



**HAL**  
open science

# MASK SPOOFING IN FACE RECOGNITION AND COUNTERMEASURES

Neslihan Kose, Jean-Luc Dugelay

► **To cite this version:**

Neslihan Kose, Jean-Luc Dugelay. MASK SPOOFING IN FACE RECOGNITION AND COUNTERMEASURES. *Image and Vision Computing*, 2014, 32 (10), pp.779-789. 10.1016/j.imavis.2014.06.003 . hal-03906926

**HAL Id: hal-03906926**

**<https://hal.science/hal-03906926v1>**

Submitted on 19 Dec 2022

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# MASK SPOOFING IN FACE RECOGNITION AND COUNTERMEASURES

Neslihan Kose, *Student Member, IEEE*, Jean-Luc Dugelay, *Fellow, IEEE*  
*Multimedia Department EURECOM Sophia-Antipolis France*

---

## Abstract

In this paper, initially, the impact of mask spoofing on face recognition is analyzed. For this purpose, one baseline technique is selected for both 2D and 3D face recognition. Next, novel countermeasures, which are based on the analysis of different shape, texture and reflectance characteristics of real faces and mask faces, are proposed to detect mask spoofing. In this paper, countermeasures are developed using both 2D data (texture images) and 3D data (3D scans) available in the mask database. The results show that each of the proposed countermeasures is successful in detecting mask spoofing, and the fusion of these countermeasures further improves the results compared to using a single countermeasure. Since there is no publicly available mask database, studies on mask spoofing are limited. This paper provides significant results by proposing novel countermeasures to protect face recognition systems against mask spoofing.

*Keywords:* spoofing, mask attacks, countermeasure, face recognition

---

*Email address:* [neslihan.kose@eurecom.fr](mailto:neslihan.kose@eurecom.fr), [jean-luc.dugelay@eurecom.fr](mailto:jean-luc.dugelay@eurecom.fr)  
(Neslihan Kose, *Student Member, IEEE*, Jean-Luc Dugelay, *Fellow, IEEE*)

---

## 1 1. INTRODUCTION

2 In a spoofing attempt, a person tries to masquerade as another person  
3 and thereby, tries to gain access to recognition system. Face recognition is  
4 used in domains such as surveillance and access control. Since face data can  
5 be acquired easily in a contactless manner, spoofing is a real threat for face  
6 recognition systems.

7 The most common spoofing attacks are photograph and video attacks  
8 due to their convenience and low cost. Based on the observations that 2D  
9 face recognition (FR) systems are vulnerable to these attacks, researchers  
10 started to work on countermeasures to reduce their impact on recognition  
11 performances.

12 Proposed countermeasures against photo and video attacks are mainly  
13 based on liveness detection, motion analysis and texture analysis. Counter-  
14 measures based on liveness detection examine movements such as eye blinking  
15 [24] or lip movements [9]. In the literature, there are several countermeasures  
16 based on motion analysis [4, 22]. These countermeasures rely on the fact that  
17 the movement of a 2D plane is different compared to the movement of a 3D  
18 object. In [4], under the assumption that the test region is a 2D plane, the  
19 authors obtain a reference field from the actual optical flow field data. Then  
20 the degree of differences between the two fields is used to distinguish between  
21 a 3D face and a 2D photograph. In [22], a set of facial points are located auto-  
22 matically and their geometric invariants are used to detect attacks. The last

23 group is countermeasures based on texture analysis. In [3] and [19], images  
24 are examined to find printing artifacts and blurring, respectively. In [14], dif-  
25 ferent contrast and texture characteristics of photographs and real faces are  
26 analyzed to detect spoofing. Furthermore in [20, 21], micro-texture analysis  
27 is proposed to detect 2D attacks. The study [7] includes brief information  
28 about different types of 2D face countermeasures, which were developed for  
29 a competition on countermeasures against 2D facial spoofing attacks. Six  
30 teams participated to this competition. These teams are AMILAB, CASIA,  
31 IDIAP, SIANI, UNICAMP and UOULU. All teams used one or multiple  
32 clues obtained clearly from motion analysis, texture analysis and liveness  
33 detection. The CASIA team presented a method with the combination of  
34 motion and texture analysis techniques, and the method also allows switch-  
35 ing between detection schemes based on the scene context. The AMILAB  
36 and the UNICAMP teams used all motion analysis, texture analysis and live-  
37 ness detection in deriving the detection scheme. IDIAP and UOULU used  
38 texture analysis method and obtained zero percent Equal Error Rate (EER)  
39 on development set and zero percent Half Total Error Rate (HTER) on test  
40 set. This leads to the conclusion that, the attack videos in the database used  
41 for this competition (i.e. PRINT-ATTACK Database [26]) mainly consist of  
42 detectable texture patterns.

43 When 3D masks are introduced as attacks, some of the countermeasures  
44 proposed for the detection of 2D attacks are no longer applicable. The study  
45 of Kollreider et al. [13] shows that a face recognition system relying on eye

46 blinking and lip movements can be defeated by using photographic masks  
47 wrapped over face with eyes and mouth regions cut out. Also, since motion  
48 based countermeasures depend on different movements of 2D and 3D surfaces,  
49 they are not applicable when masks are used instead of photos or videos. It  
50 appears that the detection of 3D mask attacks is more challenging compared  
51 to the detection of 2D facial attacks.

52 3D mask attacks to FR systems is a considerably new subject. The main  
53 reason for the delay in mask spoofing studies is due to the unavailability of  
54 public mask databases. To our knowledge, in the literature, there are two  
55 countermeasure studies against 3D mask attacks [11, 29] excluding our stud-  
56 ies. These two studies are based on reflectance analysis. They utilize 2D data  
57 (texture images) in their approach to detect 3D mask attacks. Kim et al.  
58 [11] exploit the reflectance disparity based on albedo between real faces and  
59 mask materials (silicon, latex or skinjell). The feature vector, which is used  
60 in their approach for mask detection, consists of radiance measurements of  
61 the forehead region under 850 and 685 nm illuminations. They report 97.78%  
62 accuracy for mask detection. In [11], the experiments are done directly on  
63 the mask materials not on the real facial masks. Thus, it is not possible to  
64 report spoofing performances of the masks used. The measurements are done  
65 at exactly 30 cm and on the forehead region for mask detection. The require-  
66 ment for an exact distance and occlusion possibility in the forehead during  
67 the measurements are the limitations of this method. In [29], multi-spectral  
68 reflectance analysis is proposed. After measuring the albedo curves of facial

69 skin and mask materials with varying distances, two discriminative wave-  
70 lengths (850 and 1450 nm) are selected. Finally, a Support Vector Machine  
71 (SVM) [8] classifier is used to discriminate between real and fake samples.  
72 Experiments are conducted on a database of 20 masks from different materi-  
73 als (4 plastic, 6 silica gel, 4 paper pulp, 4 plaster and 2 sponge). The results  
74 show that this method can achieve 89.18% accuracy. The superiorities of  
75 [29] compared to [11] are the elimination of range limitation and the usage  
76 of real facial masks. However, spoofing performances of the masks are still  
77 not reported. In order to contribute to this compelling research problem and  
78 fill the missing portions of the existing studies, we have proposed several  
79 countermeasure techniques against 3D mask attacks in [15], [18] and [17].

80 The spoofing performances of the masks used and the countermeasure  
81 which uses 3D data (3D scan) instead of 2D data (texture image) as input  
82 to detect mask spoofing were first analyzed in our previous studies [16, 15],  
83 respectively, using the mask database which was prepared within the context  
84 of the European Union (EU) research project TABULA RASA [28]. The  
85 mask database used in the present study and our previous studies [15, 16,  
86 18, 17] was created by MORPHO [23]. This database includes many high-  
87 quality mask samples. It consists of 3D masks of 16 real subjects. The scans  
88 of subjects were acquired by a 3D scanner, and the masks were manufactured  
89 using a 3D printer. In addition to texture images, it includes 3D scans for  
90 both real and mask samples. Thanks to the nature of this database, we  
91 were able to evaluate the impact of mask spoofing on both 2D and 3D face

92 recognition, and to develop our countermeasures using both 2D and 3D data.

93 If a 3D mask is not able to spoof a recognition system, it is not a suc-  
94 cessful attack, and there is no need to develop a countermeasure against it.  
95 Therefore, in [16], we analyzed how well the spoofing performances of the  
96 masks used in our studies are. The results of this study show that the masks  
97 used have very similar texture and especially 3D face shape characteristics to  
98 their target faces. They are very successful to spoof face recognition systems.  
99 To the best of our knowledge, spoofing performances of the masks used were  
100 the first to be analyzed in this study. In [15], we proposed to apply micro-  
101 texture analysis on both texture and depth images, and obtained 88.12%  
102 and 86% accuracy, respectively, for the classification of mask and real faces.  
103 The novelty of this work is that it was the first time 3D data was utilized  
104 to discriminate mask and real samples. In our next study [18], which is the  
105 continuation of [15], we studied fusing the information extracted from both  
106 the texture and depth images, and obtained a higher classification accuracy  
107 of 93.5%. In addition to the increase in performance, it was the first time the  
108 performances of face recognition systems were analyzed with/without mask  
109 attacks and with/without the proposed countermeasure integrated to the  
110 recognition systems in [18]. By this way, it is possible to observe the positive  
111 impact of countermeasure on recognition performances in presence of mask  
112 attacks. Finally, in [17], we proposed a countermeasure based on reflectance  
113 analysis using the texture images in the same database. We obtained 94%  
114 classification accuracy in [17], which was the best score we had obtained so

115 far.

116 In this paper, after showing the impact of attacks on the selected recog-  
117 nition systems; we provide an overview on our spoofing detection approaches  
118 which were introduced in the studies [15], [18] and [17]. We extend the  
119 works explained in these studies with some improvements, additional analy-  
120 sis, comparisons of performances of diverse countermeasures using the same  
121 protocol, and with a detailed analysis of the fusion scenarios. Additionally, a  
122 novel countermeasure is proposed in the present paper. In [15], micro-texture  
123 analysis is applied on texture images, whereas in this paper, we apply micro-  
124 texture analysis on reflectance components of texture images as a new coun-  
125 termeasure. We observe that higher accuracy is obtained using reflectance  
126 component instead of texture image (original image) itself. This proves that  
127 reflectance image provides more appropriate information than original im-  
128 age to discriminate mask and real samples. In the present study, we obtain  
129 98.99% classification accuracy, which is the best accuracy that have been re-  
130 ported in the literature for mask spoofing detection up to now. Also, in the  
131 present paper, we integrate the countermeasure with the best performance  
132 to the selected 3D FR system in order to show the positive impact of the  
133 countermeasure on the system under mask attacks, directly.

134 The paper is organized as follows: Section 2 gives brief information on the  
135 mask database which is used in this study. Section 3 presents the selected 2D  
136 and 3D FR systems, and then evaluates the impact of mask spoofing on these  
137 systems. Section 4 gives brief information about the techniques that were



138 used to develop the proposed countermeasures. Section 5 explains each of  
139 the proposed countermeasures. Section 6 shows the experiments and results  
140 of all the proposed countermeasures together with the fusion scenarios for  
141 comparison purposes. Finally, conclusions are provided in Section 7.

## 142 2. THE MASK DATABASE

143 A mask used for 3D face spoofing purposes has to show very similar  
144 3D shape characteristics to the target face to be considered as a successful  
145 attack. The mask database used in this study was prepared to fulfill this  
146 objective. Initially, scans of the subjects in the mask database were taken  
147 by a 3D scanner which uses a structured light technology in order to obtain  
148 similar face shape characteristics to the target person. Then the 3D model  
149 (3D mesh) of each subject was sent to a 3D printer and the masks were  
150 manufactured by Sculpteo 3D Printing [27]. The material used for the masks  
151 is polychrome mineral powder, which is a 3D printing standard.

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

156 In the mask database, 20 subjects appear in total. The masks were  
157 manufactured for 16 of these subjects. For DB-r, an average of 10 scans  
158 of each subject were acquired. For DB-m, an average of 10 scans of each  
159 subject wearing either his/her own mask or masks of the other subjects that



Figure 1: Example from the mask database which is created by [23]. 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.

160 appear in the same database were acquired. Finally, 200 real face acquisitions  
 161 from 20 subjects and 198 mask acquisitions from 16 masks are used for the  
 162 evaluations of this study. Figure 1 shows one example from this database for  
 163 a real face access and the corresponding mask attack access.

164 In the mask database, DB-r and DB-m are partitioned in train and test  
 165 sets. 8 subjects out of 16 subjects whose masks are manufactured, and 2  
 166 subjects out of 4 subjects whose masks are not manufactured are selected  
 167 for DB-r. The samples of the selected subjects are assigned to the test set of  
 168 DB-r, while the rest is used for the train set of DB-r. For DB-m, the mask  
 169 attack accesses to the corresponding identities in the test set of DB-r are  
 170 involved in the test set of DB-m, while the rest is used for the train set of  
 171 DB-m. There is no overlap between the train and test sets, which makes the  
 172 spoofing detection more challenging. Finally, there are 100 samples in each  
 173 of the client (real accesses) test and train sets, and 99 samples in each of the

174 impostor (mask attack accesses) test and train sets.

### 175 **3. IMPACT ANALYSIS OF MASK SPOOFING ON FACE RECOG-** 176 **NITION**

177 In this section, initially, we explain the pre-processing applied for the  
178 selected 3D and 2D FR techniques. Next, we give the details about these  
179 recognition techniques. Finally, we evaluate the impact of spoofing mask  
180 attacks on both 3D and 2D face recognition.

#### 181 *3.1. Pre-Processing for the Selected Face Recognition Systems*

182 The pre-processing for the selected 3D FR system is based on the method  
183 given in [10]. In order to crop the face region, the tip of the nose is detected,  
184 and the facial surface is cropped by a sphere with radius 80 mm, centered 10  
185 mm away from the nose tip in +z direction. Note that the face looks towards  
186 +z direction. Next, the spikes are removed by thresholding, and then a hole  
187 filling procedure is applied. Finally, a bilateral smoothing filter is used to  
188 remove white noise while preserving edges. These pre-processed 3D scans  
189 (only shape, without texture) are used as input for 3D face recognition.

190 For 2D face recognition, the texture images in the mask database are  
191 cropped as shown in Figure 1, and resized into  $64 \times 64$  images. In this study,  
192 we aim to show first how vulnerable the systems are to spoofing mask attacks  
193 by evaluating the performances of the selected systems with/without attacks,  
194 and then how a countermeasure improves the performance in presence of

195 mask attacks by evaluating the performances of these systems with/without  
196 countermeasure. For micro-texture analysis applied inside the proposed  
197 countermeasures, which is explained in Section 4, images are resized into  
198  $64 \times 64$  as proposed in [20]. Thus, although this resizing parameter may  
199 reduce the baseline performance in 2D, since our aim in this study is not  
200 to report high baseline performance, we preferred to use the same images as  
201 input for both the 2D baseline evaluation and proposed countermeasures.

202 Figure 1 shows an example for the texture images and 3D scans which  
203 are used in 2D and 3D evaluations, respectively.

### 204 *3.2. Selected Face Recognition Systems*

205 The 3D FR system selected as baseline for this study is introduced in [10].  
206 It is also selected as baseline system in TABULA RASA project [28]. It uses  
207 the pre-processed 3D mesh of the face as input. Three landmark points are  
208 previously annotated at the nose tip and outer eye corners for each sample  
209 in the database. Initially, a linear transformation is computed in a least  
210 squares sense (LSS), based on two sets of landmarks (landmarks of generic  
211 model and subject’s face). Least squares means that the overall solution  
212 minimizes the sum of the squares of the errors made in the results of every  
213 single equation. The best fit in the LSS here is calculated by minimizing  
214 the squared distance between the point sets of generic model and subject’s  
215 face. For this purpose, the obtained transformation that includes rotation,  
216 translation and isotropic scaling is applied onto the generic model, aligning it

217 with the subject's face. Next, the alignment is further improved by Iterative  
 218 Closest Point (ICP) method [5]. Afterwards, 140 previously selected points  
 219 on the generic model are coupled with the closest vertices on the face under  
 220 analysis, and Thin Plate Spline (TPS) [6] warping is applied on the generic  
 221 model resulting in warping parameters (WP) of size  $140 \times 3$ . WPs that  
 222 represent the deviations from the common structure are given to the classifier  
 223 for recognition. Finally, the distance between two face models is computed  
 224 by taking the median of cosine distances between the corresponding feature  
 225 vectors (WP), and recognition rates are computed. Figure 2 shows the feature  
 226 extraction scheme on a sample model using this method, which is named WP.

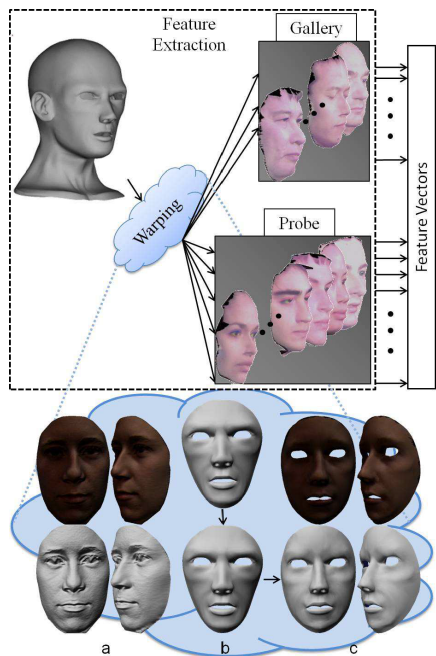


Figure 2: The feature extraction scheme and an illustration on a sample model: (a) The subject's face with and without texture (b) generic model before and after alignment (c) generic model after warping with and without texture. This figure is taken from [10].

227 For 2D face recognition, Local Binary Patterns (LBP) [1] is selected as  
228 baseline. The success of LBP in face description is due to the discriminative  
229 power, computational simplicity of the operator, and its robustness to mono-  
230 tonic gray scale changes caused by, for example, illumination variations. The  
231 use of histograms as features also makes the LBP approach robust to face  
232 misalignment and pose variations to some extent. For 2D FR, we use the  
233 operator  $LBP_{8,2}^{u2}$  on  $8 \times 8$  blocks. The similarity between each image pair is  
234 computed with chi-square distance metric. Performances are evaluated using  
235 the similarity scores between image pairs.

### 236 3.2.1. Evaluation on 2D and 3D Face Recognition

237 In this part, the evaluations are done for 2 modes. The first mode is  
238 the baseline mode: a standard biometric system with no spoofing and no  
239 countermeasure. The baseline performance is evaluated using DB-r. Perfor-  
240 mance is evaluated by verification all vs. all. Access from every identity in  
241 DB-r is tested against all other models in DB-r. The performance is mea-  
242 sured by observing the rate of users rejected when authenticating against  
243 their own template (False Rejection Rate - FRR) and by the rate of users  
244 accepted when authenticating against someone else's template (False Accep-  
245 tance Rate - FAR). The second mode is the evaluation of FR systems under  
246 mask attacks (baseline under attacks in Figure 3). Both DB-r and DB-m  
247 are used. When spoofing attacks are applied, performance is expected to  
248 degrade. In this mode, the FAR corresponds to the rate of attacks that are

249 accepted by the system when spoofed. The FRR corresponds to the rate of  
 250 real-access attempts that are incorrectly dismissed by the system as attacks.

251 For the evaluations regarding 2D and 3D FR systems here, only test set  
 252 is used. Train set is used for classifier training inside the proposed counter-  
 253 measures. Figure 3 shows the behavior of the 3D and 2D baseline systems  
 254 with/without attacks. All results are presented in terms of detection error  
 255 trade-of (DET) profiles which illustrate the behavior of a system as the de-  
 256 cision threshold is changed, i.e. how the false rejection rate varies according  
 257 to the false acceptance rate.

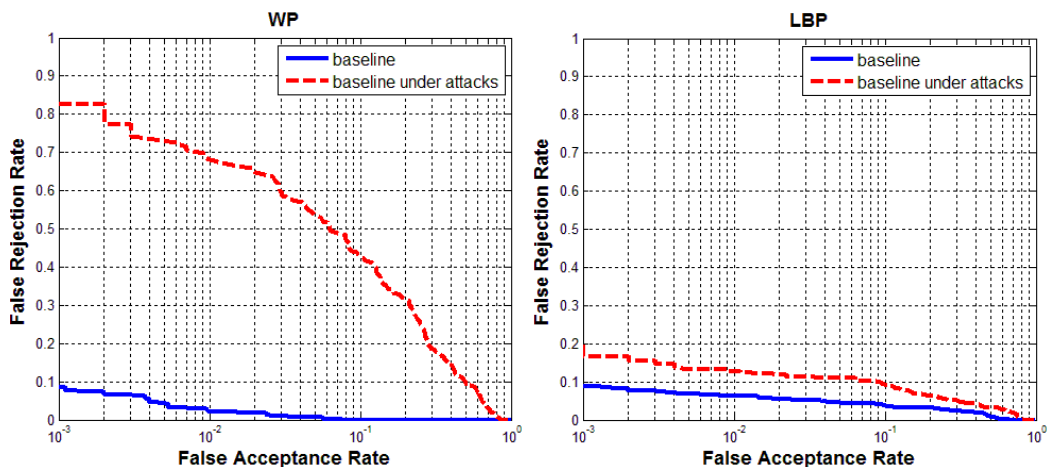


Figure 3: The DET Curves of the 3D and 2D face baseline biometric system with/without mask attacks, respectively.

258 Figure 3 shows that:

- 259 • Although the mask attacks are successful to spoof both 2D and 3D  
 260 FR systems, the 3D FR system is more vulnerable to mask attacks  
 261 compared to the 2D FR system (area between red and blue curves is

- 262 much more for 3D compared to 2D FR system).
- 263 • Equal Error Rate (EER) at the baseline mode increases from 1.8% to  
264 25.1% for 3D and from 4.7% to 9.9% for 2D FR system under attacks.
  - 265 • 3D shape characteristics of a real face and corresponding mask at-  
266 tack are more similar compared to their texture characteristics. Hence,  
267 analysis on texture may reveal more information to detect mask attacks  
268 compared to analysis on 3D shape characteristic.
  - 269 • Robustness against mask spoofing is observed to be both method and  
270 modality dependent as also concluded in [16].
  - 271 • FR systems are vulnerable to spoofing mask attacks hence, counter-  
272 measures are necessary to reduce their impact on face recognition.

273 For the baseline mode evaluations, we used the test set of DB-r, which  
274 contains 100 real samples from 10 subjects. In this study, we also report  
275 the baseline performances of the selected systems on the Face Recognition  
276 Grand Challenge Database (FRGC) v1.0 [25] database in order to check  
277 if the selected systems still provide satisfactory baseline performances with  
278 more number of subjects. The scans of the subjects in the mask database  
279 were acquired with a high quality laser scanner (technology of MORPHO).  
280 The FRGC database was also prepared using the high quality laser scanner  
281 Minolta. Therefore, the scan quality in the FRGC is quite similar to the scan  
282 quality in our mask database. Furthermore, FRGC v1.0 includes 943 samples



283 from 275 subjects and more challenging compared to the DB-r of the mask  
 284 database. Table 1 shows the EER, verification rate at 0.001 FAR and rank-1  
 285 identification rates computed with the selected systems using both the mask  
 286 database (the DB-r of the mask database) and the FRGC database.

Table 1: EER, Verification Rate at 0.001 FAR and Rank-1 Identification Rate (IR) for the 3D and 2D baseline systems using the Mask Database (MD) and the FRGC Database.

<b>Techniques</b>	<b>WP</b>		<b>LBP</b>	
	(MD)	(FRGC)	(MD)	(FRGC)
EER (%)	1.79	2.41	4.68	2.96
VR at 0.001 FAR (%)	91.33	87.70	90.89	90.03
IR (%)	100	94.01	98.89	94.50

287 Table 1 shows that slightly better performances are obtained in terms of  
 288 identification and verification using the mask database compared to the ones  
 289 obtained using the FRGC. For each FR technique, EER computed for the two  
 290 databases are quite similar. Although there is a high increase in the number  
 291 of subjects/samples when the FRGC is used for the evaluation, the perfor-  
 292 mances of the selected baseline systems on the FRGC are still satisfactory,  
 293 even quite similar to the results obtained using the mask database. These  
 294 results show that the selected systems provide significant performances hence  
 295 they are appropriate for this study, and the number of subjects/samples in  
 296 the mask database is sufficient enough to obtain consistent results in this  
 297 study.

## 298 4. TECHNIQUES USED INSIDE THE PROPOSED COUNTER- 299 MEASURES

300 Mask attack is a 3D attack that can be used to spoof both 2D and 3D  
301 FR systems. Most of the existing 3D scanners do not provide only 3D scan,  
302 they also capture texture image. Figure 1 shows an example for the two  
303 outputs of a scanner. Thus, when there is only one camera for 2D and one  
304 scanner for 3D FR system, a countermeasure which uses texture images as  
305 input can be used to protect both 2D and 3D FR systems if texture images  
306 are provided as default output of a scanner. In the present study, we propose  
307 four countermeasures against 3D mask attacks, which use either the depth  
308 maps or texture images as input (Figure 5).

309 In this section, we first explain the pre-processing applied for the proposed  
310 countermeasures. Then, we give a detailed information about the techniques  
311 that were used to develop the proposed countermeasures.

### 312 *4.1. Pre-Processing for the Countermeasures*

313 There are slight alignment differences between faces in the mask database.  
314 For the countermeasures, initially, all 3D faces in DB-r3 and DB-m3 are  
315 aligned to a generic face using LSS alignment, which makes the alignment of  
316 all faces identical.

317 In this study, we want to benefit from the information that the mask  
318 surface is smoother than the real face surface to detect mask attacks. There-  
319 fore, the depth maps are estimated from the raw aligned 3D scans. Next, 2D

320 cropping is applied to extract face region from both the texture images and  
321 depth maps. Then all images are resized into  $64 \times 64$  grayscale image.

322 In our previous studies [15, 17], we used normalized images. We notice  
323 that normalization has a positive impact in performances when the counter-  
324 measure is applied on depth maps, whereas it reduces performances slightly  
325 when applied on texture images. In the present paper, we did not apply nor-  
326 malization, and we improved our cropping code compared to the ones used  
327 in our previous studies. The final version of the texture images and depth  
328 maps used for the proposed countermeasures are shown in the second and  
329 fourth columns of Figure 1, respectively.

#### 330 *4.2. The Variational Retinex Algorithm*

331 In the present paper, the image is decomposed into reflectance and il-  
332 lumination components using the variational retinex algorithm explained in  
333 the studies [2, 12].

334 In this subsection, we first give information about minimizing energy  
335 functions. Then, we explain the variational retinex algorithm [2, 12].

##### 336 *4.2.1. Minimizing Energy Functions*

337 The concept of minimizing the energy of a given system is used in image  
338 processing. Minimizing energy functions often includes solving partial differ-  
339 ential equations, more specifically, Euler-Lagrange differential equations.

340 In the Euler-Lagrange problem, we usually have a continuous real-valued  
341 function  $y = f(x)$  with continuous derivative  $y' = df/dx$ . Considering  $x, y,$

342 and  $y'$  as three independent variables, a new function  $g(x, y, y')$  is defined.  
 343 Using this function, the energy function is defined as:  $E = \int g(x, y, y')dx$ .  
 344 The energy function  $E$  has a minimal value if Euler-Lagrange equation:

$$\frac{\partial g}{\partial y} - \frac{\partial}{\partial x} \left( \frac{\partial g}{\partial y'} \right) = 0 \quad (1)$$

345 is satisfied. The left hand side of this equation is denoted as  $\nabla E$ . Here  $f$  is  
 346 introduced as a function of one independent variable  $x$ , the same concept is  
 347 applied when  $f$  is a function of  $n$  independent variables:  $x_1, x_2, \dots, x_n$ . In  
 348 particular, when  $u = f(x, y)$ , function of two independent variables  $x$  and  $y$ ,  
 349 Euler- Lagrange equation becomes:

$$\nabla E = \frac{\partial g}{\partial u} - \frac{\partial}{\partial x} \left( \frac{\partial g}{\partial u_x} \right) - \frac{\partial}{\partial y} \left( \frac{\partial g}{\partial u_y} \right) = 0 \quad (2)$$

350 The variational retinex algorithm is developed by defining and minimizing  
 351 an energy function.

#### 352 4.2.2. The Variational Retinex Algorithm

353 An image can be considered as a two dimensional function  $S(x, y)$ , where  
 354  $(x, y)$  denotes a pixel on the image. The value of the function  $S = S(x, y)$   
 355 represents the intensity of the light at the pixel  $(x, y)$ . As stated in [2], the  
 356 intensity  $S$  may be characterized by two components which are;

- 357 • the amount of source illumination falling on the object, the illumination  
 358 component  $L(x, y)$ .

359 • the amount of illumination reflected by the object, the reflectance com-  
360 ponent  $R(x, y)$ .

361  $S(x, y)$  is computed using the illumination and reflectance components as  
362 shown in Eq. (3).

$$S(x, y) = L(x, y) \times R(x, y) \quad (3)$$

363 In [2], it is stated that if images are assumed to be composed of illumina-  
364 tion and reflectance components, generating the retinex effect means being  
365 able to separate one component from another. A first step taken by most  
366 algorithms in such sort of problems is the conversion to the logarithmic do-  
367 main by  $s = \log(S)$ ,  $l = \log(L)$ , and  $r = \log(R)$ . In the logarithmic domain  
368 the relation between these three images becomes:  $s = l + r$ .

369 In [12], Kimmel et al. make the following assumptions:

- 370 1. The logarithmic illumination  $l$  varies spatially smoothly.
- 371 2. The logarithmic reflectance  $r$  consists of constant or smooth parts and  
372 discontinuous jump parts.
- 373 3.  $l$  is greater than or equal to the logarithmic intensity  $s$  ( $l \geq s$ ).
- 374 4.  $l$  is close to  $s$  (i.e.  $l$  does not deviate far away from  $s$ ).

375 Based on the assumptions listed above, in the studies [2, 12], the energy  
376 function is defined as follows:

$$E(l) = \int (|\nabla l|^2 + \alpha(l - s)^2 + \beta|\nabla(l - s)|^2) dx dy \quad (4)$$

377 where  $\alpha$  and  $\beta$  are positive constants. Since  $S$  is the given image,  $s$  here is  
378 constant. In this equation;

- 379 • The first penalty term ( $|\nabla l|^2$ ) forces spatial smoothness on  $l$ .
- 380 • The second penalty term  $(l - s)^2$  forces a proximity between  $l$  and  $s$ .  
381 The difference between these images is exactly  $r$ , which means that the  
382 norm of  $r$  should be small. Simultaneously, it forces the solution  $l$  to  
383 be  $l \geq s$ . In [12], it is stated that in practice this term should be weak  
384 enough not to attract  $l$  down too much towards  $s$ . This is why the  
385 parameter  $\alpha$  should be very small.
- 386 • The third term forces  $r$  to be spatially smooth. In [12], it is stated  
387 that the parameter  $\beta$  should be a very small value to preserve the  
388 discontinuous jumps of  $r$ . Note that spatially smooth  $r$  contradicts  
389 spatially smooth  $l$  since  $r + l = s$ . However in practice adding this  
390 penalty term kicks in mainly on sharp edges and handles situations  
391 where the illumination is not smooth (as well as cases of direct light  
392 sources and specularities).

393 The integrand of this energy function is:

$$\begin{aligned} g(l, l_x, l_y) &= |\nabla l|^2 + \alpha(l - s)^2 + \beta|\nabla(l - s)|^2 \\ &= (l_x^2 + l_y^2) + \alpha(l - s)^2 + \beta((l_x - s_x)^2 + (l_y - s_y)^2) \end{aligned} \tag{5}$$

394 Euler-Lagrange equation becomes:

$$\begin{aligned}
\nabla E &= \frac{\partial g}{\partial l} - \frac{\partial}{\partial x} \left( \frac{\partial g}{\partial l_x} \right) - \frac{\partial g}{\partial y} \left( \frac{\partial g}{\partial l_y} \right) \\
&= 2\alpha(l - s) - \frac{\partial}{\partial x} (2l_x + 2\beta(l_x - s_x)) - \frac{\partial}{\partial y} (2l_y + 2\beta(l_y - s_y)) \\
&= 2\alpha(l - s) - 2l_{xx} - 2\beta(l_{xx} - s_{xx}) - 2l_{yy} - 2\beta(l_{yy} - s_{yy}) \\
&= 2 [\alpha(l - s) - \Delta l - \beta\Delta(l - s)] \\
&= 0
\end{aligned} \tag{6}$$

395 which means  $\alpha(l - s) - \Delta l - \beta\Delta(l - s) = 0$ . In [2], to solve this equation, the  
396 idea of the steepest descent is applied with an auxiliary variable  $t$ :

$$\frac{dl}{dt} = -\nabla E = \Delta l + \beta\Delta(l - s) - \alpha(l - s) \tag{7}$$

397 To find a local minimum of a function using steepest descent, one takes  
398 steps proportional to the negative of the gradient of the function at the  
399 current point. In our evaluation,  $l$  is computed via steepest descent as follows:

$$l_n = l_{n-1} - dt \cdot \nabla E \tag{8}$$

400 Finally, projecting onto the constraint  $l \geq s$  is done by  $l_n = \max(l_n, s)$ .

401 In our experiments, the values 0.0001 and 0.1 are used for  $\alpha$  and  $\beta$ ,  
402 respectively, as suggested in [2, 12]. The initial value of  $l$  ( $l_0$ ) is taken as  
403  $s$ . The step size  $dt$  and the total number of iterations are selected as 0.05  
404 and 5000, respectively. After the iterations, optimum  $l$  is obtained, and

405  $r$  is computed from  $r = s - l$ . Finally, the reflectance and illumination  
 406 components are evaluated from  $R = e^r$  and  $L = e^l$ , respectively.



Figure 4: Example from the mask database which is created by [23] (a) The real face with texture, the reflectance image and the illumination image of the real face (b) Same images associated with the mask of the same person.

407 Figure 4 shows an example from the mask database for a real face and  
 408 corresponding mask attack. First column shows the original images, sec-  
 409 ond column and third column show the reflectance and illumination images,  
 410 respectively, which are computed using the variational retinex algorithm.

#### 411 4.3. *Micro-Texture Analysis Technique*

412 The micro-texture analysis, which was first proposed in [20] to detect 2D  
 413 face attacks, is used to detect 3D mask attacks here. In [20], it is applied  
 414 on texture images, whereas in this paper, we apply this technique not only  
 415 on texture images but also on depth maps estimated from 3D scans and on  
 416 reflectance components of texture images.



417 This LBP based micro-texture analysis technique emphasizes the micro-  
418 texture differences in the feature space. It aims at learning the differences  
419 between real and fake face, and designs a feature space which emphasizes  
420 those differences. The original LBP forms labels for the image pixels by  
421 thresholding the  $3 \times 3$  neighborhood of each pixel with the center value and  
422 considering the result as a binary number. The LBP operator has been  
423 extended to use neighborhoods of different sizes.  $LBP_{P,R}$  is computed such  
424 that for a given central pixel in an image, a pattern number is computed by  
425 comparing its value with those of its neighbors. In Eq. (9),  $g_c$  is the gray  
426 value of the central pixel,  $g_p$  is the value of its neighbors,  $P$  is the number of  
427 neighbors around a circle of radius  $R$ .  $LBP_{P,R}$  calculation is shown in Eq.  
428 (9) and (10):

$$LBP_{P,R} = \sum_{p=0}^{P-1} s(g_p - g_c)2^p \quad (9)$$

429

$$s(x) = \begin{cases} 1, & x \geq 0 \\ 0, & x < 0 \end{cases} \quad (10)$$

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

435 label.

436 In [20], authors claim that micro-texture details that are needed to dis-  
437 criminate a real face from face print can best be detected using combination  
438 of different LBP operators. Thus, they derive an enhanced facial repre-  
439 sentation using multi-scale LBP operators. Their proposed representation  
440 computes LBP features from  $3 \times 3$  overlapping regions to capture the spatial  
441 information and enhances the holistic description by including global LBP  
442 histograms computed over the whole image. This is done as follows: the face  
443 is cropped and resized into a  $64 \times 64$  pixel image. Then,  $LBP_{8,1}^{u2}$  operator is  
444 applied on the face image and the resulting LBP image is divided into  $3 \times 3$   
445 overlapping regions (with an overlapping size of 14 pixels). The local 59-bin  
446 histograms from each region are computed and collected into a single 531-bin  
447 histogram. Then, two other histograms are computed from the whole face  
448 image using  $LBP_{8,2}^{u2}$  and  $LBP_{16,2}^{u2}$  operators, yielding 59-bin and 243-bin his-  
449 tograms that are added to the 531-bin histogram previously computed. In  
450 [20], the length of the final enhanced feature histogram is reported as 833  
451 (i.e.  $531 + 59 + 243$ ).

#### 452 4.4. Classification Technique

453 Mask face detection is a two-class classification problem. Since SVM [8]  
454 are proven to be a powerful tool for discriminating two classes of data, we  
455 adopted an SVM classifier for this purpose. SVM finds the maximum margin  
456 hyper-plane to separate the training data in feature space and a decision for

457 a new test data  $x$  is classified. In our experiments, we adopted linear kernel  
 458 since our feature vectors are high-dimensional and are hence likely to be  
 459 linear separable.

## 460 5. THE PROPOSED COUNTERMEASURES

461 Four countermeasures are proposed in this study to discriminate mask  
 462 and real samples. Three of them use the 2D data (texture images), and the  
 463 remaining one uses the 3D data (depth maps estimated from the raw 3D  
 464 scans) available in the mask database as input.

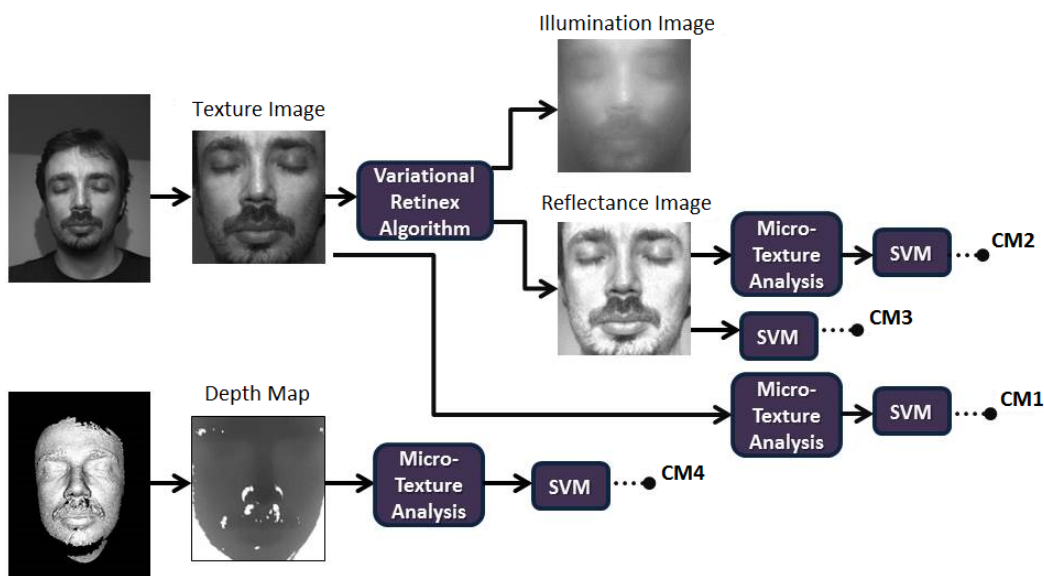


Figure 5: The flowchart of the proposed countermeasures.

465 The flowchart of the countermeasures proposed in this paper are shown in  
 466 Figure 5. In this figure, the micro-texture analysis (explained in Subsection  
 467 4.3) applied on texture images is called CM1, applied on reflectance images

468 is called CM2, applied on depth maps is called CM4, and finally the counter-  
469 measure for which the pixel intensity values on reflectance images are used  
470 directly by the classifier is called CM3 (CM denotes countermeasure).

471 CM1 and CM4 are first introduced in our study [15], and CM3 is first  
472 introduced in our study [17]. In the present paper, we provide an overview on  
473 our spoofing detection approaches introduced in the studies [15, 17]. We ex-  
474 tend the works explained in these studies with some improvements (e.g. bet-  
475 ter cropping, usage of non-normalized images instead of normalized images),  
476 additional analysis, comparisons of performances of diverse countermeasures  
477 using the same protocol, and with a detailed analysis of the fusion scenarios.  
478 Also, CM2 is first introduced in the present paper. It is a new counter-  
479 measure providing very satisfactory accuracy to classify mask and real faces.  
480 The results of CM2 show that reflectance component of an image provides  
481 more appropriate information than original image itself for mask detection.  
482 From the fusion results, we also observed that it provides complementary  
483 information on mask detection.

#### 484 5.1. CM1: Micro-Texture Analysis on Texture Images

485 Captured image from mask may visually look very similar to the image  
486 captured from live face (e.g. the texture images in the first column of Fig-  
487 ure 4). A close look at the differences between faces in DB-r2 and DB-m2  
488 reveals that their surface properties are different. For mask manufacturing  
489 3D printers are used, hence they may contain printing quality defects that

490 can be detected with micro-texture patterns. For CM1, micro-texture anal-  
491 ysis is applied on texture images, and the feature histogram of length 833 is  
492 obtained. Finally, linear SVM classifier is applied to detect mask and real  
493 faces.

### 494 *5.2. CM2: Micro-Texture Analysis on Reflectance Images*

495 For CM2, initially, the illumination and reflectance components (Figure  
496 5) of the texture images are obtained using the variational retinex algorithm  
497 introduced in Subsection 4.2. Then, micro-texture analysis is applied on  
498 reflectance components of texture images rather than texture images itself.  
499 The reason of this analysis on reflectance images is that a close look at the  
500 differences between the reflectance images of the real and mask faces reveals  
501 that the texture characteristics on their reflectance components are also dif-  
502 ferent. The feature vectors of length 833, which are obtained by applying  
503 micro-texture analysis on reflectance images, are used as input by linear SVM  
504 classifier. This feature vector gives information from the reflectance image  
505 in the image texture level.

### 506 *5.3. CM3: Pixel Intensity Analysis on Reflectance Images*

507 Our observations on the reflectance components of mask and real faces  
508 reveal that reflectance characteristics of mask and real face samples are dif-  
509 ferent especially at some specific regions of the face (eyelashes, eyebrows and  
510 moustache). Based on these observations, in this study, we use the intensity

511 values on reflectance component of each image as input for linear SVM clas-  
512 sifier. Since the intensity values on reflectance images are between 0 and 1  
513 ( $R(x, y) \in [0, 1]$ ), we stretched it to the interval  $[0, 255]$  by multiplying  $R$   
514 with 255. The reflectance component, which is in the size of  $64 \times 64$  pixel  
515 image, is reshaped as  $[1 \ 4096]$  ( $64 \times 64 = 4096$ ). The resultant vector is  
516 the feature vector providing information in the pixel intensity level. Finally,  
517 linear SVM classifier is applied to detect real and mask faces.

#### 518 *5.4. CM4: Micro-Texture Analysis on Depth Maps*

519 The 3D shape of high quality mask is also very similar to the 3D shape  
520 of corresponding real face (e.g. the 3D scans in the second column of Figure  
521 1). Our analysis on DB-r3 and DB-m3 show that the mask scan is smoother  
522 than the real face scan. Especially the parts of the face with facial hair are  
523 quite different. Since there is no real facial hair (e.g. mustache, eyebrow) on  
524 the masks, the 3D scan of mask is smoother in these parts compared to the  
525 real face scan. When high quality scanners are used for acquisition, although  
526 there is a decrease in the number of holes, it is still possible to observe some  
527 holes on the scan especially at the parts of the face with facial hair. Thus, in  
528 our study, micro-texture analysis is also applied on the depth maps which are  
529 estimated from the raw 3D scans, and the other feature histogram of length  
530 833 is obtained. Finally, linear SVM classifier is applied to detect real and  
531 mask faces.

## 532 6. EXPERIMENTS AND RESULTS

533 In this section, we first show the stand-alone classification performances  
534 of the proposed countermeasures together with the fusion scenarios. Then,  
535 we integrate the countermeasure providing the best performance to the se-  
536 lected 3D FR system in order to observe the improvement in the recognition  
537 performance of the system in presence of mask attacks.

### 538 *6.1. Stand-Alone Classification Performances of the Countermeasures*

539 In the present study, we apply the proposed countermeasures (CM1, CM2,  
540 CM3, CM4) using the same database with the same train-test sets, hence an  
541 exact comparison between these countermeasures is possible. Train set is  
542 used for classifier training. This classifier is subject to two kind of errors:

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

547 The lower these two errors, the better the performance of the counter-  
548 measures. In this section, we first evaluate the performances of the single  
549 countermeasures, and then evaluate the performances for the fusion scenar-  
550 ios. The Region of Convergence (ROC) curves in Figure 6 shows the stand-  
551 alone classification performances of the four countermeasures together with  
552 the fusion based countermeasure providing the best performance in Table 3.

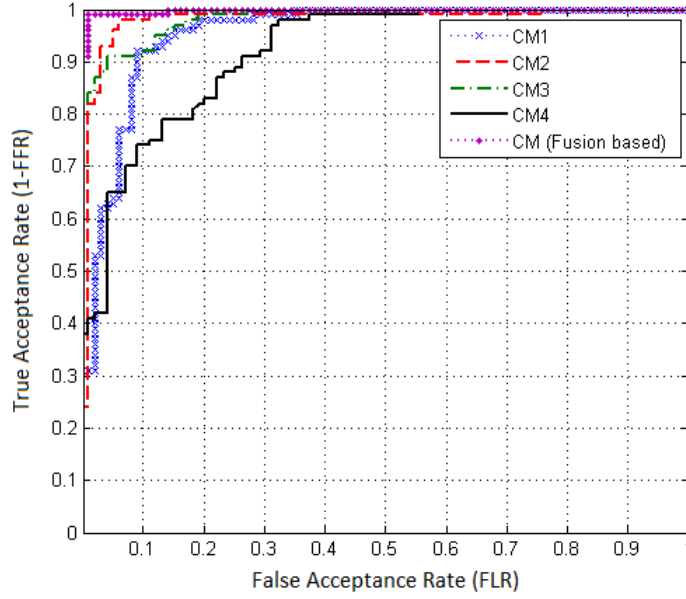


Figure 6: The Classification Performances of the Countermeasures.

553 Area Under Curve (AUC), EER and best accuracy results using CM1,  
 554 CM2, CM3, and CM4 are shown in Table 2.

Table 2: AUC, EER and Accuracy Results Using the Four Countermeasures

Countermeasures	AUC	EER(%)	Accuracy(%)
CM1	0.956	9.04	91.46
CM2	0.980	<b>5.02</b>	<b>95.98</b>
CM3	<b>0.984</b>	9.04	93.47
CM4	0.919	18.59	82.91

555 Table 2 and Figure 6 show that;

- 556 • The best performances in terms of EER and accuracy are obtained  
 557 using CM2, and the best performance in terms of AUC is obtained  
 558 using CM3.



- 559 • The best performances are obtained with the countermeasures based  
560 on reflectance analysis (CM2 and CM3) compared to the performances  
561 obtained with CM1 and CM4. This shows that reflectance character-  
562 istics of the real and mask faces in the mask database provide more  
563 appropriate information than their texture and smoothness character-  
564 istics.
- 565 • CM4, which is based on smoothness analysis, provides worse results  
566 compared to the other countermeasures. However, the performance of  
567 CM4 can be still considered as satisfactory.
- 568 • 2D data (texture images) provide more information than 3D data (depth  
569 maps) to detect mask spoofing.

570 After evaluating the performances of the single countermeasures, we an-  
571alyze the performances for the fusion scenarios.

572 For feature level fusion, the feature histograms computed from different  
573 types of images (texture, reflectance and depth) are concatenated and the  
574 classifier is applied on the resultant feature histogram. In Table 3, the fea-  
575 ture level fusion of 2 countermeasures, 3 countermeasures and finally all the  
576 4 countermeasures are reported for which the length of the final feature his-  
577 tograms are 1666, 2499 and 3332, respectively. Once the enhanced histogram  
578 is computed, a linear SVM classifier is used to determine whether the image  
579 corresponds to a live face or not.

580 For score level fusion, linear SVM classifier is applied using the features  
 581 computed from each type of images (texture, reflectance and depth) sepa-  
 582 rately, and then Z-score normalization is applied for each score group. Fi-  
 583 nally, the weighted score level fusion is used for combining the outputs of the  
 584 individual SVMs to determine whether the image corresponds to a live face  
 585 or not.

586 AUC, EER and best accuracy results are shown in Table 3 for the fusion  
 587 scenarios.

Table 3: AUC, EER and Accuracy Results for the Fusion Scenarios of the Proposed Countermeasures

Countermeasures Involved in Fusion	Feature Level Fusion			Score Level Fusion		
	AUC	EER (%)	Acc. (%)	AUC	EER (%)	Acc. (%)
CM1, CM2	0.994	3.01	97.49	0.988	3.01	97.49
CM1, CM3	0.985	9.04	93.97	0.993	5.02	96.48
CM1, CM4	0.972	9.04	92.96	0.976	8.04	94.47
CM2, CM3	0.984	9.04	93.47	0.993	5.02	96.48
CM2, CM4	0.994	3.01	97.49	0.992	4.02	96.98
CM3, CM4	0.986	8.04	92.96	0.994	5.02	95.98
CM1, CM2, CM3	0.985	9.04	93.47	<b>0.998</b>	<b>2.01</b>	<b>98.99</b>
CM1, CM2, CM4	<b>0.998</b>	<b>2.01</b>	<b>97.99</b>	0.993	5.02	96.48
CM1, CM3, CM4	0.987	8.04	93.47	0.997	2.01	97.99
CM2, CM3, CM4	0.986	8.04	93.47	0.997	2.00	98.49
CM1, CM2, CM3, CM4	0.987	8.04	93.47	0.997	2.01	98.49

588 From the reported results in Table 3, we can remark the followings:

- 589 • Both score and feature level fusion of the countermeasures improve the  
 590 performance compared to using single countermeasure. For instance,  
 591 CM4 provides a detection accuracy of 82.9% whereas when CM4 and

592 CM1 are fused, the accuracy is improved to 92.96% for feature and  
593 94.47% for score level fusion. This proves that when both the texture  
594 images and depth maps are provided by 3D scanners, more robust  
595 countermeasures can be obtained by fusion.

596 • For feature level fusion, the best performances are obtained by the  
597 fusion of CM1, CM2 and CM4. In this part, we observed that when  
598 we concatenate the same type of features (micro-texture features of  
599 length 833 for each of CM1, CM2 and CM4), we observe a significant  
600 increase in the performance. In CM3, the features are pixel intensity  
601 values (features of length 4096). Therefore, when we apply feature level  
602 fusion using CM3 with the other countermeasures, the positive impact  
603 of CM3 in the performances was not observable as shown in Table 3.

604 • For score level fusion, the best performances are obtained by the fu-  
605 sion of CM1, CM2 and CM3. All these countermeasures (CM1, CM2  
606 and CM3) uses texture images as input (reflectance image is computed  
607 from texture image). This proves that 2D data provides very beneficial  
608 information for mask spoofing detection.

609 • CM3 increases the performances when it is used in score level fusion,  
610 whereas the impact of it in feature level fusion is not remarkable.

611 • Although CM1, CM2 and CM3 provide very satisfactory results alone,  
612 the score level fusion of these countermeasures provides the best perfor-  
613 mance compared to all other scenarios in Table 2 and 3. Therefore, in

614 Figure 6, the ROC curve of this fusion based countermeasure is shown  
615 as the best one.

- 616 • Since existing 3D scanners provide both 3D scan and corresponding  
617 texture image, more robust countermeasures can be developed by fusion  
618 of these two type of outputs (2D and 3D data).

### 619 *6.2. Integration of the Countermeasure to Face Recognition Systems*

620 In this subsection, we integrate the countermeasure with the best per-  
621 formance (fusion of CM1, CM2 and CM3) to the 3D FR system selected as  
622 baseline.

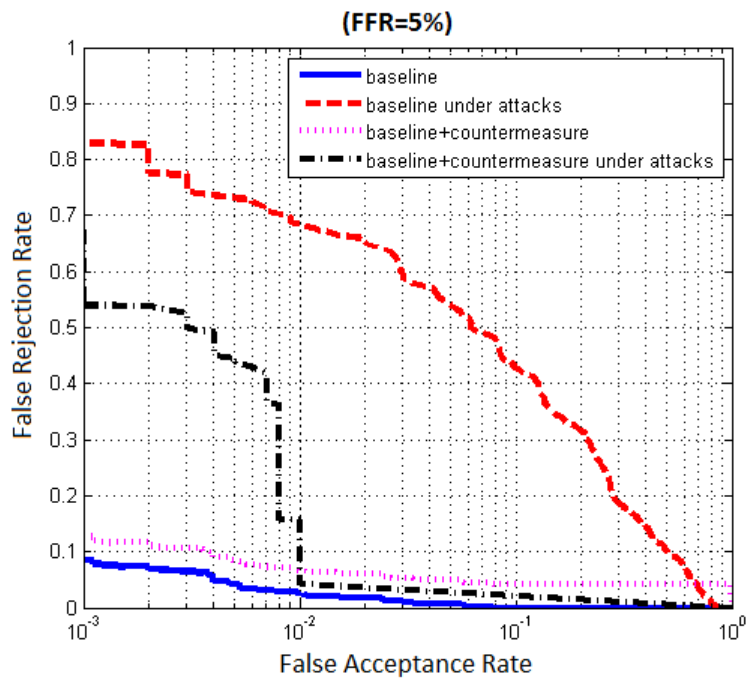
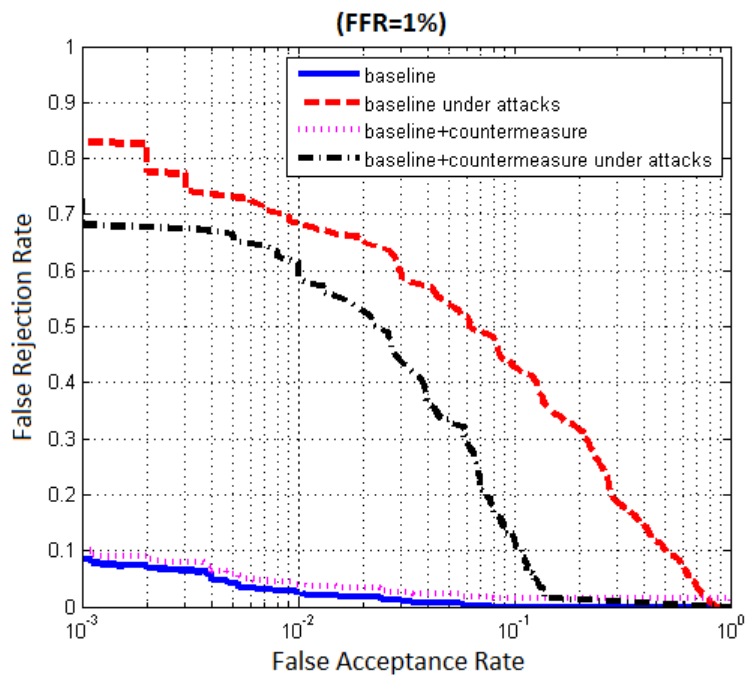
623 The evaluations are done for 4 modes. The first two modes are the base-  
624 line mode and the mode under attacks, which are explained in Subsection  
625 3.3. The third mode illustrates performance when the countermeasure is  
626 applied against the attacks, that results in an improved performance with  
627 respect to the second mode. For the samples which are detected as attack  
628 by the countermeasure, a least similarity score, which is zero in this test,  
629 is assigned to those samples in verification tests. Last mode evaluates the  
630 performance of the baseline system together with the countermeasure in the  
631 normal operation mode of system, i.e., without attacks. The inclusion of the  
632 countermeasure may degrade the baseline performance when not confronted  
633 to attack (e.g. the countermeasure may consider as fake some real users.).

634 For evaluations, we fix 3 different evaluation points at  $\text{FFR} = 1\%$ ,  $5\%$ ,  
635 and  $10\%$  (FFR and FLR were defined in the previous subsection). Once

636 fixed, we incorporate the countermeasure as a first step into the baseline bio-  
637 metric systems oriented to discard fake data, and generate the performance  
638 evaluations for the 4 modes explained above.

639 Figure 7 shows the behavior of the 3D face baseline system with/without  
640 attacks and with/without the countermeasure. The three figures represent  
641 the overall system performance under spoofing attacks when three differ-  
642 ent operating points ( $\text{FFR} = 1\%$ ,  $5\%$ , and  $10\%$ ) are used for adjusting the  
643 countermeasure.

644 It is clear from Figure 7 that the 3D FR system is vulnerable to mask  
645 attacks (more area between blue and red curves indicates more vulnerability  
646 to the attacks). Performance enhancement is obtained almost all regions of  
647 DET plots in Figure 7 when the countermeasure is introduced in presence of  
648 mask attacks (black curve compared to red curve). If we take an operating  
649 point where  $\text{FFR} = 1\%$ , then  $\text{FRR}$  of the 3D FR system under attacks  
650 drops from around  $65\%$  to around  $50\%$  at  $\text{FAR} = 2\%$ . For both of the two  
651 other plots (at  $\text{FFR} = 5\%$  and  $10\%$ ), the introduction of the countermeasure  
652 lowers  $\text{FRR}$  from around  $65\%$  to  $4\%$  and  $7\%$ , respectively, at  $\text{FAR} = 2\%$ .  
653 The performance of the countermeasure is observed to be better at  $\text{FFR} =$   
654  $5\%$  compared to the cases at  $\text{FFR} = 1\%$  and  $10\%$ . Finally, the inclusion  
655 of the countermeasure improves the results of 3D FR system under attacks,  
656 whereas it degrades baseline performances of the system when not confronted  
657 to attack (pink curve compared to blue curve).



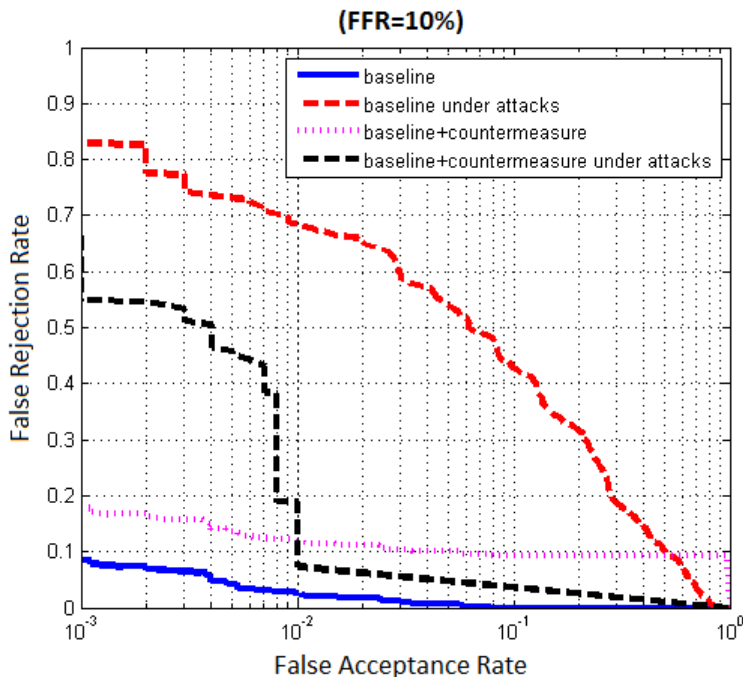


Figure 7: The DET Curves of the 3D face baseline biometric system when integrating the countermeasure.

658 **7. CONCLUSIONS**

659 In this study, a 2D+3D mask attack database is used to evaluate the per-  
 660 formances of the proposed countermeasures for the protection of face recog-  
 661 nition systems against mask attacks.

662 The novelty of this study is that it is still one of the few studies that pro-  
 663 poses countermeasures against 3D mask attacks. The analysis are done on  
 664 depth maps, texture images and reflectance components of texture images,  
 665 and 4 different countermeasures are proposed. Three of the proposed coun-  
 666 termeasures use 2D data (texture images), and the remaining one uses 3D  
 667 data (depth images) as input. These countermeasures can be used to protect

668 both 2D and 3D FR systems against mask attacks. The results of this study  
669 show that analysis on reflectance images provide the best results compared  
670 to analysis on texture and depth images. All of the 4 countermeasures pro-  
671 vide satisfactory information hence can be used as independent sources to  
672 discriminate masks from real faces. However with the fusion of these coun-  
673 termeasures, we observe a significant improvement in the performances. For  
674 instance, in this paper, a classification accuracy of 99% (almost perfect ac-  
675 curacy) is achieved for real face vs. mask face by fusing the information  
676 extracted from the reflectance images and texture images.

677 Up to now, we have analyzed several characteristics of real and mask  
678 faces, and obtained almost perfect results on this mask database. The lim-  
679 itation of our study is that we were able to test the performances of the  
680 proposed countermeasures using the masks made from one type of material,  
681 which is polychrome mineral powder. When masks made from different ma-  
682 terials are used, we may obtain different performance accuracy. Our future  
683 works are first to test the performances of the proposed countermeasures us-  
684 ing masks made from different materials in order to observe if we can still  
685 obtain satisfactory results, and then to propose new countermeasures for  
686 more challenging mask databases with higher number of subjects as soon as  
687 available.

688 **Acknowledgment:** This work has been performed by the TABULA  
689 RASA project 7th Framework Research Programme of the European Union  
690 (EU), grant agreement number: 257289. The authors would like to thank



691 the EU for the financial support and the partners within the consortium for  
692 a fruitful collaboration. For more information about the TABULA RASA  
693 consortium please visit <http://www.tabularasa-euproject.org>.

## 694 **References**

- 695 [1] T. Ahonen, A. Hadid, M. Pietikainen, Face description with local bi-  
696 nary patterns: Application to face recognition, *IEEE Trans. on Pattern*  
697 *Analysis and Machine Intelligence* (2006) 2037–2041.
- 698 [2] N. Almoussa, Variational retinex and shadow removal, Technical Report,  
699 The Mathematics Department, UCLA (2008).
- 700 [3] J. Bai, T.T. Ng, X. Gao, Y.Q. Shi, Is physics-based liveness detection  
701 truly possible with a single image?, *IEEE International Symposium on*  
702 *Circuits and Systems* (2010) 3425–3428.
- 703 [4] W. Bao, H. Li, N. Li, et al., A liveness detection method for face recogni-  
704 tion based on optical flow field, *Int. Conf. on Image Analysis and Signal*  
705 *Processing (IASP)* (2009) 233–236.
- 706 [5] P. Besl, N. McKay, A method for registration of 3-d shapes, *IEEE Trans.*  
707 *on Pattern Analysis and Machine Intelligence* (1992) 239–256.
- 708 [6] F.L. Bookstein, Principal warps: Thin-plate splines and decomposition  
709 of deformations, *IEEE Trans. Pattern Analysis and Machine Intelligence*  
710 (1989) 567–585.

- 711 [7] M.M. Chakka, A. Anjos, S. Marcel, et al., Competition on counter mea-  
712 sures to 2-d facial spoofing attacks, IEEE IAPR Int. Joint Conference  
713 on Biometrics (IJCB) (2011) 1–6.
- 714 [8] C.C. Chang, C.J. Lin, Libsvm : a library for support vector machines,  
715 ACM Transactions on Intelligent Systems and Technology (2011) 2:27:1–  
716 27:27.
- 717 [9] G. Chetty, M. Wagner, Multi-level liveness verification for face-voice  
718 biometric authentication, Biometrics Symposium: Special Session on  
719 Research at the Biometric Consortium Conference (2006) 1–6.
- 720 [10] N. Erdogmus, J.L. Dugelay, On discriminative properties of tps warp-  
721 ing parameters for 3d face recognition, IEEE Int. Conf. on Informatics,  
722 Electronics and Vision (ICIEV) (2012) 225–230.
- 723 [11] Y. Kim, J. Na, S. Yoon, J. Yi, Masked fake face detection using radiance  
724 measurements, Journal of the Optical Society of America A (2009) 760–  
725 766.
- 726 [12] R. Kimmel, M. Elad, D. Shaked, R. Keshet, I. Sobel, A variational  
727 framework for retinex, International Journal of Computer Vision (2003)  
728 7–23.
- 729 [13] K. Kollreider, H. Fronthaler, J. Bigun, Verifying liveness by multiple  
730 experts in face biometrics, IEEE Computer Society Conference on Com-  
731 puter Vision and Pattern Recognition Workshops (2008) 1–6.

- 732 [14] N. Kose, J.L. Dugelay, Classification of captured and recaptured images  
733 to detect photograph spoofing, IEEE Int. Conf. on Informatics, Elec-  
734 tronics and Vision (ICIEV) (2012) 1027–1032.
- 735 [15] N. Kose, J.L. Dugelay, Countermeasure for the protection of face recog-  
736 nition systems against mask attacks, IEEE Automatic Face and Gesture  
737 Recognition (FG) (2013) 1–6.
- 738 [16] N. Kose, J.L. Dugelay, On the vulnerability of face recognition systems  
739 to spoofing mask attacks, IEEE Int. Conf. on Acoustics, Speech, and  
740 Signal Processing (ICASSP) (2013) 2357–2361.
- 741 [17] N. Kose, J.L. Dugelay, Reflectance analysis based countermeasure tech-  
742 nique to detect face mask attacks, Int. Conf. on Digital Signal Processing  
743 (DSP) (2013) 1–6.
- 744 [18] N. Kose, J.L. Dugelay, Shape and texture based countermeasure to pro-  
745 tect face recognition systems against mask attacks, IEEE Computer  
746 Vision and Pattern Recognition Workshop, on Biometrics (CVPRW)  
747 (2013) 111–116.
- 748 [19] J. Li, Y. Wang, T. Tan, et al., Live face detection based on the analysis  
749 of fourier spectra, Proceedings of the SPIE (2004) 296–303.
- 750 [20] J. Maatta, A. Hadid, M. Pietikainen, Face spoofing detection from single  
751 images using micro-texture analysis, Proc. of IEEE Int. Joint Conf. on  
752 Biometrics (IJCB) (2011) 1–7.

- 753 [21] J. Maatta, A. Hadid, M. Pietikainen, Face spoofing detection from single  
754 images using texture and local shape analysis, IET Biometrics (2012)  
755 3–10.
- 756 [22] M.D. Marsico, M. Nappi, D. Riccio, J.L. Dugelay, Moving face spoofing  
757 detection via 3d projective invariants, IAPR Int. Conf. on Biometrics  
758 (2012) 73–78.
- 759 [23] MORPHO, <http://www.morpho.com/>, 2010.
- 760 [24] G. Pan, L. Sun, Z. Wu, S. Lao, Eyeblick-based antispoofing in face  
761 recognition from a generic webcam, IEEE Int. Conf. on Computer  
762 Vision (ICCV) (2007) 1–8.
- 763 [25] P.J. Phillips, P.J. Flynn, T. Scruggs, K. Bowyer, J. Chang, K. Hoff-  
764 man, J. Marques, J. Min, W. Worek, Overview of the face recognition  
765 grand challenge, IEEE Computer Society Conf. on Computer Vision and  
766 Pattern Recognition (2005) 947–954.
- 767 [26] PRINT-ATTACK Database, <http://www.idiap.ch/dataset/printattack>,  
768 2010.
- 769 [27] Sculpteo 3D Printing, <http://www.sculpteo.com/en/>, 2009.
- 770 [28] TABULA RASA Project, <http://www.tabularasa-euproject.org/>, 2010.
- 771 [29] Z. Zhang, D. Yi, et.al, Face liveness detection by learning multispectral

772 reflectance distributions, IEEE Automatic Face and Gesture Recogni-  
773 tion (FG) (2011) 436–441.