Egyptian Informatics Journal 19 (2018) 75–82

Contents lists available at ScienceDirect

Egyptian Informatics Journal

journal homepage: www.sciencedirect.com

Review

The detection of spoofing by 3D mask in a 2D identity recognition system

Bensenane Hamdan *, Keche Mokhtar

Laboratoire Signals and Images, Dept. of Electronique, Université des Sciences et de la Technologie d'Oran Mohamed Boudiaf, USTO-MB, BP 1505, 3100 Oran, Algeria

ARTICLE INFO

Article history:
Received 2 November 2016
Revised 29 July 2017
Accepted 4 October 2017
Available online 20 October 2017

Keywords:
Anti-spoofing
Angular Radial Transformation (ART)
Linear Discriminant Analysis (LDA)
Support vector Machine (SVM)
Nearest Neighbor Classifier (NNC)

ABSTRACT

Nowadays face recognition systems are facing a new problem after having won the challenge of reliability. The problem is that these systems have become vulnerable to attacks by identity theft. In order to deceive the recognition systems hackers use several methods, such as the use of face images or videos of people belonging to the system database. Luckily, this type of attack is thwarted by the use of adapted systems. But unfortunately another type of attack that uses 3D face masks appeared. This type of attack is very efficient, since as will be shown, a high percentage of hackers who use 3D masks can mislead a good facial recognition system, like the one used in our investigation. In this paper, a new method is proposed for the detection of hackers that use 3D masks to deceive face recognition systems. This method uses the Angular Radial Transformation (ART) to extract pertinent features that are fed into a classifier to decide whether the captured image represents a face image. The performance of the proposed method was evaluated using a public 3D Mask Attack Database (3DMAD). The obtained results show the efficiency of the proposed method, since it can reduce the error rate in discriminating between a real face and a face mask down to 0.90%.

Contents

1. Introduction . . . 76
2. The recognition phase . . . 77
   2.1. Extraction of characteristics . . . 77
   2.2. Classification for recognition . . . 77
3. The verification phase . . . 79
   3.1. Extraction of facial features using the ART . . . 79
   3.2. The ART projection basis . . . 79
   3.3. Classification in the verification step . . . 80
   3.4. Maximum likelihood classifier . . . 80
4. Experimentation . . . 80
   4.1. The 3DMAD database . . . 80
   4.2. Pretreatment of the images in the database . . . 80

* Corresponding author.
E-mail addresses: hamdan.bensenane@univ-usto.dz, bensenane1300@gmail.com (B. Hamdan), m_keche@yahoo.com (K. Mokhtar).

Peer review under responsibility of Faculty of Computers and Information, Cairo University.

Production and hosting by Elsevier

https://doi.org/10.1016/j.eij.2017.10.001

4.3. Protocol . . . 80
4.3.1. The recognition phase . . . 80
4.3.2. The verification phase . . . 81
5. Results and discussion . . . 81
5.1. Recognition results . . . 81
5.2. The verification results . . . 81
6. Conclusion . . . 82
References . . . 82

1. Introduction

Today security has become an international concern. Among the fields where the security is a concern, one can cite, as examples: access control to computers, e-commerce, identification based banking, public transport, etc.

A biometric system is essentially a pattern recognition system that uses biometric data of individuals. Depending on the context of the application, a biometric system may operate in the learning mode, verification mode or identification mode. The choice of using facial recognition as a biometric modality is motivated by the fact that it is contactless, natural, well accepted and requires only a very inexpensive sensor (Webcam) that is virtually available on all electronic devices. Furthermore, it requires a small cooperation from the users during the acquisition phase of the facial features.

Automatic face recognition involves two main steps: extraction of facial features and classification. Unfortunately, all the advantages of facial recognition systems have fallen into the water with easy pirating of facial characteristics. The experiments showed that hackers can easily fool facial recognition systems in the acquisition phase of facial features with a simple photo or video recording of the face.

In the case of identity theft by a picture, liveliness detection (eye blinking, facial micro movements...) can distinguish a real face from a picture and thus definitely neutralize this type of hacking.

For video based hacking, the usual approach to detect the attack is to analyze the motion in the scene by examining how objects move in front of the sensor. The movements of the planar objects like screens differ greatly from those of a true face.

Another method, proposed by Bai et al. [1], consists of finding printing artifacts and/or blurring of the texture of the face image to distinguish between a real face and a stroke. For the same purpose, Li et al. [2] proposed a technique based on the analysis of 2-D Fourier spectra. All these attack detection methods have failed to deal with identity usurpation with 3D mask. Indeed, blink detection of eyes and lips movements can be overcome simply by using high resolution printing masks of the eyes and regions of the mouth.

Several research works were carried out to distinguish between a real face and a face mask. The most commonly used approaches rely on distinguishing between the human skin and the facial mask material, thanks to the difference between their light reflecting factors. To this end, the reflectance disparity based on the albedo between facial skin and face mask materials (silicon, latex, etc.) is exploited.

In [3], a 2D characteristic vector composed of 2 radiance measurements under beams of light (685–850) nm is used to detect a fake face, through Linear Discriminant Analysis (LDA). An accuracy of 97.78% was reported. The fact that for mask detection, the measurements of radiation should be acquired at 30 cm on the forehead region, in addition the possibility of occlusion in the forehead and light range limitations, make this method impractical.

Similarly, Zhang et al. [4] proposed a multi-spectral analysis for fake face detection. After measuring the skin albedo curves of the face and mask materials with varying distances, two discriminating wavelengths (850–1450-nm) were selected to train a Support vector Machine (SVM) classifier for discriminating between genuine and false attempts. The experiments were performed on a base of 20 masks of different materials: 4 plastic, 6 silica gel, 4 pulp, 4 plaster and 2 sponge. The results show that the correct classification can attain 89.18% accuracy.

The authors of this experiment did not do their studies with masks that are replicas of real subjects. On the contrary, Kose and Dugelay [5] carried out their work with a database of printed masks of about 16 real subjects. The analyses of the facial features were carried out by a 3D scanner after the masks were realized by means of a 3D printing service. In addition to the texture images, the database also comprises the two samples with real face and face mask for each person. The authors propose a method based on various linear binary pattern (LBP) techniques, for feature extraction using two image types (color and depth) and they claim 88.12% and 86% accuracy both types of images.

In this article we propose a face recognition system that includes a new approach to distinguish between a real face and a face with mask. As shown in Fig. 1, the proposed system consists of two step, a recognition step followed by a verification step, to detect impostors with mask 3D.

The advantage of the proposed approach is that it can be used by any system of recognition that uses RGB images from a simple webcam, unlike other approaches, such as the ones proposed in [3] and [5] that uses special sensors for the acquisition.

The approaches proposed by Kim et al., [3] and Zhang et al., [4] measure the reflected light by probing light waves on the face, which can be damaging and harmful to users' health.

On the contrary, our technique can be used without any risk to the user's health. For the verification stage, Kose and Dugelay [5] use two types of images, a depth one and a RGB one, captured with an adequate acquisition camera. In contrast, our method uses only RGB images that could be acquired by a simple Webcam.

The ART has been widely used in several algorithms, such as logo recognition [8], video surveillance systems [9], face detection [10] and Region-based descriptor in MPEG-7 [11]. It has also been

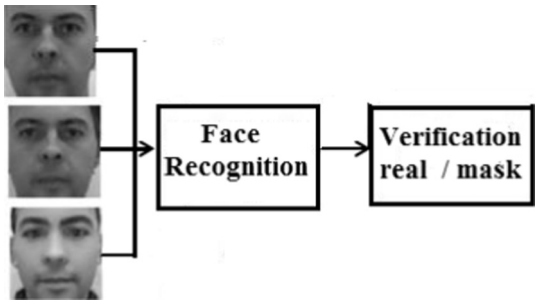

Fig. 1. The principle of detecting impostors in a face recognition system.

used for characteristics extraction, in the general face recognition system that we have proposed in [6]. The results obtained with this system were very good for some databases, composed of non-inclined faces, such as the ORL, the faces94, and the faces96 databases[22,23]. However, this system, like any other general face recognition system, is vulnerable to attacks that try to deceive it, especially those accomplished with 3D masks.

In this paper, we propose a new simple and efficient verification method, as a countermeasure to 3D mask attacks in a face recognition system. This method uses the Angular Radial Transformation (ART) to extract a feature vector from the whole image and input it to a Maximum Likelihood (ML) classifier, for discriminating between true and fake faces.

To improve the results, we have used optimized values/methods. Data optimization methods are widely used in several other domains, which include the precipitation analysis by using optimized neural networks [12], developpement of water lifting devices [13], and water engineering in general [14,15].

To validate the proposed system and after authorization of the owners, we used the 3D Mask Attack Database (3DMAD) available in [16]; this database was developed by the laboratory (Idiap research institute).

Unlike the methods proposed in [3,4], the proposed method presents no risk to the user's health. It has also the advantage of being less complex and of using a lower cost camera, such as a web camera, compared it main competitor method, proposed in [7].

The rest of the paper is organized as follows. Section two describes the used face recognition method. The following section presents the proposed verification method for spoofing by 3D masks detection. In section four the performance of the proposed method are evaluated. Finally, in the last section some conclusions and perspectives are given.

2. The recognition phase

As we have already mentioned, a recognition system consists of two phases: the characteristics extraction phase and the classification phase (Fig. 2).

2.1. Extraction of characteristics

Many 2D techniques have been developed in recent years, for the extraction of characteristics, in face recognition systems. Among them, Eigenfaces (PCA), developed by Turk and Pentland [17], and Linear Discriminant Analysis, proposed by Swets and Weng [18]. LDA is a technique particularly prized by the biometric researchers community.

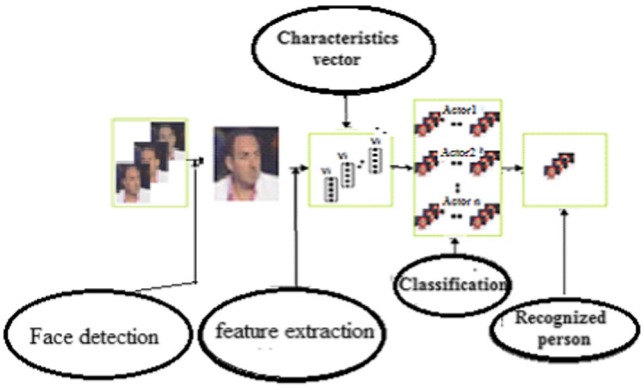

Fig. 2. Schematic of a general face recognition system.

The aim of the LDA is to reduce the size of the image, represented in the form of a vector, but with the preservation of the discriminatory information. In this method, the training set of images you must first organized into several classes: a class by a subject and multiple images per class. The organization of the 3DMAD database obeys this rule, which pushed us to adopt this method to extract the characteristics. LDA analysis the eigenvectors of the data dispersion matrix, aiming to maximize the inter-class variations while minimizing the intra-class variations. This reduces to find an optimal W projection base that maximizes the intra-class dispersion, related to the matrix Sw, and minimize the inter-class dispersion, related to the matrix Sb. It can be shown that this is equivalent to find W, which minimizes the Fisher optimization criterion J(W):

$$W = \operatorname{argmax}(J(W)) = \frac{|W^T S_b W|}{|W^T S_w W|} \tag{1}$$

The solution of Eq. (1) can be found by applying the generalized eigenvalue technique, as demonstrated in [19]:

$$Sb \cdot W = \lambda \cdot W \cdot Sw \cdot W \tag{2}$$

This problem reduces to a research of the eigenvectors of matrix $S_w^{-1} \cdot Sb$. By projecting the original image using W; a new projected image is obtained. The size of this image is solely equal to the number of classes −1, compared to (N × N), which is the size of the original image.

2.2. Classification for recognition

Two classification methods were tested, for the recognition of face images. The first one named Nearest Neighbor Classifier (NNC), uses a simple Euclidean distance to find the nearest neighbor. The second one is more efficient, but more complex; it uses SVM to improve the recognition rate.

These two methods are very used in biometrics, mainly in face recognition, as in the works of Cox et al. [20] and Abdul Muqeet and Holamb [21]. They have already been used to obtain good results with other databases, such as ORL [22], Essex Grimace [23], Yale [24] and Sterling face [25],

The Euclidean distance is a special case of the more general Minkowski distance of order p. For two vectors $X = (x_1, x_2, \ldots, x_N)$ and $Y = (y_1, y_2, \ldots, y_N)$, this distance is defined as:

$$Lp = \left(\sum_{i=1}^{N} |x_i - y_i|^p \right)^{1/p} \tag{3}$$

The Euclidean distance is:

$$L2 = \sqrt{\sum_{i=1}^{N} |x_i - y_i|^2} \tag{4}$$

The SVM is a well-known classifier that was used in many fields, among which face recognition [26]. It is used to decide to which class a sample belongs, that is to say to solve the problem of discrimination.

To find a solution to this problem, we must seek a function which associates each input vector, x, with its output class, y.

$$y = h(x) \tag{5}$$

A supervised learning algorithm such as the SVM aims to learn the h (x) function from a learning set:

$\{(x_0, y_0), \ldots, (x_k, y_k)\}$, with $x_k \subset \Re^N$ and $y_k \in \{-1, 1\}$.

The class is given by Y, and is defined as:

$$Y = \operatorname{Sign}(y_i \cdot h(x_i)) \tag{6}$$

Also:

$$\begin{cases} Y = 1 & \text{if } y_i \cdot h(x_i) \geqslant 0 \\ Y = -1 & \text{if } y_i \cdot h(x_i) < 0 \end{cases} \quad (7)$$

In practice, we can find an unlimited number of hyperplanes that separate two classes, but there is a single optimal hyperplane that has the largest margin (distance) from the elements of the two classes (Fig. 3).

Finding the optimal hyperplane is a quadratic optimization problem of dimension p (number of vectors) under constraints. The solution of this problem may be found with the conventional Lagrange method. The optimal Lagrange multipliers, α_k^*, are used to form the optimal hyperplane equation:

$$h(x) = \sum_{k=1}^{p} \alpha_k^* \cdot y_k \cdot x_k \cdot x + w_0 \quad (8)$$

When the data are not linearly separable the problem can be solved by using the kernel functions; the equation of the separation hyperplane is expressed, in this case, as follows:

$$h(x) = \sum_{k=1}^{p} \alpha_k^* \cdot y_k \cdot K(x_k \cdot x) + w_0 \quad (9)$$

The use of kernel functions $k(x \cdot x')$ reduces the computational complexity.

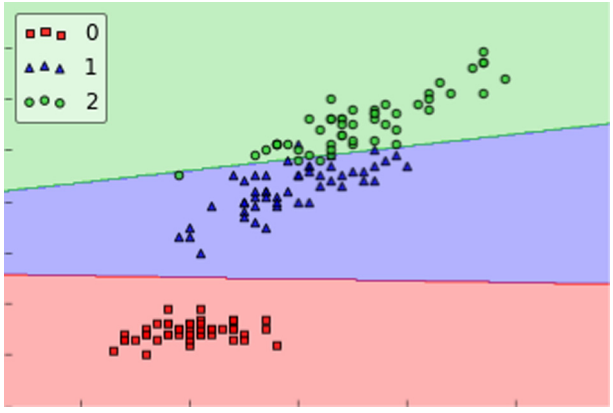

Fig. 3. The SVM classifier with linear kernel for linearly separable data.

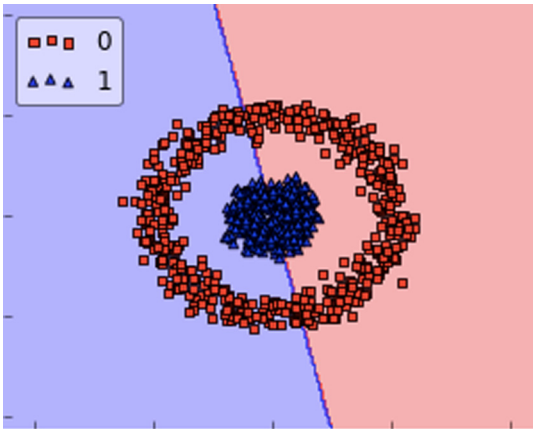

Fig. 4. The SVM classifier with a linear kernel fails to separate non-linearly separable data.

The most usual kernel functions are:

- Linear: k (x, x') = x · x'
- Polynomial: k (x, x') = (x × x')^d or k(x, x') = (c + x · x')^d
- Gaussian/RBF (Radial Basis Function): $k(x,x') = e^{-|x-x'|^2/2\cdot\sigma^2}$

The linear kernel works well if the data is linearly separable, as shown in Fig. 3. However, if the data is not linearly separable, a linear kernel cannot be used, as shown in Fig. 4.

Fig. 5 shows that by using a nonlinear Kernel, such as the RBF, the data in Fig. 4 can be separated.

A linear kernel is a parametric model, whereas an RBF kernel is not.

The complexity of the latter grows with the size of the learning set. Furthermore, it is more expensive to form an RBF kernel, since more parameters have to be adjusted.

In the case where the two preceding kernels do not give satisfactory results, the data may be projected in a new space, where they can be separated; this can be achieved by using the polynomial kernels, as seen in Fig. 6.

The polynomial kernel is less used, in practice, for reasons of efficiency (calculations and predictions). In the case where the size of the data is too large there is no rule for choosing the SVM kernel.

The only way is to perform simulations and choose the kernel that gave the best results.

The SVM is originally a binary classification (two classes). Several methods may be used for its extension to the multi-class classification problem (M> 2 classes). The one-versus-all method (OVA) uses M binary classifiers, each classifier compares one class

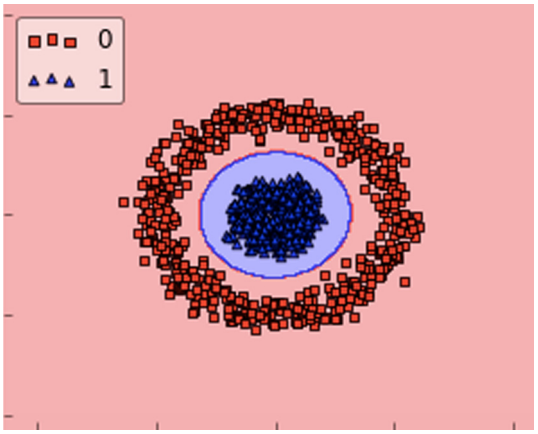

Fig. 5. The SVM classifier with a nonlinear RBF kernel.

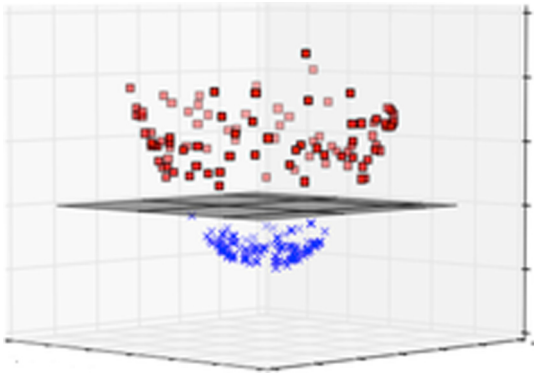

Fig. 6. The SVM classifier with a polynomial kernel.

to the rest, whereas the one-versus-one method uses M·(M − 1)/2 binary classifiers, each classifier compares one class to another class.

The evaluation of our face recognition revealed that it achieves a high recognition rate for real subjects (near 97%), but at the same time it has a high spoofing rate of subjects with masks (near 65%). This means that it can be easily misled by people with 3D masks of real faces.

To remedy this problem, a verification step is necessary for rejecting a face mask.

After several tests, we found that the decomposition of images with ART and the classification by the Maximum Likelihood (ML) technique, allows discriminating between a real face and a 3D mask.

3. The verification phase

For the verification step we designed a new method that uses a binary classifier to discriminate between a real face and 3D mask. The inputs to the classifier are the images decomposed with the ART.

This transformation allows discriminating between the light intensities reflected by a real face and a 3D mask.

3.1. Extraction of facial features using the ART

Extracting feature vectors of images using ART was proposed in [27]. ART is an orthogonal projection of an image on a radial basis; it is defined by:

$$W_{m,n} = \int_0^{2\pi} \int_0^1 f(r,\theta) \cdot V_{nm}(r,\theta) dr d\theta \tag{10}$$

To apply the above equation to an image I (x, y), it should be discretized as follows:

$$W_{m,n} = \sum_{r\leqslant 1} \sum_{\theta\leqslant 2\pi} I(r,\theta).[V_{nm}(r,\theta)]^* \tag{11}$$

Where $I(r,\theta)$ represents the image intensityin polar coordinates, and $V_{n,m}(r,\theta)$ represents the orthogonal radial basis function.

3.2. The ART projection basis

The radial basic function in ART is the product of two functions:

$$V_{n,m}(r,\theta) = A_m(\theta) \cdot R_n(r) \tag{12}$$

$A_m(\Theta)$ is an exponential function that ensures the invariance in rotation:

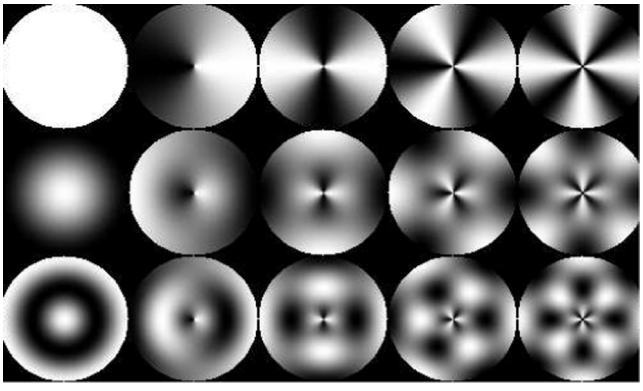

Fig. 7. The 2D ART projection basis, for n = 0:2 and m = 0:4.

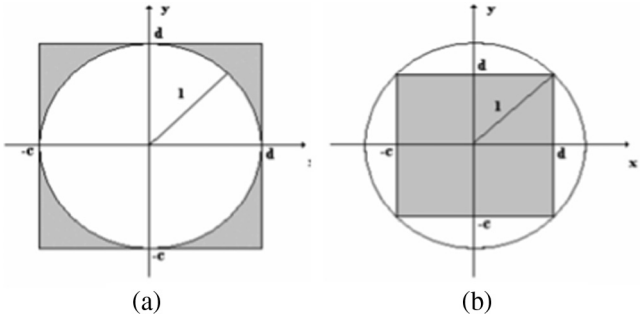

Fig. 8. Recalculation of the function I(x,y) (Gray rectangle) in the unit circle: (a) c = −1 and d = 1, (b) c = $-1/\sqrt{2}$ and d = $1/\sqrt{2}$.

$$A_m(\theta) = \frac{1}{2\pi} \cdot \exp(jm\theta) \tag{13}$$

The radial basis function $R_n(r)$ is a cosine function, defined by:

$$R_n(r) = \begin{cases} 1 & \text{for } n = 0 \\ 2 \cdot \cos(n \cdot r.\theta & \text{for } n \neq 0 \end{cases} \tag{14}$$

Fig. 7 represents the 2D ART projection basis, for n = 0: 2 and m = 0: 4.

The function I(x,y) should be calculated inside the unit circle to preserve the orthogonality of the bases$V_{m,n}$.

For this purpose, I(x,y) must be expressed in polar coordinates (r, θ), with the center of the image being the center of the circle (Fig. 8). The transformation of Cartesian coordinates into polar coordinates is given by the following equations:

$$\begin{cases} x = r \cdot \cos\theta \\ y = r \cdot \sin\theta \end{cases} \quad \text{and} \quad \begin{cases} r = \sqrt{x^2 + y^2} \\ \theta = \tan^{-1}(\frac{y}{x}) \end{cases} \tag{15}$$

The shape of the image I(x,y) is rectangular or square, which is incompatible with the shape of the unit circle. This requires making a choice between the elimination of some points of the image (especially the corners) or the introduction of points that do not belong to the original image. The above relationship may be written again as follows:

$$\begin{cases} x_i = c + \frac{i(d-c)}{N-1} \\ y_j = d - \frac{j(d-c)}{M-1} \end{cases} \quad \text{and} \quad \begin{cases} r_{ij} = \sqrt{x_i^2 + y_j^2} \\ \theta_{ij} = \tan - 1(\frac{y_j}{x_i}) \end{cases} \tag{16}$$

i and j are the coordinates of a point in the original image, x_i and y_j are the new coordinates of this same point in the new coordinate system (the unit circle), M and N are the horizontal and vertical extents of this image, respectively, and c and d are the parameters that allow to choose between recalculating the function I(x,y) entirely ($c = -\sqrt{1/2}$ and $d = \sqrt{1/2}$) or partially ($c = -1$ and $d = 1$) inside the unit circle, as shown in Fig. 8.

The circular moments obtained by the orthogonal projection of the face image onto the radial basis function can be expressed as a feature vector of a face image in several ways:

(1) $\{\Re(P_{nm}), \Im(P_{nm})\}$: a one-dimensional complex number is converted into a two dimensional real number.
(2) $P^2_{n,m}$: the amplitude, which is the absolute value of the complex number.
(3) $\arg(P_{nm})$: the phase or the argument of the complex number.
(4) $\{||P_{n,m}||^2, \arg(P_{nm})\}$: the amplitude and the argument of the complex number.

In our work we combine the real and imaginary part in one feature vector. This representation allows preserving the phase and avoids complex calculations, compared to other conversions.

To further reduce the size of the vectors from the ART, while preserving the information, we combined the ART and the LDA. The LDA allows a well separation between the classes that will enhance the performance of the classification.

3.3. Classification in the verification step

For the classification in the verification step, two methods were tested too. The first one is the simple NNC, already described in the recognition step. The second method is the maximum likelihood classifier, which is one of the most popular methods for images classification. The ML classifier calculates the probability that a given feature vector belongs to each class, real face or mask in our case, and assigns to this vector the class with the highest probability. For this, the probability density functions of a feature vector conditional to the different classes must be known. A statistical analysis of the distribution of the feature vectors revealed that these density functions can be well approximated by a normal distribution. Therefore, each class is fully characterized by the corresponding mean vector and covariance matrix, which are estimated offline from the training data. The fact the LDA is used for the extraction of the feature vectors minimizes the intra-class variation and maximizes the inter-class variations, as already said. This will make the probabilities that a feature vector belongs to the two classes well different, and allows therefore for a good discrimination between real faces and fake faces (masks). This argument, with the fact that it is less complex than the SVM, advocated for the choice of the ML classifier.

3.4. Maximum likelihood classifier

As stated above, to implement the ML classifier, the mean vectors and the covariance matrices of the classes must be estimated.

The mean vector, M_i and the covariance matrix C_i of class I with $i \in (1, 2)$, are estimated as follows:

$$M_i = \frac{1}{M} \cdot \sum_{j=1}^{M} x_{i,j} \tag{17}$$

$$C_i = \sum_{i=1}^{M} (X_{i,j} - M_i)(X_{i,j} - M_i)^T \tag{18}$$

where $X_{i,j}$ denotes the j^{th} sample feature vector, from class i.

To compute Eqs. (17) and (18), for each class, M Images were selected randomly, among the images of the subjects used in the reconnaissance phase.

The probability, P (T), that a vector T belongs to the class i, is computed as follows:

$$P_i(T) = \frac{p(i,T)}{\sum_{i=1}^{2} p(i,T)} \tag{19}$$

$$p(i,T) = \frac{\exp^{-\frac{1}{2}(T-M_i)^T \cdot Ci^{-1} \cdot (T-M_i)}}{2\pi |C_i|^{1/2}} \tag{20}$$

Usually, in the ML classifier, the vector T is assigned to the class with the highest probability. However, this may result in unequal error rates. Therefore, in order to enhance the immunity of the proposed method to attacks by 3D masks, the classifier decides that an image represents a true face only if the probability of the true faces class, calculated using Eqs. (19) and (20), is higher than a threshold that is higher than 0.7. This threshold is the one that ensures an Equal Error Rate (EER). It is determined by using a validation set, which is also composed of two subsets, made up of images that are different from those used in the training subsets.

4. Experimentation

4.1. The 3DMAD database

The 3D Mask Attack database (3DMAD) is composed of 17 different real subjects and their attack mask, recorded by the Microsoft Kinect sensor in the form of video.

For each subject 15 videos were recorded, each video sequence is composed of 300 frames. Each frame contains an image of a single face (real or mask), with a size equal to 640 × 480 pixels. The first ten sequences for each subject contain records of the real face of the subject, and the last five sequences contain records with a mask of the same subject.

To use this database in our work, a pretreatment was first applied to all the images that it contains.

4.2. Pretreatment of the images in the database

The pretreatment applied to all the images in the database consists of several steps:

- The first step converts the RGB color image of each frame to an image in gray levels, as shown in Fig. 9.
- The second step locates the facial area in an image by automatic detection, using the Haar algorithm developed in [28], as shown in Fig. 10.
- The third step crops the face region in the image and resizes it to a 64 × 64 picture,
- as shown in Fig. 11.

At the end of this stage we get our data base of 17 subjects: 3000 real face images and 1500 images of faces with masks, for each subject.

4.3. Protocol

4.3.1. The recognition phase

To evaluate our recognition system that uses the LDA algorithm with one of two classification methods, we have chosen 12 of the 17 subjects and left the other 5 subjects to be used as probable impostors.

For each subject among the 12 subjects 20 images of real faces were selected randomly to construct the training set. From this set, the projection matrix was calculated and used to obtain the feature vector of any image. Two other sets were also formed:

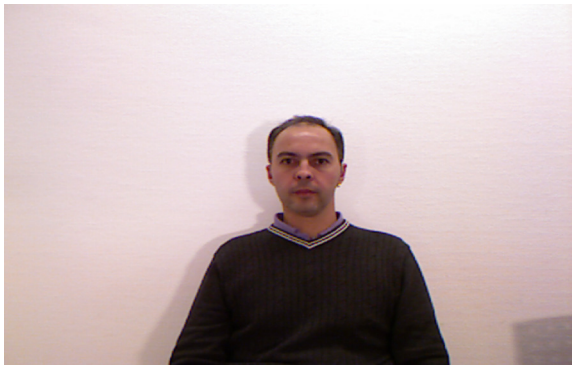

Fig. 9. RGB Color image.

Fig. 10. Face detection.

Fig. 11. Extracted face image (64 × 64).

- A validation set composed of 120 frames of the same 12 subjects of the training phase, to fix the decision threshold for the NNC classifier by calculating the Equal Error Rate (ERR).
- A test set composed of three subsets:
  - The First subset consists of 240 images of the 12 subjects used in the training phase, to calculate the False Rejection Rate (FRR) and the Recognition Rate.
  - The second subset consists of 120 images of the 5 subjects not used in the training phase to calculate the False Acceptance Rate (FAR).
  - The third subset consists of 120 images with mask of the same 12 subjects of the training phase, to calculate the Spoof False Acceptance Rate (SFAR).

4.3.2. The verification phase

To evaluate our verification system that uses the ART algorithm with the NNC classifier or the ML classifier, we constructed a training set that consists of two classes a set with subjects with real faces and another set with masks. For each class, 100 images of the same 12 subjects chosen in the recognition phase were selected randomly.

Besides the training set, two other sets were formed, a validation set and a test set, each of these sets is composed of two subsets: a subset of subjects with real faces and a subset of subjects with masks.

The subset of subjects with real faces is used to calculate the False Fake Rate (FFR), where the real accesses are classified as mask attacks. On the other hand, the subset of subjects with mask is used to calculate the False Living Rate (FLR), where the mask attacks are classified as real accesses.

The performances are measured with the Half Total Error Rate's (HTERs) which is the average of this two error rates. HTERs are calculated at the EER threshold computed using the validation set.

In The test set, for each subset, 200 images of the 12 same subjects used in the recognition phase, were selected randomly.

Table1
The RR and the HTER of the LDA algorithm with different databases.

Database	Recognition Rate (RR) (%)	HTER (%)
3DMAD LDA+SVM	97.41	3.31
3DMAD ddddddLDA+NNC	96.30	5.05
ORL	98.72	1.27
Essex Grimace	91.20	4.64
Yale	97.91	2.10
Sterling face	97.68	2.32

The validation set is composed of 100 other images of the same 12 subjects of the recognition phase, used to calculate the EER rate in order to fix the decision threshold for NNC and ML classifiers.

5. Results and discussion

5.1. Recognition results

The criteria used to evaluate the performance of the used recognition system are the Recognition Rate (RR) and the Half Total Error Rate (HTER), which is the mean of the FAR and the FRR.

Table1 lists the RR and the HTER, obtained by the LDA, using the 3DMAD database, and other databases (ORL, Essex Grimace, Yale, and Sterling face), as reported in [20].

For the classification by SVM, a linear kernel was chosen, because it gave us the best results [6]. From the results presented in the above table, it can be said that face recognition by LDA gives good results, for all databases, including the 3DMAD database. We now, test the immunity of this method to spoofing by 3D masks.

Table 2, below, gives the SFAR, i.e. the rate of faces with mask that are accepted as true faces, using either the NNC or the SVM classifier. It can be observed that this SFAR is very high, which means that this recognition system is vulnerable to mask attacks and that a verification step is necessary to reject fake faces.

5.2. The verification results

To assess the performance of the proposed ART based verification method, we use the HTER, calculated using both the validation and the test sets. Table 3, below, gives the obtained HTERs for the two classification methods and the two sets:

For a further assessment of the performance of the proposed anti-spoofing method, we compare in Table 4 its verification rate with that of another method [7] that used LBP, with classification by LDA, and the 3DMAD database.

Table 2
The SFAR rate of the LDA algorithm with the NNC and the SVM classifiers.

	Validation set (%)	Test set (%)
LDA+NNC	57.24	55.29
LDA+SVM	64.21	63.12

Table 3
Performance of the proposed ART based verification method, with the NNC and ML classifiers.

	Validation set (%)	Test set (%)
ART+NNC	1.43	1.51
ART+ML	**0.89**	**0.91**

The method with the best result in bold.

Table 4
Comparison between the proposed verification method and the LBP+LDA method. From the results presented in the above table, it can be stated that the proposed ART +ML method performs slightly better than the LBP method.

Classification	HTERs (%)
ART+ML	**0.91**
LBP+LDA	0.95
ART+NNC	1.51

The method with the best result in bold.

6. Conclusion

The use of 3D masks to deceive a recognition system has become an easier and cheaper technology with advances in technology of 3D printing.

In this paper, a new method to discriminate between a real face and a face with mask is proposed. This method uses the moments based on ART to extract the feature vectors and the ML as classifier.

The performances of our method were evaluated using the database 3DMAD. It is actually the only database that gives images of the subjects with real and masked faces.

The obtained results demonstrate the effectiveness of the proposed method to detect attacks with 3D masks.

The other advantage of this method is its simplicity. In particular, RGB images acquired with a low cost webcam can be used as inputs to the system. This allows the application to run on all systems that use a webcam for acquisition.

In perspective, it is envisaged to realize a recognition system, which can detect the persons wearing a 3D mask and signal them, without the necessity of a verification step.

References

[1] Bai, J., Ng, T.-T., Gao, X., Shi, Y.-Q., 2010. Is physics-based liveness detection truly possible with a single image? In: IEEE International Symposium on Circuits and Systems, pp. 3425–3428.

[2] Li, J., Wang, Y., Tan, T., Jain, A.K., 2004. Live face detection based on the analysis of Fourier spectra, SPIE 5404, Biometric Technology for Human Identification, pp. 296–303.

[3] Kim Y, Na J, Yoon S, Yi J. Masked fake face detection using radiance measurements. J. Opt. Soc. Am. 2009;26(4):760–6.

[4] Zhang, Z., Yi, D., Lei, Z., Li, S., 2011. Face, liveness detection by learning multispectral reflectance distributions. In: IEEE International Conference on Automatic Face Gesture Recognition and Workshops, pp. 436–441.

[5] Kose, N., Dugelay, J.L., 2013. Countermeasure for the protection of face recognition systems against mask attacks. In: IEEE International Conference on Automatic Face and Gesture Recognition.

[6] Hamdan B, Mokhtar K. Face recognition using angular radial transform. J. King Saud Univ. Comput. Inform. Sci. 2016. doi: https://doi.org/10.1016/j.jksuci.2016.10.006.

[7] Nesli E, Sébastien M. Spoofing in 2d face recognition with 3d masks and anti-spoofing with kinect. IEEE 6th International Conference on Biometrics: Theory, Applications and Systems (BTAS'13); 2013. p. 1-8.

[8] Wahdan O, Omar K, Nasrudin F. Logo recognition system using angular radial transform descriptors. J. Comput. Sci. 2011;7:1416–22.

[9] Lee SH, Sharma S, Sang L, Park J, Park YG. An Intelligent Video Security System using Object Tracking and Shape Recognition, vol. 6915. Berlin Heidelberg: ACVIS LNCS Springer-Verlag; 2011.

[10] Fang, J., Qui, G., 2003. Human Face Detection using Angular Radial Transform and Support Vector Machines, International Conference on Image Processing, vol. 1, pp. 69–72.

[11] Bober M. MPEG-7 visual shape descriptors. IEEE Trans. Circuits Syst. Video Technol. 2001;11(6):716–9.

[12] Valipour M. Optimization of neural networks for precipitation analysis in a humid region to detect drought and wet year alarms. J. Meteorol. Appl. 2016;23(1):91–100.

[13] Yannopoulos SI, Lyberatos G, Theodossiou N, Li W, Valipour M, Tamburrino A, Angelakis AN. Evolution of water lifting devices (pumps) over the centuries worldwide. Water J. 2015;7(9):5031–60.

[14] Valipour M, Singh VP. Global experiences on wastewater irrigation: challenges and prospects, the series. Water Sci. Technol. Lib. 2016;72:289–327.

[15] Valipour M, Sefidkouhi MAG, Raeini−Sarjaz M. Selecting the best model to estimate potential evapotranspiration with respect to climate change and magnitudes of extreme events. Agric. Water Manage. 2017;180(A):50–60.

[16] Availabe at: <http://www.idiap.ch/dataset/3dmad>.

[17] Turk M, Pentland A. Eigenfaces for recognition. J. Cognitive Neurosci. 1991;3(1):71–86.

[18] Swets DL, Weng J. Using discriminant eigenfeatures for image retrieval. IEEE Trans Pattern Anal Mach Int 1996;18:831–6.

[19] Belhumeur P, Hespanha J, Kriegman D. Eigenfaces vs. fisherfaces: recognition using class specific linear projection. IEEE Trans. Pattern Anal. Mach. Intell. 1997;19:711–20.

[20] Cox, I., Ghosn, J., Yianilos, P., 1996. Feature-based face recognition using mixture-distance. In: Proceedings of the Conference on Computer Vision and Pattern Recognition, CVPR'96, pp. 209–216.

[21] Abdul Muqeet M, Holambe RS. Local appearance-based face recognition using adaptive directional wavelet transform. J. King Saud Univ. Comput. Inform. Sci. 2016. doi: https://doi.org/10.1016/j.jksuci.2016.12.008.

[22] Available at: <http://www.uk.research.att.com/pub/data/att_faces.zip>.

[23] Available at: <http://www.cswww.essex.ac.uk/mv/allfaces/grimace.zip>.

[24] Available at: <http://www.cvc.yale.edu/projects/yalefaces/yalefaces.html>.

[25] Available at: <http://pics.psych.stir.ac.uk/>, University of Stirling online database.

[26] Huang, J., Shao, X., Wechsler, H., 1998. Face pose discrimination using support vector machines (SVM). In: Proceedings of the 14th International Conference on Pattern Recognition ICPR'98, vol. 1, pp. 154–156.

[27] Hu MK. Visual Pattern Recognition by Moment Invariants. IRE Trans. Inform. Theor. 1961;49:179–87.

[28] Mohan A, Papageorgiou C, Poggio T. Example-based object detection in images by components. PAMI 2001;23(4):349–61.