J Math Imaging Vis
DOI 10.1007/s10851-010-0252-0

Three-Dimensional Occlusion Detection and Restoration of Partially Occluded Faces

Alessandro Colombo · Claudio Cusano · Raimondo Schettini

Abstract This paper presents an innovative three dimensional occlusion detection and restoration strategy for the recognition of three dimensional faces partially occluded by unforeseen, extraneous objects. The detection method considers occlusions as local deformations of the face that correspond to perturbations in a space designed to represent non-occluded faces. Once detected, occlusions represent missing information, or "holes" in the faces. The restoration module exploits the information provided by the non-occluded part of the face to recover the whole face, using an appropriate basis for the space in which non-occluded faces lie. The restoration strategy does not depend on the method used to detect occlusions and can also be applied to restore faces in the presence of noise and missing pixels due to acquisition inaccuracies. The strategy has been experimented on the occluded acquisitions taken from the Bosphorus 3D face database. A method for the generation of real-looking occlusions is also presented. Artificial occlusions, applied to the UND database, allowed for an in-depth analysis of the capabilities of our approach. Experimental results demonstrate the robustness and feasibility of our approach.

Keywords Three-dimensional face detection · Three-dimensional face recognition · Face occlusions · Gappy principal component analysis · Global registration · Face restoration

A. Colombo (✉) · C. Cusano · R. Schettini
DISCo (Dipartimento di Informatica, Sistemistica e Comunicazione), Università degli Studi di Milano–Bicocca, Viale Sarca 336, 20126 Milano, Italy
e-mail: colomboal@disco.unimib.it

C. Cusano
e-mail: cusano@disco.unimib.it

R. Schettini
e-mail: schettini@disco.unimib.it

Published online: 21 December 2010

1 Introduction

The growing availability of three dimensional imaging systems has paved the way to the use of 3D face models for face recognition (see [8]). Although current automatic recognition systems have reached a respectable level of reliability, [1, 6, 11, 12, 15, 30, 33, 35, 39], their success is limited by the conditions many practical applications impose. three dimensional face recognition in the presence of varied pose and facial expression, unforeseen occlusions, or high acquisition noise that may cause large and complex-shaped holes (missing 3D data), is still a largely unsolved problem. In particular, a recognition system that cannot reliably deal with partial occlusions cannot recognize uncooperative subjects and consequently is not suitable for many applications. Moreover, even when the cooperation of the subjects is not a problem (e.g. for an authentication system) these systems may inconvenience users, who must remove their eyeglasses, and check that their hair does not hide their face.

In this paper we present an innovative three dimensional detection and restoration strategy for the recognition of three dimensional faces which may be partially occluded by unforeseen, extraneous objects. No a-priori knowledge about the occluding objects is required. These may be glasses, hats, scarves and the like, and differ greatly in shape or size, introducing a high level of variability in appearance. The restoration strategy is independent of the method used to detect occlusions and can also be applied to restore faces in the presence of noise and missing pixels due to acquisition inaccuracies.

The greater part of state of the art work in face recognition does not consider the occlusion issue; only a few approaches which consider 2D images can be found in the literature. When the occluding object is known, a specific strategy can be developed for its detection and for reliable

face recognition. For example, Park et al. have proposed a method for removing glasses from a frontal image of the human face [37]. They first detect the regions occluded by the glasses and then generate a natural looking facial image without glasses using Principal Component Analysis (PCA) reconstruction.

A more general solution is needed, however, when the occlusions are unforeseen and the characteristics of the occluding objects are unconstrained. The problem has been addressed using local approaches which divide the face into parts which are independently compared. The final outcome is determined by a voting step. For example, Martinez has divided each face into local regions which are analyzed separately [31, 32]. Kim et al. have proposed a part-based local representation method called Locally Salient Independent Component Analysis (LS-ICA) [27]. Gökberk et al. [25] investigated the effects of deception attacks (occlusions and camouflage) on two basic face recognition systems: a PCA-based system and a Gabor wavelet-based recognizer. They proposed a distance called Asymmetric Trimmed Distance Measure that selects and combines only the regions where the similarities (between the probe and the gallery faces) are high. A different approach has been investigated by Tarrés and Rama [42]. Instead of searching for local non-occluded features, they try to eliminate some features which may hinder recognition accuracy in the presence of occlusions or changes in expression. Zhang et al. [44] proposed a method based on local Gabor Binary patterns. They divided the faces into rectangular region and, on the basis of their graylevel histograms, estimate the probability of occlusion. Park et al. [36] presented a method for face description based on an attributed relational graph. According to their experimental results, this description is quite robust with respect to various facial expression changes, varying illumination, and partial occlusions.

Mo et al. [34] approached the problem of the reconstruction of occluded faces by using an appearance model which represent shape and texture variations: given a target face an exact a map of the occluded regions, they select a subset of the training faces that are relatively similar to the target face. The target face is then represented by a positive linear combination of the chosen faces. Weights are chosen by evaluating the match of unoccluded regions of the faces. Once known parts of the face are fit, the same linear mixture is used to extrapolate the unknown regions.

An approach based on morphable models has been presented by De Smet et al. [21]. They describes an algorithm which iteratively estimates the parameters of a 3D morphable face model to approximate the appearance of a face in a 2D image. Simultaneously, a visibility map is computed which segments the image into visible and occluded regions. Another approach based on morphable models has been investigated by Hwang and Lee [26].

The problem of detecting occlusions in 2D faces has been investigated by Lin and Tang [29]. They derived a Bayesian formulation unifying the occlusion detection and recovery stages. They also developed a quality assessment model which drives both the detection and recovery processes.

We and other research groups have recently decided to work with 3D images of faces because 3D data is invariant with respect to changes in lighting conditions, and the variability in pose and scale can be greatly reduced by geometric normalization [4, 9, 14, 16, 17, 19].

The problem of occlusions has been already addressed in the context of 3D object retrieval [41], but although occlusions are one of the major sources of intra-class variability in 3D faces few methods exist, as far as we know, for the three dimensional recognition of partially occluded faces.

Alyüz et al. proposed a part based method, which uses average regional models to independently match (via ICP) various parts of the face [2]. They experimented various fusion techniques to integrate the similarity between the face parts and the corresponding average regional models. They reported a significant improvement in recognition rates with respect to global ICP matching in the cases of facial expressions and occlusions (hand, hair, or eyeglasses). However, it must be noted that their experiments relied on manually annotated feature points.

The innovative strategy illustrated here approaches this problem by performing a restoration of the faces: the three dimensional occluded regions are detected, and the non-occluded regions are then used to recover the missing information. Our occlusion detection method, inspired by the work of Skocaj and Leonardis [41], considers three dimensional occlusions as local deformations of the face that correspond to perturbations in a face space designed to represent non-occluded faces. Once detected, occlusions represent missing information which can be seen as "holes" in the faces. The restoration module recovers the whole face by exploiting the information provided by the non-occluded part of the face, and using a basis that is appropriate for the face space in which the non-occluded faces lie.

Our strategy can therefore be employed in any three dimensional and/or multimodal face recognition system, between the detection/normalization stage and that of feature extraction (Fig. 1). For instance, see [20] for a 3D face detection and normalization procedure which could be used in the presence of occlusions (although, at least the eyes, or an eye and the tip of the nose must be visible). The occlusion detector may be directly exploited by a suitable recognition algorithm capable of matching only the parts detected as uncovered (see, for example, Bronstein et al. [10]). However, we think that the restoration approach presents several advantages. The restored faces may be analyzed by any 3D recognition algorithm, whether holistic or feature-based, decoupling the problem of occlusions from that of matching. More advanced recognition algorithms may take into

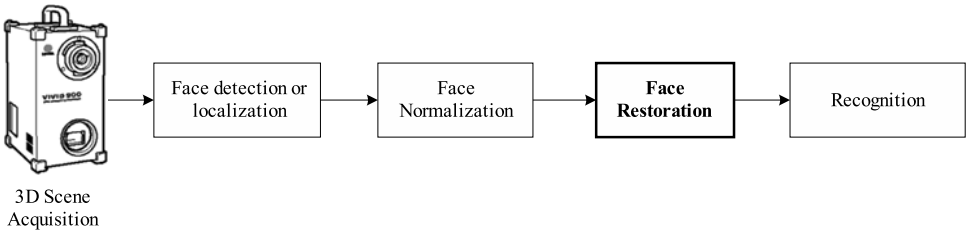

Fig. 1 The face restoration module can be used in any face recognition system before recognition and after normalization

account the information provided by the occlusion detector by assigning a different degree of reliability to the features extracted from the non-occluded and from the restored regions. Moreover, strategies based on multiple recognition algorithms can easily incorporate the restoration module in order to deal with partially occluded faces. Even in the presence of an occlusion-aware recognition system, restoration may still be useful for human operators for the visualization, monitoring, and browsing of recognition events.

Performance evaluation is a crucial issue in 3D face recognition, and it is not trivial to design a sound experimental procedure in the case of occlusions. Moreover, even if a large number of acquisitions of occluded faces is available, the assessment of the accuracy of the methods is problematic: the ground-truth (i.e. an indication of the occluded parts of the faces) can be determined only by a tedious and error-prone manual procedure, and information about the regions covered by the occlusions is not available. The first version of our algorithm (presented in [18]) has been evaluated on a small dataset of subjects acquired by a laser scanner with the face covered by various occluding objects. Motivated by the difficulties encountered in assessing the performance of the method with real occlusions, in this work we decided to synthesize realistic acquisitions using an artificial generator of occlusions. An occluded face is generated by placing an acquisition of a real object on a non-occluded face. Since we used only 3D information, we avoid the problem of generation of a realistic texture. The results are that a human observer cannot easily distinguish between our artificially occluded faces and acquisitions with real occlusions. The use of artificial occlusions presents several advantages:

- a large, publicly available database of faces may be used and therefore the results obtained may be compared with the state of the art in recognition of non-occluded 3D faces;
- a great variety of occlusions may be generated, varying in shape, position, and size, allowing for an in-depth evaluation under different conditions;
- since the mask of occluded regions is immediately available (without manual annotation) we can accurately measure the precision in detecting occlusions;
- the fidelity of the faces reconstructed by the restoration module can be evaluated using the information provided by the original non-occluded faces;
- 3D faces can be accurately normalized before the addition of the occlusion. This allows the evaluation of the algorithms in a way that excludes normalization errors.

The strategy has been evaluated using the UND database [13, 24] artificially occluded by various three dimensional objects. We evaluated three aspects of the restoration strategy: the accuracy of the occlusion detection, the restoration error, and the impact of restoration on recognition performance.

We also performed recognition experiments on real occlusions taken from the Bosphorus database [40]. In this case, we used the occlusion invariant 3D face detection algorithm described in [20] to build a full-automatic detection and recognition pipeline which is able to handle the presence of occluding objects.

Our methods for the detection and localization of 3D occlusions and for the restoration of partially occluded faces are described in Sect. 2. Section 3 describes the dataset used in our experimentation and presents the results obtained. Finally, Sect. 4 reports our conclusions and introduces our plans for future improvements.

2 3D Detection of Occlusions

The detection strategy consists of two major steps. First the most evident occlusions are detected, then a refined occlusion mask is determined and used to restore the occluded faces.

2.1 Initial 3D Detection of Occlusions

The three dimensional faces are encoded as range images, that is, images whose pixels are labeled with the coordinates of a point in 3D space. The range images have been obtained by an orthographic projection of the acquired scenes on a plane, in such a way that the x and y coordinates of adjacent pixels differ by a fixed value. This representation (often referred as to "2.5D") allows the application of 2D algorithms by considering the z-coordinate as a pixel value, and x and y coordinates as the indices of a matrix. For surfaces like human faces, the orthographic projection retains almost all the information about three dimensional shape.

We consider any part of the acquired 3D scene that does not look like part of a face and lies between the acquisition device and the acquired face to be a generic occlusion (occluding objects may not touch the face). In other words, sets of points which do not fit the model of a non-occluded 3D face are classified as occlusions.

In the present implementation the 3D face model is obtained by the popular eigenface approach based on principal component analysis [28, 43]. Each face is represented by a linear combination of a small set of images called eigenfaces:

$$\mathbf{x} + \mathbf{e} = \mathbf{m} + \sum_{i=1}^{N} y_i \mathbf{v}_i, \tag{1}$$

where $\mathbf{x}$ is the input range image (encoded as a d-dimensional vector), $\mathbf{e}$ is the approximation error, $\mathbf{m}$ is the range image of the mean face, $\mathbf{v}_i$ are the N eigenfaces considered, and y_i are the coefficients of the linear combination. The mean face and the eigenfaces are selected by maximizing the amount of variability retained in a training set of non-occluded face range images. Each coefficient y_i is obtained by projecting the vector $(\mathbf{x} - \mathbf{m})$ onto the corresponding eigenface $\mathbf{v}_i$. Because of the orthographic projection and the normalization step only the z coordinate of the pixels need to be considered. An alternative PCA-based approach consists in using all the three spatial coordinates to encode the faces. This approach would require a dense point-to-point correspondence between each face and a reference face, instead of the orthographic projection. Blanz and Vetter, for instance, defined a 3D morphable model where correspondences are established by a 3D variant of the optical flow [7].

The quantity $\|\mathbf{e}\|$ is called the Distance From Features Space (DFFS) and can be considered a measure of "faceness" [38]. Occluded faces cannot be well represented by the linear combination of the computed eigenfaces. Therefore the DFFS of occluded faces is expected to be quite large and can be used to reveal the presence of occlusions, although not their location. Differences in the pixels of the reconstructed and the original image (vector $\mathbf{e}$) are likely to be more pronounced where the face is occluded.

A preliminary mask $\mathbf{M}$ of occlusions can be obtained by thresholding vector $\mathbf{e}$:

$$\mathbf{M}_i = \begin{cases} 1 & \text{if } \mathbf{e}_i > \tau, \\ 0 & \text{if } \mathbf{e}_i \leq \tau, \end{cases} \tag{2}$$

where τ is a threshold that must take into account the resolution of the imaging device, acquisition noise, and the accuracy achieved in face detection and pose normalization (Sect. 3 reports an analysis on the influence of τ on occlusion detection performance).

Since occlusions must be located between the acquisition device and the acquired face, their z coordinates will be smaller than those of the reconstructed face, resulting in positive components of $\mathbf{e}$ (we consider z as increasing with the distance from the acquisition device). Unfortunately, the reconstruction error of the non-occluded regions is influenced by the occlusions, which may determine an inaccurate choice of the coefficients y_i in (1). However the effect is significant only for occlusions which are very far from the face, or very large in size. On the basis of this intuition we adopted a two-phase approach: first those points that are very likely to be occluded, and which can introduce excessive perturbations in the PCA reconstruction are detected (we denote the resulting mask as $\mathbf{B}$); these points are excluded from the second step, which searches for all the occluded points (mask $\mathbf{M}$). Occluding objects which are far from the face are detected on the basis of the difference between the (possibly) occluded face $\mathbf{x}$ and the mean (non-occluded) face $\mathbf{m}$. Components of $\mathbf{m} - \mathbf{x}$ which are positive and large enough can be considered part of an occlusion. A mask $\mathbf{B}$ of these occlusions is obtained as follows:

$$\mathbf{B}_i = \begin{cases} 1 & \text{if } \mathbf{m}_i - \mathbf{x}_i > \nu, \\ 0 & \text{if } \mathbf{m}_i - \mathbf{x}_i \leq \nu, \end{cases} \tag{3}$$

ν threshold must be tolerant with respect to face variability in the data set to be processed. We decided to keep this threshold very high, so that it detects only really prominent occlusions including a negligible fraction of false positives. On our experimental data (see Sect. 3), it is very unlikely that any face differs more than 15 mm from the mean face. Consequently we set threshold ν at that value.

2.2 Refined 3D Detection of Occlusions and Face Restoration

A more accurate estimate of occlusion mask $\mathbf{M}$ can be obtained by excluding the pixels selected in mask $\mathbf{B}$ from the computation of the reconstruction error $\mathbf{e}$.

The key idea here is to use only non-occluded parts of the face to estimate the distance between the face space and the processed face. More precisely, given a 3D face $\mathbf{x}$, and the corresponding mask $\mathbf{B}$, the coefficients y_i in (1) are determined again excluding the components $\mathbf{x}_i$ for which $\mathbf{B}_i = 1$. To this end, we apply Gappy Principal Component Analysis (GPCA), a variation of Principal Component Analysis, designed for the analysis of incomplete patterns [23]. The incomplete (occluded) face $\mathbf{x}$ is expressed as the linear combination of the eigenfaces, as in (1), where the coefficients y_i are computed considering only the available information (the non-occluded parts) in minimizing the distance of the processed face from the face space. We let $(\mathbf{u}, \mathbf{v})_{\mathbf{B}}$ be the gappy inner product, depending on the occlusion mask $\mathbf{B}$:

$$(\mathbf{u}, \mathbf{v})_{\mathbf{B}} = \sum_{i=1}^{d} \mathbf{u}_i \mathbf{v}_i (1 - \mathbf{B}_i). \tag{4}$$

The corresponding gappy norm $\|\mathbf{v}\|_{\mathbf{B}}$ is defined as $\sqrt{(\mathbf{v},\mathbf{v})_{\mathbf{B}}}$.

The coefficients y_i which encode the processed face in the face space are chosen to minimize $\|\mathbf{e}\|^2_{\mathbf{B}}$, the reconstruction errors on the non-occluded parts:

$$\begin{aligned}\|\mathbf{e}\|^2_{\mathbf{B}} &= \left\| -\mathbf{x}+\mathbf{m}+\sum_{i=1}^{N} y_i \mathbf{v}_i \right\|^2_{\mathbf{B}} \\ &= \|\mathbf{x}-\mathbf{m}\|^2_{\mathbf{B}} - 2\sum_{i=1}^{N} y_i (\mathbf{x}-\mathbf{m},\mathbf{v}_i)_{\mathbf{B}} \\ &\quad + \sum_{i=1}^{N}\sum_{j=1}^{N} y_i y_j (\mathbf{v}_i,\mathbf{v}_j)_{\mathbf{B}}. \end{aligned} \tag{5}$$

Differentiating with respect to each y_i, and requiring that the partial derivatives be null, we obtain a system of N linear equations, where N is the number of eigenfaces retained:

$$\frac{\partial \|\mathbf{e}\|^2_{\mathbf{B}}}{\partial y_i} = -2(\mathbf{x}-\mathbf{m},\mathbf{v}_i)_{\mathbf{B}} + 2\sum_{j=1}^{N} y_j (\mathbf{v}_j,\mathbf{v}_i)_{\mathbf{B}} = 0, \tag{6}$$

for i in $\{1,\ldots,N\}$. The coefficients obtained by solving (6) can now be substituted in (1) to determine a more accurate reconstruction error.

The same technique may be applied to deal with cases where the input 3D face is incomplete. Some acquisition devices, in fact, cannot acquire certain kinds of surfaces: eyebrows and hair, for example, are not always captured by laser scanners, creating holes in the acquired faces. These holes can be encoded in a mask $\mathbf{H}$, so that the gappy products $(\cdot,\cdot)_{\mathbf{B}}$ may be replaced in (6) by $(\cdot,\cdot)_{\mathbf{B}\vee\mathbf{H}}$, where $\mathbf{B}\vee\mathbf{H}$ is the component-wise logical OR of holes and occlusions. It should be noted that at this point color information can be exploited if available: points having a color which is not likely to be found on a face could be marked as occlusions. In our future work we plan to investigate this issue.

On the basis of the new reconstruction error $\mathbf{e}$ obtained using the coefficients y_i estimated by solving the system in (6), the occlusion mask $\mathbf{M}$ can now be determined more precisely. Here again, when the reconstruction error is high, the corresponding pixel is considered occluded. A variation of (2) is used:

$$\mathbf{M}_i = \begin{cases} 1 & \text{if } (\mathbf{e}_i > \tau) \vee \mathbf{B}_i = 1 \vee \mathbf{H}_i = 1, \\ 0 & \text{otherwise.} \end{cases} \tag{7}$$

As a final step, morphological filters can be used to clean the occlusion mask, enforce the locality of the occlusion, and discard tiny regions. This post processing step could be improved by adopting a more complex (and computationally expensive) approach; for instance, in [21] Markov Random Fields are used for a similar purpose. If the resulting mask $\mathbf{M}$ is empty, the face is considered non-occluded, and can be directly processed for recognition. Otherwise $\mathbf{M}$ replaces $\mathbf{B}$ in (6):

$$\frac{\partial \|\mathbf{e}\|^2_{\mathbf{M}}}{\partial y_i} = -2(\mathbf{x}-\mathbf{m},\mathbf{v}_i)_{\mathbf{M}} + 2\sum_{j=1}^{N} y_j (\mathbf{v}_j,\mathbf{v}_i)_{\mathbf{M}} = 0, \tag{8}$$

for i in $\{1,\ldots,N\}$. This linear system (8) yields the coefficients y_i which employed as indicated in (1) allow the restoration of the input face. The restored face can be used as is, alternatively it could be blended with the parts of the original face which have not been detected as occluded; we chose to not apply such a blending to avoid the formation of artifacts near the border between occluded and non-occluded points.

At this stage, any recognition algorithm, whether holistic or feature-based, can then be applied to recognize the restored face. The whole procedure is summarized as a flow chart in Fig. 2. Figure 3 shows an example of occlusion detection and restoration of the face of a subject wearing glasses.

3 Experimental Results

We decided to test our approach in two ways: first, we experimented artificial occlusions on a publicly available UND database [13, 24]; then we experimented real occlusions from the Bosphorus DB [40]. The first experiment allowed us to perform an in-depth analysis of the characteristics of our approach, including its capability to detect the occlusions, to restore the faces, and to recognized the restored faces. This is made possible by the use of artificial occlusions, for which we exactly know the ground truth. Since artificially occluded faces can be precisely normalized (by considering the non-occluded version of the faces) we can also assess the performance of the proposed approach in a way that excludes normalization errors. In addition, we also considered a fully automatic pipeline which includes modules for an occlusion tolerant detection and normalization of faces. We used this automatic pipeline also in the second experiment, where we evaluated our approach on real occlusions.

3.1 Artificial Occlusions Applied to the UND Database

The use of artificially generated occlusions is a common practice in this field. Typically occlusions are generated by corrupting the original image introducing a patch of some fixed value [27, 32, 42] or blobs of noisy pixels [41]. In our opinion this type of data does not represent well the complexity of real occlusions and the results cannot be directly generalized to the real case. The obvious solution of using

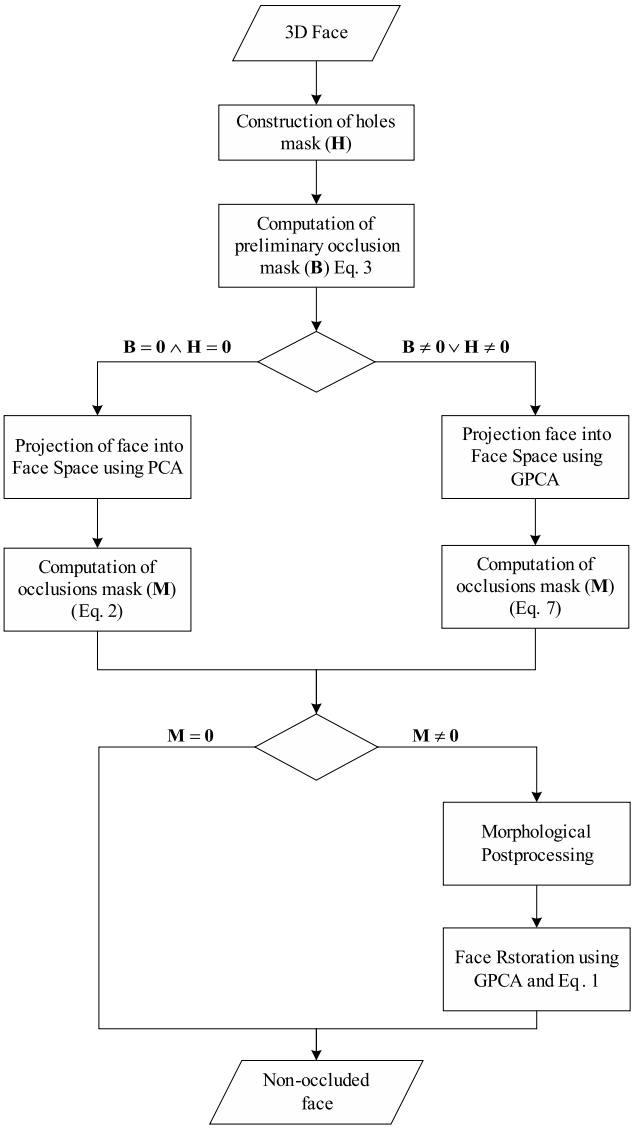

Fig. 2 Diagram of the face restoration module

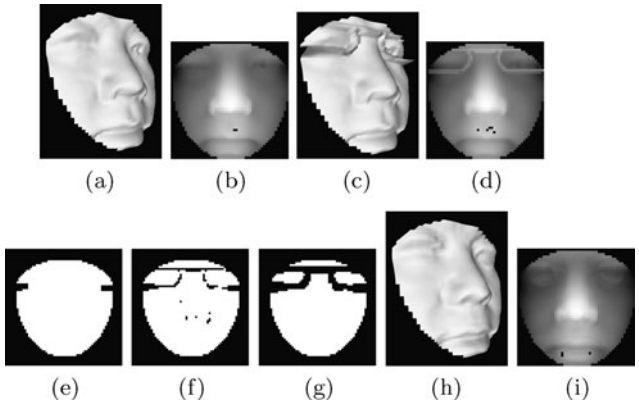

Fig. 3 Occlusion detection and restoration on a face artificially occluded by glasses: (**a**, **b**) original non-occluded face; (**c**, **d**) artificially occluded face; (**h**, **i**) restored face; (**e**) Preliminary detected occlusions; (**f**) occlusions detected after the application of (2); (**g**) Final occlusion mask after the application of morphological operators

acquisition with real occlusions may be problematic: the acquisition of a large number of occluded faces and their annotation with a ground truth are costly and time intensive operations. Moreover, several in-depth analysis may not be possible because information about the regions covered by the occlusions is not available (at least with the most common acquisition devices). We decided to handle this problem using artificial occlusions applied on a large database of non-occluded faces: the UND database [13, 24]. Since we used only 3D information, it is possible to produce artificially occluded faces which are hard to distinguish from faces with real occlusions even by a human observer. This can be achieved because, with respect to the 2D case, only shape information need to be reproduced disregarding the complexity of light interactions. This solution combines the advantages of synthetic occlusions without losing a believable indication about the performance in real scenarios.

The UND DB is composed of 951 2D+3D frontal acquisitions captured from 277 subjects. We selected the training set by randomly picking three acquisitions for each of the 158 subjects in the database who are present at least three times. The training set consists then of 474 faces, we used the remaining 477 acquisitions for the test set. Note that there are subjects in the test set who are not included in the training set. Moreover, there are subjects who are present in the training set but not in the test set. More in detail, 119 subjects have less than three acquisitions and have been included only in the test set; 39 subjects have exactly three acquisitions and have been included only in the training set; the remaining 119 subjects have more than three acquisitions and are present in both the training and the test sets.

Artificial occlusions were generated acquiring real-world objects like eyeglasses, hats, scarves etc. and placing them in front of each face of the test set. The objects were acquired using the Minolta Vivid 900 laser range scanner. Note that the same acquisition device has been used to acquire both faces and occluding objects so that they present the same characteristics. The final position of the occluding objects was computed starting from the UND facial features annotation and adding a random perturbation to the position as well as the orientation. For each class of object we defined the ad-hoc starting position and perturbation magnitude; i.e. for example eyeglasses are placed in the eye region and are subjected to small perturbations. The face and the occluding object are both rasterized (according to the same orthographic projection) using the z-buffer algorithm. Figure 4 illustrate the registration and occlusion procedure. Figure 5 shows a small sample of occluded faces. The fraction of occluded face greatly varies in size, as illustrated by the histogram in Fig. 6. Figure 7 reports the distribution of the distance be-

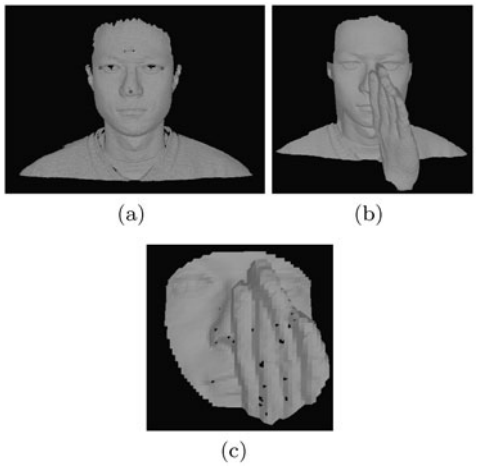

Fig. 4 The normalization procedure used in the experimentation. (**a**) The position of the fiducial points is used to place the face in a standard position. (**b**) The same acquisition artificially occluded. Note that the face image is also preprocessed, i.e smoothing and small holes interpolation is applied. (**c**) The occluded image is rotated, translated, and cropped in order to retain only the central portion of the face

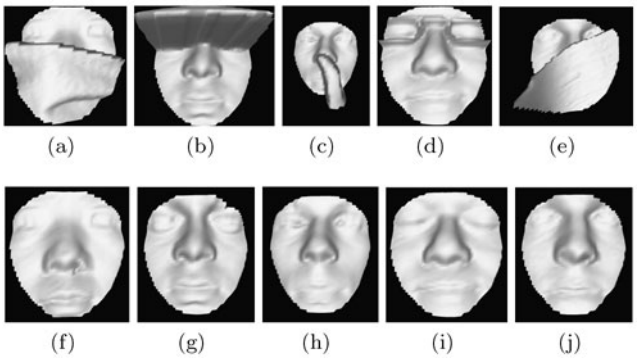

Fig. 5 Artificially occluded faces by a scarf, cap, finger, glasses, and a magazine (**a**–**e**), and the corresponding original images (**f**–**j**)

tween occluded points and the corresponding non-occluded faces included in the test set.

To evaluate face recognition strategies independently of other processing modules, it is common the use of manual or semi-automatic detection and normalization steps [21, 27, 30, 42]. We have applied a semi-automatic normalization procedure to place the faces to be processed in a standard position (additional results obtained with fully automatic pipeline are reported at the end of this section). The UND database annotation specifies several fiducial points; in particular we considered the tip of the nose and the outside corners of the eyes. The registration was performed as follows: the corners of the eyes were fitted by a line in 3D space. The face was rotated to align the line with the horizontal (x) axis. A second rotation was used to register the last degree of freedom: the normal of the plane passing through the tip of the nose and the eyes line was rotated until the angle between the plane of equation $z = 0$

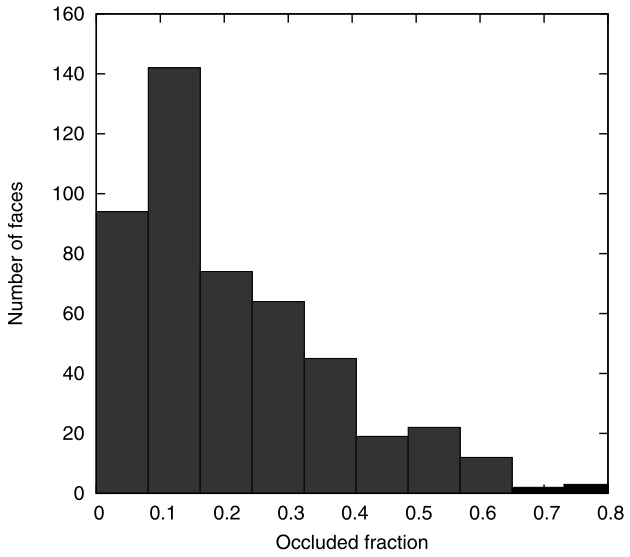

Fig. 6 Histogram of the fraction of face covered by the artificial occlusions included in the test set

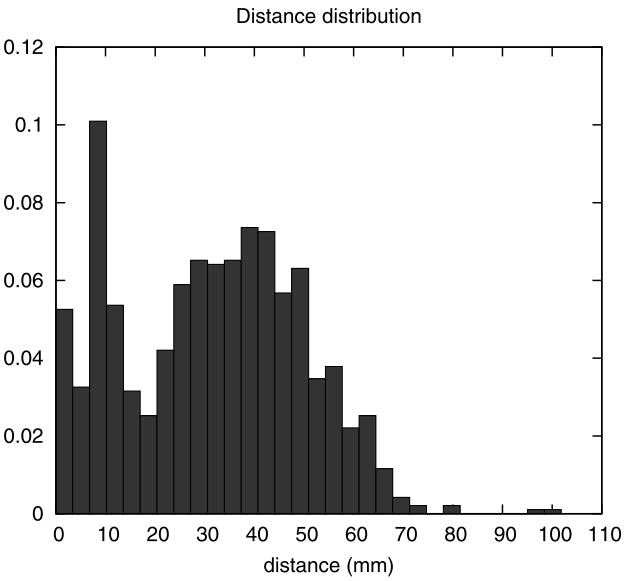

Fig. 7 Distribution of the distance between occluded points and the corresponding non-occluded faces included in the test set

and the normal itself was equal to 45 degrees. Finally, the face was translated placing the tip of the nose at the origin of the reference system. Since three points are not sufficient to obtain a good registration, we used a mean face template and a global registration algorithm to refine the final position. The mean face was computed using acquisitions from the training set normalized as described before. The ICP algorithm [5] was applied to register the non-occluded face surface with the mean face. An orthographic projection has been applied to the registered images, converting them into a matrix of 51×51 samples. Finally, the image was cropped

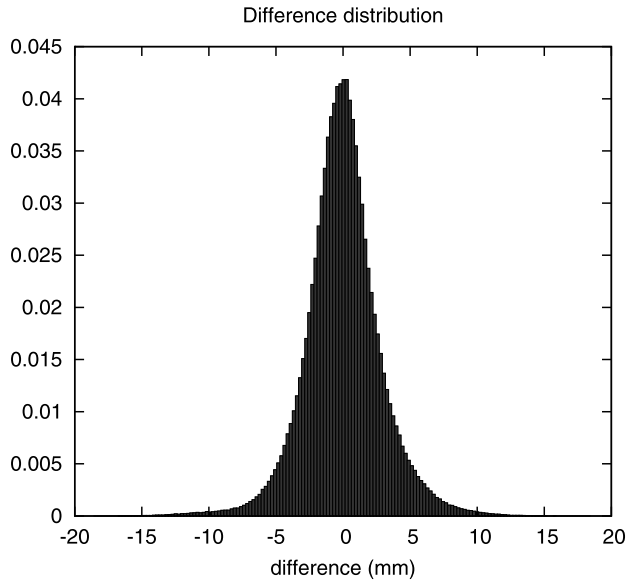

Fig. 8 Histogram of the distribution of the pixel by pixel difference between training faces and the mean face

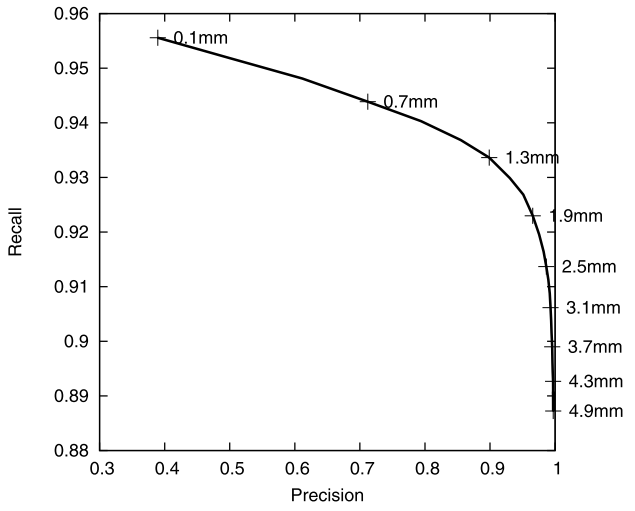

Fig. 9 Precision vs. recall corresponding to the pixels detected as occluded. The plot has been obtained varying the value of the threshold τ. Some values of the threshold have been marked on the plot

(by excluding the points which lie outside a fixed binary mask) to retain only the central portion of the face.

We evaluated three aspects of the restoration strategy: the accuracy of the occlusion detection, the restoration error, and the impact of restoration on identity verification performance. For all the experiments, the same PCA basis has been used. The basis has been built on the training set and consists of 121 eigenfaces (obtained retaining 99% of the variance in the training set).

3.1.1 Occlusion Detection Accuracy

First we considered how the values of the thresholds ν (3) and τ (2) influence the accuracy in the detection of occlusions. We used the training set to estimate the mean face $\mathbf{m}$ and to evaluate the distribution of the differences between the pixels of non-occluded faces and those of the mean face. Taking into account such a distribution, which is reported in Fig. 8, we set the threshold ν to be 15 mm which is more than five times the estimated standard deviation of the distribution ($\sigma = 2.849$ mm). Less than 0.01% of the pixels of the training images exceeds the mean face by that value.

To evaluate the accuracy of the occlusion detection method we ran the procedure on an artificially occluded version of the training set varying the threshold τ. Since we know exactly which parts of the faces were occluded, we can compute the precision (fraction of true positives among the pixels detected as occluded) and the recall (fraction of the occluded pixels which have been detected). The results we obtained are reported in Fig. 9. We also evaluated the occlusion detection method on non-occluded faces. Figure 10 shows the false detection rate (i.e. the fraction of total points

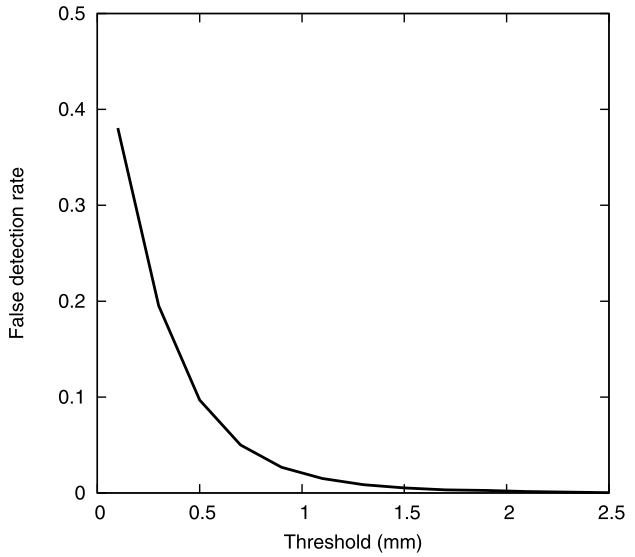

Fig. 10 False detection rate (fraction of total points which have been erroneously detected as occluded) obtained on the training set (UND DB), as a function of the threshold τ

which have been erroneously detected as occluded) obtained on the non-occluded version of the training set for different values of τ. On the basis of these measures we considered the value of 1.9 mm as a good compromise for the threshold τ which has been set to that value for the other experiments. This value corresponds to an estimated precision of about 96.4%, a recall of 92.3%, and a false positive rate of about 0.3%. The same value of the threshold corresponds to 96.5% of precision and 92.1% of recall on the test set.

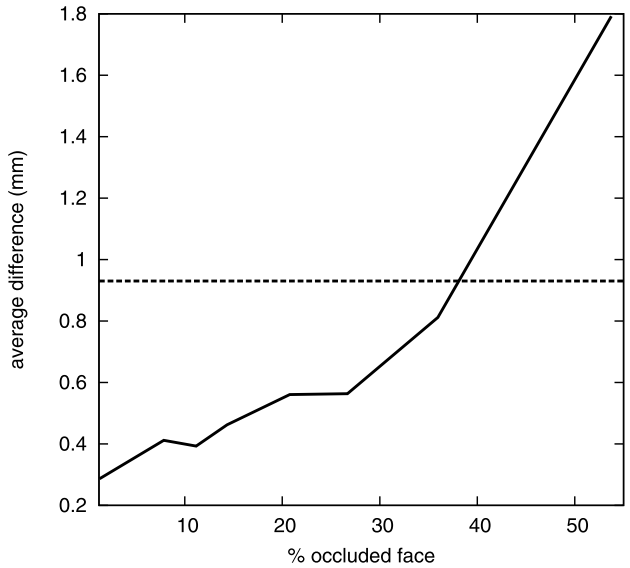

Fig. 11 Average absolute difference of the z coordinate between original and restored pixels, as a function of the fraction of occluded face. The *dashed line* represents the average difference between two non-occluded acquisitions of the same subject

3.1.2 Restoration Accuracy

As a measure of the accuracy of the restoration method we considered the pixel by pixel absolute differences between the original (non-occluded) and the restored faces. More in detail, we considered the output of the restoration procedure applied to each occluded face in the test set and the corresponding original (i.e. non-occluded) acquisition. For each pair of restored and original faces we computed the average absolute difference between the z coordinates of corresponding points. Since the restoration accuracy is expected to be highly correlated to the extent of the occlusions we report the results obtained as a function of the area of occluded regions (see Fig. 11). As illustrated by the examples in Figs. 12, 13, and 14, two major factors affect the quality of the restored faces: errors in the detection of the occlusions (in particular detection misses), and the extent of the occlusions.

To understand how actually significant the restoration error is we estimated the average absolute z difference between the pixels of pairs of non-occluded acquisitions of the same subject. On the non-occluded faces of the test set we measured an average difference of about 0.93 mm. It seems that the error introduced by the restoration process on moderately occluded faces (less than 40% of occluded area) is comparable to imaging inaccuracies and facial expression variations.

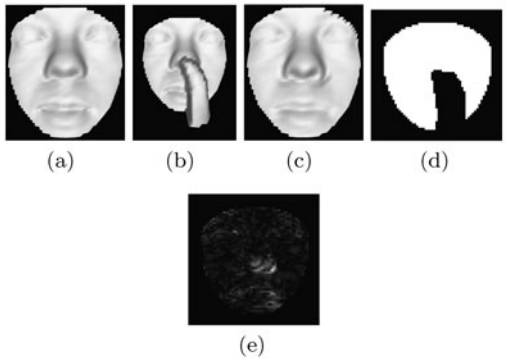

Fig. 12 Restoration of a face partially covered by a finger: (**a**) original non-occluded face; (**b**) artificially occluded face; (**c**) restored face; (**d**) detected occlusions; (**e**) reconstruction error, reported as a gray-level image (dark pixels correspond to low reconstruction error)

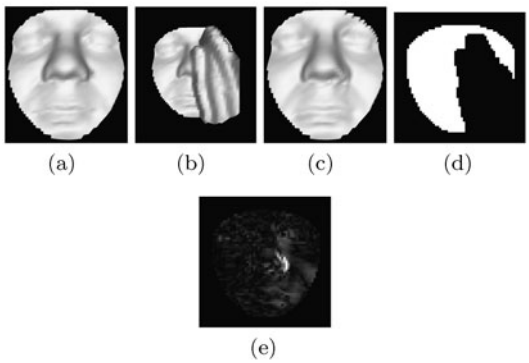

Fig. 13 Restoration of a face artificially occluded by a hand: (**a**) original non-occluded face; (**b**) artificially occluded face; (**c**) restored face; (**d**) detected occlusions; (**e**) reconstruction error, reported as a gray-level image (dark pixels correspond to low reconstruction error)

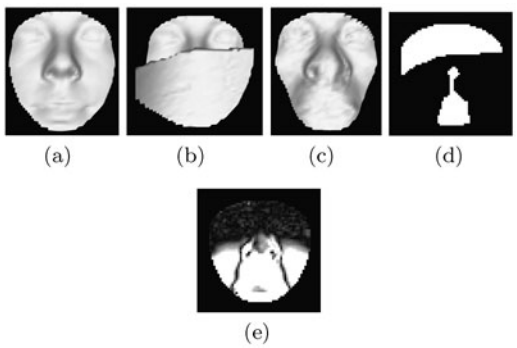

Fig. 14 Example of a failure on a face occluded by a magazine: (**a**) original non-occluded face; (**b**) artificially occluded face; (**c**) restored face; (**d**) detected occlusions; (**e**) reconstruction error, reported as a gray-level image (dark pixels correspond to low reconstruction error)

3.1.3 Recognition Accuracy

Since the restoration error may be unevenly distributed over the faces, recognition performance could be affected more

than expected. In order to ensure that a restored face may be reliably recognized, we applied the same holistic recognition method to the original, the occluded, and the restored faces. Restoration has been applied considering the occlusion mask as computed by the occlusion detector, and an "exact" occlusion mask which includes all the points where the original and the occluded faces differ. We considered a holistic recognition method to ensure that all the pixels, including those which have been restored, are taken into account. We have used the Fisherfaces method [3] here because is very well known and understood, but any other method may be applied. We considered the identity verification scenario in which the system must determine whether the identity claimed by the subject corresponds to his real identity. The training set (non-occluded faces) was used to train the Fisherfaces and populate the gallery of known subjects. To simulate the identity verification scenario each test image has been proposed to the system, each time claiming a different identity (once for each subject included in the training set). Figure 15 shows the Receiver Operator Characteristic curves (which report the False Acceptance Rates and the False Rejection Rates produced with the variation of the acceptance threshold) obtained on non-occluded, occluded, and restored faces (with detected and "exact" occlusions). As expected, the recognition of occluded faces is very difficult: the Equal Error Rate (EER) for occluded faces is in fact, 0.488 which is slightly better than random guessing. The application of our restoration strategy significantly improves the EER to a more acceptable 0.062, which is close to the 0.020 EER obtained on the original faces. Moreover, since very few pixels are misclassified as occluded, recognition performance on non-occluded faces is not affected by restoration. In the case of restoration with exact occlusions we obtained an EER of 0.022 and a ROC curve which is very close to that obtained on non-occluded faces. This fact suggests that a more accurate occlusion detector could significantly improve our strategy.

We also considered the identification scenario in which the system must determine the identity of the subjects. We assumed a closed-world hypothesis, that is, all the faces submitted to the system must correspond to subjects enrolled in the training set. This reduced the test set to 306 acquisitions. The results are coherent with those obtained in the verification scenario. For non-occluded faces we obtained an Identification Rate (IR, the percentage of correctly identified faces) of 99.35%, which corresponds to only two errors. The same IR has been obtained on occluded faces restored with the exact occlusion mask. On the faces restored with the proposed occlusion detector we obtained an IR of 96.08% (12 errors). As expected, a very low identification rate (27.45%) was obtained on artificially occluded faces.

The impact of restoration is related to the extent of the occlusion. We partitioned the test set in three groups: the

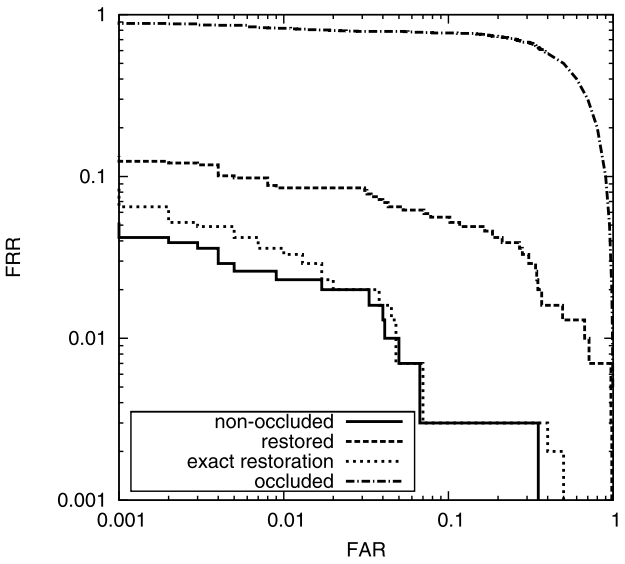

Fig. 15 Receiver Operator Characteristic (ROC) curves for the evaluation of the restoration strategy in the identity verification scenario. The curves (in logarithmic scale) refer to the performances obtained on the original faces, the artificially occluded, and the restored faces (UND DB). Restoration has been performed using the occlusions detected by our method and using an exact occlusion mask (points where the original and the occluded faces differ)

first group contains the 263 faces with less than 20% of the total area occluded ("small" occlusions); the second group contains 148 faces with "medium" occlusions (the occluded area is between 20% and 40%); the third group is composed of the remaining 66 faces ("large" occlusions, ranging from 40% to 82% of the area occluded). Figure 16 reports the ROC curves obtained using the three groups as a test set after the application of the occlusion detection and restoration procedure. In the first case we obtained an Equal Error Rate of 0.048 which, considering the toughness of the problem, is slightly worse than the error rate measured on non-occluded faces. These results demonstrate that the adoption of 3D data is a promising way to deal with the problem of occlusions in face recognition.

Even in the case of medium or large occlusions we obtained better results than those obtained on the non-restored faces (EER of 0.103 and 0.222 respectively).

In Table 1 (first column) we report the identification rates subdivided by the occluding objects type. As it can be seen, the performances are quite similar for all the object types; slightly lower identification rates have been registered for faces occluded by objects located near the mouth (scarf, newspaper, and cup).

3.1.4 Full-Automatic Recognition Pipeline

According to Ekenel and Stiefelhagen [22] normalization is one crucial aspect for accurate recognition of partially oc-

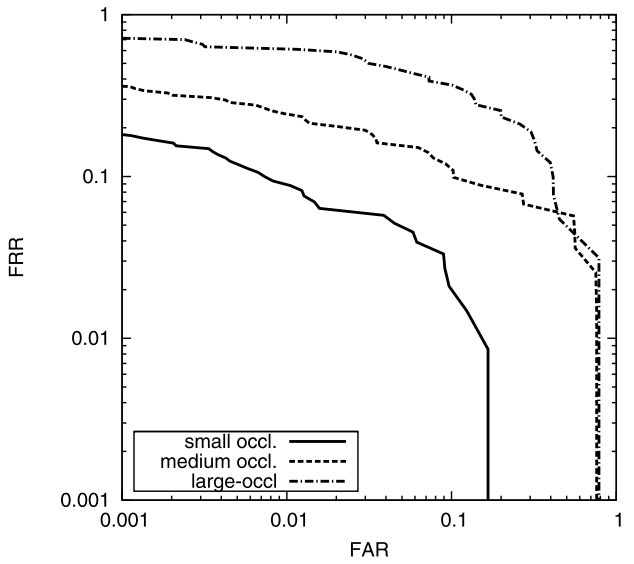

Fig. 16 ROC curves (in logarithmic scale) obtained on three subsets of the test set (UND DB). Faces have been subdivided in the three groups on the basis of the extent of the occlusions. The results have been evaluated on the groups of faces after the application of the occlusion detection and restoration procedure

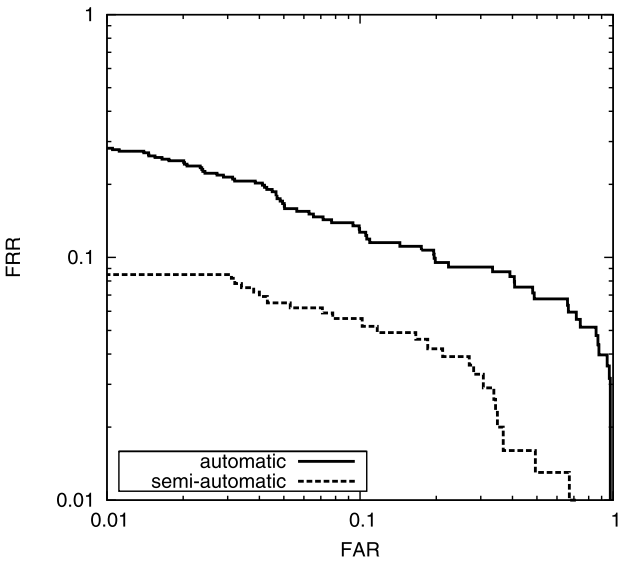

Fig. 17 ROC curves (in logarithmic scale) obtained on the faces of the test set after automatic detection and normalization (UND DB). The plot has been computed on the 401 correctly detected faces, excluding the 76 which have been missed by the face detector

Table 1 Identification performances subdivided by occluding object type obtained on the artificially occluded UND DB by the proposed approach, with manual and automatic normalization

Occlusion type	Manual normalization	Automatic normalization
Cup	100%	81%
Scissors	100%	100%
Eyeglasses	100%	100%
Hat	100%	63.16%
Hand	97.18%	96.49%
Phone	100%	93.75%
Newspaper	92%	87.5%
Scarf	95.92%	46.34%

cluded 2D faces. Therefore, it is important to verify the impact of registration errors to the proposed approach and to verify if, and to what extent, the claim reported in [22] holds for 3D faces as well.

We used a full automatic detection and normalization procedure proposed in [20]. This approach is robust with respect to partial occlusions if at least the eyes, or the nose and one eye are clearly visible. Briefly, the algorithm uses invariant properties of the surfaces to segment the eyes and the nose. Candidate faces are then registered using an ICP-based approach aimed to avoid those samples which belong to the occluding objects. The final face versus non-face discrimination is based on a holistic analysis of the regions of the surface which are considered to be non-occluded. The normalization stage is based on ICP initialized with the position of the detected non-occluded features. More in detail, the detection and normalization algorithm works as follows:

- mean and Gaussian curvatures of the scene are computed;
- candidate eyes and noses are extracted by thresholding the curvatures and by classifying the regions (saddles, elliptical convex regions, elliptical concave regions, etc...);
- candidate faces are generated by considering suitable pairs or triplets of candidates eyes and noses;
- each candidate is registered on the mean face using a version of the ICP algorithm specifically tuned to handle the presence of occluding objects;
- registered candidates are classified as faces or non faces using GPCA where probably occluded points are discarded on the basis of their distance to the mean face.

We applied our restoration approach to the artificially occluded UND dataset after automatic detection and normalization. Out of the 477 images of the test set 76 faces, usually covered by large occlusions, have been missed. Clearly, in evaluating the full pipeline the 76 missed faces should be considered part of the overall system failures; in the following we will focus on recognition errors. Figure 17 show the ROC curve we obtained on the 401 correctly detected faces, compared with that obtained with semi-automatic normalization. The EER is of 0.115. Taking into account the type of detection errors, we consider safe to compare this result with the performance obtained before on faces covered by small and medium occlusions (0.048 and 0.103 of EER, respectively). The performance of the complete pipeline are still reasonable: faces with occlusions of moderate size are very often correctly detected, normalized, and recognized.

In Table 1 we report the identification rates subdivided by the occluding object types using both the automatic and manual pipeline. As it can be seen, the major performance drops correspond to faces occluded by hats or by objects locates near the mouth (scarfs, cups, newspapers).

3.2 Real Occlusions from the Bosphorus DB

In order to experiment our approach on real occlusions, we adopted the Bosphorus 3D face database [40].

The database consists of 105 subjects in various poses, expressions and occlusion conditions. There are 60 men and 45 women in total, mostly between 25 and 35 years old. The majority of the subjects are Caucasian. Also, 27 professional actors/actresses are incorporated in the database. Up to 54 face scans are available per subject, but 34 of these subjects have 31 scans. Thus, the number of total face scans is 4652. Each scan has been manually labelled for 24 facial landmark points such as nose tip, inner eye corners, etc, provided that they are visible in the given scan. Occlusions are classified in four types: mouth, eye, eyeglasses, and hair. For the occlusion of eyes and mouth, subjects were allowed to choose a natural pose for themselves; for example, as if they were rubbing their eyes or as if they were surprised by putting their hands over their mouth. For occlusion with eyeglasses, a pool of different eyeglasses has been collected so that each subject could select at random one of them. Among the male subjects 18 had beard/mustache and 15 had short facial hair. If subjects' hair was long enough, their faces were also scanned with hair partly occluding the face. In general, the subject to subject variation of occlusions is more pronounced as compared to expression variations. For instance, while one subject occludes his mouth with the whole hand, another one may occlude it with one finger only; or hair occlusion on the forehead may vary a lot in tassel size and location.

We enrolled the 299 acquisitions with neutral expressions and frontal pose (3 acquisition for subject, when available) and we used the set of 381 occluded acquisitions for the test set. All the 299 acquisitions for the gallery have been correctly detected and normalized while 21 occluded faces have been missed by our automatic face detector [20], producing a recognition test set of 360 faces. The verification accuracy obtained over the 360 automatically detected and normalized faces is depicted in Fig. 18, where the ROC curve is compared against the curve obtained on the occluded UND DB using the same pipeline, as previously described. The obtained EER is equal to 0.109 which is slightly better if compared to the EER of 0.115 obtained on the UND DB; the ROC curves indicate that the performance are very close.

In general, we consider the acquisitions from the Bosphorus DB to be no more challenging than the artificial occlusions we generated for the UND DB. However, after a manual inspection of the Bosphorus DB, we found some corrupted acquisitions among the set of occluded ones. As reported by the authors of the Boshporus DB, some corrupted acquisitions have been included in the dataset:

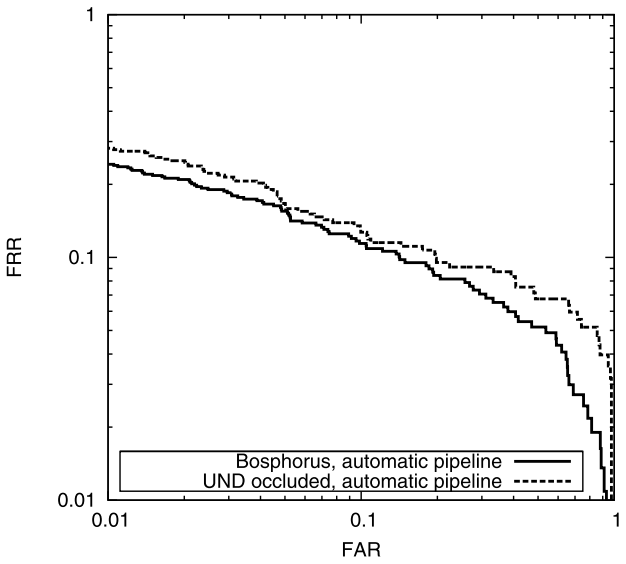

Fig. 18 ROC curves (in logarithmic scale) obtained on the faces of the test sets from the occluded UND database and from the Bosphorus database after automatic detection and normalization. The plot corresponding to the results on the Boshporus DB has been computed on the 360 correctly detected faces, excluding the 21 which have been missed by the face detector. The plot for the UND DB has been computed on the 401 correctly detected faces, excluding the 76 which have been missed by the face detector

- subject movements during acquisitions: faces that were deemed to be seriously faulty were re-captured, but small movements (for instance due to breathing or muscle contractions during expressions) have been tolerated;
- data on hair and facial hair, such as beard and eyebrows, generally causes spiky noise; spiky surfaces arise also over the eyes;
- since data are captured from single views with this system, self-occlusions occur causing holes in the facial data, and incomplete or distorted facial contour;
- when mouth is open, or in occluded face scans, the data present discontinuities.

The assumption that occlusions lie in front of the faces is not always satisfied in these cases. In Fig. 19 we report some examples of corrupted acquisitions.

Comparing the Boshporus DB with the artificial occlusions generated for the UND Database we observed that the latter present more variability in terms of location on the face and area of the occluded region.

We compared our approach with a part-based method. In particular, we considered the results reported by Alyüz et al. [2], since they used the same database (even if not exactly

Table 2 Identification performances subdivided by occlusion type obtained on the Bosphorus DB by the proposed approach (automatic) and by the method described in [2] (manual)

Method	Acquisitions	Occlusion type			
		Eye	Mouth	Eyeglasses	Hair
Alyuz et al. [2] (manual)	381	93.62%	93.62%	97.87%	89.66%
proposed approach (automatic)	360	91.18%	74.75%	94.23%	90.47%

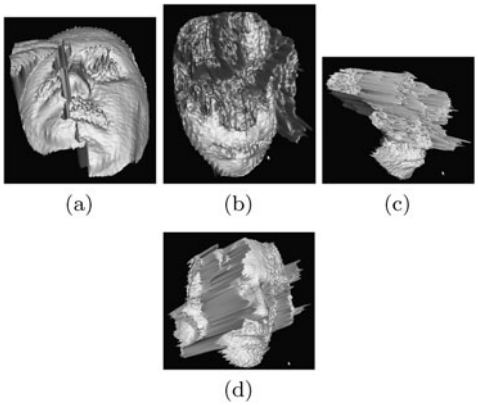

Fig. 19 Examples of corrupted acquisitions from the Bosphorus Database

the same test acquisitions: in [2] 381 acquisitions of 81 subjects are used while this work ha been evaluated on 360 acquisitions of 105 subjects). Briefly their method works as follows:

- the input face is manually annotated with the position of non-occluded landmarks;
- the face is roughly aligned by Procrustes analysis on the landmarks;
- the face is registered on the mean face by ICP;
- the mean face is subdivided into eight regions on the basis of a predefined partition;
- each region of the mean face is registered to the aligned test face defining, at the same time, a corresponding region on the test face;
- each region of the test face is compared, by an approximated volumetric difference, with the corresponding region of gallery faces;
- regions corresponding to a bad ICP registration (defined by a suitable threshold) are considered as occluded, and are therefore discarded;
- at this point the non-occluded regions correspond to nearest neighbor classifiers, which are then combined using fusion methods such as the sum of the scores, the product, Borda count ...

This method is based on a preliminary manual selection of landmarks; therefore, the recognition performances are not affected by potential errors that an automatic face detection and normalization procedure may introduce. On the contrary, our procedure is completely automatic. We have investigated in Sect. 3.1.4 the performance degradation due to our automatic normalization procedure. In light of these results we think that the recognition performances reported in Table 2 can be positively considered. In general, the two methods obtained a similar identification rate, with the notable exception of faces where the occlusion covers the mouth. We think that this performance drop is caused by the normalization error which usually occurs with large objects covering the lower part of the face. This explanation is supported by the results obtained on the UND database (see Table 1 and Sect. 3.1.4).

4 Conclusions and Future Work

We have presented here an innovative strategy for the restoration of partially occluded 3D faces. Since it is independent of the recognition algorithm, our strategy can improve the reliability with respect to occlusions of any 3D recognition system, even when low computational resources are available (occlusion detection and restoration require an average of 0.24 seconds to process a face on a modern personal computer). The restoration strategy is independent of the method used to detect occlusions and can also be applied to restore faces in the presence of noise, self-occlusions, and acquisition artifacts.

Experimentation has been conducted on an artificially occluded dataset which we plan to make available (or, at least, we will release the tool for automatic occlusions generation) and on the real occluded acquisitions from the Bosphorus database. We plan to acquire a larger number of faces with real occlusions to further demonstrate the applicability of our method. Since occlusions in 3D face recognition is an unexplored topic, we cannot state that the restoration strategy is the optimal solution in terms of recognition accuracy. Even if restoration will be proven to be suboptimal, occlusion detection would still be a useful and valuable component in 3D recognition system pipelines.

Other authors reported an improved invariance with respect to facial expressions when dealing with occlusions in face recognition (see [2], for instance). Our method is not expected to provide high accuracy in the case of emphasized

facial expressions. However, we plan to investigate a similar restoration approach: expressions would be classified and "neutralized" before recognition.

Acknowledgements We would like to thank the University of Notre Dame (in particular Prof. Patrick J. Flynn) and the University of Boğaziçi for making the data available for our experiments.

References

1. Achermann, B., Bunke, H.: Classifying range images of human faces with Hausdorff distance. In: Proc. 15th Int'l Conf. Pattern Recognition, pp. 809–813 (2000)
2. Alyüz, N., Gökberk, B., Akarun, L.: A 3d face recognition system for expression and occlusion invariance. In: Proceedings of 2nd IEEE International Conference on Biometrics: Theory, Applications and Systems, pp. 1–7 (2008)
3. Belhumeur, P., Hespanha, J., Kriegman, D.: Eigenfaces vs. fisherfaces: Recognition using class specific linear projection. IEEE Trans. Pattern Anal. Mach. Intell. **19**(7), 711–720 (1997)
4. Berretti, S., Del Bimbo, A., Pala, P.: Description and retrieval of 3d face models using iso-geodesic stripes. In: Proc. 8th ACM Int'l Workshop Multimedia information Retrieval, pp. 13–22 (2006)
5. Besl, P.J., McKay, N.D.: A method for registration of 3-d shapes. IEEE Trans. Pattern Anal. Mach. Intell. **14**(2), 239–256 (1992)
6. Beumier, C., Acheroy, M.: Automatic 3d face authentication. Image Vis. Comput. **18**(4), 315–321 (2000)
7. Blanz, V., Vetter, T.: Face recognition based on fitting a 3d morphable model. IEEE Trans. Pattern Anal. Mach. Intell. **25**(9), 1063–1074 (2003)
8. Bowyer, K.W., Chang, K.I., Flynn, P.J.: A survey of approaches and challenges in 3d and multi-modal 3d + 2d face recognition. Comput. Vis. Image Underst. **101**(1), 1–15 (2006)
9. Bronstein, A.M., Bronstein, M.M., Kimmel, R.: Expression-invariant 3d face recognition. In: Proceedings of Audio- and Video-Based Biometric Person Authentication, pp. 62–70 (2004)
10. Bronstein, A.M., Bronstein, M.M., Kimmel, R.: Efficient computation of isometry-invariant distances between surfaces. SIAM J. Sci. Comput. **28**(5), 1812–1836 (2006)
11. Bronstein, M.M., Bronstein, A.M., Kimmel, R.: Three-dimensional face recognition. Int. J. Comput. Vis. **64**(1), 5–30 (2005)
12. Cartoux, J.Y., Lapreste, J.T., Richetin, M.: Face authentification or recognition by profile extraction from range images. In: Proc. IEEE CS Workshop Interpretation of 3D Scenes, pp. 194–199 (1989)
13. Chang, K.I., Bowyer, K.W., Flynn, P.J.: Face recognition using 2d and 3d facial data. In: ACM Workshop on Multimodal User Application, pp. 25–32 (2003)
14. Chang, K.I., Bowyer, K.W., Flynn, P.J.: An evaluation of multimodal 2d + 3d face biometrics. IEEE Trans. Pattern Anal. Mach. Intell. **27**(4), 619–624 (2005)
15. Chang, K.I., Bowyer, K.W., Flynn, P.J.: Multiple nose region matching for 3d face recognition under varying facial expression. IEEE Trans. Pattern Anal. Mach. Intell. **28**(10), 1695–1700 (2006)
16. Colombo, A., Cusano, C., Schettini, R.: A 3d face recognition system using curvature-based detection and holistic multimodal classification. In: Proc. 4th Int'l Symp. on Image and Signal Processing and Analysis, pp. 179–184 (2005)
17. Colombo, A., Cusano, C., Schettini, R.: 3d face detection using curvature analysis. Pattern Recognit. **39**(3), 444–455 (2006)
18. Colombo, A., Cusano, C., Schettini, R.: Detection and restoration of occlusions for 3D face recognition. In: Proc. of IEEE International Conference on Multimedia & Expo, pp. 1541–1544 (2006)
19. Colombo, A., Cusano, C., Schettini, R.: Face^3 a 2d+3d robust face recognition system. In: Proceedings of 14th IEEE International Conference on Image Analysis and Processing, pp. 393–398 (2007)
20. Colombo, A., Cusano, C., Schettini, R.: Gappy pca classification for occlusion tolerant 3d face detection. J. Math. Imaging Vis. **35**(3), 193–207 (2009)
21. De Smet, M., Franses, R., Van Gool, L.: A generalized em approach for 3d model based face recognition under occlusions. In: Proc. IEEE Conference on Computer Vision and Pattern Recognition, vol. 2, pp. 1423–1430 (2006)
22. Ekenel, H.K., Stiefelhagen, R.: Why is facial occlusion a challenging problem. In: Advances in Biometrics, pp. 299–308 (2009)
23. Everson, R., Sirovich, L.: Karhunen-Loève procedure for gappy data. J. Opt. Soc. Am. A **12**(8), 1657–1664 (1995)
24. Flynn, P.J., Bowyer, K.W., Phillips, P.J.: Assessment of time dependency in face recognition: an initial study. In: Audio- and Video-Based Biometric Person Authentication, pp. 44–51 (2003)
25. Gökberk, B., Akarun, L., Aksan, B.: How to deceive a face recognizer. In: Proceedings of the Biometrics: Challenges arising from Theory to Practice Workshop (2004)
26. Hwang, B., Lee, S.: Reconstruction of partially damaged face images based on a morphable face model. IEEE Trans. Pattern Anal. Mach. Intell. **25**(3), 365–372 (2003)
27. Kim, J., Choi, J., Yi, J., Turk, M.: Effective representation using ica for face recognition robust to local distortion and partial occlusion. IEEE Trans. Pattern Anal. Mach. Intell. **27**(12), 1977–1981 (2005)
28. Kirby, M., Sirovich, L.: Application of the Karhunen-Loeve procedure for the characterization of human faces. IEEE Trans. Pattern Anal. Mach. Intell. **12**(1), 103–108 (1990)
29. Lin, D., Tang, X.: Quality-driven face occlusion detection and recovery. In: Proc. IEEE Conf. on Computer Vision and Pattern Recognition, pp. 1–7 (2007)
30. Lu, X., Jain, A.K., Colbry, D.: Matching 2.5d face scans to 3d models. IEEE Trans. Pattern Anal. Mach. Intell. **28**(1), 31–43 (2006)
31. Martinez, A.M.: Recognition of partially occluded and/or imprecisely localized faces using a probabilistic approach. In: Proc. IEEE Conf. Computer Vision and Pattern Recognition, vol. 1, pp. 712–717 (2000)
32. Martinez, A.M.: Recognizing imprecisely localized, partially occluded and expression variant faces from a single sample per class. IEEE Trans. Pattern Anal. Mach. Intell. **24**(6), 748–763 (2002)
33. Medioni, G., Waupotitsch, R.R.: Face recognition and modeling in 3d. In: Proc. IEEE Int'l Workshop Analysis and Modeling of Faces and Gestures, pp. 232–233 (2003)
34. Mo, Z., Lewis, J.P., Neumann, U.: Face inpainting with local linear representations. In: British Machine Vision Conference, vol. 1, pp. 347–356 (2004)
35. Nagamine, T., Uemura, T., Masuda, I.: 3d facial image analysis for human identification. In: Proc. Int'l Conf. Pattern Recognition, pp. 324–327 (1992)
36. Park, B., Lee, K., Lee, S.: Face recognition using face-arg matching. IEEE Trans. Pattern Anal. Mach. Intell. **27**(12), 1982–1988 (2005)
37. Park, J.S., Oh, Y.H., Ahn, S.C., Lee, S.W.: Glasses removal from facial image using recursive error compensation. IEEE Trans. Pattern Anal. Mach. Intell. **27**(5), 805–811 (2005)
38. Pentland, A., Moghaddam, B., Starner, T.: View-based and modular eigenspaces for face recognition. In: Proc. IEEE Conf. Computer Vision and Pattern Recognition, pp. 84–91 (1994)
39. Samir, C., Srivastava, A., Daoudi, M.: Three-dimensional face recognition using shapes of facial curves. IEEE Trans. Pattern Anal. Mach. Intell. **28**(11), 1858–1863 (2006)

40. Savran, A., Alyüz, N., Dibeklioglu, H., Çeliktutan, O., Gökberk, B., Akarun, L., Sankur, B.: Bosphorus database for 3d face analysis. In: The First COST 2101 Workshop on Biometrics and Identity Management, pp. 47–56 (2008)
41. Skocaj, D., Leonardis, A.: Robust recognition and pose determination of 3-d objects using range images in eigenspace. In: Third International Conference on 3-D Digital Imaging and Modeling, pp. 171–178 (2001)
42. Tarrés, F., Rama, A.: A novel method for face recognition under partial occlusion or facial expression variations. In: Proc. 47th Int'l Symp. ELMAR, pp. 163–166 (2005)
43. Turk, M., Pentland, A.: Eigenfaces for recognition. J. Cogn. Neurosci. **3**(1), 71–86 (1991)
44. Zhang, W., Shan, S., Chen, X., Gao, W.: Local Gabor binary patterns based on Kullback Óleibler divergence for partially occluded face recognition. IEEE Signal Process. Lett. **14**(11), 875–878 (2007)

Alessandro Colombo took his degree in Computer Science in 2004 and his Ph.D. in Computer Science in 2008 at DISCo, Department of Information Science, Systems Theory, and Communication at the University of Milano-Bicocca. He is currently a post-doc researcher at the Imaging and Vision Laboratory (www.ivl.disco.unimib.it). His research interests cover the processing, analysis and synthesis of 2D and 3D images. In particular, he focused his research on 2D/3D object and face detection and recognition.

Claudio Cusano is a post-doc researcher at DISCo (Department of Information Science, Systems Theory, and Communication), of the University of Milano-Bicocca, where he took his Ph.D. in Computer Science. Since April 2001 he has been a fellow of the the ITC Institute of the Italian National Research Council. The main topics of his current research concern 2D and 3D imaging, with a particular focus on image analysis and classification, and on face recognition.

Raimondo Schettini is an associate professor at the University of Milano Bicocca (Italy). He is Vice-Director of the Department of Informatics, Systems and Communication, and head of Imaging and Vision Lab (www.ivl.disco.unimib.it). He has been associated with Italian National Research Council (CNR) since 1987 where he has leaded the Color Imaging lab from 1990 to 2002. He has been team leader in several research projects and published more than 200 refereed papers and six patents about image processing, analysis and classification.