Towards a Better Understanding of the Performance of Latent Fingerprint Recognition in Realistic Forensic Conditions

Maria Puertas[1], Daniel Ramos[1], Julian Fierrez[1], Javier Ortega-Garcia[1] and Nicomedes Exposito[2]

[1]ATVS-Biometric Recognition Group, Universidad Autonoma de Madrid, Spain
[2]Departamento de Identificacion. Servicio de Criminalistica de la Guardia Civil, Ministerio del Interior, Spain.
{maria.puertas, daniel.ramos}@uam.es

Abstract— **This work studies the performance of a state-of-the-art fingerprint recognition technology, in several practical scenarios of interest in forensic casework. First, the differences in performance between manual and automatic minutiae extraction for latent fingerprints are presented. Then, automatic minutiae extraction is analyzed using three different types of fingerprints: latent, rolled and plain. The experiments are carried out using a database of latent fingermarks and fingerprint impressions from real forensic cases. The results show high performance degradation in automatic minutiae extraction compared to manual extraction by human experts. Moreover, high degradation in performance on latent fingermarks can be observed in comparison to fingerprint impressions.**

Keywords-latent fingerprints; feature extraction; AFIS; forensic identification; biometrics.

I. INTRODUCTION

In forensics, fingerprint images can be classified in two categories: fingerprint impressions and latent fingermarks. Fingerprint impressions are obtained either by scanning the inked impressions on paper or by using scanning devices [1]. The acquisition of this kind of prints is controlled by an expert to make sure that the images have good quality. Impressions are also divided in two types: rolled and plain. Rolled impressions are obtained by rolling the finger from one side to the other. Plain impressions are those in which the finger is pressed down but not rolled. Fingerprint impressions are usually contained in ten-print cards. A ten-print card contains the rolled and plain impressions of the ten fingers of a person. The reason to include both types of impressions is that while rolled fingerprints contain larger size and a higher number of minutiae, plain fingerprints are less distorted and often have clearer ridges [1]. The other group of fingerprints is latent fingermarks. These marks are accidentally left in a crime scene and then recovered and scanned [2].

Nowadays, state-of-the-art fingerprint recognition systems for good quality fingerprint images have an acceptable level of performance [3] [4] [5]. However, when dealing with latent fingermarks, there is still a lot of research to be done. Low quality, incompletion and distortion are typical problems that forensic Automated Fingerprint Recognition Systems (AFIS) have to face when extracting features from latent fingermarks. As many of the available AFIS systems are mainly minutiae-based, errors in minutiae extraction play a critical role in performance. Unfortunately, such kind of errors are frequent with latent fingerprints, which almost always contain less clarity, less content and less undistorted information than fingerprints taken under controlled conditions [6]. Figure 1 shows an example of the three different types of fingerprints.
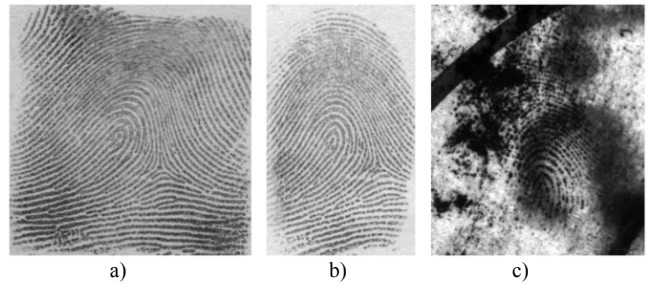

Figure 1. Three types of fingerprint images: a) Rolled fingerprint, b) Plain fingerprint, c) Latent fingerprint. (Extracted from [1])

Automated Fingerprint Identification Systems (AFIS) allow the search of the fingermarks among millions of ten-print cards usually producing a ranked list of top candidates based on similarity scores [7].

The matcher calculates these scores mainly based on the comparison of minutiae features and their spatial relationship [8].

As a consequence of the previously discussed quality problems in latent fingermarks, the common protocol in forensics is to manually mark the minutiae before launching a search. When the search finishes, the list of top candidates is usually manually reviewed by experts in search of a match [9].

Given the importance of the minutiae extraction process with latent fingerprints in forensic applications, in this work we compare and analyze the performance of automatic and manual latent feature extraction in forensic fingerprint recognition. Moreover, we compare and analyze the performance of automatic minutiae extraction in three types of fingerprints: latent, rolled and plain.

The selection of a proper realistic experimental set-up is critical in forensics. Therefore, we have used as query a set of 50 latent fingerprints from real, solved cases of the Identification Department of Spanish Guardia Civil. Also, the 100 impressions used as queries in the case of rolled and plain fingerprints have been taken from ten-print cards stored in the Guardia Civil database. Finally, the experiments reflect a realistic scenario, because all the queries are compared to the full database of ten-print cards in Guardia Civil database (around 2.5 million cards).

The rest of the paper is organized as follows: Section 2 presents the experimental framework used in this work. Section 3 reports and discusses the experimental results, and conclusions are finally drawn in Section 4.

II.   EXPERIMENTAL FRAMEWORK

A. Fingerprint Recognition System

A state-of-the-art fingerprint recognition system has been used (one of the top ranked in NIST FpVTE 2003 [3], and subsequent evaluation campaigns like NIST PFT[1] [10]). This system works in identification mode [8]. For each input fingerprint, the system compares it to a ten-print card database. The output consists of a list of the 15 impressions of the ten-print card database that achieve the highest matching scores.

B. Fingerprint Databases

All the fingerprints used in this work belong to the Spanish Guardia Civil database. For the experiments, two different sets of fingerprints have been used.

The first set consists of 50 latent fingermarks from real forensic cases. All the fingermarks in this set have been previously manually identified with their sources by human fingerprint experts of Guardia Civil, and therefore the identity of each mark will be assumed to be known in our experiments. The second set is composed by 100 right index impressions from 50 convicted individuals. For each individual there are 2 impressions: one rolled fingerprint and one plain fingerprint.

Both sets of prints are compared to a database that contains about 2.5 million ten-print cards. For each fingermark or fingerprint used as a query, it is guaranteed that at least one ten-print card from the individual at hand is available in the searched database.

The small size of the sets of fingerprints used for the experiments is due to the limitations in obtaining and handling fingerprint data from real cases. This limitation reduces de statistical significance of the results obtained in the experiments. However, the sample size is big enough to observe trends in the performance with different minutiae extraction techniques and types of images.

C. Experiments

In this work, two types of experiments have been carried out. In the first type, the matcher compares the set of latent fingerprints to the ten-print card database. In the second scenario, the set of fingerprint impressions is compared to the ten-print card database.

Four different tests have been carried out with the set of latent fingermarks:
1. Latent fingerprints with all the minutiae extracted automatically by the system.
2. A human fingerprint expert manually marks all the minutiae that he or she can find in the fingerprints.
3. The human expert selects the best compact subset of 12 minutiae for each fingerprint according to his experience and only those are used for the matching.
4. The best subset of 8 minutiae is manually selected in every fingerprint among the 12 in the previous experiment, before sending the fingerprint to the matcher.

In the case of the fingerprint impressions, we have run the experiments separating the fingerprints in two groups: rolled and plain. For both sets of impressions, only the automatic feature extraction has been used, following the forensic protocol used in casework for this kind of prints.

In all the experiments of this work, the minutiae in the ten-print card database are automatically extracted by the system.

III.   RESULTS

In this section, experimental results are shown using CMC (Cumulative Match Characteristic) curves, which are often used to show the matching performance of identification systems [11]. A CMC curve plots the rank-n identification rate against n, for n=1, 2, … , M, where M=15 in our case. The rank-n identification rate indicates the proportion of times the genuine identity appears among the top n matches.
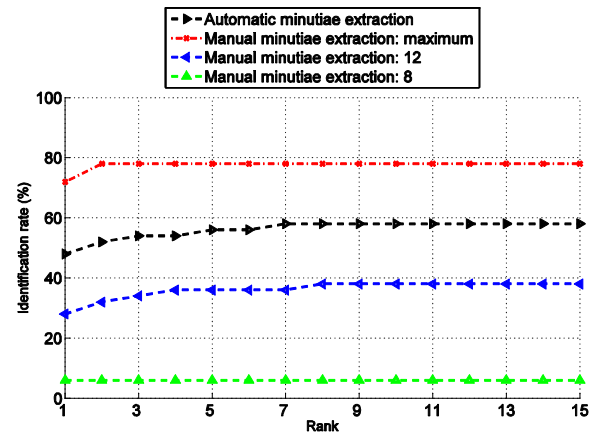

Figure 2. CMC curves for the latent queries.

TABLE I. FIRST CANDIDATE VS. NOT-APPEARANCE RATES FOR LATENT QUERIES.

	Max. Manual	Max Automatic	12 Manual	8 Manual
Rank 1	72%	48%	28%	6%
Rank>15	22%	42%	62%	94%

A. Manual vs. automatic minutiae extraction in latent fingerprints

Results for latent fingermarks as queries against the ten-print database are shown in figure 2. It can be observed that manual minutiae selection leads to much better results than automatic selection, as it was expected. However, when the manual selection is limited to 12 minutiae, the performance dramatically drops. It is shown that automatic extraction provides 48% of first position genuine candidates, when the manually selected 12 minutiae only rates 28% first position genuine candidates. It can be also observed that matching with only 8 manually-selected minutiae shows to have the worst performance, as it was expected. We found that only 6% of the genuine fingerprints rated first on the lists of candidates and 94% of them never appeared in the lists of candidates.

These results using latent queries against the ten-print database are summarized in Table 1. It is important to remark that the average number of minutiae that the system marked in the latent prints was 31.2 while the average number of minutiae marked by the experts was 25.2. This shows that due to the poor quality of latent images, the system tends to mark minutiae in points where actually there are not any, leading to a worse performance in the matcher.

B. Automatic minutiae extraction with latent, rolled and plain fingerprints

In this section, results using impressions compared to the ten-print database are shown, and then compared to latent results when minutiae are automatically extracted.

As it can be observed in figure 3, matching with rolled impressions leads to 100% identification rate for rank 1, which means that all the genuine fingerprints appeared as the first candidate in all the experiments. For the plain impressions, the results are slightly worse, as in 94% of the cases, the target print appeared as the first candidate. However, in 4% of the experiments with plain impressions, the genuine fingerprint did not appear in the list of candidates. The decrease in performance of plain with respect to rolled impressions is mainly due to the fact that rolled impressions tend to present a much higher number of minutiae, also presenting a much wider fingerprint area. Worth noting, while the average number of minutiae that the system found in rolled impressions is 83.8, in plain impressions it is only 44.16.
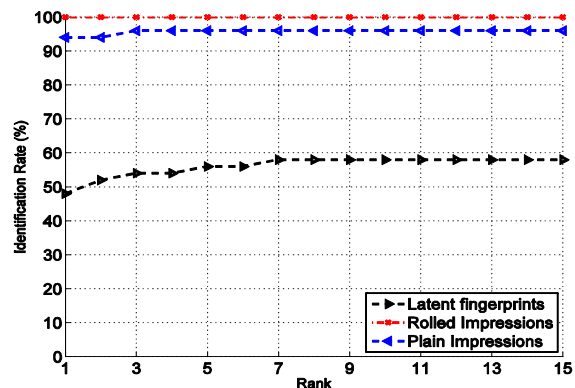

Figure 3. CMC curves for rolled, plain and latent queries.

Figure 3 also shows the results using latent queries with automatic minutiae selection. It is shown that the performance is much worse for latent fingerprints than for the case of plain or rolled impressions.

In table 2, the performance of the three types of matching experiments is shown in terms of percentage of first-rank appearance and non appearance of target identity.

The big difference in performance among latent fingermarks and fingerprint impressions can be attributed not only to a decrease in the image quality, but also to the average number of minutiae for the latent set, which was 31.2, much lower than in the case of impressions (i.e. 83.8 for rolled and 44.16 for plain).

TABLE II. FIRST CANDIDATE VS. NOT-APPEARANCE RATES FOR GENUINE IMPRESSIONS AND LATENT FINGERPRINTS

	Rolled	Plain	Latent
Rank 1	100%	94%	48%
Rank > 15	0%	4%	42%

IV. CONCLUSIONS AND FUTURE WORK

This work presented a comparative study of the performance of biometric recognition with latent fingermarks and rolled and plain fingerprint impressions from real forensic cases. A database from the Identification Department of Spanish Guardia Civil has been used for that purpose. We presented and discussed the results of two types of experiments regarding fingerprint feature extraction and matching. The aim of the first experiment was to compare manual and automatic feature extraction in latent fingerprints. The second experiment consisted of comparing automatic feature extraction for latent marks and rolled and plain impressions.

In the first scenario, when comparing manual and automatic minutiae extraction in latent fingerprints, it is clear that the identification accuracy increases when the

feature selection is manually made by an expert, which was an expected result. However, it is significant that automatic minutiae selection works much better than manual selection when selecting a limited amount of minutiae (12 in our experiments).

In the second scenario, we have compared the matching performance for automatic feature extraction in latent prints, rolled and plain impressions. In this case, it has been shown that automatic minutiae extraction leads to a high performance on the matching step when dealing with good quality fingerprints, such as rolled and plain impressions. However, the performance in latent fingerprints is much worse due to the wrong minutiae extraction made by the system.

Although the experimental scenario considered is very realistic (state-of-the-art AFIS and the ten-print database of Guardia Civil comprising around 2.5 million templates), it is important to remark that these results are obtained with a small set of fingerprint queries. This is mainly due to the limitations in obtaining and handling data from real forensic cases. For this reason, the statistical significance of the results must be taken with care. However, results are relevant in order to observe trends in the performance with different minutiae extraction techniques and types of images, and pioneering in the sense of scientific understanding of this important technology for forensics.

Future work includes measuring the quality of latent marks in order to better understand its influence in performance [12]. Also, the extension of the sets of queries is planned in order to have more significant results.

REFERENCES

[1] A. K. Jain and J. Feng, "Latent Fingerprint Matching", *IEEE Transs. on Pattern Analysis and Machine Intelligence*, 2010

[2] C. Champod, C. Lennard, P. Margot, and M. Stoilovic. *Fingerprints and Other Ridge Skin Impressions.* CRC Press, 2004.

[3] C. Wilson, R. Austin Hicklin et al. "Fingerprint Vendor Technology Evaluation 2003", *Technical Report NISTIR 7123*, 2004.

[4] J. Fierrez-Aguilar, Y. Chen, J. Ortega-Garcia and A. K. Jain. "Incorporating image quality in multi-algorithm fingerprint verification", *Proc. IAPR ICB.* Springer LNCS-3832, pp 213-220, January 2006.

[5] H. Fronthaler, K. Kollreider, et al. "Fingerprint Image Quality Estimation and its Application to Multi-Algorithm Verification", *IEEE Trans. on Information Forensics and Security,* Vol. 3, n. 2, pp. 331-338, June 2008.

[6] D.R. Asbaugh. *Quantitative-Qualitative Friction Ridge Analysis: An Introduction to Basic and Advanced Ridgeology.* The CRC Press, Boca Raton, FL, 1999.

[7] N.M. Egli, C. Champod, and P. Margot. "Evidence evaluation in fingerprint comparison and automated fingerprint identification systems: Modelling within finger variability", *Forensic Science Intl*, 167:189-195, 2007.

[8] N. Ratha and R. Bolle. *Automatic Fingerprint Recognition Systems.* Springer, 2003.

[9] D. Dessimoz and C. Champod. "Linkages Between Biometrics and Forensic Science", *Handbook of Biometrics.* (ed.) A.K. Jain at al. Springer, 2007.

[10] C. Watson et al. "Two Finger Matching With Vendor SDK Matchers", *Technical Report NISTIR 7249*, 2005.

[11] R.M. Bolle, J.H. Connell et al. "The relation between the ROC curve and the CMC", *Proc. IEEE AutoID*, 2005.

[12] F. Alonso-Fernandez, J. Fierrez et al. "A comparative study of fingerprint image-quality estimation methods", *IEEE Trans. on Information Forensics and Security*, Vol.2, n.4, pp. 734-743, December 2007.