ON THE VULNERABILITY OF FACE RECOGNITION SYSTEMS TO SPOOFING MASK ATTACKS

Neslihan Kose, Jean-Luc Dugelay

Multimedia Department, EURECOM, Sophia-Antipolis, France
{neslihan.kose, jean-luc.dugelay}@eurecom.fr

ABSTRACT

There are several types of spoofing attacks to face recognition systems such as photograph, video or mask attacks. To the best of our knowledge, the impact of mask spoofing on face recognition has not been analyzed yet. The reason for this delay is mainly due to the unavailability of public mask attacks databases. In this study, we use a 2D+3D mask database which was prepared for a research project in which the authors are all involved. This paper provides new results by demonstrating the impact of mask attacks on 2D, 2.5D and 3D face recognition systems. The results show that face recognition systems are vulnerable to mask attacks, thus countermeasures have to be developed to reduce the impact of mask attacks on face recognition. The results also show that 2D texture analysis provides more information than 3D face shape analysis in order to develop a countermeasure against high-quality mask attacks.

Index Terms— spoofing; face mask; face recognition

1. INTRODUCTION

In a spoofing attempt, a person tries to masquerade as another person and thereby, tries to gain access to the system. Based on the observations that 2D face recognition systems are vulnerable to spoofing attacks, researchers started to work on countermeasures to reduce the impact of spoofing attacks on face recognition performances. There have been studies on countermeasures to detect photograph and video spoofing, which are 2D face attacks [1 - 3].

Mask attacks to face recognition systems, which are 3D face attacks, is a considerably new subject. To the best our knowledge, the impact of mask attacks on face recognition has not been analyzed yet. The main reason for this delay is due to the unavailability of public mask attacks databases. In this paper, it is the first time that the impact of mask spoofing is analyzed on face recognition using the mask database which was prepared within the context of the European Union (EU) research project TABULA RASA [4].

The preparation of a mask attacks database is much more difficult and expensive than the preparation of photo or video attacks databases. Initially, to prepare a high quality mask, a 3D scanner is necessary to obtain the 3D model of the target person, which are generally high-cost devices. The procedure continues with manufacturing of the masks which is also an expensive procedure.

Figure 1. Example sample for fabric mask. In the second column, the mask is worn on the face. The picture is taken from [5].

The mask attacks database which is used in this study was created by MORPHO [6]. Since the database includes many high-quality mask samples, it is possible to detect the performances of face recognition systems, accurately, under mask attacks. The mask database consists of both the 3D scans and the corresponding 2D texture images. Thanks to the nature of this database, in this paper, we are able to conduct the benchmark evaluations for each of 2D, 2.5D and 3D face recognition. The aim of this study is not to propose a new face recognition method, but instead to show the impact of mask attacks on existing face recognition methods.

The paper is organized as follows: Section 2 gives brief information on the mask database which is used in this study. Section 3 explains the face recognition systems which are selected to test the performance of these systems under mask attacks. Section 4 shows the experiments and results. Finally, conclusions are provided in Section 5.

2. THE MASK DATABASE

A mask is an object normally worn on the face, typically for protection, performance or entertainment. Additionally, masks can also be used for spoofing purposes.

There are several ways of mask manufacturing. Mask of a person can be prepared even by using papers.The company 'Thats My Face' [5] provides colored masks (Fig. 1). For each ethnicity, the company has a standard 3D face model and masks are manufactured by mapping one frontal and one profile picture of the target person on this model. However, since the model is based on an ethnic shape, it does not show exact 3D face shape characteristic of the target person.

The mask which is used for 3D face spoofing purposes has to show very similar 3D face shape characteristics of the target face to be considered as a successful attack. The mask database used in this study was prepared for this purpose. To obtain similar face shape characteristics of the target person,

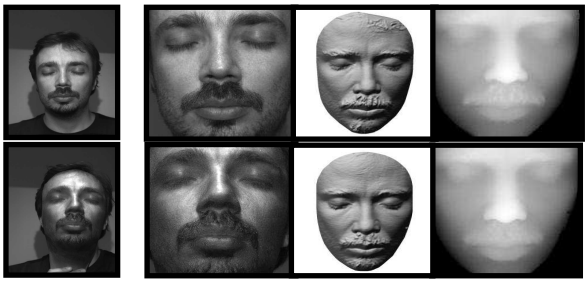

Figure 2. Example from the mask database which is created by [6]. From left to right (upper row) The real face, the texture image, the 3D scan after preprocessing, and the depth map estimated from this scan (lower row) same images for the corresponding mask attack.

initially, scans of the subjects in the mask database were taken by a 3D scanner which uses a structured light technology. Then the 3D model (3D mesh, the output of acquisition) of each subject was sent to the 3D printer and the masks were manufactured by Sculpteo 3D Printing [7].

In the mask attacks database, 20 subjects appear in total. The masks were manufactured for only 16 of these subjects. In this database, these 16 subjects appear with both their own mask and also with the masks of other people. The remaining 4 subjects appear with the masks of the other 16 subjects. For each subject, 10 scans are taken for the original person (real accesses) and almost 10 scans are taken for the person who wears either his/her own mask or masks of the other subjects that appear in the same database (mask attack accesses). Finally, 200 real face acquisitions and 199 mask acquisitions are used for the evaluations of this study. Fig. 2 shows one example from this mask attacks database for a real face access and the corresponding mask attack access.

3. SELECTED FACE RECOGNITION SYSTEMS

3.1. Pre-Processing

The pre-processing in this study is based on the method given in [8]. In order to crop the face region, the tip of the nose is detected and the facial surface is cropped by a sphere with radius 80mm, centered 10mm away from the nose tip in +z direction. Note that the face looks into +z direction. Next, the spikes are removed by thresholding and then the hole filling procedure is applied. Finally, a bilateral smoothing filter is used to remove white noise while preserving edges.

For the sake of clarity, the database of real accesses will be referred as DB-r and the database of mask accesses will be referred as DB-m in the rest of this study. In evaluations, the pre-processed 3D scans (only shape, without texture) are used for 3D, and the depth maps that are estimated from these pre-processed scans (which are previously aligned to a generic model) are used for 2.5D face recognition. The texture images in the mask database are used to measure the performance on 2D face recognition. Both the depth maps and the texture images are cropped as shown in Fig. 2 and resized into 96×96 images. (Fig. 2 shows an example for the texture images, the 3D scans and the depth maps which are used in 2D, 3D and 2.5D evaluations, respectively.)

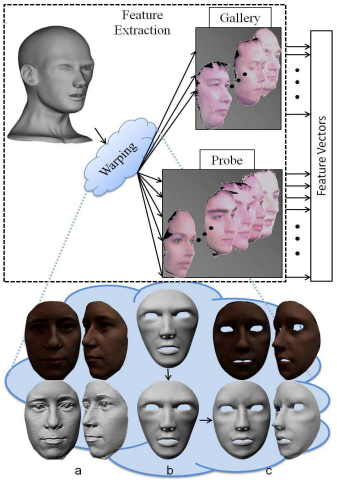

Figure 3. The feature extraction scheme and an illustration on a sample model: (a) The subject's face with and without texture (b) generic model before and after alignment (c) generic model after warping with and without texture. This figure is taken from [8].

3.2. Short Description on Selected Recognition Systems

The first system selected for this study is introduced in [8]. It is also selected as the baseline system in TABULA RASA project [4]. It uses the pre-processed 3D mesh of the face as input. Initially, a linear transformation is computed in a least square sense, based on two sets of landmarks (landmarks of the generic model and the subject's face). The landmark points are previously annotated at the nose tip and outer eye corners for each sample in the database. The best fit mapping is calculated by minimizing the squared distance (LSS) between the point sets of generic model and subject's face.Then the obtained transformation that includes rotation, translation and isotropic scaling is applied onto the generic model, aligning it with the subject's face.Next, the alignment is further improved by Iterative Closest Point (ICP) method [9].Afterwards 140 previously selected points on the generic model are coupled with the closest vertices on the face under analysis and Thin Plate Spline (TPS) [10] warping is applied on the generic model resulting in warping parameters (WP) of size 140x3. WPs that represent the deviations from the common structure are given to the classifier for recognition. Finally, the distance between two face models is computed by taking the median of cosine distances between the corresponding feature vectors (WP) and verification rates are computed. Fig. 3 shows the feature extraction scheme on a sample model using this method, which is named as WP.

Most of the existing 2D face recognition techniques can be applied on depth maps. In this study, LBP is used for both 2.5D and 2D face recognition. LBP descriptor provides state-of-the-art results in representing and recognizing face patterns [11]. The success of LBP in face description is due to the discriminative power and computational simplicity of the operator, and its robustness to monotonic gray scale changes caused by, for example, illumination variations. The use of histograms as features also makes the LBP approach robust to face misalignment and pose variations to some

extent. Originally LBP was designed to summarize local primitives on 2D texture image. In this paper, we also report its results on depth maps, so as to provide the benchmark for evaluation of advanced LBP variants [12,13] on depth maps. In both 2D and 2.5D, we use the operator $LBP^{u2}_{8;2}$ on 8×8 blocks. The similarity between each image pair is computed with chi-square distance metric which is shown in Eq. (1) as:

$$D(A,B) = \sum_{i=1}^{N} \frac{(A_i - B_i)^2}{A_i + B_i} \quad (1)$$

The performance evaluations are done using these similarity scores between image pairs. LBP applied on depth maps will be referred as LBP-depth, in the rest of this paper.

4. EXPERIMENTS AND RESULTS

In this part, the experiments are done for two scenarios to show the impact of mask attacks on face recognition. In the first scenario, the baseline performances of the systems are computed, which is called as the normal operation mode. In the second scenario, the performances of these systems under spoofing attacks are computed, which is called as the mode under spoofing attacks.

In face recognition, performance can be evaluated on both verification and identification [14]. In face verification, a claimed identity is validated based on the image/3D mesh of a face, and it either accepts or rejects the identity claim. If the similarity score is above a certain threshold, the user's identity is verified. In face identification, the references for all faces in the database are examined and the one with the best similarity score denotes the class of the input.

The selected techniques do not need training. Therefore, the whole database is used to test recognition performances.

- **Scenario 1: The Normal Operation Mode (Baseline)**

It is possible to measure the baseline performance of face recognition systems by using the real accesses data (DB-r) available at the mask attacks database. This step reports how well such a system behaves in normal operation mode.

In the normal operation mode, for verification tests, the systems use DB-r for both enrolment and authentication trials. This means that all vs. all comparison is done. Access from every identity in the DB-r is tested against all other models in the DB-r to report the baseline performance. The performance is measured objectively by observing the rate of users rejected when authenticating against their own template (False Rejection Rate - FRR) and by the rate of users accepted when authenticating against someone else's template (False Acceptance Rate - FAR). The scores generated against matched clients is considered true client accesses, while all others are impostors. For identification tests, the first sample of every subject is used as reference in the gallery and the rest is used for testing in the probe set.

- **Scenario 2: The Mode Under Spoofing Attacks**

In this mode, the enrollment is again achieved using DB-r, similar to the normal operation mode. However, this time the authentication trials are carried out by using DB-m instead of DB-r. A successful attack is accomplished when the system confuses a spoofing attempt with the corresponding matched user template.

The identity which is hidden under the mask (the real identity of the attacker) and the identity on the mask (the target identity) may be different. In this part, the identities of mask attackers are assumed as their target identities. (e.g. A_B: represents the identity A who wears the mask of identity B. Here, the target identity is B. In Scenario 2, the identity of A_B is assumed as the identity B). Also, whoever wears the mask, even if the person wears his/her own mask, it is considered as a spoofing attack. In this scenario, the mask attack attempts are assumed as real attempts and the same procedure in Scenario 1 is applied to compute FAR and FRR. This scenario is realized to see if the masks are recognized as their target identities by recognition systems.

4.1. Evaluation on 2D, 2.5D and 3D Face Recognition

In this part, the Detection-Error-Trade-Off (DET) curves, which describe the relationship between false acceptance rate and false rejection rate, are given for the two scenarios on the same figure to show the vulnerability of face recognition systems to mask attacks.

Since the mask attempts are assumed as real attempts in Scenario 2; if the masks spoof the systems perfectly, two curves in Fig. 4 should be as close as possible. The curves close to each other demonstrate that the masks are perceived as their target identities by the systems, which is the purpose of mask spoofing. Fig. 4 shows that for all systems, there is a decrease in the performances under spoofing attacks. The decrease in the performances is observed to be more for LBP compared to LBP-depth and WP. However, the performances under spoofing attacks can be still considered as high for all three systems.

The Equal Error Rates (EER is the value at FAR=FRR) for the two modes in Fig. 4 are reported in Table I. Lower EER indicates better performance. This table shows that the best performance is obtained with WP. It is clear that there is an increase in EER under spoofing attacks for all systems. However, EER under spoofing attacks still indicate significant recognition performances. This proves that the identities of the mask attackers are mostly recognized as their target identities by the systems.

The rank-1 identification rate of the baseline system and also the rate of successful mask attacks are given in Table II, for all three systems.

The rate of successful attacks is evaluated by dividing "the number of mask attacks in the probe set (the set of test trials) which are identified as their target identities by the system" to "the total number of mask attacks in the probe set". (If A_B is identified as B, it is considered as a successful attack.). According to Table II, WP (3D face shape analysis) provides the best identification rate of the baseline, however it is also the most vulnerable system to the spoofing attacks. It is clear that most of the mask attacks are identified as their target identities which means that the system is not robust against mask attacks. Baseline performances of LBP and

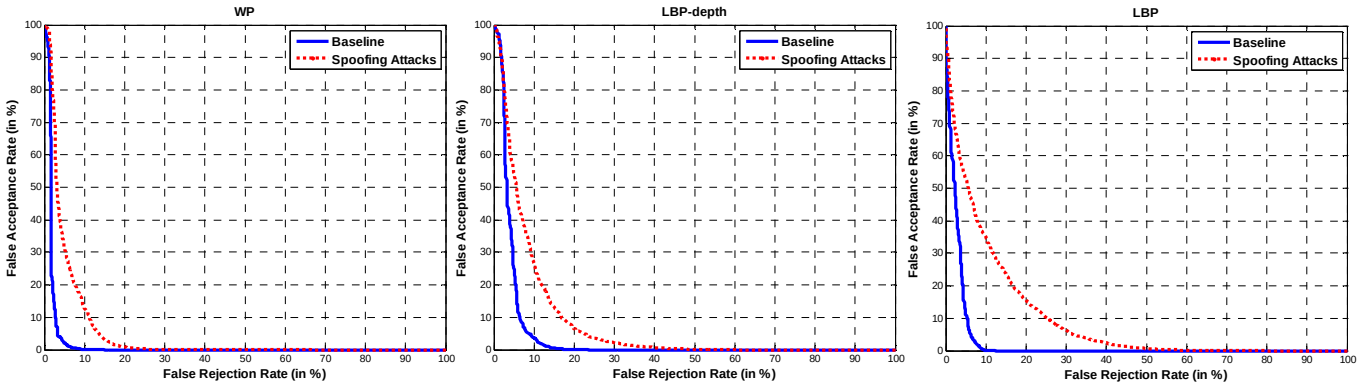

Figure 4. Performance rates computed for both the baseline and the performance under spoofing attacks using WP, LBP, LBP-depth.

Table I. Equal Error Rates for the Normal Operation Mode (Baseline) and the Mode under Spoofing Attacks

Equal Error Rates (%)	WP	LBP-depth	LBP
Baseline	03.85%	07.27%	05.90%
Spoofing Attacks	10.88%	14.26%	18.08%

Table II. The rank-1 identification rates of the baseline and the rate of successful mask attacks

Rates (%)	WP	LBP-depth	LBP
Rank-1 Identification	99.44%	92.78%	94.44%
Successful Attacks	91.46%	88.94%	72.87%

LBP-depth are observed to be slight worse than WP. It is again clear that most of the attacks are successful with LBP-depth which uses 3D information as input. The results of LBP show that although it provides similar baseline performance to other two techniques, it is more robust against mask attacks compared to WP and LBP-depth.

4.2. Comparison of the Impact of Mask Attacks on 2D, 2.5D and 3D Face Recognition

To measure the vulnerability of face recognition systems to mask attacks, one state-of-the art face recognition method is used for each 2D, 2.5D and 3D face recognition evaluations.

The results show that LBP applied on texture images provides slight better performance, but is observed to be more robust against mask attacks compared to LBP-depth. WP, which uses 3D mesh (without texture) as input, is the most vulnerable face recognition system to mask attacks. Since LBP is a texture analysis based technique, we can say that LBP method which is applied on texture images is observed to be more robust against mask attacks compared to 3D and 2.5D face recognition methods, which analyze 3D face shape characteristic. Although the best identification rate of the baseline is obtained with WP, which is independent of texture, pose and illumination changes, it is the most vulnerable system to mask attacks. The reason is that the masks in the mask database have very similar 3D shape characteristics of their target identities, which makes the masks perceived as their target identities by 3D and 2.5D face recognition systems.The results of this study reveal that;

- Techniques which are based on 3D face shape analysis may improve the performances significantly in face recognition, however in the presence of mask attacks, they are the most vulnerable systems to mask spoofing.
- It is much easy to differentiate a high-quality mask from a real face by using the texture information instead of using 3D face shape information.

5. CONCLUSIONS

In this study, a 2D+3D face mask attack database is used which was prepared for TABULA RASA research project. It is used to evaluate the performances of the state-of-the art face recognition techniques under spoofing attacks.

The novelty of this study is, it is the first time that the impact of mask spoofing is analyzed on 2D, 2.5D and 3D face recognition. Since it is possible to measure the baseline performances using the real accesses in the mask database, comparison between the baseline performance and the performance under spoofing attacks is possible, which is a significant advantage of this study. The results in our study show that the face recognition systems are vulnerable to spoofing mask attacks. Robustness against mask spoofing is observed to be both method and modality dependent. The systems which are based on 3D face shape analysis is observed to be the most vulnerable systems to mask attacks. Therefore, we can say that 2D face recognition systems are observed to be more robust against mask spoofing compared to 3D and 2.5D face recognition systems. Since standard techniques are vulnerable to mask attacks, robust algorithms are necessary to mitigate the effects of spoofing on face recognition. In this study, we observe that analysis on texture may reveal more information to detect mask attacks compared to analysis on 3D face shape characteristic. Our future work is to develop countermeasure techniques which uses both 2D and 3D data to detect mask attacks.

ACKNOWLEDGMENT

This work has been performed by the TABULA RASA project 7th Framework Research Programme of the European Union (EU), grant agreement number: 257289. The authors would like to thank the EU for the financial support and the partners within the consortium for a fruitful collaboration. For more information about the TABULA RASA consortium please visit http://www.tabularasa-euproject.org.

6. REFERENCES

[1] M-M. Chakka, A. Anjos, S. Marcel, et al., "Competition on counter measures to 2-d facial spoofing attacks," in IEEE IAPR Int. Joint Conference on Biometrics, IJCB, 2011.

[2] A. Hadid, M. Pietikäinen, "Face spoofing detection from single images using texture and local shape analysis," in IET Biometrics, vol. 1, March 2012, pp. 3–10.

[3] N. Kose, J.-L. Dugelay, "Classification of Captured and Recaptured Images to Detect Photograph Spoofing," in IEEE IAPR International Conference on Informatics, Electronics & Vision, ICIEV, May 2012.

[4] http://www.tabularasa-euproject.org/

[5] http://www.thatsmyface.com/Products/products.html

[6] http://www.morpho.com/

[7] http://www.sculpteo.com/en/

[8] N. Erdogmus, J.-L. Dugelay, "On Discriminative Properties of TPS Warping Parameters for 3D Face Recognition," in International Conference on Informatics, Electronics & Vision 2012, ICIEV 2012, May 2012.

[9] P. Besl and N. McKay, "A Method For Registration of 3-D Shapes," in IEEE Transactions on Pattern Analysis and Machine Intelligence,vol.14, no. 2, pp. 239-256, 1992.

[10] F. L. Bookstein, "Principal warps: Thin-Plate Splines and Decomposition of Deformations," in IEEE Transactions Pattern Analysis and Machine Intelligence, vol. 11, pp. 567–585, 1989.

[11] T. Ahonen, A. Hadid, and M. Pietikainen, "Face description with local binary patterns: Application to face recognition," in IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 28, no. 12, pp. 2037–2041, dec. 2006.

[12] Y. Huang, Y. Wang, and T. Tan, "Combining statistics of geometrical and correlative features for 3d face recognition," in Proc. BMVC, 2006, pp. 90.1–90.10, doi:10.5244/C.20.90.

[13] T. Huynh, R. Min, and J.-L. Dugelay, "An efficient LBP-based descriptor for facial depth images applied to gender recognition using RGB-D face data," in ACCV 2012, Workshop on Computer Vision with Local Binary Pattern Variants, Daejeon, Korea, November 5-9, 2012.

[14] S. Z. Li and A. K. Jain, Handbook of Face Recognition, Second Edition, Springer, 2011 (ISBN 978-0-85729-931-4).