2011 International Conference on Information and Network Technology
IPCSIT vol.4 (2011) © (2011) IACSIT Press, Singapore

Spoof Attacks on Multimodal Biometric Systems

Zahid Akhtar [1,+], Sandeep Kale [2] and Nasir Alfarid [3]

[1] Dept. of Electrical and Electronic Engineering, University of Cagliari, Italy
[2] Dept. of Electronic Science, University of Pune, India
[3] Cognizant Technology Solutions, India

Abstract. Biometrics, referred as the science of recognizing an individual based on his or her physical or behavioral traits, has been widely employed as a security system in the awake of latest security issues. However, recent researches have shown that many biometric traits are vulnerable to spoof attacks. In addition, a latest results have questioned that, contrary to a common claim, multimodal systems can be cracked by spoofing *only one* trait. Those results were obtained using *simulated* spoof attacks, under the unrealistic assumption that the spoofed and genuine samples are identical, turned out to be the same outputs. We further investigate this significant security issue, focusing on behavior of fixed and trained score fusion rules, using *real* spoof attack samples under different spoof attack scenarios. Preliminary empirical results on real biometric systems made up of face, fingerprint and iris with twelve score fusion rules confirm that multimodal biometric systems are not intrinsically robust against spoof attacks as believed so far. In particular, most widely used fixed rules can be less robust, even if the quality of fake biometric trait is low. The false acceptance rate increases substantially under spoof attacks which means that an attacker might wrongly get authenticated by spoofing a subset of traits. In all considered spoofing scenarios, we also found that trained rules are more accurate, flexible and robust against spoof attacks as compare to fixed one.

Keywords: Biometrics, Multimodal biometric system, Score fusion rules, Spoof attacks.

1. Introduction

Biometrics using fingerprint, face, voice and iris etc. based recognition systems to identify an individual, has been accepted as a legitimate technology. Each biometric trait should pose attributes like uniqueness, and hard to circumvent [1]. Sadly, recent researches have shown that an attacker can lift and replicate the biometric traits, which later can be used to attack on biometric systems [2-4]. As a result, multimodal biometric systems have been proposed to increase the recognition accuracy as well as security against attacks as compared to the unimodal biometric systems that make them up. Several empirical evidences have shown that they are effective to accuracy. It is claimed that multimodal systems are more robust against spoof attacks, since evading several systems is more difficult than evading just *one* [1]. This claim implies that to evade a multimodal system it is necessary to evade *all* fused individual systems *simultaneously*. However, there is no experimental evidence to support this assumption, with the exception of [5], where some evidence was provided that a multimodal biometric system can be fooled by spoofing one of the individual matchers. However, the experiment in [5] was carried out by simulating spoof attacks under the assumption that spoofed and genuine traits are identical which produces same output matching scores, which is not true in real world [6-7]. Hence, it is of great interest to investigate the robustness of multimodal biometric systems against real spoof attacks under more realistic scenarios where a attacker is not able to fabricate a perfect replica of biometric trait. In this work we further contribute to this goal using real spoof attack samples unlike [5]. We analyze the security of multimodal biometric systems by focusing on spoof attacks at the sensor level, which are so far the ones raising the most of interest in the biometric community [8]. They are carried by submitting

[+] Corresponding author. Tel.: + 393294913022 ; fax: + 390706755782.
E-mail address: z.momin@diee.unica.it

to the system a biometric replica, like a fake fingerprint [2].

In this paper, we further empirically evaluate the probability of evading a multimodal system by real spoofing of the individual biometrics, by considering three multimodal biometric systems made up of face, fingerprint and iris matchers. We specially analyze the behavior of four fixed and eight trained score fusion rules with spoofing scenarios when: i) only the fingerprint trait has been spoofed; ii) only face has been spoofed; iii) only iris has been spoofed; iv) both, fingerprint and face, have been spoofed; v) both, fingerprint and iris have been spoofed; vi) both, face and iris has been spoofed. We compare the probability of an impostor being authenticated as genuine user under these scenarios.

Our results show that multimodal biometric systems may be more vulnerable to spoof attacks when the most accurate biometric trait is spoofed. It is known that trained rules are more flexible and in principle more accurate than fixed ones [1]. We provide empirical evidence that trained rules can be more robust, as well, against spoof attacks as compare to fixed rules.

2. State-of-the-art and Goals

Due to great acceptance of biometric security systems, their issues about resilience against attacks are also raising. Several researchers are studying the susceptibility of biometric systems, the potential attack mechanisms with their counteractions. As mentioned in [9], a generic biometric system has eight vulnerable points that can be exploited by an adversary to get unauthorized access. But, faking biometric input to the system, known as spoof attack, is a raising concern. Spoof attack is related to the sensor, and is also called as ”direct attack”. The stolen or lifted biometric trait can be utilized to reproduce synthetic samples which can eventually be used to attack biometric systems. For instance, 60% fake fingerprints reproduced using gum and gelatin were accepted as legitimate user by the system in [2]. One possible counteraction suggested in literature is liveness detection (vitality testing) [4], but no method is fully matured yet.

In general, multimodal biometric systems are considered as intrinsically robust against spoof attacks under the assumption that to evade them adversary needs to spoof *all* biometric traits *simultaneously*. However, there is no theoretical or empirical support for this claim in literature. Only in [5] multimodal system made up of face and fingerprint matcher were studied and showed, using *simulated* spoof attack, that it is possible to crack multimodal system by spoofing only one of the matchers, under unrealistic hypothesis that genuine and spoofed traits are identical. Such outcome are quite interesting and inspire to investigate the issue of robustness and performance of multimodal biometric systems under spoof attacks in real scenarios with *real* spoof attacks, which is goal of this paper. In this work we address the following three issues.

1) Can multimodal biometric systems be evaded by spoofing *only one* trait?. Result presented in [5] with simulated spoof attack using two fusion rules showed that it is possible. We further analyze the issue on three multimodal systems with *real* spoof attacks and twelve score fusion rules (four fixed and eight trained rules).

2) What is the behavior of several trained and fixed score fusion rules, when a multimodal biometric system is under spoof attacks, with scenarios mentioned in the introduction section?. We started with the fact that trained score fusion rules are considered as flexible and accurate [1] and found that they can be also more robust against spoof attack than trained rules.

3) To achieve a best trade-off between accuracy and robustness against spoof attacks, which is the most effective score fusion scheme, a system designer should select?. Selection of the fusion rule depends on factors such as accuracy requirements, availability of training data, computational cost, and the validity of simplifying assumptions. We consider the problem of designing a system when i) the number of training samples are limited, then preferable fusion scheme is the fixed rules, in terms of computational cost and complexity.; ii) the number of training samples are large enough, then trained rule is better choice to approximate the operating parameters of the system, to optimize both verification accuracy and robustness.

3. Experimental Setup

Since, no biometric data set composed of real spoof attack samples are available publicly. Hence, we used PolyU HRF DBII fingerprint [10], Essex face [11] and IIT Delhi iris [12] data sets and

produced the real spoof attack samples. The data set used in the experiments contains face, fingerprint and iris images of 100 individuals, with 10 genuine and 10 fake (spoof attacks) samples per individual for each modality.

Spoofed face and iris samples were replicated with a ”photo attack” method. We put the photo of each individual in front of capture device, displayed on laptop screen. While, spoofed fingerprint samples were created by the same method carried out in [13] to create fake palmprint images. For each individual, we created 10 spoofed face, fingerprint and iris samples.

The face, iris and fingerprint recognition systems used for the experiments were implemented using the Principle Component Analysis (PCA) [14], the Iris Code [15] and the minutiae-based [16] methods, respectively. All the scores were normalized using the hyperbolic tangent method [1]. Using the face, fingerprint and iris unimodal systems, we created three multimodal biometric systems by pairing in all possible ways the face system with the fingerprint and the iris systems. The resulted three multimodal systems are named as Face-Fingerprint, Face-Iris and Fingerprint-Iris multimodal biometric systems.

To study effect of spoof attacks on score fusion rules, we estimated the decision thresholds and parameters of trained fusion rules on whole data set to operate the systems under optimal configuration. In particular, the decision thresholds and parameters of trained rules were evaluated on original dataset (without spoof attacks) while the performance of the systems under attacks was evaluated by replacing impostor score with respective spoofed scores.

3.1. Score Fusion Rules

We evaluated four fixed and eight trained score fusion rules. Let s_1, s_2, s_3 and s be the scores of face, fingerprint, iris matchers and the fused score, respectively.

A) Fixed Rules:

1) **Sum:** The fused score for the set of N matchers by sum rule is computed as follows: $s = \sum_{i=1}^{N} s_i$

2) **Product:** The product rule computes the fused score for the set of N matchers as: $s = \prod_{i=1}^{N} s_i$

3) **Min:** The fused score using min rule for N individual matchers is calculated as: $s = \min_{i=1}^{N} s_i$

4) **Max:** The fused score using max rule for N matchers is estimated as: $s = \max_{i=1}^{N} s_i$

B) Trained Rules:

1) **Weighted Sum:** The weights (w) for the weighted sum rule can be computed using linear discriminant analysis (LDA)[17]. The aim of using LDA based fusion rule is to obtain fused scores with minimum within-class and maximum between-class variations. The fused scores are computed as: $s = \sum_{i=1}^{N} w_i s_i$

2) **Exponential Sum:** The fused score for N matchers is obtained as follows [18]: $s = \sum_{i=1}^{N} w_i \exp(s_i)$

3) **Tan-hyperbolic Sum:** The fused score is computed as follows [18]: $s = \sum_{i=1}^{N} w_i \tanh(s_i)$

The weights were for exponential sum and tan-hyperbolic sum were computed as in weighted sum rule.

4) **Perceptron:** The perceptron-based fusion rule for N matchers can be implemented as follows [19]: $s = 1 \Big/ 1 + \exp\left[-(w_0 + \sum_{i=1}^{N} w_i s_i)\right]$. The weights were computed by a gradient descent algorithm with a least-squares loss function.

5) **Weighted Product:** The fused score of the set of N matchers using weighted product rule which is also known as logarithm opinion pool, is computed as follows [18]: $s = \prod_{i=1}^{N} s_i^{w_i}$. The weights $w_i \in [0,1]$ has been computed by maximizing the system performance on the chosen operational points. This rule accounts the varying discrimination ability and reliability of each matcher.

6) **Exponential Product:** The fused score of the set of N matchers using exponential product rule is obtained as follows [20]: $s = \prod_{i=1}^{N} w_i \exp(s_i)$. The weights were calculated as in weighted product rule.

7) **Tan-hyperbolic Product:** The fused score of the set of N matchers using tan-hyperbolic product rule is obtained as follows [21]: $s = \prod_{i=1}^{N} w_i \tanh(s_i)$. The weights were calculated as in weighted product rule.

8) **Likelihood Ratio Rule (LLR):** This rule computes the fused score as follows:

$$s = \prod_{i=1}^{N} p(s_i \mid G_i) \Big/ \prod_{i=1}^{N} p(s_i \mid I_i)$$

where $p(\cdot|G)$ and $p(\cdot|I)$ are the matching score's probability density function (PDF) of genuine and impostor users, respectively. LLR is referred as the optimal fusion rule, when all the PDFs are estimated accurately [22]. We used Gaussian to model the genuine and impostor score distributions.

The false acceptance rate (FAR) is the percentage of impostor being accepted as genuine users. To investigate issues 1, 2 and 3 of section 2, we evaluated the increase of FAR due to spoof attacks at 0%, 0.01% and 0.1% FAR operating points, resulting the lowest threshold values that produce FAR on training data set equal to operational points, respectively.

4. Experimental Results

We report the results obtained on Face-Fingerprint and Fingerprint-Iris multimodal systems in Tables I, and II, respectively, using sum, min, weighted sum (with LDA), exponential product and LLR score fusion rules. The results obtained on Face-Iris system and with product and max rules were qualitatively very similar to sum rule, the results with perceptron and weighted product rules were similar to weighted sum rule while the results with exponential sum, tan-hyperbolic sum, tan-hyperbolic product rules were qualitatively very similar to exponential product rule, hence are not reported due to lack of space.

Consider issue 1 of section 2, our results on all multimodal systems with *real* spoof attacks clearly show that the answer is "yes", which is further support to the results obtained with *simulation* in [5]. For instance, from the Table I, it can been seen that even at 0.l% FAR operating point the FAR under attack attained values up to 63.91%, when *only one* trait was spoofed.

Consider now the issue 2 of the section 2. We observed that the four considered fixed score fusion rules perform in the same pattern. The increase of FAR under spoof attacks was remarkably high, when the best accurate matcher is spoofed, even as the quality of spoof traits were low. The less accurate matcher spoofing rather leads to less increase of FAR. For example, Fingerprint-Iris multimodal biometric system using sum rule, in Table II, when the most accurate matcher (iris here), is spoofed the FAR attained under attack is 48.74% while in case of fingerprint (less accurate in this case) matcher spoofing, the FAR under attack is 6.03%. Similar performance can be observed in all systems. As, the decision thresholds were computed on whole data set, this means when the decision thresholds are computed on training data set, even a small difference in estimation of operating point due to practically observed problem of concept drift will degrade the performance of multimodal biometric systems very much against spoof attacks.

With regard to trained rules, we observed the same phenomenon of spoofing accurate matcher leads to high increase of FAR as compare to spoofing less accurate matcher. In general, trained rules are considered as more accurate than fixed score fusion rules. Our results show that they are not only more accurate, flexible but also more robust as compare to fixed one against spoof attacks, in all considered different spoofing attack scenarios, because of their peer ability to tune and contribute more in training procedure. For instance, Face-Fingerprint multimodal biometric system, Table I, at 0.01% the FARs under attacks when: only the Fingerprint is spoofed; both, Fingerprint and face are spoofed, using min rule (fixed rule) are 51.00% and 60.80% while using LLR rule (trained rule) are 3.51% and 40.00%, respectively.

Operating Point	**Sum**			**Min**			**Weighted Sum**			**Exp. Product**			**LLR**		
	Fing. Sp.	Face Sp.	Both Sp.	Fing. Sp.	Face Sp.	Both Sp.	Fing. Sp.	Face Sp.	Both Sp.	Fing. Sp.	Face Sp.	Both Sp.	Fing. Sp.	Face Sp.	Both Sp.
0% FAR	41.39	1.69	53.60	40.10	1.44	52.04	39.87	1.00	48.11	7.24	1.00	40.40	2.79	1.01	33.48
0.01% FAR	52.14	3.09	63.67	51.00	3.00	60.80	50.75	2.30	57.89	9.87	2.27	47.23	3.51	2.23	40.00
0.1% FAR	63.91	5.05	76.50	62.46	4.80	71.31	59.84	3.66	63.93	13.82	3.43	59.01	4.59	2.96	49.99

Table I: FAR(%) of the Face-Fingerprint system, with the sum, min, weighted sum, exponential product and LLR rules, when either the fingerprint, the face or both the fingerprint and face are spoofed, at three operating points.

Operating	**Sum**			**Min**			**Weighted Sum**			**Exp. Product**			**LLR**		
	Iris	Fing.	Both	Iris	Fing.	Both	Iris	Fing	Both	Iris	Fing.	Both	Iris	Fing.	Both

Point	Sp.	Sp.	Sp.	Sp.	Sp.	Sp.	Sp.	Sp.	Sp.	Sp.	Sp.	Sp.	Sp.	Sp.	Sp.
0% FAR	34.12	1.89	51.32	33.4	1.88	49.16	12.58	1.01	40.06	10.67	1.00	39.0	32.89	1.08	47.23
0.01% FAR	40.91	3.67	58.20	39.3	3.17	56.90	21.85	2.78	49.83	18.11	2.59	46.4	38.02	2.40	55.21
0.1% FAR	48.74	6.03	69.90	47.6	5.55	65.56	37.81	3.90	61.35	34.08	2.86	58.5	45.51	3.71	61.90

Table II: FAR(%) of the Fingerprint-Iris system, with the sum, min, weighted sum, exponential product and LLR rules, when either the iris, the fingerprint, or both the iris and the fingerprint are spoofed, at three operating points.

Consider lastly the issue 3 of section 2. When less number of training samples are available then fixed rules should be preferred [1]. In spite of the fact that our results showed that they are less robust, we argue that their robustness against spoof attacks can be increased by coupling them with liveness detection method in order to obtain the required performance of the system with less computational cost. Trained rules strongly depend on the size of training data [22]. As, the reported results showed that trained rules outperformed the fixed rules under spoof attacks. Their resilience can be improved further by two way: first, using liveness detection method, second, training them the methods like in [5].

5. Conclusion

Our empirical investigation on *real* spoof attack samples verify that multimodal biometric systems are not intrinsically robust against spoof attacks contrary to the common belief, providing the further confirmation to the results obtained with *simulation* in [5]. They can be cracked by spoofing *only one* biometric trait. In particular, the most used fixed score fusion rules determined to be very vulnerable to spoof attacks while the trained score fusion rules are more resilient in all different spoof attack scenarios. Spoofing most accurate matcher creates serious security breaches.

A possible way to improve the security against spoof attacks is to devise *ad-hoc* score fusion rules or secure training algorithms, which also undertake possibility of attacks. Two rules were presented in [5], even though the parameters, probability of trait is attempted to spoof and that attempt succeeds, could be relatively difficult to tune in practice as pointed out in [5].

6. Acknowledgements

The authors thanks Ajay Kumar and Lei Zhang of Department of Computing, The Hong Kong Polytechnic University for providing IITD Delhi Iris and PolyU HRF Fingerprint Databases, respectively.

7. References

[1] A. Ross, K. Nandakumar, A.K. Jain. *Handbook of Multibiometrics*, Springer, 2006.

[2] T. Matsumoto, H. Matsumoto, K. Yamada, S. Hoshino. Impact of Artificial Gummy Fingers on Fingerprint Systems. In: *Proc. of SPIE on Optical Security and Counterfeit Deterrence Tech. IV*. 2007, vol. 4677, pp. 275-289.

[3] X. He, Y. Lu and P. Shi. A Fake Iris Detection Method Based on FFT and Quality Assessment. *Proc. Chinese Conf. on Pattern Recognition*. 2008, pp. 316-319.

[4] Y. Kim, J. Na, S. Yoon, Juneho Yi. Masked Fake Face Detection using Radiance Measurements. *J. Opt. Soc. Am. A*. 2009, 26 (4): 760-766.

[5] R.N. Rodrigues, L.L. Ling, V. Govindaraju. Robustness of Multimodal Biometric Methods against Spoof Attacks. *J. of Visual Languages and Computing*. 2009, 20 (3): 169-179.

[6] J. Galbally, R. Cappelli, A. Lumini, Guillermo G., D. Maltoni, J. Fierrez, Javier O., D. Maio. An evaluation of direct attacks using fake fingers generated from ISO templates. *Patt. Rec. Letters*. 2010, 31 (8): pp. 725-732.

[7] Girija Chetty, W. Michael. Multi-Level Liveness Verification for Face-Voice Biometric Authentication. In: *Proc. of Biometrics Symposium*. 2006, pp. 1-6.

[8] A.K. Jain, K. N., A. Nagar. Biometric template security. *EURASIP J. on Adv. in Sig. Proc.*. 2008: 1-17.

[9] N. Ratha, J.Connell, R. Bolle. An analysis of minutiae matching strength. In: *Proc. of AVBPA*. 2001, pp. 223-228.

[10] PolyU HRF Database, *http://www.comp.polyu.edu.hk/biometrics/HRF/HRF.htm.*

[11] University of Essex Face Database, *http://dces.essex.ac.uk/mv/allfaces/index.html.*

[12] IIT Delhi Iris Database version 1.0, *http://web.iitd.ac.in/biometrics/DatabaseIris.htm.*

[13] David Zhang, Vivek Kanhangad, Nan Luo, Ajay Kumar. Robust Palmprint Verification Using 2D and 3D Features. *Pattern Recognition*. 2010, 43 (1): 358-368.

[14] K. Kim. Face Recognition using Principal Component Analysis. *Dept. of Comp. Sci., Univ. of Maryland*. 2000.

[15] J. Daugman. How iris recognition works. *IEEE Trans. Circuits Syst. V. Technol.*. 2004, 14 (1): 21-30.

[16] D. Maltoni, D. Maio, A.K. Jain, and S. Prabhakar. *Handbook of Fingerprint Recognition*. Springer, 2003.

[17] Duda R., Hart P., Stork D. *Pattern Classification*. John Wiley Inc., 2001.

[18] A. Kumar, Vivek K., D. Zhang. A New Framework for Adaptive Multimodal Biometrics Management. *IEEE Trans. On Information Forensics and Security*. 2010, 5 (1): 92-102.

[19] A.K. Jain, S. Prabhakar, S. Chen. Combining Multiple Matchers for a High Security Fingerprint Verification System. *Pattern Recognition Letters*. 1999, 20 (11-13): 1371-1379.

[20] Kyong I. Chang Kevin W. Bowyer, Patrick J. Flynn, Xin Chen. Multibiometrics Using Facial Appearance, Shape and Temperature. In: *Proc. of IEEE Int. Conf. on Automatic Face and Gesture Recognition*. 2004, pp. 43-48.

[21] Kar-Ann Toh, Wei-Yun Yau. Combination of Hyperbolic Functions for Multimodal Biometrics Data Fusion. *IEEE Trans.on Sys., Man, and Cybernetics*. 2004, 34 (2): 1196-1209.

[22] Nandakumar, K., Chen, Y., Dass, S.C., Jain, A.K.. Likelihood Ratio Based Biometric Score Fusion. *IEEE Trans. on PAMI*. 2008, 30 (2): 342-347.