Experimental Results on Fingerprint Liveness Detection

Luca Ghiani, Paolo Denti, and Gian Luca Marcialis

Department of Electrical and Electronic Engineering,
University of Cagliari,
Piazza d'Armi, Cagliari, Italy
{luca.ghiani,marcialis}@diee.unica.it, paolode@hotmail.com

Abstract. Fingerprint liveness detection is aimed to detect if a fingerprint image, sensed by an electronic device, belongs to an alive fingertip or to be an artificial replica of it. Recent studies have shown that a fingerprint can be replicated and, if a clever attacker tries to evade the system, this is an issue. Accordingly, several countermeasures in terms of fingerprint liveness detection algorithms have been proposed, but never compared on a benchmark data set, internationally accepted by the research community. In this paper, we present some recent experimental results on several state-of-the-art fingerprint liveness detection algorithms on the datasets available at Second International Fingerprint Liveness Detection Competition (LivDet 2011). The results we proposed help assessing which are the more effective approaches used so far.

1 Introduction

Biometry measures allow us to perform the identification of a person based on his physical (fingerprints, face, iris) or behavioural (gait, signature) attributes and to establish an identity based on who that person is, rather than what he/she possesses (e.g. a card that can be lost or stolen) or remembers (e.g. a password that can be forgotten) [1]. Nowadays, more than ever, it is very important to be able to tell if an individual is authorized to perform actions like entering a facility, access privileged information or even cross a border. Therefore, biometric systems are considered to be more reliable for the recognition of a person than traditional methods.

During the past decade we have seen a significant growth in biometric research resulting in the development of innovative sensors, novel feature extraction and matching algorithms. A biometric system is a pattern recognition system that acquires biometric data from an individual, extracts a features set from the data, compares these features against those stored in a database and executes an action based on the comparison result.

Fingerprints are the most used, oldest and well-known biometric measurements [2]. Everybody has them and each one is different from the other since ridges formation is a combination of factors both environmental and genetic. Fingerprints are composed of epidermic ridges and valleys that usually run in

F.J. Perales, R.B. Fisher, and T.B. Moeslund (Eds.): AMDO 2012, LNCS 7378, pp. 210–218, 2012.
© Springer-Verlag Berlin Heidelberg 2012

parallel. On the images obtained through sensors ridges are dark lines while the valleys are bright lines. Obviously their uniqueness depends on the type and number of features extracted (basically less features means less details and therefore less information obtained).

However, as recently shown, fingerprints can be forged and that is the subject of this work. The main fingerprints characteristic is that, each one of them is considered unique. If fingerprints can be replicated, this uniqueness is no longer valid. It is possible to create an artificial replica through several methods and using several materials, and the related images can be indistinguishable from alive ones (see for example Figures 1 and 2). Therefore, the development of liveness detection techniques is important to try to distinguish if a fingerprint image is coming from an alive person or from a replica. Liveness detection seeks additional data to verify if a biometric measure is authentic. Fingerprint liveness detection, with either hardware-based or software-based systems, is used to check if a presented fingerprint originates from a live person or an artificial finger [3]. It is based on the principle that additional information can be obtained from the data acquired by a standard verification system. This additional data can be used to verify if an image is authentic.

To detect liveness, hardware-based systems use additional sensors to gain measurements outside of the fingerprint image itself while the software-based ones use image processing algorithms to gather information directly from the collected fingerprint. These systems classify images as either live or fake.

(a) Digital (b) Biometrika (c) Italdata (d) Sagem

Fig. 1. Examples of live fingerprints acquired with the 4 sensors

In this paper, we focus our attention on the software-based approaches, which are cheapest than hardware-based, since these require additional and invasive hardware to measure the liveness directly from the fingertip of people. Instead, software-based must detect liveness from features extracted from the fingerprint images captured by the sensor. In other words, the liveness detection problem is treated as a pattern recognition problem, where a set of features must be selected in order to train an appropriate classifier. Although several feature sets have been proposed to this aim, it is difficult to assess the state-of-the-art appropriately, due to the lack of data sets. In order to cover this gap, we considered a large set

(a) Digital (b) Biometrika (c) Italdata (d) Sagem

Fig. 2. Examples of fake fingerprint acquired with the 4 sensors

of liveness measurements at the state-of-the-art and tested to the data sets made available for the two editions of the fingerprint liveness detection competition, held in 2009 and 2011 [4, 5]. We compared the state-of-the-art results with the LivDet competition results and made a first point about the real performances of these approaches on realistic data sets.

This paper is organized as follows. Section 2 briefly describes the investigated algorithms. Section 3 shows the experimental results and performance with respect to the LivDet results. Section 4 concludes the paper.

2 Investigated Algorithms

In fingerprint liveness detection there are two different types of algorithms: the live-based ones look for characteristics of a living person's fingerprint, while the fake-based ones take advantage from the fact that, during the fake production, some details are lost and some defects are introduced.

We studied the fingerprint liveness detection state-of-the-art and we selected two live-based (pores detection and ridges wavelet) and six fake-based algorithms (local binary patterns, power spectrum, wavelet energy, valleys wavelet, curvelet energy and curvelet GLCM).

- LBP: local binary patterns were first employed for two-dimensional textures analysis and excellent results were obtained due to their invariance with respect to grey level, orientation and rotation. It extracts certain uniform patterns corresponding to micro-features in the image. The histogram of these uniform patterns occurrence is capable of characterize the image as it combines structural (it identify structures like lines and borders) and statistical (micro-structures distribution) approaches [6].
- Pores detection: since the pores presence in live fingerprints determines the perspiration effect, the pores detection algorithm analyzes pores distribution in order to discriminate between fake and live fingerprint images. By scanning the image along the fingerprint ridges[7] it extracts the pores number and the average distance between pores.

- Power spectrum: Coli *et al.* [8] analyzed fingerprints images in terms of high frequency information loss. In the artificial fingerprint creation, the ridge-valley periodicity is not altered by the reproduction process but some micro-characteristics are less defined. Consequently, high frequency details can be removed or strongly reduced. It is possible to analyze these details by computing the image Fourier transform modulus also called power spectrum.
- Wavelet energy signature: wavelet decomposition of an image [9] lead to the creation of four sub-bands: the approximation sub-band containing global low frequency information, and three detail sub-bands containing high frequency information. The image is decomposed in 4 levels using 3 different wavelet filters (Haar, Daubechies (db4) and Biorthogonal (bior2.2)) and the approximation image is not considered, hence the sub-bands number is $3 \times 4 = 12$.
- Ridges wavelet: after his extraction, a fingerprint skeleton can be used as a mask to obtain the gray level values along the ridges and these values are united into a signal. A wavelet multiresolution decomposition is applied to that signal with seven decomposition levels [10].
- Valleys wavelet: in this case the skeleton of the valleys is obtained. As for the ridges wavelet analysis, the skeleton is used as a mask to extract a signal representing the gray level values along the valleys. As for the ridges wavelet analysis a wavelet multiresolution decomposition is applied to that signal with seven decomposition levels [11].
- Curvelet [12] transform partitions curves into a collection of ridge fragments and then uses ridgelet transform to represent each of them. It is very efficient for representing edges and other singularities along curves due to his high directional sensitivity and his high anisotropy. We consider two different curvelet signature:
  - Curvelet energy signature: the energies of the 18 sub-bands are measured by computing means and variances of curvelet coefficients.
  - Curvelet co-occurrence signature: for each of the 18 sub-bands, the GLCM (Gray Level Co-occurrence Matrix) is calculated together with 10 corresponding features.

3 Experimental Results

3.1 The LivDet 2011 Data Sets

Two editions of the International Fingerprint Liveness Detection Competition (LivDet) have been held in 2009 and in 2011 [4, 5].

In particular, the four LivDet 2011 datasets consist of images acquired with four different devices; Biometrika, Digital Persona, Italdata and Sagem. There are 4000 images for each of these devices, 2000 live images and 2000 spoof images.

All the fake fingerprints have been created with the consensual method, by following these steps: the volunteer releases his fingerprint on a mould of plasticine or silicon-like material; the chosen material is poured or applied over the

mould and after a certain time interval, this cast is removed from the mould, and can be used as fingerprint replica. Each replica in sensed by the sensor, that provide a "fake fingerprint image".

Spoof materials used were gelatine, latex, PlayDoh, silicone and wood glue for Digital Persona and Sagem; gelatine, latex, liquid silicon, silicone and wood glue for Biometrika and Italdata (400 of each of 5 spoof materials in both cases).

Each dataset of 4000 images per scanner was divided into two equal parts, training and testing. The first part had to be used to train the algorithm, and the second part to test them on independent data. In this paper, we follow the same protocol in order to compare the results with those of the competition.

3.2 Results

The classifier output is a posterior probabilty estimation of the live class given the feature set, thus included in the $[0, 1]$ real interval. This can be considered as a measurement of the "liveness" of the data received as input. Once a threshold t is selected, if this probability s is higher than t, the fingerprint is considered as alive, otherwise it is considered as fake. These results can be presented with a ROC curve: by varying the t value from 0 to 1, we can plot the correspondent percentage of fake fingerprints misclassified as alive ones, and the percentage of live fingerprints misclassified as fake ones. According to the LivDet terminology, these values are called FerrFake and FerrLive, respectively [4, 5].

ROC curves are shown in Figures 3, 4, 5, 6. We notice that usually the results with the Italdata sensor are worse than the others (except for the pores detection and valleys wavelet algorithms) and that, for each of the four datasets, the best performance is always the one obtained with the LBP algorithm. However, in general, the performance on LivDet 2011 data set is worse than that reported in several papers [7–12], for many algorithms. The results worsening can be due to the quality improvement of the artificial fingers produced for the LivDet competition. As a matter of fact, some feature sets, as the one based on the power spectrum or on the wavelet transform, aimed to measure the image liveness at high frequencies, show several limitation in the performance. On the contrary, a texture classification algorithm as the LBP based one, preserves its discriminatory power. This is probably because its capability in describing some image peculiarity continues to work no matter what the quality of the fakes is. We are currently better investigating the properties of this feature set, in order to understand the motivation of his "robustness" with respect to the fake quality. Anyway, this is certainly an important advantage, which, if confirmed theoretically and experimentally, could set that the best way to improve a liveness detection algorithm is to follow the LBP paradigm.

Table 1 shows the rate at which each algorithm produces a false acceptance of a spoof image (FerrFake) and Table 2 shows the rate at which each algorithm produces a false rejection of a live subject (FerrLive). The threshold t used is set to 0.5, that is, the Bayesian error is computed according to the posterior

probabilities estimated by the SVM. These results are paired with those obtained by the three LivDet 2011 participants (in the last three rows of both tables). Once again the LBP emerges as the best algorithm even when its results are compared with those of the three participants. In fact, if the LBP's FerrFake value is higher than that of LivDet 2011 participants (for example, Dermalog), the corresponding FerrLive value is much lower. This also holds for the FerrLive value, and clearly shows that, on overall, the Bayesian error of LBP is the lowest one. The only performance really comparable with that of LBP is the one of the "Federico" participant, but only for the Digital sensor. On average, we notice that all algorithms exhibit, with some variations, the same performance, which makes current fingerprint liveness detection algorithms unacceptable if integrated on a fingerprint matching system as a separated module, due to their high FerrLive value which is not adequately counterbalanced by a low enough FerrFake value.

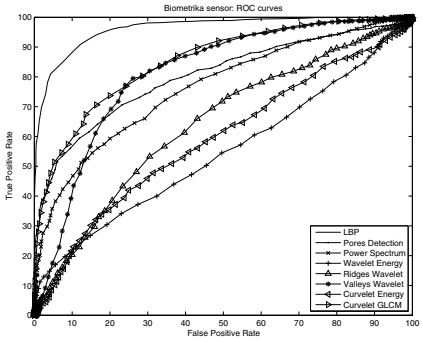

Fig. 3. Biometrika sensor ROC curves

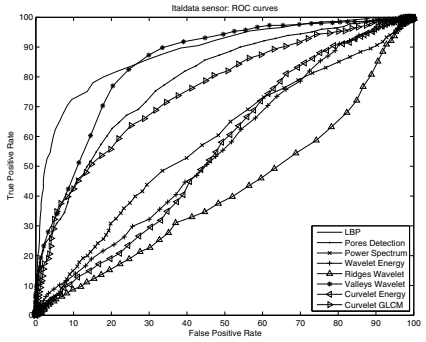

Fig. 4. Italdata sensor ROC curves

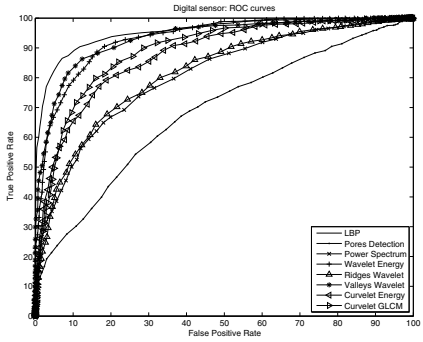

Fig. 5. Digital sensor ROC curves

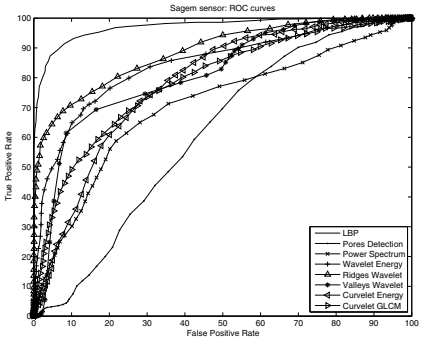

Fig. 6. Sagem sensor ROC curves

Table 1. FerrFake calculated with the Biometrika, Italdata, Digital and Sagem sensors

	Biometrika	Italdata	Digital	Sagem
LBP [6]	16.40	15.10	8.70	4.34
Pores Detection [7]	27.80	22.00	30.50	49.90
Power Spectrum [8]	23.90	29.40	23.50	21.81
Wavelet Energy [9]	73.00	51.80	15.10	16.22
Ridges Wavelet [10]	47.10	63.10	37.00	18.15
Valleys Wavelet [11]	48.60	39.10	12.40	55.12
Curvelet Energy [12]	55.10	40.70	27.40	39.58
Curvelet GLCM [12]	16.40	25.20	22.00	25.00
Dermalog [5]	29.00	28.50	6.20	12.50
Federico [5]	42.00	40.10	11.60	13.10
CASIA [5]	38.10	2.80	34.70	23.60

Table 2. FerrLive calculated with the Biometrika, Italdata, Digital and Sagem sensors

	Biometrika	Italdata	Digital	Sagem
LBP [6]	5.90	22.00	12.60	12.70
Pores Detection [7]	26.90	35.30	41.70	30.40
Power Spectrum [8]	37.40	56.20	30.80	41.20
Wavelet Energy [9]	27.40	41.80	13.00	27.90
Ridges Wavelet [10]	30.50	50.80	18.10	22.90
Valleys Wavelet [11]	9.40	8.20	13.70	9.00
Curvelet Energy [12]	35.30	55.10	16.40	17.50
Curvelet GLCM [12]	29.40	36.30	14.70	31.10
Dermalog [5]	11.00	15.10	66.00	15.10
Federico [5]	38.00	39.90	6.20	13.80
CASIA [5]	29.70	50.60	16.10	22.10

4 Conclusions

In this paper, we investigated the performance of several fingerprint liveness detection algorithms, by testing them using the protocol and data sets adopted in the LivDet 2011 competition, where the quality of fake fingerprint was certainly higher than that of the previous edition.

Like a "cops and robbers game", it must be considered that the improvement of the liveness detection algorithms will be also followed by an improvement of the falsification techniques. In this continuous struggle it is mandatory to introduce novel feature sets, intrinsically robust to the quality of the fingerprint replica.

Reported analysis clearly suggests that there is much work to do on fingerprint liveness detection, in order to design algorithms and features sets whose integration with fingerprint verification systems can be considered "acceptable" in terms of overall verification performance.

Our future works are related to these mentioned problems: the study of materials peculiarities, and the spoof creation process by non consensual methods, and focusing on a robust enough fingerprint liveness detection algorithm to be integrated on fingerprint verification systems.

Acknowledgments. This work was partly supported by the Tabula Rasa project, 7th FRP of the European Union (EU), grant agreement number: 257289; by the PRIN 2008 project "Biometric Guards - Electronic guards for protection and security of biometric systems" funded by the Italian Ministry of University and Scientific Research (MIUR). The authors also thank Zahid Akthar and Valerio Mura for collecting part of the LivDet 2011 data sets.

References

1. Jain, A.K., Flynn, P., Ross, A.: Handbook of Biometrics. Springer (2007) ISBN 978-0-387-71040-2
2. Maltoni, D., Maio, D., Jain, A.K., Prabhakar, S.: Handbook of Fingerprint Recognition. Springer, New York (2003) ISBN 0387954317
3. Coli, P., Marcialis, G.L., Roli, F.: Vitality Detection from Fingerprint Images: A Critical Survey. In: Lee, S.-W., Li, S.Z. (eds.) ICB 2007. LNCS, vol. 4642, pp. 722–731. Springer, Heidelberg (2007) ISBN 978-3-540-74548-8, doi:10.1007/978-3-540-74549-5_76
4. Marcialis, G.L., Lewicke, A., Tan, B., Coli, P., Grimberg, D., Congiu, A., Tidu, A., Roli, F., Schuckers, S.: First International Fingerprint Liveness Detection Competition—LivDet 2009. In: Foggia, P., Sansone, C., Vento, M. (eds.) ICIAP 2009. LNCS, vol. 5716, pp. 12–23. Springer, Heidelberg (2009)
5. Yambay, D., Ghiani, L., Denti, P., Marcialis, G.L., Roli, F., Schuckers, S.: LivDet 2011 - Fingerprint Liveness Detection Competition 2011. In: 5th IAPR/IEEE Int. Conf. on Biometrics, New Delhi (India), March 29, April 1 (in press, 2012)
6. Ojala, T., Pietikäinen, M., Mäenpää, T.: Multiresolution gray-scale and rotation invariant texture classification with local binary patterns. IEEE Trans. on Pattern Analysis and Machine Intelligence, 971–987 (2002), doi:10.1109/TPAMI.2002.1017623
7. Marcialis, G.L., Roli, F., Tidu, A.: Analysis of Fingerprint Pores for Vitality Detection. In: Proc. of 20th IEEE/IAPR International Conference on Pattern Recognition (ICPR 2010), Instanbul (Turkey), August 23-26, pp. 1289–1292 (2010) ISBN 978-1-4244-7542-1, doi:10.1109/ICPR.2010.321
8. Coli, P., Marcialis, G.L., Roli, F.: Power spectrum-based fingerprint vitality detection. In: Tistarelli, M., Maltoni, D. (eds.) IEEE Int. Workshop on Automatic Identification Advanced Technologies AutoID 2007, Alghero (Italy), June 7-8, pp. 169–173 (2007)

9. Nikam, S.B., Agarwal, S.: Texture and Wavelet-Based Spoof Fingerprint Detection for Fingerprint Biometric Systems. In: First International Conference on Emerging Trends in Engineering and Technology, ICETET 2008, Nagpur, Maharashtra, July 16-18, pp. 675–680 (2008) ISBN 978-0-7695-3267-7, doi:10.1109/ICETET.2008.134
10. Tan, B., Schuckers, S.: Liveness Detection for Fingerprint Scanners Based on the Statistics of Wavelet Signal Processing. In: Computer Vision and Pattern Recognition Workshop, CVPRW 2006, June 17-22, pp. 26–26 (2006) ISBN 0-7695-2646-2, doi:10.1109/CVPRW.2006.120
11. Tan, B., Schuckers, S.: New approach for liveness detection in fingerprint scanners based on valley noise analysis. J. Electron. Imaging 17, 011009 (2008), doi:10.1117/1.2885133
12. Nikam, S.B., Agarwal, S.: Fingerprint Liveness Detection Using Curvelet Energy and Co-occurrence Signatures. In: Fifth International Conference on Computer Graphics, Imaging and Visualization 2008 IEEE, pp. 217–222 (2008) ISBN 978-0-7695-3359-9, doi:10.1109/CGIV.2008.9