√
Access Control	User Identification	Security: Confidentiality, Authorisation	√		√	√	√
Electronic Signature	User Verification	Security: User Authentication	√	√	√	√	

Figure 2. Evaluation matrix extended by user classes

It can be seen that intention aspects vary widely, depending on the application domain. For example, while in document recognition, only users with a clear and declared intention and users, who accidentally claim a non-authentic identity can be expected, electronic signature applications will be exposed to all types of intentions, except accidental forgeries.

For the further discussions, we will focus on aspects of forgery production within the design of evaluation systems and we will present an implementation for handwriting forgeries based on a more granular model for the sub class of attackers.

3. ATTACKS TO SIGNATURE VERIFICATION

As shown in the previous chapter, biometrics can be used for a wide scope of applications for users with potentially different intensions. In handwriting based biometric systems, this obviously has an impact on the kind of attacks to be expected and thus needed to be simulated during the test.

Every biometric system can be attacked in a number of ways. In the context of this work, we limit our views to attacks in terms of forgeries, while we do not refer to other aspects like electrical, physical or mechanical replay attacks to the verification system. Handwriting forgeries, being the result of a behavioral activity, differ significantly in quality, depending on the degree of training and effort. While there are a number of possibilities for a grading of forgery quality (e.g. [4]), we use a systematically reproducible classification previously presented in [5], which classifies into 3 plus 1 degrees (blind, low-force, brut-force and accidental) of forgeries in handwriting. Blind attackers do only have a textual knowledge about the writing content (e.g. precise spelling of the signature of an authentic person). Producers of low-force forgeries are in possession of a blue-print, offline representation of the original handwritten sample and can trace the signature image during the writing process. In addition to this, brut-force attackers will have opportunity to observe the dynamics of the writing process such as velocity and timing. As compared to these three classes of intended forgeries, the last category of accidental forgeries describes the process of attempting a verification of arbitrary, non-authentic writing sample against some other reference. Accidental forgeries can be generated in a very straightforward manner from any biometric database by implementing random choice; therefore we will limit our considerations to the three intended forgery categories. With time constraints for the production of low-force and brut-force forgeries as part of the test arrangement, the above classification was chosen as it is adequate for the design of a systematic forgery support tool.

4. SYSTEM FOR FORGERY TESTING

It has been shown that behavioral biometric systems can be exposed to forgeries of varying qualities. As assessments of the accuracy of such systems can only be determined empirically by performing field test, adequate test tools have to be designed in order to satisfy the requirement to produce such forgeries. Additionally, an evaluation tool has to comply with the following main conditions:

- detailed recording of system parameters such as sensor type and manufacture, location and time of test
- relevant attributes of the test person (age, language, left/right-handed, gender, ethnics, etc.)
- semantic class used (e.g. short/long signature, other semantics as described in [6])
- additional annotations of reference data, based on objective and/or subjective) observations (e.g. significance areas/periods of the writing sample)
- if forgery: attack strength (blind/low/brut-force)
- intension of the writer during the writing process as described in the previous chapter
- all data need to recorded unfiltered in its raw format to allow reproducibility

Over the past years, the authors and others[1] have developed an evaluation system under consideration of these design aspects and the system currently consists of three main components:

- a data kernel implemented in a MySQL database system, that supports storage of samples of writing signals, scanned images for other biometric methods and test environment data according to the above requirements
- a Microsoft Windows™-based user interface developed in Borland Delphi™ including forgery support functionality and verification algorithm plug-ins
- a Palm-OS™ based user interface for signature capture including forgery support functionality

[1] The authors would like to acknowledge the contribution of J. Daum, M. Haisch and F. Ramann in context with their master thesis' to the evaluation system project in the years 2000 - 2002.

Figure 3 gives an overview of the basic relational data model. The event data table collects a set of writing processes in a particular environment, tables sample and sampledata keep references to the biometric raw data and sample annotations stores additional annotations in writing samples. Particularly noticeable is the fact that besides other information, the data model allows storage of the attributes on semantics, intention, sensor, biometric method used, spatial and temporal annotations, independently of the biometric method.

Although until now, the evaluation system has only been used to evaluate algorithms on the single modality of handwriting, the evaluation system is not limited to this one method. Currently, work is continuing to extend the system by additional biometric methods. The aim is to support multi-modal evaluations such as cross-modal correlations of biometrics in future with a plug-in interface for additional verification algorithms.

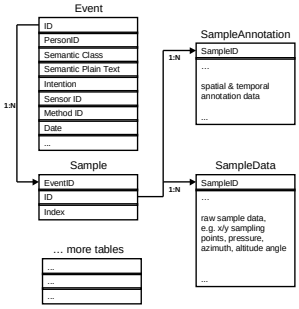


Figure 3. Data model outline of the test system

5. USER INTERFACE DESCRIPTION

User interfaces in the described evaluation system need to be furnished with functionality to support low-force and brut force forgeries of hand written online samples, as described earlier. In order to simplify the test procedures for the test subjects, our intention is to give automated forgery support without human supervision. To achieve this, we have focused on the following design concept for the three forgery classes:

- Blind Forgery: forgers write on a blank surface with only textual knowledge (neither visual nor dynamic) about the signature or writing sequence they are trying to counterfeit
- Low-Force Forgery: forgers get a blueprint of the handwriting projected on the writing surface, which they may trace. However no dynamic information provided
- Brut-Force Forgery: an animated pointer projects the real-time the writing sequence onto the writing pad. The forger may observe the sequence and follow the pointer

Additional design aspects include programmable boundaries in respect to minimum and maximum time, determination of original writing samples to be counterfeit in each test and demands for appropriate hardware, being able to project and capture writing sequences simultaneously. It was decided to take into account two different hardware devices with quite different technical parameters and to develop two separate interfaces for the system. The first interface is built on a rather simple digitizer technology of a mobile personal digital assistant (PDA), whereas for the second implementation a high resolution digitizer-LCD display was chosen. The two interfaces will be described in detail in the following sections.

5.1. Mobile device user interface

The mobile interface for the evaluation system is based on a Palm® V PDA equipped with a Motorola® MC68EZ328 CPU clocked at 20 MHz and 8 Mbytes of RAM, running Palm OS version 4.0. Both display and digitizer resolutions are equally 160x160 pixels at a maximum sampling rate of app. 70 Hz.

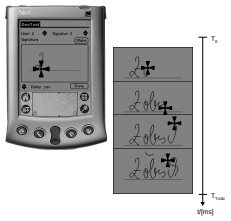


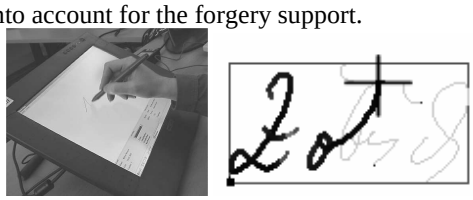
Figure 4. Real-time writing projection on PDA interface

The handwriting capture interface has been developed using GNU C/C++ compiler for Palm OS. Besides other functionality, a test mode is integrated, which consists of a programmable dialog sequence, where the subject is asked to provide authentic writing samples and/or forgeries. While for authentic writing samples and blind forgeries, the writing background will remain blank except for the trace up to the actual pen position, the interface will behave differently in the forgery support mode. For low-force forgeries, a programmable sequence of original writing samples will be projected onto the display background and the user is asked to generate a counterfeit. In brut-force mode, an animated, reticule-style pointer generates the writing trace in real-time, allowing the user to place the pen in the center of the reticule and follow the movement synchronously. Figure 4 visualizes snap shots of this replay process for time t with $T_0 \leq t \leq T_{Total}$, with T_0 and T_{Total} being the time stamps for begin and end of the writing process respectively.

5.2. PC User Interface

The PC User Interface has been developed using Borland Delphi™ software development environment for Microsoft Windows® operation systems. For forgery testing in our evaluation framework, the software is run on a Windows 2000® system based on an industrial PC system with a CPU clocked at 2 GHz and equipped with a

WACOM® Cintiq15X, an interactive pen display. This device provides a display resolution of 1024x768 pixels and a digitizer resolution of 508 lpi at a maximum sampling rate of 205 Hz. While the concept to support test persons in low-force and brut-force forgeries is identical to the PDA user interface, namely by projecting either the blue-print or the real time writing sequence onto the writing area, the PC user interface provides two main additional functions for large-scale testing. Firstly, the system is equipped with a dialog system, which conducts a complete test run interactively with the test subject. Test runs can be configured in a programmable manner, where each test can consist of the phases, familiarity, enrollment, verifications and attacks. While in the familiarity phase, the users can get used to the system voluntarily with no time constraints and predetermined sequence, all other phases can have constraints regarding minimum / maximum writing samples and time. Secondly, during the enrollment phase, besides the plain text and the semantic class, relevance areas are being recorded in addition to the biometric reference. Reference areas denote one or more user-defined spatial rectangular sections annotated by authentic users, allowing those individuals to mark areas that subjectively appear to be of particular relevance to them. This information is acquired with the perspective of a future localization for the verification process and not taken into account for the forgery support.



Figures 5a and 5b. PC User Interface and Real-Time Writing Projection

During the attack phase, users will be provided with information regarding semantic content, offline representation of the writing sample and real time replay functionality, depending on the attack strength. Currently the system is designed to support three forgery strengths: blind forgery (only plain text known by forger), low-force and brut-force attacks, as describer earlier.

6. FIRST EXPERIMENTAL RESULTS

In our first experiment, we have performed forgery tests by four users with a total of 82 forgery samples. The verification algorithm was of very basic nature, only taking into account the average quadratic deviation of horizontal and vertical writing signals. Tests to evaluate the quality of skilled forgeries have been performed with four different threshold values and results in terms of False-Acceptance-Rates (FAR) are shown in table 1.

Attack Strength / Threshold	Low	Medium	High	Very High
Blind	0,00%	0,00%	0,00%	0,00%
Low-force	26,67%	6,67%	3,33%	0,00%
Brut-force	46,67%	30,00%	6,67%	0,00%

Table 1. FAR of Forgeries for three Attack strengths

7. SUMMARY AND FUTURE WORK

We have presented a new tool for systematic evaluation of handwriting verification algorithms towards their robustness for forgeries. The system is implemented on two computer platforms and first tests have indicated that this concept actually can support subjects in generating forgeries of different quality. Based on the motivating results, we will perform tests on a larger-scale in terms of test subjects and algorithms in order to have statistically more significant results. Furthermore, we will evaluate localized feature determination based on the manual annotation of relevance areas. Actual results are available on http://wwwiti.cs.uni-magdeburg.de/iti_amsl/.

Finally, the overall observation that automated replaying of signatures can be used to train individual's handwriting styles confirms the requirement of strong protection of biometric reference data.

8. REFERENCES

[1] A. Brömme, M. Kronberg, O. Ellenbeck and O. Kasch, "A Conceptual Framework for Testing Biometric Algorithms within Operating Systems", *ACM Symposium on Applied Computing SAC 2002*, Madrid, Spain, 2002

[2] The Biometric Evaluation Methodology Working Group, "Common Methodology for Information Technology Security Evaluation - Biometric Evaluation Methodology Supplement", www.cesg.gov.uk/technology/biometrics/media/BEM_10.pdf, 2002

[3] F. Ramann, C. Vielhauer, R. Steinmetz , "Biometric Applications based on Handwriting", IEEE *International Conference on Multimedia and Expo 2002 (ICME)*, Lausanne, Switzerland, 2002

[4] G. Lassmann, "Some Results on Robustness, Security and Usability of Biometric Systems", IEEE *International Conference on Multimedia and Expo 2002 (ICME)*, Lausanne, Switzerland, 2002

[5] C. Vielhauer, R. Steinmetz, "Sicherheitsaspekte biometrischer Verfahren: Klassifizierung von sicherheitsrelevanten Vorfällen und wesentlicher Größen zur Beurteilung der Funktionssicherheit", *7. Deutscher IT-Sicherheitskongress des BSI*, Bonn, Germany, 2001

[6] C. Vielhauer; R. Steinmetz, "Approaches to biometric watermarks for owner authentication", Proceedings of SPIE Vol. 4314, San Jose, U.S.A, 2001