ON THE ROBUSTNESS OF FINGERPRINT LIVENESS DETECTION ALGORITHMS AGAINST NEW MATERIALS USED FOR SPOOFING

Emanuela Marasco and Carlo Sansone
Dipartimento di Informatica e Sistemistica, University of Napoli Federico II, via Claudio 21 80125, Napoli, Italy

Keywords: Biometrics, Liveness detection.

Abstract: Fingerprint biometric systems may be deceived by attacks at sensor level that use fake fingers. A secure fingerprint scanner is required to possess the ability to determine if the image comes from a living individual or not. Recently, several liveness detection approaches have been proposed to address this problem. At present performances of the existing software-based solutions have been assessed with different sensors and using small data sets. Moreover, it is assumed that fake fingerprints are produced by adopting the same materials used for training the system. This paper looks at the cases where the test spoof finger is made by employing a material new for the fingerprint sensor. We propose an experimental comparison among the current fingerprint liveness detection approaches accomplished by adopting materials for training different than those used for testing. Experiments have been performed by using standard databases taken from the LivDet09 Competition.

1 INTRODUCTION

Fingerprint biometric recognition systems are considered the most efficient and widely adopted technique for secure authentication. Recent research has shown that it is not difficult to deceive a fingerprint recognition system using fake fingertips (Maltoni et al., 2003). Common spoofing techniques realize fake fingers employing inexpensive materials, e.g., *gelatin*, *clay*, *play-doh* and *silicon*. Then, it is required that a fingerprint sensor has the ability to determine if an input biometric signal comes from a live person or not (Schuckers, 2002). Recently, in order to address this problem, several approaches for liveness detection have been proposed. The state-of-the-art concerning these solutions consists of two main categories: software-based and hardware-based approaches. The methods belonging to the first category make use of additional hardware to measure temperature and heartbeat characteristics resulting in an expensive implementation. The software-based approaches exploit intrinsic vitality properties that are extracted directly from the fingerprint images acquired by the sensor resulting in a cheaper device. At present, the main software-based vitality detection approaches analyze the skin perspiration through the pores, the elastic properties of the skin and the morphology of the fingerprint (Coli et al., 2007) (Schuckers et al., 2006). The main limitation of most of those methods resides in the device-dependence, in fact the resolution of the fingerprint image varies across different technologies, and subsequently, the extracted characteristics of vitality are not universal enough (Coli et al., 2008). Moreover, in the previous works the systems are tested under the assumption that fake fingerprints are realized by employing one of the materials used for training. This assumption makes optimistic the results showed by the authors. It is worth noting that this aspect is a challenging problem in fingerprint liveness detection, since nowadays materials used for fraudulent spoof attacks are going to become very sophisticated.

So, in order to assess the influence of this aspect on the performance of liveness detection algorithms, in this paper we analyze the cases where spoof fingers are realized with materials that are new for the fingerprint scanner. We propose an experimental comparison among the current fingerprint liveness detection approaches accomplished by adopting materials for training different than those adopting for testing. Experiments were carried out by using standard databases taken from Liveness Detection Competition 2009 (LivDet09). The paper is organized as follows. Section 2 presents the main steps of each methods we have considered for our study. Section 3 describes the datasets we employed to carry out our

Marasco E. and Sansone C..
ON THE ROBUSTNESS OF FINGERPRINT LIVENESS DETECTION ALGORITHMS AGAINST NEW MATERIALS USED FOR SPOOFING.
DOI: 10.5220/0003270505530558
In *Proceedings of the International Conference on Bio-inspired Systems and Signal Processing* (MPBS-2011), pages 553-558
ISBN: 978-989-8425-35-5

experiments. Section 4 reports results about the robustness of the approaches to new materials used for spoofing. Section 5 draws our conclusions.

2 A TAXONOMY OF THE EXISTING FINGERPRINT LIVENESS DETECTION APPROACHES

Liveness methods belong to two main categories. The first one exploits characteristics as the temperature of the finger, the electrical conductivity of the skin and the pulse oximetry. They can be detected by using additional hardware in conjunction with the biometric sensor. This makes costly the device. The second category performs an extra software process of the biometric sample in order to detect the vitality information directly from the fingerprint images. In this paper, we focus on this second category of approaches, known as *software-based* (Schuckers et al., 2006). The existing software-based solutions may include *dynamic* or *static* methods.

2.1 Dynamic Approaches

Dynamic features derive from the analysis of multiple frames of the same fingers. A typical *dynamic* property of a live finger is the perspiration phenomenon that starts from the pores and evolves in time across the ridges. This distinctive spatial moisture pattern can be detected by observing multiple fingerprint images acquired in two appropriate different times. An interesting method based on perspiration changes in live fingers was presented by Abhyankar and Schuckers in (Abhyankar and Schuckers, 2009). In this method, the changing perspiration pattern is isolated through a wavelet analysis of the entire fingerprint image. For an image processing algorithm, to quantify the sweating pattern is challenging. Since this pattern is a physiological phenomenon, it is variable across subjects. Further, it presents a certain sensitivity to the environment, the pressure of the finger, the time interval and the initial moisture content of the skin (Derakhshani et al., 2003). Its effectiveness requires an efficient extraction of the evolving pattern from images.

2.2 Static Approaches

Static features can be extracted from a single fingerprint impression or as difference between different impressions. Generally, static measurements may be altered by factors as the pressure of the finger on the scanner surface.

According to the taxonomy proposed in (Coli et al., 2008), features extracted by different impressions can be skin deformation-based or morphology-based, while features extracted by a single impression can be perspiration-based or morphology-based. Elastic deformations due to the contact, the pressure and the rotation of the fingertip on the plane surface of the sensor, are more evident in fake fingerprints made using artificial materials than in live fingerprints. Deformation-based methods detect liveness by comparing these distortions through static features (Chen et al., 2005). The elastic behavior of live and fake fingers has been analyzed by extracting a specific set of minutiae points. The second type of static features using multiple impressions relies on a morphologic investigation which exploits the thickness of the ridges that is modified after producing the fingerprint replica.

Methods which exploit intrinsic properties of a single impression study the skin perspiration phenomenon. The vitality indication can be found by using Wavelet Transform and Fast Fourier Transform (Coli et al., 2007). Wavelet analysis is able to capture the non-regular shape typical of the ridges in an image acquired from a live finger. Images taken from artificial fingers show a more regular shape. Fourier Transform is employed to study the regular periodicity of pores on the ridges in live fingerprints. Such a regularity is not present in signals corresponding to spoof fingerprints.

Liveness detection methods which search for morphological characteristics of fingerprint images, are significantly efficient when based on the surface coarseness.

3 EXISTING METHODS EXPLOITED FOR OUR STUDY

In this section, we describe the methods employed for our comparative analysis. Firstly, we describe the three static *morphology-based* methods which exploit a single fingerprint image for vitality information extraction.

Moon et *al.* (Moon et al., 2005) proposed a method based on analyzing the surface coarseness in high resolution (1000dpi)fingertip images. It has been observed that the surface of a fake finger is much coarser than that one of the human skin. The coarseness feature is measured by computing the standard deviation of the residual noise of the fingerprint image. This algorithm is fast and convenient but it works well only in presence of an high resolution sensor (1000dpi, while the common commercial sensors present a resolution of about 500dpi) (Coli et al., 2007).

An interesting texture-based approach using a single fingerprint image was proposed by (Nikam and Agarwal, 2009). They analyzed liveness of a fingerprint image by using the gray level associated to the fingerprint pixels. The gray level distribution in a fingerprint image changes when the physical structure changes. Then, real and fake fingerprint images are expected to present different textural properties. In fact, due to the presence of sweat pores and the perspiration phenomenon, authentic fingerprints exhibit non-uniformity of gray levels along ridges, while due to the characteristics of artificial material surface, such as gelatin or silicon, spoof fingers show high uniformity of gray levels along ridges.

In (Abhyankar and Schuckers, 2006), Abhyankar and Schuckers proposed an approach based on multiresolution texture analysis and the inter-ridge frequency analysis of fingerprint images. They used different texture features to quantify how the gray level distribution in a fingerprint image changes when the physical structure changes. First order statistics model the gray level distribution of the single pixels by using histograms, while second order statistics refer to the joint gray level function between pair of pixels. Two secondary features were used, Cluster Shade and Cluster Prominence, based on the cooccurrence matrix. All these features have been combined with features derived from fingerprint local-ridge frequency analysis.

Secondly, we describe a static method which combines characteristics describing the morphology of the fingerprint and characteristics describing the perspiration phenomenon (Marasco and Sansone, 2010). The approach relies on static features derived from the visual texture of the fingerprint image. In particular, first order statistics and residual noise standard deviation are exploited as *morphology-based* features, while ratios between gray level values and individual pore spacing are are exploited as *perspiration-based* features.

The *standard deviation of the residual noise* measures the coarseness of the fingerprint image. Materials used to make fake fingers such as *silicon* or *gelatin* consist of organic molecules which tend to agglomerate, thus the surface of a fake finger is generally coarser than a live one (Moon et al., 2005). The residual noise indicates the difference between the original and the de-noised image, in which the noise components are due to the coarseness of the fake finger surface (Abhyankar and Schuckers, 2006). In fact, according to the approach proposed by Moon et *al.*(Moon et al., 2005), the surface coarseness has been treated as a kind of gaussian white noise added to the image.

First order statistics measure the likelihood of observing a gray value at a randomly-chosen location in the image. The gray level associated to each pixel is exploited to determine a vitality degree of the fingerprint image. They can be computed from the histogram of pixel intensities in the image. The goal is to quantify the variations of the gray level distribution when the physical structure changes. The distinction between a fake and a live finger is based on the difference of these statistics.

Individual pore spacing characteristics are extracted after analyzing the occurrence of pores that causes a gray value variability in the fingerprint image. In (Marasco and Sansone, 2010), according to the algorithm proposed in (Derakhshani et al., 2003), the 2-dimensional fingerprint image was mapped to 1-dimensional signal which represents the gray-level values along the ridges. The gray-level variations in the signal correspond to variations in moisture due to the pores and the presence of perspiration. By transforming the signal in the Fourier domain lets to measure this static variability in gray-level along the ridges. In particular, the focus is on frequencies corresponding to the spacial frequencies of the pores. The FFT was computed and the total energy associated to the spacial frequency of the pores was obtained as static feature. The coefficients of interest are from 11 to 33, since these values correspond to the spacial frequencies (0.4 - 1.2 mm) of pores.

Intensity-based features are based on the assumption that, the spoof and cadaver fingerprints images are distributed in the dark (<150), among the 256 different possible intensities (Tan and Schuckers, 2005). They have computed two particular features: *i) gray level 1 ratio*, corresponding to the ratio between the number of pixels having a gray level belonging to the range (150, 253) and the number of pixels having a gray level belonging to the range (1, 149); *ii) gray level 2 ratio*, corresponding to the ratio between the number of pixels having a gray level belonging to the range (246, 256) and the number of pixels having a gray level belonging to the range (1, 245). Moreover, they have analyzed the uniformity of gray levels along ridge lines and the contrast between valleys and ridges. Real fingerprints exhibit non-uniformity of gray levels and high ridge/valley contrast values. Then, the general variation in gray-level values of in a spoof fingerprint is less than a live one. To capture this information the gradient of the gray-level matrix of the image have been computed, too.

Finally, we describe a *perspiration-based* method using both static and dynamic features. Tan and Schuckers (Tan and Schuckers, 2005) have experimented the joint contribution of dynamic and static features. They studied the perspiration phenomenon from the intensity distribution perspective, by observing that live fingers present a distinctive contrast between white (>250, ASCII gray level range 0:255) and dark (<20) gray level, while spoof images have very small contrast difference. The decision rules to perform liveness classification is generated after considering static and dynamic features. The static features used in this work are based on the following parameters:

$$S1 = \frac{sum(151:254)}{sum(0:150)} \quad (1)$$

and

$$S2 = sum(151:254) \quad (2)$$

The dynamic features are based on the difference in the histogram distribution between zero and fifth second that is larger in live finger compared to spoof subjects. In the live fingers, perspiration makes dry (white) regions between the pores moister (darker) in time. This approach may present some limitations in cases of fingers too dry or too moist and other perspiration disorders.

4 EXPERIMENTAL RESULTS

Our experimental phase was carried out by using two databases composed by live and spoof fingerprint images. Each one refers to a different sensor (i.e., *CrossMatch* and *Identix*). They have been taken from the Liveness Detection Competition 2009 (Marcialis et al., 2009) and each one of them is composed by two subsets, one for training and the other one for testing the algorithm. Details about the data collection are shown in the tables 1 and 2. In both the cases, the subjects using for training are different with respect to those considered for testing. Table 3 reports details about the sensors used for LivDet 2009 Competition. Note that the LivDet dataset has also another database (namely, *Biometrika*) whose spoof fingerprints have been produced by employing only one material (Silicon). So, for carrying out the current comparison, we have adopted only *CrossMatch* and *Identix* databases in which spoof fingerprints have been made by employing three different materials.

The classification performance evaluation was performed by adopting the same parameters used during the LivDet09, defined as follows:

- *Ferrlive*: rate of misclassified live fingerprints.

Table 1: Datasets of images used for training.

Database	Sub	Live	Fake	Frames
Identix	35	375	375	0 and 2 sec
CrossMatch	63	500	500	0 and 2 sec

Table 2: Datasets of images used for testing.

Database	Sub	Live	Fake	Frames
Identix	125	1125	1125	0 and 2 sec
CrossMatch	191	1500	1500	0 and 2 sec

Table 3: Fingerprint sensors used for LivDet 2009.

Scanners	Model No.	(dpi)	Size
Biometrika	FX2000	569	(312x372)
Identix	DFR2100	686	(720x720)
CrossMatch	Verifier 300 LC	500	(480x640)

- *Ferrfake*: rate of misclassified fake fingerprints.

In particular, performance is measured by using the value *e* averaged on the two database *CrossMatch* and *Identix*. The value *e* is computed as follows:

$$e = \frac{Ferrlive + Ferrfake}{2} \quad (3)$$

In our first experiments, each system was trained by using features extracted from fake samples made with all the materials available in each database. In particular, in both *Identix* and *CrossMatch* databases, the materials employed are *Gelatin*, *Silicon* and *Play-Doh*. Then, we have carried out a further evaluation, in order to study the robustness of the existing liveness detection approaches with respect to *unknown* materials used for producing fake fingers. In this experiment, each system was trained by using spoof fingerprints realized with all but one of the available materials, while the excluded material was used for testing. Table 4 reports the performance of the method proposed by Marasco and Sansone. In presence of high resolution images, taken from the *Identix* database, the testing performed using *Gelatin* and *Silicon*, when the training is performed by employing fake fingers made in Play-Doh, gives rise to a good spoofing recognition rate. Table 5 shows that the method proposed by Moon et *al.* wrongly classifies the majority of the fake fingerprints taken from *CrossMatch* database, while for a higher resolution factor, such a method presents a better behavior in presence of *unknown* materials using for spoofing. Table 6 and 7 show that the variation in fake materials does not significantly affect the performance of both Nikam-Agarwal and Abhyankar-Schuckers approaches, when the training set is only composed by samples made with *Gelatin*. On the contrary, as reported in Table 8, the performance of the

Table 4: Performance of the method proposed by Marasco and Sansone on CrossMatch and Identix databases.

	CrossMatch			Identix		
	Gelatin	*Play – Doh*	*Silicon*	*Gelatin*	*Play – Doh*	*Silicon*
Ferrlive	6.5%	5.7%	12.6%	3.8%	19.2%	9.7%
Ferrfake	25.9%	16.7%	10.0%	42.3%	5.5%	30.6%
e	16.2%	11.2%	11.3%	23.05%	12.35%	20.15%

Table 5: Performance of the method proposed by Moon et *al.* on CrossMatch and Identix databases.

	CrossMatch			Identix		
	Gelatin	*Play – Doh*	*Silicon*	*Gelatin*	*Play – Doh*	*Silicon*
Ferrlive	12.30%	15.00%	35.70%	45.20%	79.60%	40.80%
Ferrfake	63.10%	61.80%	47.30%	31.80%	4.20%	36.80%
e	37.70%	38.40%	41.50%	38.50%	41.90%	38.80%

Table 6: Performance of the method proposed by Nikam and Agarwal on CrossMatch and Identix databases.

	CrossMatch			Identix		
	Gelatin	*Play – Doh*	*Silicon*	*Gelatin*	*Play – Doh*	*Silicon*
Ferrlive	27.20%	43.70%	24.20%	23.50%	29.30%	20.00%
Ferrfake	22.00%	32.90%	31.60%	16.00%	28.80%	31.50%
e	24.60%	38.30%	27.90%	19.75%	29.05%	25.75%

Table 7: Performance of the method proposed by Abhyankar and Schuckers on CrossMatch and Identix databases.

	CrossMatch			Identix		
	Gelatin	*Play – Doh*	*Silicon*	*Gelatin*	*Play – Doh*	*Silicon*
Ferrlive	45.80%	29.80%	58.60%	65.50%	61.60%	37.90%
Ferrfake	12.20%	24.40%	17.00%	2.40%	46.40%	27.70%
e	29.00%	27.10%	37.80%	33.45%	54.00%	32.80%

Table 8: Performance of the method proposed by Tan and Schuckers on CrossMatch and Identix databases.

	CrossMatch			Identix		
	Gelatin	*Play – Doh*	*Silicon*	*Gelatin*	*Play – Doh*	*Silicon*
Ferrlive	38.60%	24.40%	54.80%	64.10%	36.00%	38.80%
Ferrfake	32.20%	39.20%	43.00%	28.70%	42.40%	13.20%
e	35.40%	31.80%	48.90%	46.40%	39.20%	26.00%

Table 9: Performance of the analyzed approaches in terms of the average error *e* on Identix and CrossMatch databases.

	Marasco-Sansone	Moon et *al.*	Nikam-Agarwal	Abhyankar-Schuckers	Tan-Schuckers
Gelatin	19.63%	38.10%	22.18%	31.23%	40.90%
Play-Doh	11.78%	40.15%	33.68%	40.55%	35.50%
Silicon	15.73%	40.15%	26.83%	35.30%	37.45%
Avg	**15.71%**	**39.47%**	**27.53%**	**35.79%**	**37.45%**
All materials	**12.45%**	**30.85%**	**24.53%**	**39.37%**	**29.20%**

Tan-Schuckers method seems quite dependent on the material as well as on the considered dataset.

As resumed in Table 9, when the material used to attack the system is not known during the training, most of the algorithms decrease in performance. This confirms our claim that the performance of liveness detection algorithms reported by the authors typically represents an overestimate of that obtainable in real scenarios. Among the considered methods, the one based on a single feature (Moon et al., 2005) is the most dependent on the use of *unknown* materials for testing. Also the dynamic method proposed in (Tan and Schuckers, 2005) had a significant decrement in performance when classifying fake fingerprints realized with materials different from those present in the training set. The other methods are instead more robust, and the one proposed by Marasco and Sansone, which is based on a combination of multiple features, exhibited the best average error *e* when the material used for testing is *unknown* at training time.

5 CONCLUSIONS

In this paper, we have analyzed the impact of a new material on the performance of the existing liveness detection algorithms. This analysis has been performed by considering three different sensors.

Our experiments showed that the performance of liveness detection approaches in which only few features are exploited, significantly decreases in presence of spoof attacks realized by employing materials different from those using for training the system. This weakness can be reduced by combining multiple vitality features. In particular, the more robust approach was given by the joint usage of morphology- and perspiration-based features.

ACKNOWLEDGEMENTS

The authors would like to thank Prof. Stephanie Schuckers of the Clarkson University (USA) for her useful talks, Prof. Luisa Verdoliva of the University of Naples (Italy), for supporting the wavelet analysis techniques and Prof. Davide Maltoni of the University of Bologna (Italy), for his valuable comments about the existing core detection algorithms.

REFERENCES

Abhyankar, A. and Schuckers, S. (2006). Fingerprint liveness detection using local ridge frequencies and multiresolution texture analysis techniques. *IEEE International Conference on Image Processing*, pages 321–324.

Abhyankar, A. and Schuckers, S. (2009). Integrating a wavelet based perspiration liveness check with fingerprint recognition. *Pattern Recognition*, 42:452–464.

Chen, Y., Jain, A., and Dass, S. (2005). Fingerprint deformation for spoof detection. *Biometric Symposium*.

Coli, P., Marcialis, G., and Roli, F. (2007). Vitality detection from fingerprint images: a critical survey. *Lecture Notes in Computer Science*, 4642:722–731.

Coli, P., Marcialis, G., and Roli, F. (2008). Fingerprint silicon replicas: static and dynamic features for vitality detection using an optical capture device. *International Journal of Image and Graphics (IJIG)*, 8(4):495–512.

Derakhshani, R., Schuckers, S., Hornak, L., and O'Gorman, L. (2003). Determination of vitality from non-invasive biomedical measurement for use in fingerprint scanners. *Pattern Recognition*, 36:383–396.

Maltoni, D., Maio, D., Jain, A., and Prabhakar, S. (2003). *Handbook of Fingerprint Recognition*. Springer.

Marasco, E. and Sansone, C. (2010). An anti-spoofing technique using multiple textural features in fingerprint scanners. *IEEE Workshop on Biometric Measurements and Systems for Security and Medical Applications (BioMs)*, pages 8–14.

Marcialis, G., Lewicke, A., Tan, B., Coli, P., Grimberg, D., Congiu, A., Tidu, A., Roli, F., and Schuckers, S. (2009). First international fingerprint liveness detection competition - livdet 2009. *Lecture Notes in Computer Science*, 5716:12–23.

Matsumoto, T., Matsumoto, H., Yamada, K., and Hoshino, S. (2002). Impact of artificial gummy fingers on fingerprint systems. *Optical Security and Counterfait Deterrence Techniques IV*, 4677:275–289.

Moon, Y. S., Chen, J. S., Chan, K. C., So., K., and Woo, K. S. (2005). Wavelet based fingerprint liveness detection. *Electronic Letters*, 41(20):1112–1113.

Nikam, S. B. and Agarwal, S. (2009). Curvelet-based fingerprint anti-spoofing. *Signal, Image and Video Processing*, 4(1):75–87.

Schuckers, S. (2002). Spoofing and anti-spoofing measures. *Information Security Technical Report*, 7(4):56–62.

Schuckers, S., Derakhshani, R., Parthasaradhi, S., and Hornak, L. (2006). Liveness detection in biometric devices.

Tan, B. and Schuckers, S. (2005). Liveness detection using an intensity based approach in fingerprint scanner. In *Proceedings of Biometrics Symposium (BSYM)*.