Humans in the Loop: Study of Semi-Automatic Signature Recognition Based on Attributes

Derlin Morocho
Electric and Electronic Department
Universidad de las Fuerzas Armadas ESPE
Sangolquí, Ecuador
dmorocho@espe.edu.ec

Aythami Morales
BiDA Lab, Biometrics and Data Pattern Analytics Laboratory
Universidad Autónoma de Madrid
Madrid, España
aythami.morales@uam.es

Julian Fierrez
BiDA Lab, Biometrics and Data Pattern Analytics Laboratory
Universidad Autónoma de Madrid
Madrid, España
julian.fierrez@uam.es

Javier Ortega-Garcia
BiDA Lab, Biometrics and Data Pattern Analytics Laboratory
Universidad Autónoma de Madrid
Madrid, España
javier.ortega@uam.es

***Abstract*—The present work analyzes performance, abilities and contributions of the human being (layman) in semi-automatic signature recognition systems. During the last decade the performance of Automatic Signature Verification systems have been improved based on new machine learning techniques and better knowledge about intraclass and interclass variability of signers. However, there is still room for improvements and some real world applications demands lower error rates. This work analyzes collaborative tools such as crowdsourcing and human-assisted schemes developed to improve Automatic Signature Verification systems. The performance of humans in semi-automatic recognition tasks is directly related to the information provided during the comparisons. How humans can help automatic systems goes from direct forgery detection to semi-automatic attribute labeling. In this work, we present recent advances, analyzing their performance according to the same experimental protocol. The results suggest the potential of comparative attributes as a way to improve Automatic Signature Verification systems.**

***Keywords*— signature recognition, biometrics, ramdon, simulated, layman.**

I. INTRODUCTION

In the actual society, the user verification is a requirement in many applications as forensics, international border crossing, financial transactions, and computer security. The biometric recognition is a very broad field of research, where a lot of different research areas are generated, all focused on the analysis and evaluation of human physiological and behavioral traits for automatic recognition applications.

The signature is a globally accepted behavioral biometric modality, and has been used for centuries by different cultures. The signature is a human task composed affected by neuromotor characteristics of the signer [1] [2]. In addition, it is necessary to consider the socio-cultural influence such as the Western and Asian styles. For a long time, forensic analysis has allowed us to examine and evaluate the signatures, through forensic analysis experts who determine authenticity. At present, there are automatic signature verification systems, as an aid to Examiners of Forensic Documents (FDE) [3-5].

The handwritten signature is well accepted as an authentication method within society, in the legal and commercial field, since it is the personal characteristic that allows to certify the identification of an individual or to give authenticity to legal documents [4]. Semi-Automatic signature recognition systems have several applications where the human performs signature recognition with neither supervision nor experience, such as banking transactions, product sales, parcel/courier delivery, and public notary. In most of these applications, humans only verify transaction log, but are limited in signature validation, since they are people who have no experience in forensic document analysis (FDE). In these scenarios the human interaction for the evaluation of the signature is very important for the decision making in a short time, this biometric trait has a high intra-variability own of the human. The purpose of this paper is to present case studies and experiments carried out in this line of research, where human interaction can be applied in many scenarios of practical importance.

Development of automatic systems is minimizing the confidence of human skills. The performance of Automatic Signature Verification systems (ASV) have been improved during the last decade. However, some applications demands lower error rates and there is still room for improvements (see Fig.1). Fig. 1 shows two examples where automatic comparisons obtained through a state-of-the-art ASV system fail. However, we must keep in mind that human beings possess the innate ability of perception, which consists in exploring behavior, performance and abilities one another. What actions and to what extent can assist the ASV systems.

In previous research [6-9], Forensic Document Examiners (FDE) efficiency studies are performed in the recognition of signatures, however the performance and behavior of the human with no experience in FDE (layman), generates an increasing interest. Previous studies on human performance focus on the recognition of signatures through crowdsourcing [10-12], where they analyze laymen's performance. Human performance is measured by analyzing the opinion of the laymen based on visual comparisons from a genuine set of signatures and unmarked samples including genuine and counterfeit signatures.

978-1-5386-1585-0/17/$31.00 ©2017 IEEE

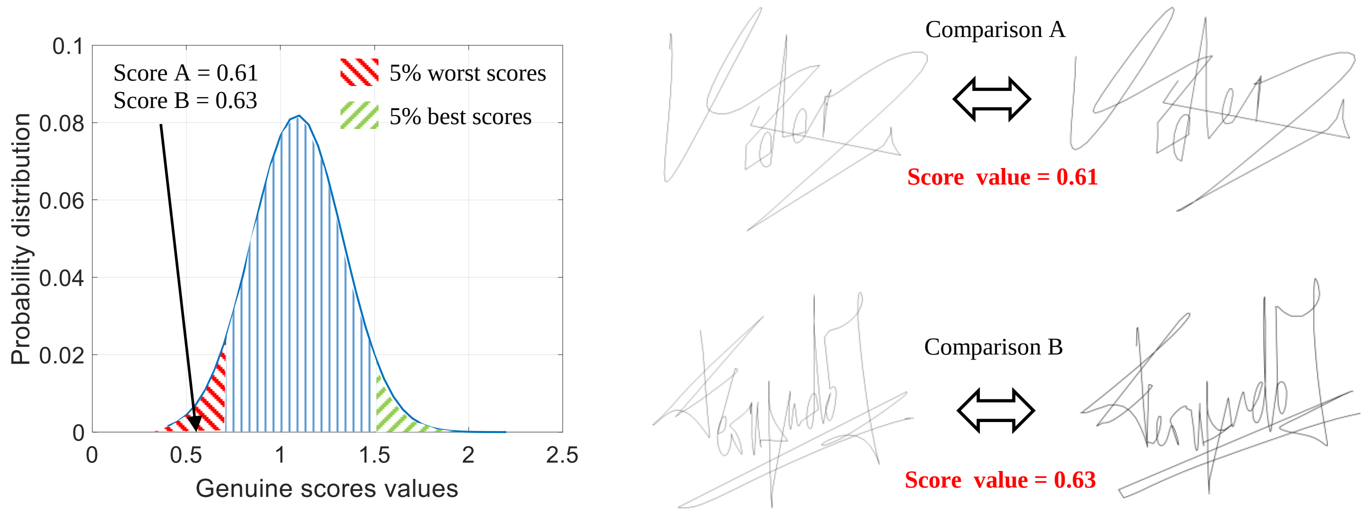


Fig. 1 Left: Probability distribution of classification scores obtained using the Dynamic Time Warping automatic signature recognition algorithm and the experimental protocol proposed in [22]. Right: Two examples of genuine comparisons with low classification scores but good visual similarity.

These studies have made it possible to establish a human performance baseline, but there is a lack of knowledge in factors that influence human scores. Is the human a good signature recognizer? Is the human stable in the signature labeling process? What parameters help improve the signature recognition in the labeling process? There are so many openended questions related to human capabilities for signature recognition.

The present work analyzes human interventions to improve semi-automatic signature recognition. The rest of the work is organized as follows: Section 2 presents the state of the art of work related to human performance through crowdsourcing and manual signature labeling. Section 3 presents the analysis of research experiments related to human performance. Finally, Section 4 summarizes the conclusions.

II. Human Intervention for Biometric Recognition

Research studies in the field of biometrics applied to human-assisted systems allow us to identify the real abilities of the human, and the capabilities of the automated systems [9-11, 13-17]. The use of human annotations in automatic system biometric recognition has provided encouraging results in the literature [14]. The annotation of attributes performed by humans has emerged as a way to improve automatic systems in face recognition [12, 18] or gait [19].

In references [20, 21], soft biometrics are used for the description of human faces and people. The development of their work focuses on the fact that people naturally use labels and physical attribute estimates to describe other people. The results obtained were extremely satisfactory, it was concluded that the absolute body descriptions to identify individuals resulted in an accuracy of identification of 48%, because the absolute labels proved to be a bad form of description, bound to subjectivity and interference. On the other hand the comparative labels proved to be less subjective than the traditional forms of description and are preferred by the majority of specialists, obtaining for this particular work an accuracy of 99.3%.

Soft biometrics in human recognition is studied in [20], in which a multimodal biometric system that uses face and fingerprint as the main features and gender, ethnicity and height as soft features; study in which experiments carried out on a database of 263 users show that the recognition performance of the primary biometric system can be significantly improved by making use of soft biometric information. This is achieved if the soft biometric features are complementary to the Primary biometric features.

Crowdsourcing allows the intervention of the human being in a variety of human tasks. Specifically, crowdsourcing for biometric applications has been in the field of face recognition [30], posture or gait [19], as far as biometric security [29].

III. Semi-Automatic Signature Recognition Based on Human Intervention

The literature on signature recognition shows how the performance of systems have been improved over time [7, 8, 16, 27, 28]. Human performance obtained in [6, 8, 9, 17] suggests that laymen find it difficult to correctly recognize the authenticity of signatures. However, it is well accepted that FDEs can achieve competitive results based on their specialized training and experience.

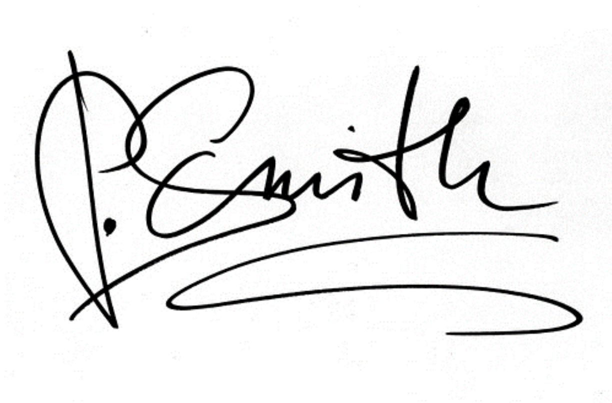
Research work developed in [13, 22] have explored the human-assisted signature recognition, including a baseline performance of human signature recognition and the analysis of the semi-automatic schemes based on attribute labeling. Their results suggest the potential of these recognition schemes in applications involving human intervention. The performance of offline signature recognition systems is lower than on-line systems [23] and, therefore, attribute-based matching could help overcome the limitations of offline correspondence.

In [10, 11, 22], authors present a studies to measure human performance based on crowdsourcing responses in comparison to Automatic Signature Verification systems. The works have evaluated several crowdsourcing interfaces with the aim of

978-1-5386-1585-0/17/$31.00 ©2017 IEEE

Absolute attributes

Shape = Calligraphic
Legibility = Concentrated
Order = Spaced
Punctuation = Yes
Flourish symmetry = partially
Flourish weight = wide
Flourish shape = rounded



Comparative attributes

Shape
Calligraphic = 4
Vertical = 1
Horizontal = 2
Round = 2
Order
Clear = 3
Confused = 2
Concentrated = 3
Spacing = 1

Fig. 2 Absolute attributes [22] versus Comparative attributes [33]

measuring human performance and its comparative observations to discriminate and evaluate a signature.

The article [14] explores the human ability to recognize the authenticity of signatures. The crowdsourcing is used to analyze the different factors affecting the performance of humans without Forensic Document Examiner experience. The human responses are used to analyze the performance of humans according to each of the scenarios and main factors. The experiments comprise 240 signatures from BiosecureID public database and responses from more than 400 people. The results suggest the difficulties associated to these tasks, with special attention to the false acceptance of forgeries with performances ranging from 50% to 75%. However, the results suggest that human perception is biased by the signature characteristics and different performances are obtained depending of the signer. Finally the combination of human ratings clearly outperform the individual performance. This study focuses its objectives on improving the human performance in the recognition of signatures through visual aids by introducing characteristic traits, the same ones that are selected from works inspired by FDEs [24, 25, 26].

The studies performed in [10,11,14] reveal human performance in signature forgery detection. These works analyze the correlation between the features denoted as important by users and their performance classifying genuine and forged signatures. The results show that most of the users focus on high level attributes of the signature such as letter style and shape. However, the average error rate are high with False Acceptance Rates up to 25% at False Rejection Rate around 30%. The study suggests that results obtained concluded that human performance depends on parameters such as quantity and quality of information available during the classification task. More information provided to the laymen, lower error rate was obtained. In addition, the result obtained in [14] suggest the importance of the signer characteristics. As can be seen in Figure 3, the average performance of the laymen varies from False Rejection Rates equal to 19% in the best case (forgers easy to detect) to 53% in the worst (forgers with larger ability).

Table I summarizes the performance of the 400 workers (combined by averaging their responses) and 240 genuine and forged signatures employed to measure their performance. In

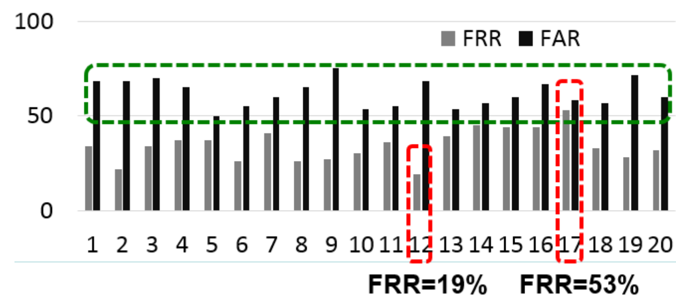


Fig. 3 Performance obtained averaging the performance by 400 laymen for a classification task involving 20 different signers [14].

addition, the table includes the performance obtained by an Automatic Signature Verifier (ASV) system based on offline characteristics [19] using a similar protocol used for the humans (4 training samples and 8 test samples for each signer).

The articles [13, 22] explore human intervention to improve Automatic Signature Verification (ASV). These works analyzed how human actions can be used to complement automatic systems. Instead of previous intervention at classification level, intervention at feature extraction level was evaluated using a self-developed tool for the manual annotation of signature attributes inspired in Forensic Document Experts analysis. The experiments include the two most popular signature authentication scenarios based on both online (dynamic time sequences including position and pressure) and offline (static images) information. The results showed that features manually labelled by human can be used to improve automatic systems in both scenarios. The improvements range from 6% to 61% depending of the conditions of the experiments.

Article [13] focuses their experiments in determining which features or attributes of a signature can improve performance in automatic recognition systems. The result reported in [13] show that most common features labelled by laymen for signature recognition tasks. In addition, the results showed that it is possible to observe the combination of automatic systems and semi-automatic systems with features labelled by laymen can be used to improve the recognition performance.

978-1-5386-1585-0/17/$31.00 ©2017 IEEE

TABLE I EVALUATION OF HUMAN PERFORMANCE (EER) WITH INTERVENTION AT CLASSIFICATION LEVEL VS. OFF-LINE AUTOMATIC SYSTEM (SIMULATED FORGERIES)

	Simulated Forgeries
Individual human performance	32.2%
Combination of human responses	13.8%
ASV based on offline features [22]	20.27%

TABLE II EVALUATION OF HUMAN PERFORMANCE (EER) WITH INTERVENTION AT FEATURE EXTRACTION LEVEL VS. AUTOMATIC SYSTEMS

	Random Forgeries	Simulated Forgeries
ASV based on online features [22]	1.9%	6.9%
ASV based on offline features [22]	4.72%	20.27%
Absolute Attributes [22]	6.89%	24.22%
Comparative Attributes [33]	5.57%	21.20%

Finally, [33] proposes a semi-automatic recognition scheme based on comparative attributes labelled by humans. The comparative attributes try to overcome the limitation of absolute attributes [22] by converting absolute labels (e.g. is vertical?) into ranges (e.g. how vertical?). Table II summarizes the results presented in [33] and [22] obtained according the same experimental protocol. The results show the superior performance of comparative attributes with error rates similar to those obtained by state-of-the-art off-line ASV systems.

IV. CONCLUSIONS

This work explores the performance of humans in signature recognition based on intervention at different levels. The literature shows that FDEs can achieve performances similar to the best state-of-the-art Automatic Signature Verification systems. The performance of laymen is far to the performance obtained by FDEs and ASV systems. However, the collaborative interventions based on crowdsourcing as well as the comparative attributes have shown encouraging results. The results of show the great potential of these schemes and its potential in applications that involve human interventions..

ACKNOWLEDGMENT

A.M. is supported by a Juan de la Cierva Incorporacion (JdC-I) contract (IJCI-2015-24742) from the Spanish MINECO. This work was partially funded by CECABANK and project CogniMetrics (TEC2015-70627-R) from Spanish Government (MINECO/FEDER).

REFERENCES

[1] J. Galbally, M. Martinez-Diaz, & J. Fierrez, “Aging in biometrics: an experimental analysis on on-line signature”, PLoS ONE, vol. 8(7), pp. e69897, 2013.

[2] D. O. Hebb, “The organization of behavior: a neuropsychological theory”, Ed. John Wiley & Sons, 1949.

[3] R. Plamondon and S. N. Srihari, “On-line and off-line handwriting recognition: A comprehensive survey”, IEEE Trans. on Pattern Analysis and Machine Intelligence, vol. 22 (1), pp 63-84, 2000.

[4] D. Impedovo and G. Pirlo, “Automatic signature verification: The state of the art”, IEEE Trans. on Systems, Man, and Cybernetics (Part C), vol. 38(5), pp. 609-635, 2008.

[5] J. Fierrez and J. Ortega-Garcia, “On-line signature verification”. Handbook of Biometrics, Springer, 2008, pp. 189-209.

[6] M. I. Malik, M. Liwicki, A. Dengel, and B. Found, “Man vs. Machine: a comparative analysis for forensic signature verification”, Proc. of the 16th International Graphonomics Society Conference, 2013, pp. 9–13.

[7] L. S. Oliveira, E. Justino, C. Freitas and R. Sabourin, “The graphology applied to signature verification”, Proc. 12th Conference of the International Graphonomics Society, 2005, pp. 286-290.

[8] J. Coetzer, B.M. Herbst, J.A. Du Preez, “Off-line signature verification: A comparison between human and machine performance”, Proc. 10th Int. Workshop on Frontiers in Handwriting Recognition, 2006, pp. 481-485.

[9] M. I. Malik, M. Liwicki, A. Dengel, “Part-based automatic system in comparison to human experts for forensic signature verification”, Proc. Int. Conf. on Document Analysis and Recognition, 2013, pp. 872–876.

[10] D. Morocho, A. Morales, J. Fierrez and R. Tolosana, “Signature recognition: establishing human baseline performance via crowdsourcing”, Proc. 4th Int. Workshop on Biometrics and Forensics, 2016, pp. 1-6.

[11] D. Morocho, M. Proaño, D. Alulema, A. Morales and J. Fierrez, “Signature Recognition: Human performance analysis vs. automatic system and feature extraction via crowdsourcing”, Proc. Mexican Conference on Pattern Recognition, Springer International Publishing, 2016, pp. 324-334.

[12] L. Best-Rowden, S. Bisht, J. Klontz and A. K. Jain. “Unconstrained Face Recognition: Establishing Baseline Human Performance via Crowdsourcing”. Proc. Int. Conference on Biometrics, 2014, pp. 1-8.

[13] D. Morocho, A. Morales, J. Fierrez and R. Vera-Rodriguez, “Towards human-assisted signature recognition: improving biometric systems through attribute-based recognition”, IEEE International Conference on Identity, Security and Behavior Analysis, 2016, pp. 1-6.

[14] D. Morocho, J. Hernandez-Ortega, A. Morales, J. Fierrez, & J. Ortega-Garcia, “On the evaluation of human ratings for signature recognition”, IEEE International Carnahan Conference on Security Technology, 2016, pp. 1-5.

[15] D. Reid, M. Nixon, “Human Identification using Facial Comparative Descriptions”, Proc. Int. Conf. on Biometrics, 2013, pp. 1-7.

[16] N. Houmani, A. Mayoue, et al., “Biosecure signature evaluation campaign (BSEC2009): evaluating online signature algorithms depending on the quality of signatures”, Pattern Recognition, vol. 45, pp. 993–1003, 2012.

[17] M. I. Malik, et al. “ICDAR2013 competitions on signature verification and writer identification for on- and offline skilled forgeries (SigWiComp2013)”. Proc. of Int. Conf. on Document Analysis and Recognition, Tunisia, 2013, pp.1108-1114.

[18] N. Kumar, A. C. Berg, P. N. Belhumeur, and S. K. Nayar. “Describable visual attributes for face verification and image search”. IEEE Trans. Pattern Analysis and Machine Intelligence, vol. 33(10), pp. 1962–1977, 2011.

[19] D. Martinho-Corbishley, M. S. Nixon, Mark, J. N. Carter, “Soft Biometric Recognition from Comparative Crowdsourced Annotations”, Proc. Int. Conf. on Imaging for Crime Prevention and Detection, London, UK, 2015, pp. 1-6.

[20] A. K. Jain, K. Nandakumar, X. Lu, and U. Park, “Integrating Faces, Fingerprints, and Soft Biometric Traits for User Recognition”. Proceedings of Biometric Authentication Workshop, LNCS 3087, 2004, pp. 259-269.

[21] P. Tome, J. Fierrez, R. Vera-Rodriguez and M. Nixon, “Soft Biometrics and their Application in Person Recognition at a Distance”. IEEE Trans. on Information Forensics and Security, vol. 9(3), pp. 464-475, 2014.

[22] A. Morales, D. Morocho, J. Fierrez, & R. Vera-Rodriguez, “Signature authentication based on human intervention: performance and complementarity with automatic systems”. IET Biometrics, vol. 6(4), pp. 307-315, 2017.

[23] J. Galbally, et al. “On-Line Signature Recognition Through the Combination of Real Dynamic Data and Synthetically Generated Static Data”, Pattern Recognition, vol. 48, pp. 2921-2934, 2015.

[24] A. Barbera F., y Méndez Baquero F., “Análisis de textos manuscritos, firmas y alteraciones documentales”, Ed. Tirant lo blanch, 2005.

[25] M. Sánchez, “Peritación caligráfica”, Sol, 1987.

[26] Sánchez Terrones, “Manual del Perito Calígrafo”, 1902.

[27] T. M. Burkes, D. P. Seiger and D. Harrison, “Handwriting examination: Meeting the challenges of science and the law”, Forensic Science Communications, vol 11(4), 2009.

[28] H. Coetzer and R. Sabourin, “A human-centric off-line signature verification system”, Proc. Int. Conf. on Document Analysis and Recognition, Curitiba, Brazil, 2007, pp. 153-157.

[29] S. Panjwani, A. Prakash, “Crowdsourcing Attacks on Biometric Systems”, In: Symposium On Usable Privacy and Security (SOUPS 2014), 2014, pp. 257–269.

[30] L. Best-Rowden, H. Han, C. Otto, B. Klare, A. Jain, “Unconstrained face recognition: identifying a person of interest from a media collection”, IEEE Trans. Inform. Forensic Secur., vol. 9(12), pp. 2144–2157, 2014

[31] J. Fierrez, J. Galbally, et al., "BiosecurID: A Multimodal Biometric Database", Pattern Analysis and Applications, vol. 13(2), pp. 235-246, 2010.

[32] Biometric Recognition Group–ATVS, “Instructions for downloading BiosecurID-SONOF DB”, http:/atvs.ii.uam.es/biosecurid_sonof_db.html

[33] D. Morocho, A. Morales, J. Fierrez and R. Vera-Rodriguez, "Human-Assisted Signature Recognition based on Comparative Attributes", ICDAR International Workshop on Human-Document Interaction, Kyoto, Japan, 2017.

978-1-5386-1585-0/17/$31.00 ©2017 IEEE