Shape and Texture Based Countermeasure to Protect Face Recognition Systems Against Mask Attacks

Neslihan Kose and Jean-Luc Dugelay
Multimedia Department, EURECOM
Sophia-Antipolis, France
{neslihan.kose, jean-luc.dugelay}@eurecom.fr

Abstract

Photographs, videos or masks can be used to spoof face recognition systems. In this paper, a countermeasure is proposed to protect face recognition systems against 3D mask attacks. The reason for the lack of studies on countermeasures against mask attacks is mainly due to the unavailability of public databases dedicated to mask attack. In this study, a 2D+3D mask attacks database is used that is prepared for a research project in which the authors are all involved. The proposed countermeasure is based on the fusion of the information extracted from both the texture and the depth images in the mask database, and provides satisfactory results to protect recognition systems against mask attacks. Another contribution of this study is that the countermeasure is integrated to the selected baseline systems for 2D and 3D face recognition, which provides to analyze the performances of the systems with/without attacks and with/without the countermeasure.

1. Introduction

Spoofing attack occurs when a person tries to masquerade as another person by falsifying data and thereby gaining illegitimate access. Based on the observations that face recognition (FR) systems are vulnerable to spoofing attacks, researchers started to work on countermeasures to reduce the impact of attacks on face recognition performances. There have been studies on countermeasures to detect photo and video attacks, which are 2D face attacks [1 - 5]. However, 3D mask attacks to FR systems is a considerably new topic.To the best of our knowledge, only in [6, 7], countermeasures are proposed to detect mask attacks, however without analyzing whether the masks are able to spoof FR systems or not. The main reason for the lack of studies on mask spoofing is due to the unavailability of public databases. In this study, the mask database which is prepared within the context of European Union research project TABULA RASA is used.

The preparation of a mask spoofing database is much more difficult and expensive than the preparation of photo or video spoofing databases. The mask attacks database which is used in this study was created by MORPHO [8]. This database contains high quality mask samples and consists of both 3D scans and texture images (Fig. 1). In this paper, mask spoofing is analyzed on both 2D and 3D FR. For this purpose, one state-of-the art technique is selected for each of 2D FR and 3D FR and the proposed countermeasure is integrated to these systems.

Figure 1. Example from the mask database which is created by [8]. From left to right (upper row) The real face, the cropped texture image, the 3D scan after preprocessing, the cropped depth map estimated from the raw 3D scan (lower row) same images for the corresponding mask attack.

In [9], a micro-texture analysis technique is proposed to detect 2D attacks (e.g. photo, video), and in [6], this technique is used to detect 3D mask attacks. In [9], texture images are used as input whereas in [6], the technique is not applied only on the texture images, it is also applied on the depth maps estimated from 3D scans. The present study is also proposed to detect 3D mask attacks. The novelties of our study compared to [6] are:

- In [6], the performance of the countermeasure is analyzed on each one of the texture images and the depth maps separately, to detect mask attacks. In the present study, the performance of the countermeasure is evaluated by fusing the information extracted from each one of the texture and the depth images in the mask database both in score and feature level to obtain better performance.
- In this study, the countermeasure is integrated to the selected state-of-the art systems for 2D and 3D FR. Hence we can analyze whether the masks are able to spoof these systems or not, and the impact of the countermeasure on the performances of the systems under attacks, directly.

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 3D and 2D FR systems that are selected as baseline in this study. Section 4 explains the countermeasure. Section 5 shows the experiments and results. Finally, conclusions are provided in Section 6.

2. Mask Database

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. To obtain similar face shape characteristics of the target person, initially, scans of the subjects in the mask database were taken by a 3D scanner which uses structured light technology. Then the 3D model (3D mesh, the output of acquisition) of each subject was sent to the 3D printer and masks were manufactured by Sculpteo 3D Printing [10].

In the mask database, 20 subjects appear in total. The masks are manufactured for 16 of these subjects. In this database, these 16 subjects appear with both their own mask and also with the masks of the other people. The remaining 4 subjects appear with the masks of the other 16 subjects. For each subject, average 10 scans are taken for the original person (real accesses) and average 10 scans are taken for the person wearing either his/her own mask or masks of the other subjects that appear in the same database (mask attack accesses). Some samples had to be removed in the mask database due to their improper scans. Finally, in the present study, 200 real accesses and 198 mask attack accesses are used for the evaluations.

3. The Selected Face Recognition Systems

3.1. Pre-Processing for Face Recognition Systems

The pre-processing in this study is based on the method in [11]. 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. Next, spikes are removed by thresholding and hole filling is applied. Finally, a bilateral smoothing filter is used to remove white noise while preserving edges.

In the evaluations, the pre-processed 3D scans are used for 3D FR and the cropped texture images of the same subjects are used for 2D FR. Finally, both the depth maps which are estimated from the raw data and the texture images are used as input for the proposed countermeasure. Fig. 1 shows an example for the texture images, the pre-processed 3D scans and the depth maps estimated from the raw 3D scans of a real face access and corresponding mask attack access, which are used in the evaluations.

3.2. Short Description on the Selected FR Systems

The 3D FR system used in this study was introduced in [11] and selected as the baseline system in the project

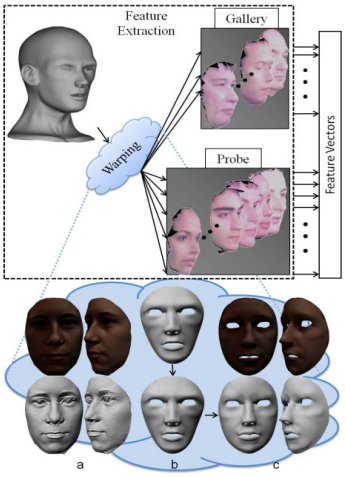

Figure 2. The feature extraction scheme and an illustration on a sample model: (a) The subject's face with and without texture (b) generic model before and after alignment (c) generic model after warping with and without texture. This figure is taken from [11].

TABULA RASA. 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. 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 [12]. 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) [13] 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. 2 shows the feature extraction on a sample model using this WP technique.

Local Binary Pattern (LBP) provides state-of-the-art results in representing and recognizing face patterns [14]. The success of LBP 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. For 2D FR system, in this study, the operator $LBP_{8;2}^{u2}$ is selected to be used on 8×8 blocks. The similarity between each image pair is computed using chi-square distance metric. The performance evaluations are done using these similarity scores between image pairs.

4. The Proposed Countermeasure Technique

Mask attack is a 3D attack that can be used to spoof both 2D and 3D FR systems. Most of the existing 3D scanners do not provide only 3D scan, they also capture texture image. Fig. 1 shows an example for the two outputs of a scanner. Thus, when there is only one camera for 2D FR and one scanner for 3D FR system, countermeasure which uses texture images as input can be used to protect both 2D and 3D FR systems if texture images are provided as default output of scanner.

In this study, we fuse the information which is extracted from both the texture and the depth images in our mask database using micro-texture analysis. Also, we integrate the countermeasure to FR systems to analyze performances of these systems with/without attacks and with/without the countermeasure a plus further from the studies [6, 7]. Both score and feature level fusion are applied in this study. Since existing 3D scanners provide both the 3D scan and the corresponding texture image, this countermeasure can be applied to protect 3D FR systems, directly.

The mask database is 2D+3D. For the sake of clarity, the database of real faces in 2D and 3D will be referred as DB-r2 and DB-r3, while the database of mask attacks will be referred as DB-m2 and DB-m3 in the rest of this paper.

4.1. Pre-Processing for the Countermeasure

There are slight alignment differences between faces in the mask database. For the countermeasure, initially all 3D faces in DB-r and DB-m are aligned to a generic face, which makes the alignment of all faces identical.

In this study, we want to benefit from the information that the mask surface is smoother than the real face surface to detect mask attacks. Therefore, the raw data is used for the countermeasure. The depth maps are estimated from the raw aligned 3D scans (e.g last column in Fig. 1). Only 2D cropping is applied to extract the face region from both the texture images and the depth maps. Then all images are resized into 64×64 grayscale image.

4.2. Micro-Texture Analysis Based Countermeasure

The micro-texture analysis method [9] is used to extract features as a part of the countermeasure proposed in our study. The novelties of this study compared to [9] are:

- The technique in [9] is used to detect 2D face print (e.g. photo, face picture on a paper) attacks whereas in our study, we use this technique to detect 3D mask attacks.
- In our study, the technique is applied both on the texture images and on the depth maps. Hence we prove that the technique is also successful on 3D data.

This LBP based technique emphasizes the micro-texture differences in the feature space. It aims at learning the differences between real and fake face, and designs a feature space which emphasizes those differences.

The original LBP forms labels for the image pixels by thresholding the 3 x 3 neighborhood of each pixel with the center value and considering the result as a binary number. The LBP operator has been extended to use neighborhoods of different sizes. $LBP_{P,R}$ is computed such that for a given central pixel in an image, a pattern number is computed by comparing its value with those of its neighbors. In Eq. (1), g_c is the gray value of the central pixel, g_p is the value of its neighbors, P is the number of neighbors around a circle of radius R. $LBP_{P,R}$ calculation is shown in Eq. (1) and (2):

$$LBP_{P,R} = \sum_{p=0}^{P-1} s(g_p - g_c)2^p, \quad (1)$$

$$s(x) = \begin{cases} 1 & x \geq 0 \\ 0 & x < 0 \end{cases} \quad (2)$$

Another extension to the original operator is the use of uniform patterns, which are verified to be the fundamental patterns of local image texture. A local binary pattern is called uniform if the binary pattern contains at most two bitwise transitions from 0 to 1 or vice versa when the bit pattern is traversed circularly. The notation is $LBP_{P,R}^{u2}$. u2 stands for using only uniform patterns and labeling all remaining patterns with a single label.

In [9], authors claim that micro-texture details that are needed to discriminate a real face from face print can best be detected using combination of different LBP operators. Thus, they derive an enhanced facial representation using multi-scale LBP operators. Their proposed representation computes LBP features from 3 x 3 overlapping regions to capture the spatial information and enhances the holistic description by including global LBP histograms computed over the whole image. This is done as follows: the face is cropped and resized into a 64x64 pixel image. Then, $LBP_{8,1}^{u2}$ operator is applied on the face image and the resulting LBP image is divided into 3x3 overlapping regions (with an overlapping size of 14 pixels). The local 59-bin histograms from each region are computed and collected into a single 531-bin histogram. Then, two other histograms are computed from the whole face image using $LBP_{8,2}^{u2}$ and $LBP_{16,2}^{u2}$ operators, yielding 59-bin and 243-bin histograms that are added to the 531-bin histogram previously computed. Hence, in [9], the length of the final enhanced feature histogram is 833 (i.e. 531+59+243).

Captured image from mask may visually look very similar to the image captured from live face (e.g. the texture images in Fig. 1). A close look at the differences between faces in DB-r2 and DB-m2 reveals that their surface properties are different. For mask manufacturing 3D printers are used, hence they may contain printing quality defects that can be detected with micro-texture patterns. In our study, the micro-texture analysis technique is first applied on the texture images in the mask database and the feature histogram of length 833 is obtained.

The 3D shape of high quality mask is also very similar to the 3D shape of the corresponding real face (e.g. the 3D scans in Fig. 1). Our analysis on DB-r3 and DB-m3 show

that the mask scan is smoother than the real face scan. Especially the parts of the face with facial hair are quite different. Since there is no real facial hair (e.g. mustache, eyebrow) on the mask, the 3D scan of the mask is smoother in these parts compared to the real face scan. High quality scanners cause less number of holes however even with the best scanners it is possible to observe some holes on the scan especially at the parts of the face with facial hair. Thus, in our study, secondly, the micro-texture analysis is applied on the depth maps which are estimated from the raw 3D scans in the mask database and the other feature histogram of length 833 is obtained.

In the present study, we apply feature and score level fusion of the information extracted from the texture and the depth images in the mask database. For feature level fusion, the two feature histograms computed from the texture and the depth images are concatenated and the classifier is applied on the resultant feature histogram. Thus, the length of the final feature histogram is 1666 (instead of 833). Once the enhanced histogram of length 1666 is computed, a linear SVM classifier [15] is used to determine whether the input image corresponds to a live face or not. For score level fusion, linear SVM classifier is applied using the texture and the depth features separately, and scores are obtained for the two groups. Then, Z-score normalization is applied for each of these groups. Finally, the weighted score level fusion is used for combining the outputs of the individual SVMs to determine whether the input image corresponds to a live face or not.

4.3. Integration of the Countermeasure to FR Systems

The evaluations in this study are done for 4 modes. The first mode is the baseline mode: a standard, state-of-the-art biometric system with no spoofing and no countermeasure. The baseline performance is evaluated using DB-r in the mask database. Performance is evaluated by verification all vs. all. Access from every identity in DB-r is tested against all other models in DB-r. The performance is measured 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). When spoofing attacks are applied, baseline performance is expected to degrade. The second mode is the evaluation of FR systems under mask attacks. Both DB-r and DB-m are used. In this mode, the FAR corresponds to the rate of attacks that are accepted by the system when spoofed. The FRR corresponds to the rate of real-access attempts that are incorrectly dismissed by the system as attacks. Third mode illustrates performance when the countermeasure is applied against the attacks, that results in an improved performance with respect to the second mode. For the samples which are detected as attack by the countermeasure, a least similarity score is assigned to those samples in verification tests. Last mode evaluates the performance of the baseline system together with the countermeasure in the normal operation mode of system, i.e., without attacks. The inclusion of the countermeasure may degrade the baseline performance when not confronted to attack. (e.g. the countermeasure may consider as fake some real users.)

In the mask database, initially DB-r and DB-m are partitioned in training and test datasets. 8 subjects out of 16 subjects whose masks are manufactured and 2 subjects out of 4 subjects whose masks are not manufactured are selected for DB-r. The samples of the selected subjects are assigned to the test set of DB-r, while the rest is used for the training set of DB-r. For DB-m, the mask attack accesses to the corresponding identities in the test set of DB-r are involved in the test set of DB-m, while the rest is used for the training set of DB-m. There is no overlap between the training and the test sets which makes the spoofing detection more challenging. This partitioning is done for both the texture images and the depth maps.

Training dataset is used for classifier training. This classifier is subject to two kind of errors:

- FLR (False Living Rate), that represents the percentage of fake data misclassified as real. (similar to FAR)
- FFR (False Fake Rate), which computes the percentage of real data assigned to the fake class. (similar to FRR)

The lower these two errors, the better the performance of the countermeasure. For evaluations, we fix 3 different evaluation points at FFR= 1%, 5%, and 10%. Once fixed, we incorporate the countermeasure as a first step into the baseline biometric systems oriented to discard fake data, and generate the performance evaluations for the 4 modes.

5. Experiments and Results

The Region of Convergence (ROC) curve in Fig. 5 shows the stand-alone classification performance of the countermeasure. Table I shows Area Under Curve (AUC) and accuracy results when the countermeasure is applied only on the texture images [6], only on the depth maps [6], and finally feature and score level fusion of the information extracted from the texture and depth images, which is the proposed approach. The results in Table I prove that both score and feature level fusion improves the results compared to using only the texture or depth images. Also, the results show that the score level fusion provides the best results in terms of both AUC and accuracy (accuracy of 93.5%). In [7], they achieved 89.2% mean detection accuracy of real face vs. mask. However an exact comparison is not possible since different database is used in [7]. Thus, in the present study, we apply the technique in [6] using the same database with the same training-test sets. The results are reported in Table I which prove that the performance is better with fusion compared to [6].

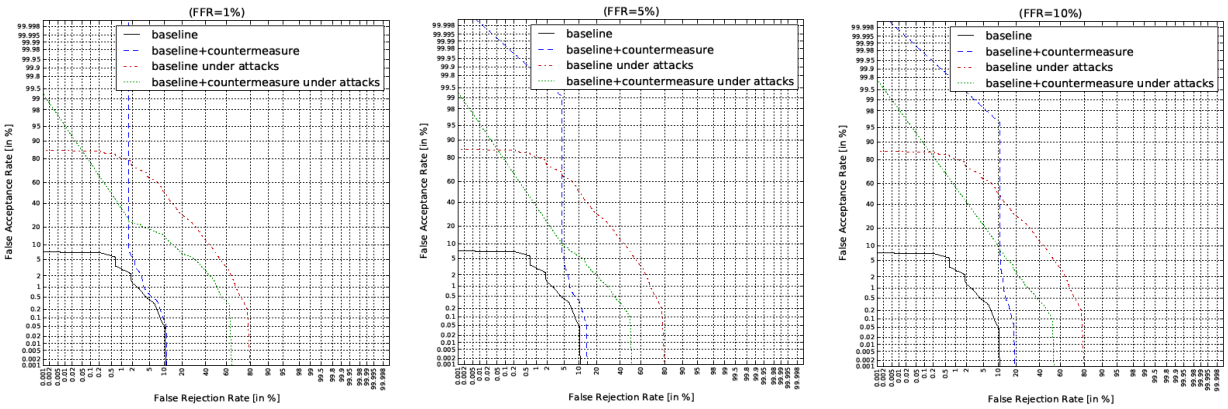

Figure 6. The DET Curves (computed with toolkit [16]) of the 3D face baseline biometric system when integrating the countermeasure.

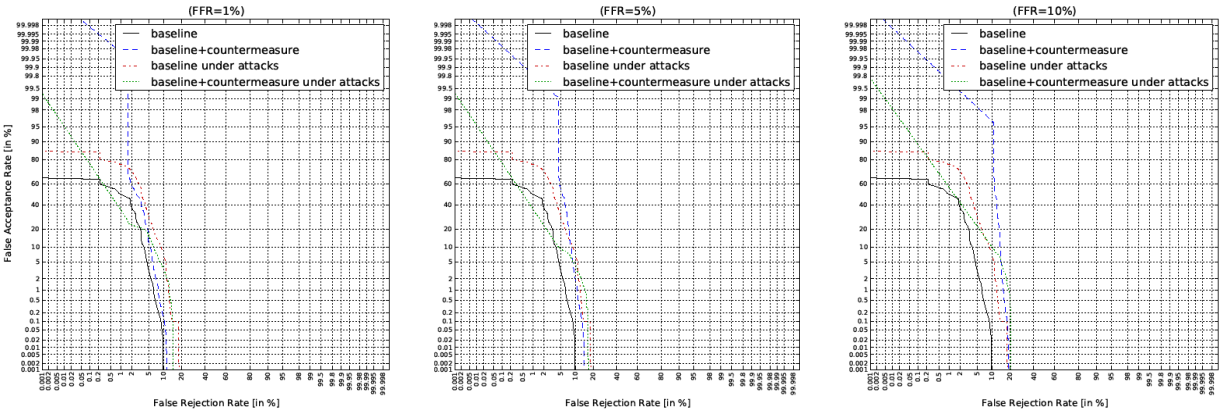

Figure 7. The DET Curves (computed with toolkit [16]) of the 2D face baseline biometric system when integrating the countermeasure.

Another contribution of this study is; the performance of the selected baseline systems for 2D FR and 3D FR is evaluated with/without mask attacks and with/without the countermeasure. By this way, firstly, we can observe whether the mask attacks are successful to spoof these systems or not, and secondly whether our countermeasure improves the performances of these systems under attacks or not. All results are presented in terms of detection error trade-of (DET) profiles which illustrate the behavior of a system as the decision threshold is changed, i.e. how the false acceptance rate varies according to the false rejection rate. The score level fusion based countermeasure is used in Fig. 5, 6 and 7, since it gives the best results.

Fig. 6 shows the behavior of the 3D face baseline system with/without attacks and with/without the proposed countermeasure. The three figures represent the overall system performance under spoofing attacks when three different operating points (FFR = 1%, 5%, and 10%) are used for adjusting the countermeasure. It is clear from Fig. 6 that the 3D FR system is vulnerable to mask attacks. (more area between black and red curves indicates more vulnerability to the attacks). In Fig. 6, it is also clear that performance enhancement is obtained almost all regions of DET plots when the countermeasure is introduced to

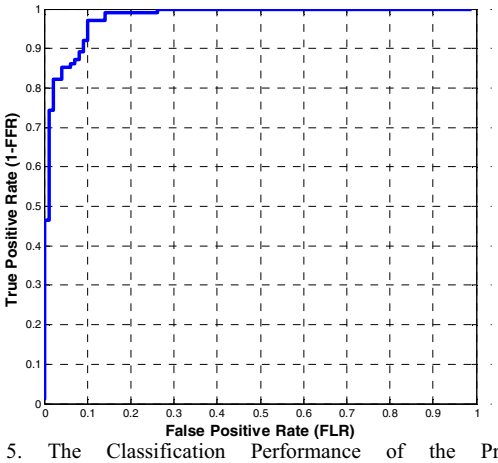

Figure 5. The Classification Performance of the Proposed Countermeasure (Score Level Fusion)

3D FR system under attacks (green curve compared to red curve). If we take an operating point where FFR=1%, then FAR of the 3D FR system under attacks drops from around 50% to around 15% at FRR= 10%. For both of the two other plots (at FFR=5% and 10%), the introduction of the countermeasure lowers FAR from around 50% to 5% and 10%, respectively, at FRR=10%. The performance of the countermeasure is slight better at FFR=5% compared

TABLE I. AREA UNDER CURVE AND BEST ACCURACY RESULTS USING THE TEXTURE IMAGES, THE DEPTH MAPS AND THE FUSION OF THEM

Micro-Texture Analysis Applied on	AUC	Accuracy (%)
Texture Images	0.956	89.4
Depth Maps	0.915	82.4
Feature Level Fusion (Proposed App.)	**0.976**	**93.0**
Score Level Fusion (Proposed App.)	**0.978**	**93.5**

to the cases at FFR=1% and 10% when the Equal Error Rates (EER) at three cases are compared. Finally, Fig. 7 shows the results of evaluations for the 2D face baseline system. Similar to the results in Fig. 6, 2D FR system is vulnerable to mask attacks and performance enhancement is obtained at most of regions of DET plots in Fig. 7 when the countermeasure is introduced to the 2D FR system under attacks. At operating point FFR=1%, FAR of 2D FR system under attacks drops from around 9% to 3% at FRR=10%. At FFR=5%, the introduction of the countermeasure lowers FAR from around 9% to around 5% and at FFR=10% FAR does not change at FRR=10%.

The plots in Fig. 6 and 7 prove that the 3D FR system, which is completely based on 3D shape analysis, is more vulnerable to mask attacks compared to the 2D FR system, which is a texture based technique (area between black and red curves is much more for the 3D compared to 2D FR system). EER at the baseline mode increases from 1.8% to 25.1% for 3D FR and from 4.7% to 9.3% for 2D FR system under attacks. This proves that the 3D face shape characteristics of masks and real faces in the mask database are more similar compared to their texture characteristics. The inclusion of the countermeasure improves the results of both 2D and 3D FR systems under attacks, whereas it degrades baseline performances of both systems when not confronted to attack (blue curve compared to black curve).

6. Conclusions

In this study, a 2D+3D mask attacks database, which is prepared for a European Union research project, is used to evaluate the performances of the proposed countermeasure for the protection of FR systems against mask attacks.

The novelty of this study is that it is one of the few studies on the topic of spoofing countermeasures against 3D mask attacks. The proposed countermeasure uses both the texture and the depth images as input. The technique can be applied only on the texture or the depth images. However, the results show that the technique provides more satisfactory results when the information from the texture and the depth images are fused. The technique can be used to protect both 2D and 3D FR systems.

In this study, it is also proved that standard FR systems are not robust to the spoofing mask attacks therefore robust algorithms are necessary to mitigate the effects of spoofing attacks. The proposed countermeasure is an LBP based approach which improves the performance of FR systems under attacks, significantly. This study shows that LBP based techniques give satisfactory results also for spoofing detection. Our future work is to develop novel countermeasures which provide even better classification accuracies 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.

References

[1] M-M. Chakka, A. Anjos, S. Marcel, et al., "Competition on counter measures to 2-d facial spoofing attacks," IEEE IAPR Int. Joint Conference on Biometrics (IJCB), pp.1-6, 2011.

[2] J. Maatta, A. Hadid, M. Pietikainen, "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," IEEE Int. Conf. on Informatics, Electronics & Vision, pp.1027-1032, 2012.

[4] J. Li, Y. Wang, T. Tan, A.K.Jain, "Live Face Detection Based on the Analysis of Fourier Spectra", SPIE, pp. 296-303, 2004.

[5] W. Bao; H. Li; N. Li; W. Jiang, "A liveness detection method for face recognition based on optical flow field," Int. Conf. on Image Analysis and Signal Processing (IASP), pp.233-236, 2009.

[6] N. Kose, J.-L. Dugelay, "Countermeasure for the Protection of Face Recognition Systems Against Mask Attacks", IEEE Automatic Face and Gesture Recognition (FG), Shanghai, April 2013.

[7] Z. Zhang, D. Yi, et.al, "Face Liveness Detection by Learning Multispectral Reflectance Distributions," IEEE Automatic Face and Gesture Recognition (FG), pp. 436-441, 2011.

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

[9] J. Maatta, A. Hadid, and M. Pietikainen, "Face spoofing detection from single images using micro-texture analysis,"Proc. of IEEE Int. Joint Conf. on Biometrics (IJCB), pp. 1-7, USA, 2011.

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

[11] N. Erdogmus, J.-L. Dugelay, "On Discriminative Properties of TPS Warping Parameters for 3D Face Recognition," IEEE Int. Conf. on Informatics, Electronics & Vision (ICIEV), pp. 225-230, 2012.

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

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

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

[15] C.-C. Chang, and C.-J. Lin. "LIBSVM: a library for support vector machines." ACM Transactions on Intelligent Systems and Technology (TIST) 2.3 (2011): 27.

[16] Score toolkit (http://publications.idiap.ch/downloads/reports/2012/Anjos_Idiap-Com-02-2012.pdf)