ON THE IMPORTANCE OF USING HIGH RESOLUTION IMAGES, THIRD LEVEL FEATURES AND SEQUENCE OF IMAGES FOR FINGERPRINT SPOOF DETECTION

Murilo Varges da Silva[*] *Aparecido Nilceu Marana* *Alessandra Aparecida Paulino*

Department of Computing, Faculty of Sciences, Sao Paulo State University (UNESP), Bauru, SP, Brazil
murilo.varges@ifsp.edu.br, {nilceu, paulinoa}@fc.unesp.br

ABSTRACT

The successful and widespread deployment of biometric systems brings on a new challenge: the spoofing, which involves presenting an artificial or fake biometric trait to the biometric systems so that unauthorized users can gain access to places and/or information. We propose a fingerprint spoof detection method that uses a combination of information available from pores, statistical features and fingerprint image quality to classify the fingerprint images into live or fake. Our spoof detection algorithm combines these three types of features to obtain an average accuracy of 97.3% on a new database (UNESP-FSDB) that contains 4,800 images of live and fake fingerprints. An analysis is performed that considers some issues such as image resolution, pressure by the user, sequence of images and level of features.

***Index Terms*—** Biometrics, spoof detection, fingerprint, pores, security.

1. INTRODUCTION

The initial interest in recognition technologies was due mainly to law enforcement agencies. More recently, the concerns about security and identity fraud have increased, therefore creating a need for biometric recognition technologies in non-forensic applications (e.g. border control, national ID, etc.) [1]. Over the years, several physical and behavioral characteristics have been explored (e.g., face, fingerprint, iris, hand vein, voice, etc.), leading to the development of new recognition technologies that have been successfully deployed.

The successful and widespread deployment of biometric systems brings on a new challenge, namely spoofing. While biometric systems are developed to secure and control access, to avoid fraud, etc., spoofing methods are developed to breach the security of biometric systems so that unauthorized users can gain access to places and/or information. Spoof detection (or liveness detection) methods have been developed to counter these types of attacks and they are essential to minimize the vulnerability of these systems.

[*]The first author is also with the Federal Institute of Education, Science and Technology of Sao Paulo, Birigui, SP, Brazil.

The importance of the problem is also highlighted by the growing interest in the liveness detection competitions such as Fingerprint Liveness Detection Competitions [2, 3, 4] and Liveness Detection Iris Competition [5].

A spoof includes the presentation to the biometric system of artificial or fake biometric traits (e.g., a silicone finger or a photograph of a face), as well as biometric traits obtained without consent from legitimate users (e.g., a dismembered finger).

The goal of this paper is to analyze and discuss some factors that might influence the performance of the liveness detection algorithm. These factors include differences in image resolution, differences in the amount of pressure that is applied during the acquisition, differences in the images acquired after a few seconds in contact with the sensor, and the use of level 3 features to improve the classification.

2. PROPOSED SPOOF DETECTION METHOD

Fingerprint features can be classified into three levels: (i) Level 1: coarse features such as orientation field, (ii) Level 2: minutiae (ends or bifurcations of the ridges), and (iii) Level 3: micro level characteristics of the fingerprints, such as pores and dots. In the liveness detection problem, the interest in the level 3 features ([6, 7, 8]) is due to the increased difficulty in faking such micro details in the spoof.

In our work, we use a combination of features. Statistical features are very simple to compute, and many researchers have reported good accuracies based on them [9]. Fingerprint image quality can also be used to improve the performance of a spoof detection method [7]. Thus, our spoof detection algorithm combines these three types of features to obtain an accuracy comparable to the methods reported in the literature, so that analyses can be performed on the factors that might influence the accuracy of the spoof detection methods.

2.1. Feature Extraction

The features used in the proposed approach are based on statistical features, fingerprint image quality and pores. For each fingerprint, a 9-dimensional feature vector $F =$

$(F_1, F_2, \ldots, F_9)$ is obtained. The features are defined as follows:

- Statistical features: these features represent the visual differences in the gray level intensities that can be observed between live and fake fingerprints [9].

$$F_1 = \sum_{n=0}^{N-1} H(n)^2 \quad \text{(Energy)} \quad (1)$$

$$F_2 = \sum_{n=0}^{N-1} H(n) \times logH(n) \quad \text{(Entropy)} \quad (2)$$

$$F_3 = \frac{\sum_{n=1}^{N-1} H(n)}{N} \quad \text{(Mean)} \quad (3)$$

$$F_4 = \sum_{n=0}^{N-1} (n-\mu)^2 H(n) \quad \text{(Variance)} \quad (4)$$

$$F_5 = \frac{1}{\sigma^3} \sum_{n=0}^{N-1} (n-\mu)^3 H(n) \quad \text{(Skewness)} \quad (5)$$

$$F_6 = \frac{1}{\sigma^4} \sum_{n=0}^{N-1} (n-\mu)^4 H(n) \quad \text{(Kurtosis)} \quad (6)$$

  where $H(n)$ is the gray level histogram of the fingerprint, N is the number of gray levels, μ is the mean and σ is the standard deviation. The histogram $H(n)$ is normalized (sum of occurrences of each gray level intensity divided by the total number of pixels) and equalized (H is the histogram after the process of histogram equalization is applied).

- NIST Fingerprint Image Quality Measure (NFIQ): this feature measures the fingerprint quality proposed in [10], which is based on image quality maps and number and quality of minutiae. NFIQ is an integer value between 1 and 5, with 1 being the highest fingerprint quality and 5 the lowest.

$$F_7 = \text{NFIQ}, \quad (7)$$

  where NFIQ $\in \{1, 2, 3, 4, 5\}$.

- Number of pores: this feature measures the total number of pores in the fingerprint using the adaptive approach proposed in [11]. This adaptive approach regulates the detection according to the direction and period of the local ridges, and it detects pores in fingerprints at multiple resolution values.

$$F_8 = \text{Number of pores}. \quad (8)$$

- Pore frequency: this feature is obtained from the analysis of the pixel intensities along the ridges [6]. The thinned fingerprint image is extracted and it is then used as a mask to obtain the intensities of the pixels along the ridges. The Fourier Transform is then extracted from this signal, and the response around pore frequency (between 11 and 33) is measured. The assumption is that live fingerprints will have a larger pore frequency since many spoof materials cannot perfectly reproduce third level characteristics of a fingerprint. Given the average of the Fourier Transform of each signal (corresponding to each fingerprint ridge), the pore frequency for each fingerprint is defined as [6]:

$$F_8 = \sum_{k=11}^{33} f(k)^2, \quad (9)$$

  where $f(k) = \frac{\sum_{i=1}^{n} \left| \sum_{p=1}^{255} S_{li}^a(p) e^{-j2\pi(k-1)(p-1)/256} \right|}{n}$, $S_{li}^a = S_{li} - mean(S_{li})$, n is the total number of ridges and S_{li} are the individual signals that represents the intensities along the ridges.

2.2. Classification

The liveness detection problem can be viewed as a two-class classification problem, where the two classes are: live and fake fingerprints. Given the 9-dimensional feature vectors for each fingerprint, Support Vector Machine (SVM) is used to classify the fingerprints into one of the two classes. Other methods were also used (Multilayer Perceptron, Optimum Path Forest, K-Nearest Neighbors) but, due to space constraints, only the results of the best performing classifier (SVM) are reported. The classification was performed by WEKA 3.7 [12] with parameters adjusted according to the suggestions in [13].

3. DATABASES

The LivDet 2013 database is a publicly available liveness detection fingerprint database that includes four subsets (Biometrika, Italdata, Crossmatch and Swipe), each one containing more than 4,000 images of live and fake fingerprints. We have applied the proposed approach on the first three subsets to verify that its performance is comparable with the performance of other methods reported in the literature. The protocol from LivDet 2013 was used, and the dataset Swipe was not considered due to the very low resolution.

We have also created a fingerprint spoof database (UNESP Fingerprint Spoof Database - UNESP-FSDB) that includes live and spoof fingerprint images in different scenarios for the purpose of analyzing some of the factors that might influence the performance of a liveness detection technique. This database was collected using the commercial fingerprint sensor CrossMatch LScan 1000T, which allows the acquisition of both normal (500 p.p.i.) and high (1000 p.p.i.) resolution fingerprint images. Following the training and testing protocols from LivDet, for each scenario, the subset being used was randomly separated into two sets of the same size, one for training and the other for testing.

The UNESP-FSDB database contains a total of 4,800 images of live and spoof fingerprints collected from 20 subjects. The database was collected from fingerprints of volunteers and from fake fingers created in a cooperative mode by the volunteer placing his/her fingers on a mold of PlayDoh, followed by the use of latex and silicone to create the spoof cast.

For each person, four spoofs were made: two of the thumb and two of the index finger, with one spoof made of latex and the other made of silicone for each finger. From these spoofs and from the live fingers, we collected the database with the following characteristics:

- Samples of live fingers: consist of 1,600 fingerprints, with images acquired from two different fingers (thumb and index), and for each finger, two different resolutions (500 p.p.i. and 1000 p.p.i.). For each resolution, the subject was asked to place his/her finger in the sensor and 10 fingerprints were captured sequentially at every second, for ten seconds (0 to 9 seconds). The subject was also asked to do this procedure twice for each resolution, one by pressing his/her finger as he/she would usually do (normal pressure), and another by increasing this pressure (high pressure).
- Samples of spoof fingers: consist of 3,200 fingerprints, with images acquired from four different spoof fingers (latex and silicone for both the thumb and the index fingers). For each finger and material at each of the two resolutions (500 p.p.i. and 1000 p.p.i.), 10 fingerprints were captured sequentially at every second (0 to 9 seconds). Again, the procedure was performed twice for each resolution, by varying the pressure that the spoof finger touched the sensor between normal and high pressure.

Figure 1 shows examples of (a) live, (b) latex and (c) silicone fingerprints, (d) live finger, and spoof fingers made with (e) latex and (f) silicone.

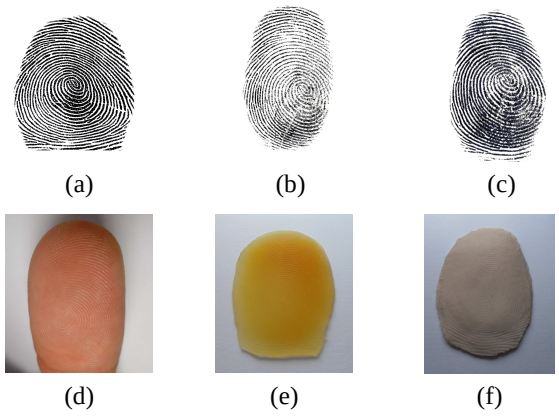

Fig. 1: Examples from UNESP-FSDB. (a) Live fingerprint, (b) latex fake fingerprint, (c) silicone fake fingerprint, (d) live finger, (e) spoof finger made with latex and (e) spoof finger made with silicone.

4. EXPERIMENTAL RESULTS AND ANALYSIS

In this section we present our experimental results on UNESP-FSDB database when different image resolutions, normal or high pressure and different acquisition times are used. Further, we present an analysis of the addition of pore information in order to improve the classification performance.

Table 1: Classification accuracies (%) of the combination approach used in our work compared to the best and worst performing algorithms submitted to the Liveness Detection Competition 2013 [14].

Subset	Combination	Best	Worst
Biometrika	82.6	98.3	67.1
Italdata	78.9	99.4	50.0
Crossmatch	64.70	68.8	44.44

The performance of liveness or spoof detection methods are usually reported using accuracy, ferrlive and ferrfake. The accuracy (or classification accuracy) of a liveness detection method is the percentage of samples that are correctly classified (either live classified as live or fake classified as fake) over all the test samples. Ferrlive is the errors that the system makes by classifying a live print as fake, and the ferrfake is the errors that the system makes by classifying a fake print as live. We will mainly discuss the classification accuracy of each scenario due to its compactness.

Table 1 shows the classification accuracy of the combination approach used in our work compared to the best and worst performing algorithms submitted to the Liveness Detection Competition 2013 [14].

4.1. Image Resolution

It is usually believed that higher resolution images can bring some gain in performance due to the greater amount of detail that can be extracted from them. On the other hand, a higher resolution might increase the amount of noise in an image. Our experiments on the UNESP-FSDB suggest that the former might not be the case for some feature sets. The classification accuracy of the spoof detection method proposed in our work reached 97.9% for 500 p.p.i. images, and it is slightly worse (96.7%) when the images at 1000 p.p.i. were used. The conclusion that images of higher resolution will decrease the performance compared to images of lower resolution cannot be drawn from our experiments because the images were not captured at the same time in different resolutions, so the variation in the errors might be just a result of the small changes that can occur during different acquisitions, e.g., area captured, amount of moisture (see Fig. 2), finger pressure, etc.. Therefore, the increase in image resolution alone does not automatically improves the performance of the spoof detection method.

4.2. Finger Pressure

In the UNESP-FSDB, the subjects were asked to apply normal pressure at the first acquisition, and they were asked to

(a) (b)

Fig. 2: (a) Dry and (b) normal live fingerprints from UNESP-FSDB (1000 p.p.i.). It can be observed that the pores are more visible for normal condition skin compared to dry skin.

Table 2: Classification accuracies (%) of the proposed method on the UNESP-FSDB when (1) no pore information is used, (2) pore frequency is used, and (3) all the nine features are used.

Resolution	(1)	(2)	(3)
500 p.p.i.	86.6	95.3	97.9
1000 p.p.i.	87.7	92.7	96.7

increase this pressure at the second acquisition. The performance on the subset of the 500 p.p.i. images increased from 96.0% to 98.7% and subset of the 1000 p.p.i. images increase from 96.7% to 97.8% when the amount of pressure in the capture process was increased. The increase was related to the live as live classification accuracy, so asking the user to increase the pressure could be one way to reduce the ferrlive.

4.3. Third Level Feature

In [8], the authors used the difference in the pore quantity between a reference image and a distorted query. Here we only used the number of pores combined with features extracted from the query fingerprint, and thus we only need the query image to decide whether it is a live or fake finger. In [7], the authors used differences in the number of pores in certain regions of the fingerprints collected 5 seconds apart without the user removing the finger from the sensor. Again, in our approach, only one query image is required. The number of pores is not a discriminative enough feature, but this information helps in increasing the performance of the detection algorithm, which is also the case with the pore frequency.

Table 2 shows the classification accuracy of the approach used here when no pore information is used, when pore frequency is used, and when all the information from pores is added to the other seven features for both the 500 p.p.i. and the 1000 p.p.i.. It can be observed that the addition of pore information increased the performance for both resolutions.

4.4. Sequence of Images

We have collected one image per second up to 9 seconds (10 images). We have divided the images in three separate groups: (i) Group 1 included images from 0 to 3 seconds, (ii) Group 2 included images from 4 to 6 seconds, and (iii) Group 3 included images from 7 to 9 seconds. Our goal by making this division was to verify whether the images from fingerprints that were in the sensor for a few seconds would yield a better performance than images captured right after the person places his/her finger in the sensor. Our experimental results show that the performance indeed increases when the time of the finger in the sensor increases, but this increase is small after 6 seconds. The accuracies in % for 500 p.p.i. images are 91.9, 96.9 and 97.9 for groups 1, 2 and 3, respectively, with the 1000 p.p.i. images presenting similar behavior (90.8, 94.4 and 94.8).

Most of the errors of our spoof detection method were made on the first captured images. This might occur because the first image is collected immediately after the user places his/her finger in the sensor, therefore the fingerprint might not be completely captured and the pressure might not be enough for a good quality fingerprint image. When designing a spoof detection system, discarding the first captured fingerprint might help improve the correct classification rate.

5. CONCLUSIONS

We have presented an analysis of different factors that might influence the performance of spoof detection algorithms such as image resolution, finger pressure and time of fingerprint image capture. Our experimental results show that higher resolution images do not automatically lead to better performance. However, high resolution fingerprint images might be more robust to variations in the amount of pressure that is applied when collecting the impression. We also showed that simple pore information actually helps the performance even when no reference image is used. In addition, our experiments show that if possible, the first capture should be discarded to obtain a better performance without significantly increasing the acquisition time.

While the features used in our method have been reported in the literature, our work differs from previous works because our main focus is on the analysis of the factors that might influence the performance of a spoof detection technique by using the database collected for this purpose. The analyses of different resolution, different pressure and extended acquisition time (up to 9 seconds), to the best of our knowledge, have not been previously performed.

We analyzed the performances based on features extracted from one image alone. As future work, we will develop algorithms that use the differences between images acquired a few seconds apart.

6. REFERENCES

[1] Davide Maltoni, Dario Maio, Anil K Jain, and Salil Prabhakar, *Handbook of Fingerprint Recognition*, Springer-Verlag, 2nd edition, 2009.

[2] "Fingerprint liveness detection competition 2009," `http://prag.diee.unica.it/LivDet09/`, 2009.

[3] "Fingerprint liveness detection competition 2011," `http://people.clarkson.edu/projects/biosal/fingerprint/index.php`, 2011.

[4] "Fingerprint liveness detection competition 2013," `http://prag.diee.unica.it/fldc/`, 2013.

[5] "Liveness detection-iris competition 2013," `http://people.clarkson.edu/projects/biosal/iris/`, 2013.

[6] Reza Derakhshani, Stephanie A C Schuckers, Larry A Hornak, and Lawrence O'Gorman, "Determination of vitality from a non-invasive biomedical measurement for use in fingerprint scanners," *Pattern Recognition*, vol. 36, no. 2, pp. 383–396, February 2003.

[7] Gian Luca Marcialis, Fabio Roli, and Alessandra Tidu, "Analysis of fingerprint pores for vitality detection," in *International Conference on Pattern Recognition ICPR*, Istanbul, August 2010, pp. 1289–1292.

[8] Marcela Espinoza and Christophe Champod, "Using the number of pores on fingerprint images to detect spoofing attacks," in *International Conference on Hand-Based Biometrics ICHB*, Hong Kong, November 2011, pp. 1–5.

[9] A. Abhyankar and Stephanie Schuckers, "Fingerprint liveness detection using local ridge frequencies and multiresolution texture analysis techniques," in *IEEE International Conference on Image Processing (ICIP)*, October 2006, pp. 321–324.

[10] Elham Tabassi, Charles L. Wilson, and Craig I. Watson, "Fingerprint image quality (NFIQ)," Tech. Rep. NISTIR-7151, National Institute of Standards and Technology NIST, August 2004.

[11] Qijun Zhao, David Zhang, Lei Zhang, and Nan Luo, "Adaptive fingerprint pore modeling and extraction," *Pattern Recognition*, vol. 43, no. 8, pp. 2833 – 2844, 2010.

[12] Mark Hall, Eibe Frank, Geoffrey Holmes, Berhhard Pfahringer, Peter Reutemann, and Ian H. Witten, "The WEKA data mining software: an update," *ACM SIGKDD Explorations Newsletter*, vol. 11, no. 1, pp. 10–18, June 2009.

[13] Diego Raphael Amancio, Cesar Henrique Comin, Dalcimar Casanova, Gonzalo Travieso, Odemir Martinez Bruno, Francisco Aparecido Rodrigues, and Luciano da Fontoura Costa, "A systematic comparison of supervised classifiers," *PLoS One*, vol. 9, no. 4, pp. 1–14, April 2014.

[14] Luca Ghiani, David Yambay, Valerio Mura, Simona Tocco, Gian Luca Marcialis, Fabio Roli, and Stephanie Schuckers, "Livdet 2013: Fingerprint liveness detection competition 2013," in *International Conference on Biometrics (ICB)*, Madrid, Spain, June 2013.