Face recognition based on 3D features: Management of the measurement uncertainty for improving the classification

Giovanni Betta [a], Domenico Capriglione [a], Michele Gasparetto [b], Emanuele Zappa [b], Consolatina Liguori [c,*], Alfredo Paolillo [c]

[a] *DIEI, University of Cassino and of Southern Lazio, Cassino, FR, Italy*
[b] *Department of Mechanical Engineering, Politecnico di Milano, via La Masa, 1, Milano, Italy*
[c] *DIIn, University of Salerno, via Giovanni Paolo II, 132, Fisciano, SA, Italy*

Received 28 October 2014
Received in revised form 17 February 2015
Accepted 20 March 2015
Available online 26 March 2015

1. Introduction

Face-based personal identification plays a crucial role in a wide variety of applications; recent advances in face-based biometrics are shown in review papers, such as [1,2]. The open issues in this field include: the possibility to work with extremely wide image databases, such as the web or social networks [3], the possibility to imple-ment the recognition on mobile devices [4], the application of recognition techniques in the cases of variable illumination conditions [5,6], the application in the cases of non-collaborative person [7,8], in the case of spoofing [9] and in the case of partially occluded face images [10].

Moreover, all the recognition approaches are based on biometric data and the effect of the uncertainty of these data on the recognition judgment is another open research field. In this work the knowledge of the uncertainty of the 3D facial features is managed to improve the recognition performances.

In a face recognition system, an unknown subject is recognized by comparing some quantities, extracted from an image of an unknown subject to be classified, with the corresponding ones extracted from a preexisting database [11–18] and [19]. Afterwards, a maximum likelihood approach is generally adopted to classify and recognize the unknown subject. However, the quantities measured for the classification are affected by uncertainty, thus generating a risk in accepting the decision. This risk could be quantified and reduced by suitably taking into account the measurement uncertainty in the comparison stage [20–22], or using fuzzy logic [23–26].

* Corresponding author.
E-mail addresses: betta@unicas.it (G. Betta), emanuele.zappa@polimi.it (E. Zappa), tliguori@unisa.it (C. Liguori).

In previous papers [27–29] the authors have proposed a novel approach to classification and recognition problems which takes into account the measurement uncertainty affecting input data in order to improve the overall reliability of such kind of processes. The proposed method is based on a probabilistic approach for the evaluation of the confidence level of system outputs and the suitable use of this information for a performance improvement in terms of correct, wrong and abstention rate. In particu-lar, the output of a recognition process consists not only in the class which the subject belongs to, but in a list of pos-sible classes, each one characterized by a confidence level (CL), obtained considering also the actual measurement uncertainty evaluated on the processed representation. Then, the knowledge of such CLs can be exploited to weight the final decision also on the basis of the actual context in which the images are acquired. As an example, a circumstance in which several classes present similar values of CL should alert the user that the results are quite questionable while a situation in which a class presents a CL significantly higher than the other ones, should assure the user that such class can be really associated to the input subject.

As a case study, it has been applied to a 2D face classification algorithm based on Linear Discriminant Analysis. The results clearly showed the advantages of the procedure in terms of relevance and augmented information contributing to the decision, as well as of reduced false decision rates. That leads to what can be considered the ultimate point in favor of the proposed method from a user's perspective, which is the improved awareness in making a recognition decision.

In this paper, the proposed methodology is extended to manage the uncertainty in the face recognition based on 3D features. The 3D localization of the significant points of the face (generally also called repere-points or land-marks) [30] requires a prior image processing to extract the features in the 2D images [31–33]. In this work the 3D features are obtained triangulating the 2D landmarks detected on the two facial images by processing a pair of stereoscopic images through an Active Appearance Models (AAM) algorithm. The AAM is an algorithm developed by Cootes et al. [30] for matching a statistical model of object shape and appearance to a new image; it is used in this work to automatically detect a set of given feature points on the images of faces. The statistical models are built during a training phase by means of a set of images and the corresponding coordinates of the desired landmarks in each image. An introduction to AAM, as well as details on its application in this work, is provided in Section 3.1.

Recognition is performed by evaluating the weighted measurement discrepancies (hereinafter called Score) between 3D geometric masks obtained from stereoscopic images and the 3D masks included in a given database.

The application of the previously proposed approach to the classification using reduced features reconstructed in the 3D space gives rise to additional problems to be dealt with, mainly concerning: the uncertainty source (e.g. the uncertainty on the stereo camera position, the uncertainty of the calibration procedure), the propagation of the uncertainty on the 3D extracted features (e.g. influence of the training phase, the composition of the uncertainty on both images), and on the Scores used in the comparison phase.

Moreover, this approach could be useful to determine the optimum values of some system parameters (such as the camera position, the weight of each feature in the score evaluation). The selection of such optimum values could help for minimizing their uncertainty contribution.

In the following, after a brief description of the image processing system used for the 3D features extraction, the classification procedure is concisely recalled and the application of the classification methodology to the exploited face recognition algorithm is shown, also highlighting the performance improvement with respect to a basic approach.

2. The proposed classification cystem

The flow chart of the proposed classification scheme is sketched in Fig. 1.

- At first, the two 2D images of the unknown subject are processed by the biometric algorithm (AAM and triangulation) in order to estimate the 3D features that describe it. Then, the Score of each subject of the database is evaluated.
- The two images are processed further by the uncertainty estimation algorithm that determines the uncertainty on the obtainable scores.
- Finally, starting from all the evaluated scores and the estimated uncertainty, the decision procedure provides a classification list where all the subjects in which the unknown can be recognized are reported together with their Confidence Level, CL.

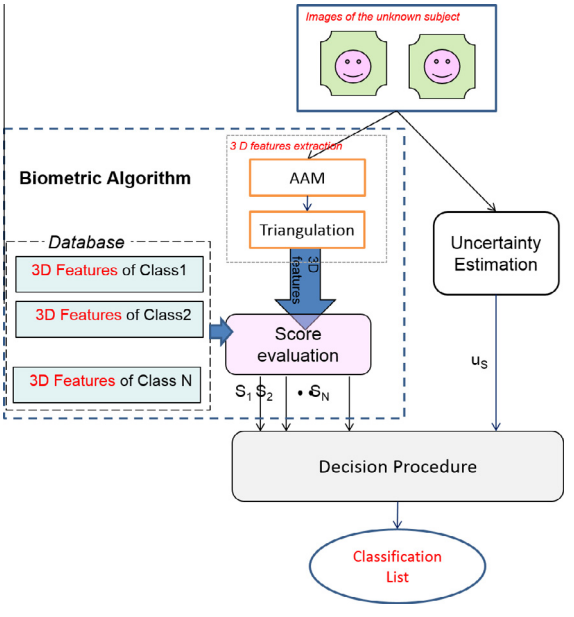

Fig. 1. The proposed classification scheme.

The architecture of the stereoscopic system is composed of two AVT Pike F-145B (Sony 2/3″ 1388 × 1038 pixel CCD progressive scan sensor) cameras, vertically aligned and placed in front of the person's face and an angle of 45° between the cameras (see Fig. 2). The image acquisition system was equipped also with a third camera (Cam 0, visible in Fig. 2) that allows to acquire frontal images of the face, however in this work only the images acquired by cam 1 and cam 2 are used.

The cameras are equipped with 25 mm focal length lenses and connected to a computer via Firewire IEEE1394 connection. The configured system allows having a field of view of approximately 300 × 400 mm, large enough to contain the subject's face placed 1000 mm away from the system.

3. The biometric algorithm

3.1. The 3D features extraction

The facial images obtained by means of a couple of stereoscopic cameras are analyzed with the AAM-API software to automatically detect the 2D coordinates of a set of landmarks. A 3D mask is obtained triangulating the two 2D masks (see Fig. 3). The 3D mask of a person to be identified is finally compared with each of the 3D masks included in a database and a Score is computed for each comparison, as detailed below.

The Active Appearance Model (AAM) technique used to automatically detect the facial features in the images, is an algorithm for feature extraction, developed by Cootes et al. [30]. The AAM has been used in literature to extract 2D biometric features and several studies have implemented this algorithm in 2D face recognition [33–35].

The operation of Shape and Appearance models can be split in two phases: the first one allows to create a morphable model, thanks to the analysis of training images, i.e. a sample of images where a set of facial features have been previously manually annotated. In the second phase, the software can automatically detect the position of the same set of facial features in face images that do not belong to the training set.

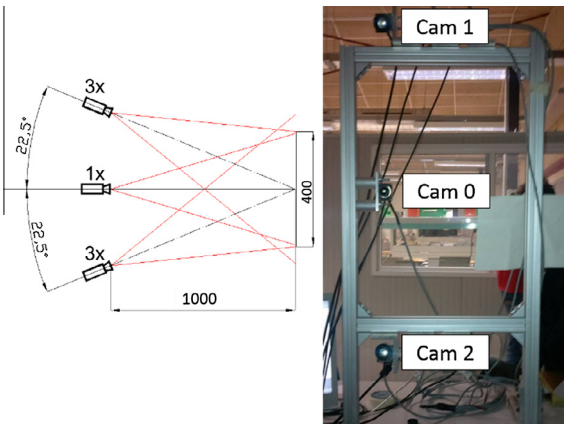

Fig. 2. The image acquisition system.

The manual annotation process of an image consists in tracing different landmarks that outline the most important facial traits on various images. In particular, 58 landmarks demarcating seven areas of the face have been used in this work: jaw, mouth, nose, eyes, and eyebrows, as shown in Fig. 3, where an example of 3D mask is also shown [29]. The choice of the 58 points is made according to the processing method carried out in previous studies [35].

Although the detailed description of the AAM algorithm is out of the scope of the present work, an outline of this algorithm is briefly provided here, while more detailed information can be found in [30] and [33].

As previously said, the AAM algorithm is based on the combined use of two different statistical models: the Shape Model and the Appearance Model.

As for the 2D Shape Model: the shape is defined as a 2D point set describing the shape of a target body. During the shape model creation, shapes traced on images of bodies belonging to the same family are submitted to the Procrustes Analysis in order to align these shapes to a common reference system and to make the application of the Principal Component Analysis (PCA) possible.

The PCA generates the shape variation basis ϕ, defined as:

$$\phi = (\phi 1 | \phi 2 \ldots | \phi t) \tag{1}$$

that can represent any x shape of analyzed bodies starting from a mean shape $\bar{x}$; vector b is a real number set that models deformable shape parameters:

$$x = \bar{x} + \phi b \tag{2}$$

As for the Appearance Model, it is defined as the texture of a portion of the target. The Appearance Model arranges all pixel intensity variations of the images on the mean shape. During the creation of this model, all the training images are transformed into images of the same shape and dimensions and the normalization of the texture is performed to avoid lighting or luminosity changes in pictures. Thereafter, appearance model is elaborated by performing PCA on the training images. Similarly as the shape model, the appearance model consists of the grey level vector of the mean appearance $\bar{g}$, the variation basis ϕ_g and a group of grey level parameters b_g:

$$g = \bar{g} + \phi_g b_g \tag{3}$$

The cameras have been calibrated using the Zhang's method [36]. The 3D coordinates of each facial feature have been estimated by means of optical triangulation relying on the facial features detected on the two stereoscopic images using the previously described AAM technique.

3.2. The database

A database was created with stereoscopic images of 117 volunteers.

The control system allowed to automatically acquire sequences of images with a user-defined interval, set in this case to 5 s, with the aim of obtaining multiple images of the same person in the same nominal position but

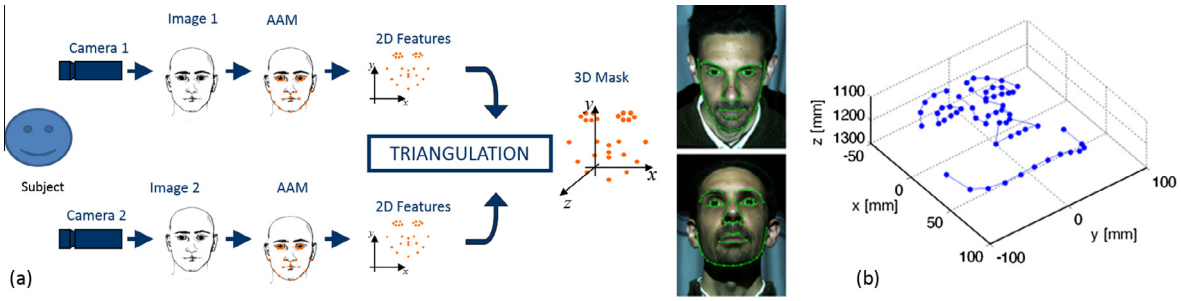

Fig. 3. (a) The schematic of biometric algorithm and (b) stereoscopic images with the detected features and the obtained 3D mask.

allowing small changes of expression that naturally happen. In particular, the database contains 5 stereoscopic image pairs for each volunteer. The first repetition was used to train the AAM model, and the other four to verify the repeatability.

3.3. Score evaluation

The recognition judgment is based on a Score that basically represents the sum of squared discrepancy between the 3D coordinates of the mask to be recognized and the corresponding coordinates of each mask in the database. Prior to that evaluation of point-to-point distances, a rototranslation is computed in order for the coordinate frame of one mask to be moved onto the coordinate frame of the other mask with a rigid motion. The rototranslation allows to compensate for differences in position and orientation of the subject with respect to the stereoscopic system in acquisitions, so that the Score obtained after rototranslation does not depend on the position and orientation of the face but depends on the mask shape only. Since the reliability in the feature localization is not constant for all the regions of the face, different weights are assigned to the various points according to their estimation repeatability.

Given a 3D mask to be recognized and the set of weights W_k ($k = 1, \ldots, n$, where n is the total number of points in the mask), the score S_i for each ith mask of the database is computed as:

$$S_i = \frac{\sum_{k=1}^{n} (W_k * (V_{k,i} - V_{k,\mathrm{ref}})^2)}{n} \tag{4}$$

where: $V_{k,i}$ are the coordinates of the kth point for the i-th individual, W_k is the weight of the kth point of the mask.

In order to estimate weights for each one of the 58 landmarks, the standard deviation between repeated measurements of the face has been calculated analyzing 5 masks obtained from repeated pictures of the same subject, for all the 117 subjects in the database and the reciprocal of the estimated standard deviation was used as a weight. Since small rotations of the face may occur between consecutive acquisitions, it is necessary to align one mask to the other one before to calculate average position and relative variance for each landmark. In this way, the variance of each point associated with each individual was obtained.

In Table 1 the average weights, obtained averaging the weights of all the 117 people in the database, are shown. Note that the weights are not associated to each single landmark but to each region of the face. The results confirm what was intuitively suggested: there is more stability on the eyes; the mouth is subject to large modifications due to facial expression changes, while the outline of the jaw presents the highest variability.

It should be underlined that, in the image acquisition, the environmental conditions (in particular the lighting) are controlled in order to ensure that the images are acquired without appreciable shadows, light reflections or motion blur; moreover the facial images are acquired with the person in frontal position and with the best possible focusing. With this attention, we can ensure that the uncertainty due to environmental conditions and focusing can be neglected in image acquisition. As will be shown in the following of the paper, variability in image brightness, focusing and motion blur will be simulated numerically in order to explore the uncertainty due to those parameters in controlled conditions.

4. Uncertainty estimation

The uncertainty on the Score, u_S, depends mainly on the uncertainty of the 2D coordinates of the face features that in their turns depend on the characteristics of the processed images. As evidenced in previous papers [28,29,38] the main influence quantities in face recognition problems can be related to luminance, defocus and motion blur. In order to quantify this uncertainty according to the ISO-GUM [37], a simple model was used to associate each quantity of influence to u_S. Denoting with u_i the contribute due to the ith quantity of influence on the Score uncertainty, we have posed:

$$u_i = f_i(\Delta_i) \tag{5}$$

Table 1
Average weights for each region of the face.

Area	Weight
Jaw	0.23
Mouth	0.51
Nose	1.10
Eyes	1.87
Eyebrow	0.96

where Δ_i is the value of the ith quantity of influence.

All of the quantities of influence are considered uncorrelated with the other ones, then the combined uncertainty on the score is evaluated as:

$$u_S = \sqrt{\sum_{i=1}^{N} u_i^2}. \quad (6)$$

where N is the number of the considered quantities of influence.

A statistical approach [30] was followed in order to estimate an uncertainty model using couples of artificial images generated by manipulating the reference images (i.e. the ones contained in the training database) in order to achieve new images characterized by the desired values of the quantities of influence. In particular, for each considered condition, the procedure is the following:

- for each subject of the database, each considered quantity of influence and each considered level of the quantity of influence, 2 new images have been generated by applying suitable digital filtering on the reference images in a controlled simulation environment;
- on the so modified images, the 3D features are evaluated applying the AAM algorithm and the triangulation;
- the Score with respect to the recorded data of the subject itself is estimated;
- finally on the Score obtained for all the subjects a statistical analysis is made, and the uncertainty is estimated.

$$(u_S)_i = \sqrt{\frac{(\mu_S^2)_i}{3} + (\sigma_S^2)_i} \quad (7)$$

where $(\mu_S)_i$ and $(\sigma_S)_i$ are respectively the mean and the sample standard deviation of the measured Scores due to the ith influence quantity.

Moreover, for each analyzed condition, the study of u_S has been made also considering the effect of the ith quantity of influence when affecting only one image at time.

4.1. Luminosity

The grey level of the original images, considered with the optimal luminosity, were modified of ±5, ±10, ±15,±20, ±30, ±50 (codes for an 8 bit representation), with respect to the reference image. Tests were made changing either the two images or only one. In Fig. 4 the corresponding results are reported. As you can see the uncertainty is almost constant for each kind of luminosity variation, consequently a constant value can be considered for modification of both image u_S can be posed equal to 0.08. Furthermore, a little lower uncer-tainty is observed if only one image is characterized by a not optimally luminosity, and it does not depend on which camera is considered (see Fig. 4). Since the diminution is not significant the previously value can be consid-ered still valid.

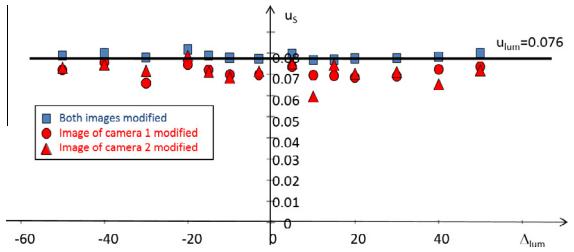

Fig. 4. Analysis of u_S versus the level of luminance.

4.2. Defocus

Following the same procedure used for the evaluation of the luminosity influence, the effects of the lens defocus were analyzed. Also in this case, since the configuration of the two cameras can be different, test were made considering three different conditions: both images, only camera one, only camera two out of focus, and with different level of defocus (see Fig. 5). The modified images are obtained by filtering the image with a Gaussian linear kernel filter characterized by a standard deviation of 3–20 pixels with respect to the reference image (image size 1388 × 1038px).

Fig. 6 shows the obtained results. As expected, when both images are out of focus the uncertainty is greater than when only one is out of focus. Moreover, the results are almost the same whatever the altered image. It is possible to use two models, one for both the images out of focus and the other one when only one image is out of focus. In both cases, a second order polynomial model well fits the observed data.

4.3. Motion blur

Generally, the motion blur is due to the movement between the subject and the camera. In the case of a fixed stereo rig, it is due to the subject movement and consequently it is present in both images. Two different kinds of motion blur are analyzed, one that simulates a motion between the subject and the camera pair along a vertical direction (Fig. 7a) and another that reproduces the motion along the horizontal direction (Fig. 7b).

The modified images are obtained by filtering the image with a directional filter characterized by a kernel length of 6–110 pixels with respect to the reference image (image size 1388 × 1038px). The filter is implemented as a convolution with a square kernel with a size proportional to the simulated speed effect of the motion. Concerning the values of the elements of the kernel, the closer to a straight segment oriented along the direction of motion, the higher their value. Fig. 8 reports the observed uncertainty and their simplified models for both type of simulated motion blur.

5. Decision procedure

In Fig. 9, a diagram of the main steps forming the decision procedure is reported: (i) at first, the probability, P_j,

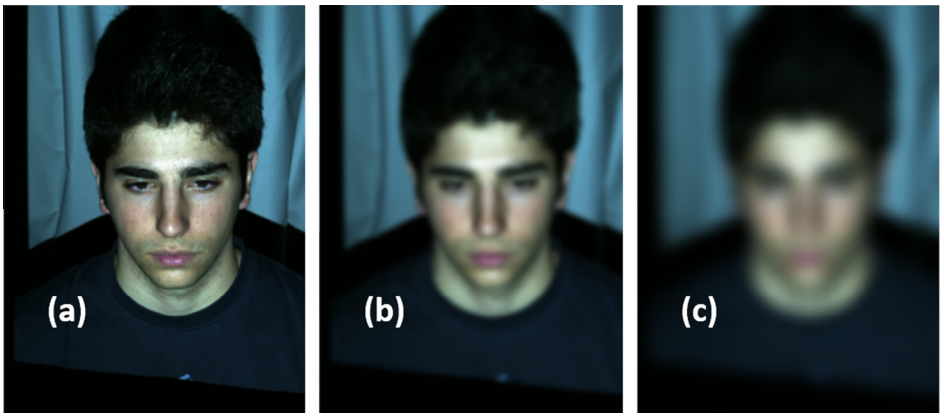

Fig. 5. Examples of filtering for achieving the defocused images due to the lens: (a) $\Delta_{focus} = 0$ (reference image), (b) $\Delta_{focus} = 10$, and (c) $\Delta_{focus} = 27$.

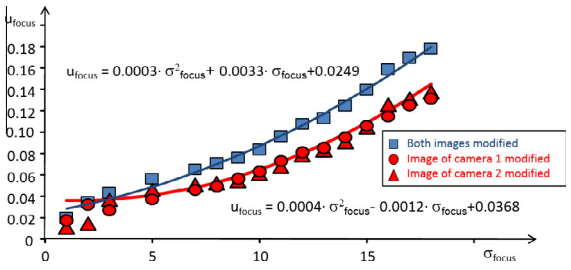

Fig. 6. Analysis of u_S versus the level of defocus due to the lens.

that an input subject is the jth subject of the database is evaluated; then, (ii) on the basis of the obtained probabilities a classification list is created with a selection of the probable subjects; finally, (iii) the confidence level for each subject (class) in the list is evaluated. In particular:

(i) P_j represents the probability that the unknown subject is the jth subject of the database (i.e. the jth class). Considering the score as a random variable, P_j represents the probability that the Score of the jth class is equal to zero given a measured value S_j. This probability is evaluated by the score probability density function, $p(s)$, with the following relationship (see Fig. 10):

$$P_j = P(S_j = 0|\overline{S}_j) = \begin{cases} 1 & \text{if} \quad \overline{S}_j \leqslant th \\ \int_{\overline{S}_j}^{\infty} p(s - \text{th})\text{ds} & \text{if} \quad \overline{S}_j > \text{th} \end{cases}$$

With this position we consider that each measured value of $\overline{S}_j$ less than two times the uncertainty (th = 2 u_S) can be considered equal to zero, and consequently its probability is equal to 1. The choice of a coverage factor equal to 2 is made for taking into account the residual imperfections of the simple model considered for estimating u_S (see Section 4). Of course, other coverage factors could be adopted, but the general proposed procedure still remains valid.

This function is applied to all the considered classes of the database, evaluating the probability of each one of them.

The Scores of the correct class for each subject of the database, evaluated in Section 4, are used in order to define $p(s)$. For each value of the influence quantities a Chi-square test is made for comparing the observed distribution with expected ones (Exponential, Beta, and Normal distributions are considered). The exponential distribution

Fig. 7. Example of filtering which produces the defocused images due to the motion blur: (a) vertical ($\theta = 10°$ and $\Delta_{motion} = 20$) and (b) horizontal ($\theta = 100°$ and $\Delta_{motion} = 20$).

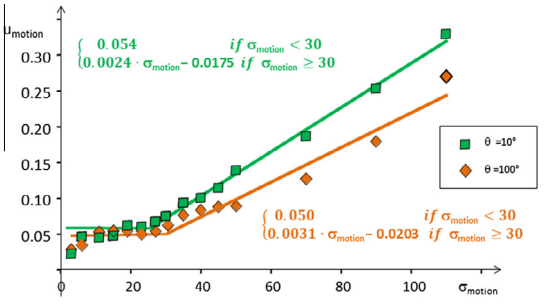

Fig. 8. Analysis of u_S versus the level of defocus due to the motion blur.

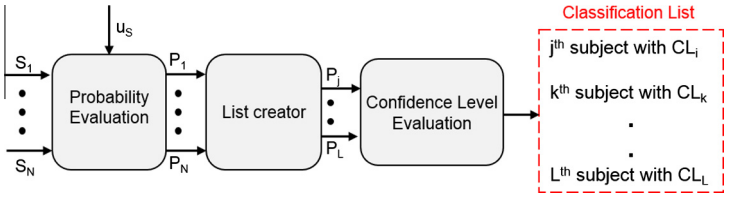

Fig. 9. Simplified block diagram of the decision procedure.

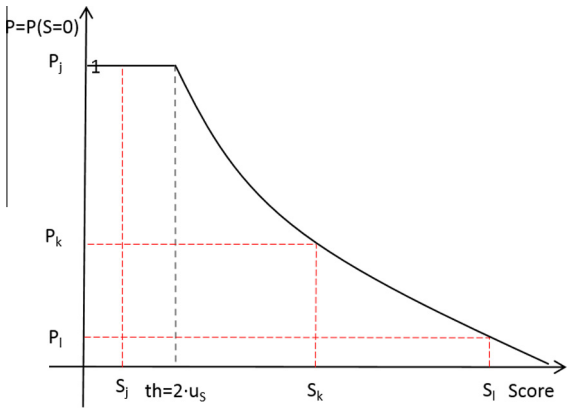

Fig. 10. Example of function for each P_j evaluation.

showed the best fit with the observed scores, for all of the uncertainty cases and for all of the values.

(ii) Starting from the so calculated probabilities, the classification list is composed of all the classes which show a probability greater than a second threshold, ***TH***. Generally, the value of ***TH*** affects the sensitivity and the selectivity of the method: high values of ***TH*** increase the selectivity but make worse the sensitivity and vice-versa. Therefore, a suitable tradeoff has to be considered. To this aim, a general approach for defining ***TH*** involves the analysis of the TP (correct classes with a probability greater than ***TH***) and false positive FP (wrong classes with a probability greater than ***TH***) versus ***TH***, looking for identifying the value of ***TH*** that meet the requirements of the specific application.

(iii) The probabilities of all the classes included in the classification list are used in evaluating a normalization factor, K, defined as follows:

$$K = \sum_{j} P_j \text{ for all } j \text{ with } P_j > \boldsymbol{TH}. \tag{9}$$

Then, the confidence level (CL) of each class belonging to the list is evaluated as the probability P_j divided by K:

$$\mathrm{CL}_j = \frac{P_j}{K} \tag{10}$$

6. Experimental results

In order to verify the performance of the proposed classification method, it is compared with a basic approach, for which a minimum value of Score is found in order to discriminate a subject belonging or not belonging the database. In Fig. 11 the observed average margins (for a particular sample is the difference between the score of the correct class and the lowest score of the other classes) are reported versus u_S. As you can see, as the uncertainty on images increases the margin decreases, and for high values of uncertainty, the margin approaches zero, thus making difficult the definition of a suitable discriminating threshold. In the following, the attained results are reported in comparison with a basic approach in order to evidence the obtainable improvement. As for the basic approach, a threshold on the score is also introduced so that the classification is accepted only if the minimum score (i.e. the score of the winning class) is lower than such threshold. The presence of such a threshold is needed in practical applications for containing the false acceptance rate in the case of input subject not belonging to the training database (in absence of such a threshold a subject not belonging to the database would be wrong classified and confused with a subject belonging to the training database even if it will exhibit an high value of score).

The threshold has been posed equal to the value of the observed lowest margin when $u_S = 0$, ($\boldsymbol{TH} = 0.25$).

The comparison is made on new sets of images obtained by modifying the original images as presented in Section 4 but using parameters different from the previous used. The obtained dataset is composed by 117×45 couples of images (9 different intensities of motion blur, 24 levels of defocus, 12 luminosity values either on both or on single image). The distribution of the obtained uncertainty is reported in Fig. 12 where the relative frequency histogram of the estimated u_S on the whole dataset is shown; this distribution reproduce well cover all the expected working conditions (both for range and frequencies).

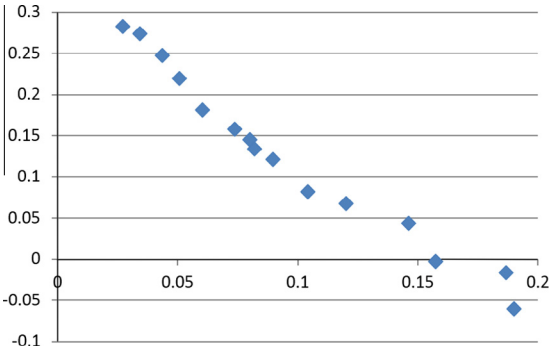

Fig. 11. Average margin versus u_S.

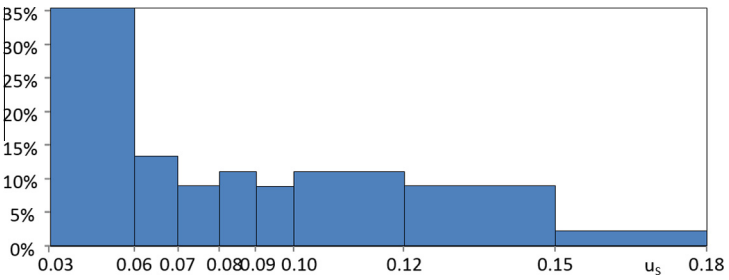

Fig. 12. Histogram of the estimated u_S.

As for ***TH***, a suitable experimental campaign carried out on the whole database has been performed for identifying the best tradeoff between sensitivity and selectivity. In a more detail, the trends of the true acceptance rate, TAR (TP/all positive) and of the false acceptance rate, FAR (FP/all negative) versus ***TH***, together with the Receiver Operating Characteristic (ROC) curve, have been analyzed. In our application the sensitivity (i.e. the TAR parameter) has to be preserved since the statistical post processing reduces the false positives. As a consequence a ***TH*** equal to 0.35 was chosen which provides a TAR greater than 0.99 and a FAR lower than 0.15.

For the proposed procedure we define the following figures of merit:

- *Correct classification*: either if the Classification list includes only the right class with CL = 1 or if the Classification list has more subjects where the right class has the highest value.
- *False reject*: If the subject is in the training database but the Classification list is empty.
- *Wrong classification*: either if the Classification has more subjects without the right class or if the right class has not the highest value of CL.

Vice versa, as for the basic approach, three cases are possible:

- *Correct classification*: if the right class has the minimum Score and it is less than the threshold (0.25).

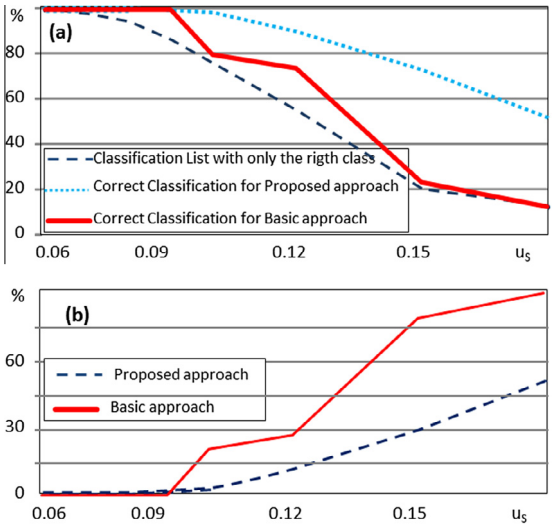

Fig. 13. Performance comparison versus the uncertainty on the Scores between basic and proposed approaches: (a) percentage of Correct Classifications and (b) percentage of False Rejections.

- *False Reject*: if no class has a Score less than the threshold.
- *Wrong classification*: if a wrong class has the minimum Score and it is less than the threshold (0.25).

In Fig. 13 and in Tables 2 and 3 the comparison between the two approaches is reported.

Table 2
Main results of the proposed approach. (The figures of merit are expressed as percentage).

	$u_S < 0.06$	$0.06 \leqslant u_S < 0.07$	$0.07 \leqslant u_S < 0.08$	$0.08 \leqslant u_S < 0.09$	$0.09 \leqslant u_S < 0.10$	$0.10 \leqslant u_S < 0.12$	$0.12 \leqslant u_S < 0.15$	$u_S \geqslant 0.15$	Whole dataset
Classification list with one class and CL = 1	100	100	99	95	75	53	40	10	86
Classification list with more subjects where the right class has the highest value	0	0	0	5	22	37	32	46	7
Abstention	0	0	1	0	3	9	26	42	6
"Wrong classification": Classification list with more subjects without the right class or where the right class has not the highest value	0	0	0	0	0	1	2	2	1

Table 3
Main results of the basic approach (The figures of merit are expressed as percentage).

	$u_S < 0.06$	$0.06 \leqslant u_S < 0.07$	$0.07 \leqslant u_S < 0.08$	$0.08 \leqslant u_S < 0.09$	$0.09 \leqslant u_S < 0.10$	$0.10 \leqslant u_S < 0.12$	$0.12 \leqslant u_S < 0.15$	$u_S \geqslant 0.15$	Whole dataset
Right classification	100	100	100	94	79	73	21	10	84
Abstention	0	0	0	6	21	26	77	87	15
Wrong classification	0	0	0	0	0	1	2	3	1

The obtained results prove that the proposed procedure allows a right classification (i.e. the right class tops the classification list with the highest CL) also in presence of high uncertainty: it is able to recover almost always the false rejections thanks to the information on the uncertainty. In detail, it reaches a correct classification in more than 97% of the cases, in presence of uncertainty up to 0.1, whilst with the basic approach, even in presence of an uncertainty equal to 0.01 the percentage becomes less than 80%, and it becomes less than 25% in presence of a very high uncertainty on the measurements. As for the wrong classification percentage, for both approaches, it is low and slightly increases with very high uncertainty.

Considering the whole dataset (see the last column of Tables 2 and 3), the performance improvement due to the proposed decision procedure is clear: we recognize the subject in the 93% of cases, obtaining abstention [39] only in the 6% of cases whilst the basic approach provides 84% and 15% of right classification percentage and false rejection percentage, respectively.

7. Concluding remarks

This paper has introduced and discussed a suitable approach to improve the classification performance of face recognition systems based on 3D features when operating in presence of measurement uncertainties.

In particular, the knowledge of measurement uncertainties involved in the measurement process and its propagation down to the classification stage is exploited to estimate a confidence level of the output results. In more detail, the output of the classification system is thought of as being composed of a list of the most probable classes in order to be associated to the input object each one characterized by a suitable confidence level. Then, such information is adopted to improve the reliability of the output results in terms of correct, wrong decision and rejection percentage.

The study has been conducted by considering both a popular algorithm based on 3D features (the AAM) and main causes of uncertainty generally affecting the performance of face recognition algorithms.

The achieved results prove that the proposed method allows the overall performance to be significantly improved with respect to a basic classification scheme, in terms of *Correct classification* and *False Rejection* percentage. More in details, the proposed method allows reaching a *Correct classification* percentage greater than 90% also if high values of uncertainty are involved.

We believe that the proposed approach can be usefully extended to be used within many other application areas such as text recognition, speech identification, and biometry based on face, fingerprint, and iris analysis.

Acknowledgements

The research Project partially described in this paper has been partially funded by a PRIN Grant by the Italian Ministry of Education, University and Research (MIUR).

The authors wish thank Dr. s Mariella Corvino and Anna Cangiano for the help given in the experimental phase.

References

[1] A.S. Raju, V. Udayashankara, Biometric person authentication: a review, in: 2014 International Conference on Contemporary Computing and Informatics (IC3I), SJCE, Mysuru, India, 27–29 November 2014.
[2] Syed M.S. Islam, Mohammed Bennamoun, Robyn A. Owens, Rowan Davies, A review of recent advances in 3D Ear- and expression-invariant, Face Biometrics Comput. Surveys 44 (3) (2012). Article 14.
[3] Enrique G. Ortiz, Brian C. Becker, Face recognition for web-scale datasets, Comput. Vision and Image Understanding 118 (2014) 153–170.
[4] Maria De Marsico, Chiara Galdi, Michele Nappi, Daniel Riccio, FIRME: face and iris recognition for mobile engagement, Image Vision Comput. 32 (2014) 1161–1172.
[5] Amirhosein Nabatchian, Esam Abdel-Raheem, Majid Ahmadi, Illumination invariant feature extraction and mutual-information-based local matching for face recognition under illumination variation and occlusion, Pattern Recognit. 44 (2011) 2576–2587.
[6] Deng-Yuan Huang, Shr-Huan Di, Wu-Chih Hu, Yi-Jen Su, Face recognition based on dual-tree complex wavelet transform under low illumination environments, IVCNZ '14, November 19–21 2014, Hamilton, New Zealand, 2014.
[7] Hu Lacey Best-Rowden, Charles Otto Han, Brendan F. Klare, Anil K. Jain, Unconstrained face recognition: identifying a person of interest from a media collection, IEEE Trans. Inf. Forensics Security 9 (12) (December 2014).
[8] Abdenour Hadid, Jean-Luc Dugelay, Matti Pietikäinen, On the use of dynamic features in face biometrics: recent advances and challenges, SIViP 5 (2011) 495–506.
[9] Zhiwei Zhang, Junjie Yan, Sifei Liu, Zhen Lei, Dong Yi, Stan Z. Li, A face antispoofing database with diverse attacks, in: Proceedings of The 5th IAPR International Conference on Biometrics (ICB 2012), New Delhi, India, 2012.
[10] Meng Yang, Zhizhao Feng, Simon C.K. Shiu, Lei Zhang, Fast and robust face recognition via coding residual map learning based adaptive masking, Pattern Recognit. 47 (2014) 535–543.
[11] Stan Z. Li, Anil K. Jain (Eds.), Handbook of Face Recognition, second ed., Springer, 2005
[12] P.J. Phillips, P.J. Flynn, T. Scruggs, K.W. Bowyer, Jin Chang, K. Hoffman, J. Marques, M. Jaesik, W. Worek, Overview of the face recognition grand challenge, in: Proceedings of IEEE Computer Vision and Pattern Recognition Conference, 2005, pp. 947–954.
[13] W. Zhao, R. Chellapa, P.J. Phillips, A. Rosefeld, Face recognition: a literature survey, J. ACM Comput. Surv. (2003) 399–458.
[14] M. Grudin, On internal representations in face recognition, Pattern Recognit. 33 (2000) 1161–1177.
[15] J. Lu, K. Plataniotis, A. Venetsanopoulos, Face recognition using kernel direct discriminant analysis algorithms, IEEE Trans. Neural Networks 14 (2003) 117–126.
[16] R. Gross, J. Shi. J. Cohn, Quo vadis Face Recognition? Report of Robotics Institute of Carnegie Mellon University, 2001.

[17] Y. Hea, L. Zhao, C. Zou, Face recognition using common faces method, Pattern Recognit. (2006) 2218–2222.
[18] B.A. Draper, K. Baek, M.S. Bartlett, J.R. Beveridgea, Recognizing faces with PCA and ICA, Comput. Vision Image Understanding (2003) 115–137.
[19] A.A. Mohammed, R. Minhas, Q.M. Jonathan Wu, M.A. Sid-Ahmed, Human face recognition based on multidimensional PCA and extreme learning machine, Pattern Recognit. 44 (2011) 2588–2597.
[20] G. Betta, D. Capriglione, F. Crenna, M. Gasparetto, C. Liguori, A. Paolillo, G.B. Rossi, E. Zappa, Face-based recognition techniques: proposals for the metrological characterization of global and feature-based approaches, Meas. Sci. Technol. 22 (12) (2011).
[21] E. Zappa, R. Testa, M. Barbesta, M. Gasparetto, Uncertainty of 3D facial features measurements and its effects on personal identification, Measurement (2014).
[22] F. Crenna, E. Zappa, L. Bovio, R. Testa, M. Gasparetto, G.B. Rossi, Implementation of perceptual aspects in a face recognition algorithm, J. Phys.: Conf. Ser. 459 (2013) 012031, http://dx.doi.org/10.1088/1742-6596/459/1/012031.
[23] G. Zhang, Face recognition based on fuzzy linear discriminant analysis, IERI Proc. 2 (Aug. 2012) 873–879.
[24] R.-C. David, C.-A. Dragos, R.-G. Bulzan, R.-E. Precup, E.M. Petriu, M.-B. Radac, An approach to fuzzy modeling of magnetic levitation systems, Int. J. Artificial Intell. 9 (A12) (October 2012) 1–18.
[25] A.S. Aisjah, S. Arifin, Maritime weather prediction using fuzzy logic in Java Sea for shipping feasibility, Int. J. Artificial Intell. 10 (S13)(Mar. 2013) 112–122.
[26] D. Zhang, Q.-G. Wang, L. Yu, H. Song, Fuzzy-model-based fault detection for a class of nonlinear systems with networked measurements, IEEE Trans. Instrum. Meas. 62 (12) (Dec. 2013) 3148–3159.
[27] G. Betta, D. Capriglione, C. Liguori, A. Paolillo, Uncertainty evaluation in face recognition algorithms, in: IEEE Proceedings of Instrumentation and Measurement Technology Conference I2MTC11, Binjiang, China, May 2011.
[28] G. Betta, D. Capriglione, M. Corvino, C. Liguori, A. Paolillo, Face based recognition algorithms: A first step toward a metrological characterization, IEEE Trans. Instrum. Meas. 62 (5) (2013) 1008–1016. art. no. 6493520.
[29] G. Betta, D. Capriglione, M. Corvino, C. Liguori, A. Paolillo, Face based recognition algorithms: The use of uncertainty in the classification, in: IEEE Proceedings of Instrumentation and Measurement Technology Conference I2MTC13, 2013, pp. 1098–1103 (art. no. 6555584).
[30] T.F. Cootes, J.E. Gareth, J.T. Christopher, Active appearance models, IEEE Trans. Pattern Anal. Mach. Intell. 23 (6) (2001) 681–685.
[31] Cong Geng, Xudong Jiang, Face recognition based on the multi-scale local image structures, Pattern Recognit. 44 (2011) 2565–2575.
[32] Zhenhua Chai, Zhenan Sun, Heydi Méndez-Vázquez, Ran He, Tieniu Tan, Gabor ordinal measures for face recognition, IEEE Trans. Inf. Forensics Security 9 (1) (January 2014).
[33] J.E. Gareth, T.F. Cootes, C.J. Taylor, Face Recognition Using Active Appearance Models, Computer Vision—ECCV'98, Springer, Berlin Heidelberg, 1998, pp. 581–595.
[34] E. Zappa, P. Mazzoleni, Reliability of personal identification base on optical 3D measurement of a few facial landmarks, Proc. Comput. Sci. 1 (1) (2010) 2769–2777.
[35] E. Zappa, P. Mazzoleni, Yumei Hai, Stereoscopy based 3D face recognition system, Proc. Comput. Sci. 1 (1) (2010) 2521–2528.
[36] Z. Zhang, A flexible new technique for camera calibration, IEEE Trans. Pattern Anal. Mach. Intell. 22 (11) (2000) 1330–1334.
[37] JCGM 100, Evaluation of measurement data — Guide to the expression of uncertainty in measurement, ISO-GUM, 2008.
[38] G. Betta, D. Capriglione, M. Gasparetto, C. Liguori, A. Paolillo, E. Zappa, Managing the uncertainty for face classification with 3D features, in: IEEE Proceedings of Instrumentation and Measurement Technology Conference I2MTC14, 2014, pp. 412–417.
[39] F. Tortorella, An optimal reject rule for binary classifiers, Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics), vol. 1876 LNCS, 2014, pp. 611–620.