ORIGINAL PAPER

CrossMark

Occlusion detection and restoration techniques for 3D face recognition: a literature review

Nicole Dagnes[1] · Enrico Vezzetti[1] · Federica Marcolin[1] · Stefano Tornincasa[1]

Abstract
Methodologies for 3D face recognition which work in the presence of occlusions are core for the current needs in the field of identification of suspects, as criminals try to take advantage of the weaknesses among the implemented security systems by camouflaging themselves and occluding their face with eyeglasses, hair, hands, or covering their face with scarves and hats. Recent occlusion detection and restoration strategies for recognition purposes of 3D partially occluded faces with unforeseen objects are here presented in a literature review. The research community has worked on face recognition systems under controlled environments, but uncontrolled conditions have been investigated in a lesser extent. The paper details the experiments and databases used to handle the problem of occlusion and the results obtained by different authors. Lastly, a comparison of various techniques is presented and some conclusions are drawn referring to the best outcomes.

Keywords Facial occlusions · 3D face analysis · Face recognition · Face detection · Facial restoration

1 Introduction

Biometric systems are automated methods of verifying or recognizing the identity of a person on the basis of some physical characteristics (i.e. fingerprint, iris, retina and hand geometry recognition, DNA matching, speaker identification or face pattern), or some aspects of behaviour, like handwriting or gait [1] . In particular, fingerprint-based systems have been proven to be effective in protecting information and resources in a large area of applications, despite the cost of the required hardware/software [2]. However, at present, the principal barrier for a wide spread of biometrics systems is determined by mistrust of improper use. Moreover, biometric systems such as those relying on fingerprints or iris recognition are more difficult to be used for security purposes in public areas, such as airports, train stations, and supermarkets for identifying individuals among the crowd, as they require aid (or consent) from the test subject.

On the contrary, systems which rely on face recognition (FR) have gained great importance and considered as a fine

compromise among other biometrics, as the measurement of intrinsic characteristics of the human face can be implemented in a non-intrusive way [3]. Hence, nowadays, human FR is widely preferred thanks to easy-to-obtain data given by contactless acquisition and to the possibility to avoid requiring consent and cooperation from the subject. Its applicability to non-cooperative scenarios makes it suitable for a range of applications such as surveillance systems.

In the past, most facial recognition software focused on the source of 2D intensity or colour images to compare or identify another 2D image from the database. FR from 2D images has been studied for over four decades [4], even though the accuracy of 2D FR is influenced by variations of poses, expressions, illuminations [5,6]. Nowadays, great breakthroughs allow automatic 2D FR systems to effectively deal with the problem of illumination and pose variation [7,8], but these issues still remain problems that must be addressed and treated, if the recent deep learning techniques are not used.

The last decades have witnessed 3D-based methodologies being emerged. Data come in the form of a set of 3D point coordinates, $p_i = [x_i, y_i, z_i]^T$, which encode the relative depths. Acquisitions obtained via 3D scanners can be used to construct 3D face models offline, i.e. point clouds, and build 3D facial recognition techniques using distinctive features of the face. Hence, 3D facial data supports the full under-

✉ Federica Marcolin
  federica.marcolin@polito.it

[1] Department of Management and Production Engineering, Politecnico di Torino, Corso Duca degli Abruzzi 24, 10129 Turin, Italy

standing of the human face features and can help to improve the performance of the current recognition systems, since 3D FR methods are immune to the pitfalls of 2D FR algorithms like lighting variations, different facial expressions, make-up, and head orientation [9]. Hence, 3D FR systems reached a notable growth thanks to their applicability to the surveillance context.

Non-cooperative and uncontrolled scenarios caused by facial occlusions are common situations which may produce a failure of standard algorithms and which may be exploited by malevolent persons, who try to camouflage themselves to trick the implemented security systems. Hence, facial changes due to occlusions are important factors that should be considered by face recognition applications [10]. Occlusions may be provoked by hair or external objects, such as glasses, hats, scarves, or by subject's hands. When the facial area is partially occluded, performances may dramatically decrease [11]. The main reason is the fact that occlusions result in loss of discriminative information. This is especially true when an upper face occlusion occurs, involving the eye and nose regions, which are known to be very discriminative [12].

The works which focus on the detection or recognition of faces in presence of occluded and missing data mainly used two types of occlusions: *natural* and *synthetic*. The first type of occlusion is caused by external objects: hair, hands, clothing material, hats, scarves, eyeglasses, and other objects. This type of occlusion is complex to handle, as the occluding object alters the facial geometry. The second one refers to artificially blockade the face with a white/black rectangular block or artificial solid inserted components. In real-world scenarios, it is common to acquire subjects wearing glasses or talking on the phone or having their hands covering their face. Hence, studies have been focused on finding a way to handle occlusions, especially upper face ones.

Most of recognition systems which handle occlusions include dedicated algorithms for the tasks of occlusion detection and localization. These methods come as the first steps to be applied for undertaking the subsequent routines: face detection, face registration, and features extraction for pattern matches. *Face detection* consists in localizing and extracting the face region from the background, which is a core technique for crowd surveillance applications. Many of the current FR techniques assume that facial data are given by frontal faces of similar sizes [13,14]. However, this assumption may not hold due to the varied natures of face appearance and environment conditions. Thus, *face registration,* or *alignment*, whose scope is to find general transformation parameters, is a necessary step to be performed. During this phase, scale variations are found and two three-dimensional shapes are aligned by rotating faces into a frontal pose, with the help of three axes. Once detection and registration are concluded, *feature extraction* is performed

followed by a *classifier phase* to determine the identity of unknown subjects or to verify their claimed identity.

This work is intended as a survey of the recent literature on 3D facial occlusions for recognition purposes. Contributions have been chosen among the years 2005 and 2017, in order to provide the most up-to-date view of the latest researches. This is also to support those researchers who have never worked on FR systems with uncontrolled conditions. A literature review on 2D occlusion detection methodologies had already been provided by Azeem et al. in 2014 [15].However, the three-dimensional occlusions management scenario has never been investigated yet.

The article is structured as follows. Section 2 gives an overview of the available 3D facial databases that researchers have used to conduct comparative performance evaluations. Section 3 presents techniques for occlusion detection and localization, including template matching, thresholding, radial and facial curves-based methods. Section 4 faces the methodologies used for occlusion removal or restoration. Section 6 concludes the work and summarizes the most recent state-of-the-art approaches and the obtained results, in order to identify the best ones for FR applications.

2 Databases

Various publicly available databases for 3D FR exist and have been employed by the research community for 3D FR and verification evaluations, but most of them contain a limited range of expressions, head poses, and face occlusions. Below, the databases of relevance which have been used to test FR methods under occlusion conditions are listed and introduced. Then, they are presented and compared in Table 1.

Most of the recent studies on facial occlusion analysis, proposed in this literature, employ at least one of the databases analysed below. The main reason for choosing these mentioned databases is that datasets, comprising several thousand images gathered from several hundred subjects, allow researchers to observe relatively small differences in performance in a statistically significant way. Moreover, the large number and difficult types of occlusions allow the possibility to manage more different types of real-life situations.

The most common 3D Face Database for occlusion studies is the *Bosphorus database* [16,17]. The typical occlusions in the Bosphorus 3D face database are: occlusion of left eye area with hand, occlusion of mouth area with hand, and occlusion with eyeglasses. Concerning the occlusion of eyes and mouth, subjects in the database have a natural pose; for example, while one subject occludes his mouth with the whole hand, another one may occlude it with one finger only. Also, different eyeglasses are present in the facial dataset. The facial locations of the occluded objects are: mouth, one eye

(for hand), both eyes (for eyeglasses) and, in the case of hair, part of the forehead or half the face. Hence, the Bosphorus database is a valuable source for training and testing recognition algorithms under adverse conditions and in the presence of facial occlusion analysis. Some face examples from the Bosphorus dataset, which are stored as structured depth maps, are shown in Fig. 1, whereas the relevant information about the dataset are recorded in Table 1.

The *UMB database* [18], outlined in Table 1, presents more challenging occlusions, as they can be caused by hair, eyeglasses, hands, hats, scarves or other objects, and their location and amount greatly vary. Hence, it is suitable to design and test methods dealing with partially occluded face and can be adopted for various kinds of investigations about face analysis in unconstrained scenarios.

The *GavabDB database* [19]is built for automatic FR experiments and other possible image applications such as pose correction and 3D face model registration. The expressions in this dataset vary considerably, including sticking out the tongue and strong facial distortions. Additionally, it has strong artefacts, i.e. distortions (holes, noise, outliers, missing regions) generated during the acquisition procedure, due to facial hair. This dataset is typical for a non-cooperative scenario.

A peculiar occluded face database, used by Zohra et al. [20], is the *EUROKOM Kinect face database* [21], consisting of different types of occlusions and neutral facial images acquired with a Kinect RGB-D camera, i.e. a sensor characterized by low data acquisition time and low-quality depth information.The database consists of different data (2-D, 2.5-D, 3-D) and multiple facial variations, including different facial expressions, illuminations, occlusions, and poses. The involved occlusions are: by paper, by hand, and by glasses.

Some works use non-occluded databases in the first steps, if they need to create facial models and set thresholds. This way, disjoined dataset are used for both training and testing phases. For this purpose, the FR Grand Challenge version 2(*FRGC v.2*) [22] is the most adopted, but also the UND and the BU3DFE are sometimes used. Over recent years, the FRGCv.2 has been the most widely used for both facial identification and verification contexts. The FR approaches which are tested only on this dataset belong to the cooperating subject scenario. An example of structured 3D point clouds from FRGCv.2 is shown in Fig. 2.

The UND dataset, compared to FRGC v.2, has fewer subjects. Some examples of UND acquisitions are shown in Fig. 3.

The BU3DFE offers more images per subject. Each subject performs seven standard expressions, which are neutral, happiness, disgust, fear, angry, surprise, and sadness. Four levels of intensity for each one is included.

Colombo et al. [23] followed a different approach which turned to be a low-cost and effective solution. They manipulated the 3D models to obtain *artificial occlusions*. The idea consists in taking an existing database of non-occluded faces (UND) and adding occluding 3D objects in each acquisition. The set of objects employed includes plausible objects, such as a scarf, a hat, two types of eyeglasses, a newspaper, hands in different configurations, a pair of scissors, whose point clouds are inserted in the facial point clouds in order to cover a part of the face. Figure 4 shows some examples of artificially occluded faces.

Table 1 outlines and details the datasets mentioned above. For each dataset, the table reports the number of 3D face images captured, the number of subjects, the quality and resolution of the images, and the variations in facial pose, expression, and occlusions.

3 Occlusion detection and localization

An explicit analysis and processing for obtaining prior information on the occlusions can significantly improve face recognition performances under severe occlusion conditions [39], even if locally emphasized algorithms can already show robustness to partial occlusions [40]. This section gives an overview of the currently adopted techniques for *automatically detecting and localizing 3D occlusions* on facial surfaces, task that has been approached by developing different strategies and leading to various levels of accuracy. Several tables (Tables 2, 3, 4, 5, 6) are embedded, to catalogue the outcomes.

Figure 5 summarizes how partial occlusions have been treated by different researchers in the 3D context.

3.1 Radial curves

Recently, significant progresses have been made in the analysis of curve shapes, obtaining efficient algorithms [41,42]. The first works which mentioned facial *curves for shape analysis* were conducted by the research group of Samir et al. and was based on level curves for FR [43–45]. However, this initial use of level curves had the limitation of passing through different facial regions, affected differently by changes in facial expressions occlusions.

To handle this problem, radial curves have been adopted. One of the most spread techniques employed for *occlusion detection* uses *radial curves*, which start form the nose tip and pass through distinct regions one-by-one. Hence, these curves, if chosen appropriately, have the potential to capture the local shape behaviour and to well represent facial surface. This representation seems natural for measuring facial deformations and is robust to challenges such as large facial expressions (especially those with open mouths), large pose variations, missing parts, and partial occlusions due to glasses, hair, etc. Besides, an advantage of these curves

Table 1 A selection of 3D face datasets employed by the recent studies on facial occlusion analysis

| Dataset	Ref.	Data	Size	Subj.	F.	M.	Sampl.	Occlusions	Expr.	Pose	2D res.	Land.	Utilization
Bosphorus	[16]	2D+3D	4666	105	45	60	29–54	*Yes* (mouth occlusion by hand, occlusion of the face with hair and eyeglasses, occlusion of the left eye and forehead regions by hands)	Yes (35)	Yes (13)	1600 × 1200 pixels	24	[17,24–34]
UMB	[18]	2D+3D	1473	143	45	98	6	*Yes* (occlusion of the face with hair, eyeglasses, hands in various configurations and positions, hats, scarves, and other miscellaneous objects)	Yes (3)	No	640 × 480 pixels	7	[24,25,27,35,36]
GavabDB	[19]	2D+3D	549	61	16	45	9	*Yes* (occlusion of the face with hair and tongue)	Yes (4)	Yes (4)	–	–	[28][29]
KinectFaceDB	[21]	2D+2.5D+3D		52	14	38	9	*Yes* (occlusion by sunglasses, occlusion by hand, occlusion by paper)	Yes (2)	Yes (2)	640 × 480 pixels	6	[20]
FRGC	[22]	2D+3D	4950	466	–	–	1–22	*No* *Hair occlusion*	Yes (6)	No	640 × 480 pixels	–	[24,25,28,29,35] [37]
BU3DFE	[38]	2D+3D	2500	100	60	40	25	*No*	Yes (6)	No	512 × 512 pixels	83	[23]
UND		2D+3D	951	277	–	–	–	*No* *Artificial occlusion*	No	No	480 × 640 pixels	4	[23]
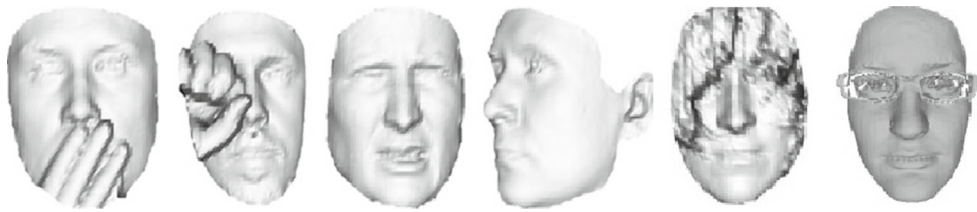
Fig. 1 3D facial models from the Bosphorus dataset with occlusions, expression variations and pose variations [16,17]

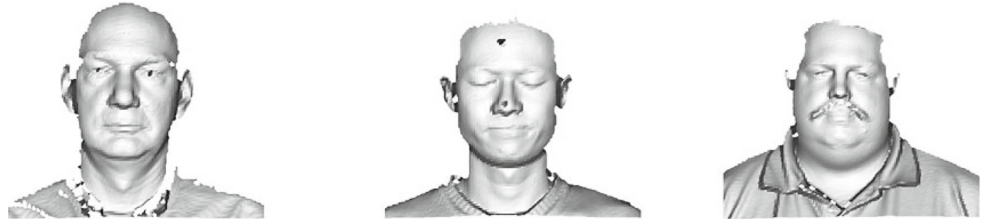**Fig. 2** 3D face models from the FRGC v2 dataset

Fig. 3 3D facial models from the UND dataset

(which start from the nose tip) is that the nose is relatively easy and efficient to detect, as it is the focal point of the face [46]. Due its potential, radial curves have been widely adopted for 3D FR purposes [47], even in the presence of facial expressions. Different variants and extensions have been proposed (Table 2); the general idea which guides these methods is to compare the facial shapes by matching the shapes of their corresponding curves.

In 2013, Drira et al. [28] represented facial surfaces with radial curves and embedded them in their method for recognizing faces partially covered by external objects or hands. However, in their previous works [47], which studied shapes of curves in the presence of facial deformations due to

changes in facial expressions, they had noticed that some curves could reduce the performance of the algorithm if they have bad quality. Hence, they selected only a subset of the all possible radial curves and discarded the others. This idea, of detecting first and then removing erroneous curves, holds in presence of facial occlusions. In case of self-occlusions indeed, some curves are partially covered or deformed by the occluded surface. Hence, *these curves have to be identified and discarded*, in order not to severely degrading the recognition performances. Only curves that have passed the quality filter that is used as discriminating factor the continuity and the shortness of curves are extracted and maintained. Hence, the first step of their approach consists in a *filtering*

Fig. 4 Some examples of artificially occluded faces starting from the original UND acquisitions [23]
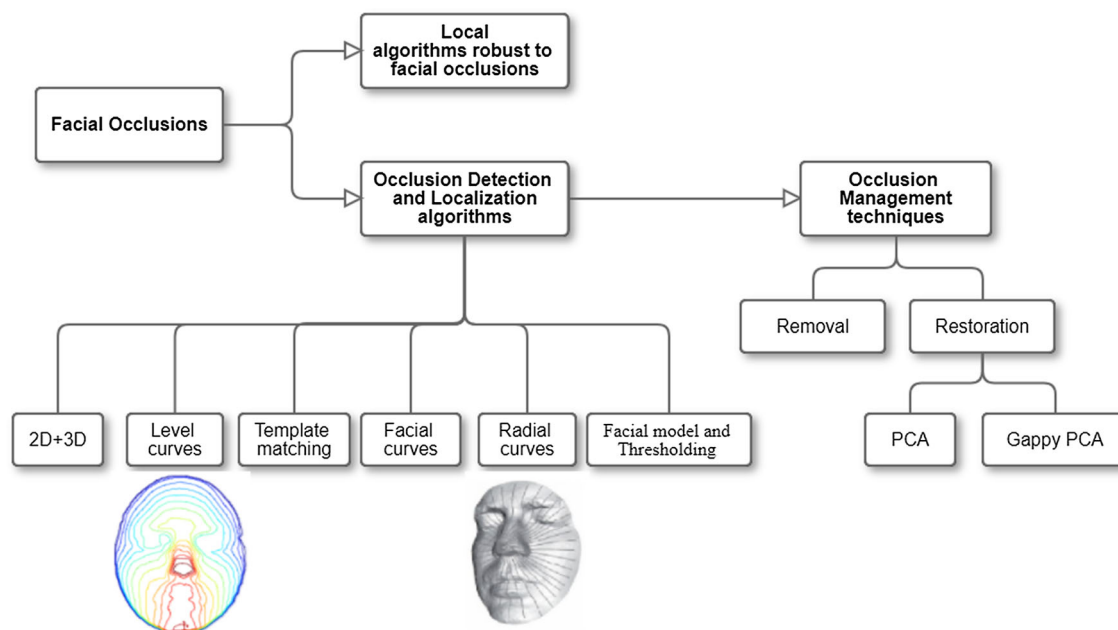
Fig. 5 Methods used to handle occlusions in 3D facial models. The first branch recalls the possibility to develop local algorithms robust to occlusions. Alternatively, an explicit analysis can be performed to obtain more information on occlusions (position and size); detection/localization of the occluded area is treated in this section. This first step can be followed by the removal or restoration of the occluded part (this is addressed in Sect. 4)

routine that detects the external objects and finds and removes curves which lie in the areas covered by external occlusions, such as glasses and hair. The points that belong to the face and those belonging to the occluding object are detected by comparing the given scan with a template scan, developed using an average of training scans that are complete, frontal, and with neutral expressions. In detail, the current face scan is matched with the template one using recursive Iterative Closest Point (ICP); the points on the scan whose values are greater than a certain threshold far from the corresponding points on the template are removed. This threshold has been determined using experimentation and is fixed for all faces. Performances have been analysed using three datasets: the FRGC v2.0dataset, involving the facial expression, the GavabDB, involving the pose variation, and the Bosphorus dataset, involving the occlusion challenge. Concerning the Bosphorus database, a neutral scan for each person is taken to form a gallery dataset of size 105; the probe set contains 381 or 360 scans with occlusions. The faces with eye occlusions resulted to be the best occluded faces recognized by the algorithm. The 98.9% of the 360 chosen faces are correctly recognized, whereas the 93.6% of accuracy is reached, if all the 381 occluded faces in the dataset are considered.

In 2014, Drira et al. [29] continue to use the radial facial curves on 3D meshes to recognize faces despite the presence of occlusion and pose variations. Similarly to their previous works, the matching of two 3D facial scans is achieved

by matching two radials string sets through string-to-string matching, showing a new way for face representation and recognition, in the presence of facial occlusions. Again, a first occlusion detection and removal step is performed that is based on recursive-ICP is developed to handle occlusions.

A similar approach was proposed in 2016 by Yu et al. [33,48]. To handle the problem of FR under partial occlusions, they proposed to use a new *attributed string matching* algorithm, as an evolution of the previous conventional sting matching approach [49]. Firstly, the 3D facial surface is encoded into an indexed collection of radial strings emanating from the nose tip. Then, a partial matching mechanism effectively eliminated the occluding parts, saving the most discriminative parts and, finally, the Dynamic Programming is used to measure the similarity between two radial strings. This way, *occluded parts are eliminated* and only the most discriminative parts are used during the matching process. Experimental results on the Bosphorus database demonstrate good performance on partially occluded data. A neutral scan for each person is selected to form a gallery set of size 105. The probe set contains 360 scans that have occlusions. The overall recognition rate (RR) is 96.3%, whereas the faces with hair occlusions are the better occluded faces recognized by the algorithm, reaching the 98.0% RR accuracy, and the glasses occluded the worst, with 94.2%.

In the same year, Bhave et al. [50] illustrated the use of radial facial curves emanating from the nose tips for shape

analysis. The facial deformations and the missing parts due to partial occlusions of glasses and hair are detected and removed, through are cursive-ICP procedure. Then, the facial surface is represented by an indexed collection of radial geodesic curves on 3D face meshes emanating from nose tip to the boundary of mesh and the facial shapes are compared through their corresponding curves.

The Table 2 summarizes the contributions which adopted radial curves for face analysis.

3.2 Facial curves

Another typology of curves is adopted for face recognition in the presence of occlusions: *facial curves*. These curves do not have the same origin, but are originated by different planes that intersect the facial surface. The most common curve is surely the central profile, i.e. the curve in the nose region. Curves in the forehead, cheeks and mouth regions are also extracted and used to compare one probe face with another. Some works have adopted facial curves for 3D FR in a non-occluded scenario [51], due to their capacity to represent the distinctiveness of a face at low cost. Also [52,53] and [54] focused on 3D FR with the use of 3D curves extracted from the geometry of the facial surface. The main issue of these methods is to extract a set of 3D curves that can be effectively used for 3D FR, i.e. the optimal set of 3D curves, reducing the amount of face data to one or a few 3D curves.

Li et al. in 2012 [37] used facial curves to compare two faces in an occluded scenario and presented an efficient 3D face detection method to handle facial expression and hair occlusion. As a matter of fact, the deformations of the facial curves can express the deformation of the region which contains the facial curve, thus revealing possible external object or missing parts. Hence, in this paper, facial curves are extracted, in or near the parts which are the most sensitive to expression or hair occlusion, and matched to evaluate the deformation of the face. Given that the deformation of the different regions of the face caused by expression and hair occlusion can be obtained through the facial curves, an adaptive region selection scheme is proposed for FR. In other words, by classifying faces in which deformation is presented, the recognition process can place more emphasis on those regions which are least affected. Beside, since the nose region is almost invariant to all expressions and generally hair occlusion will not deform it, this area is always selected during the accurate matching process. Other regions are selected only when the corresponding deformation is less than a threshold. Regions occluded by hair or moustaches are automatically discarded. Experiments have been done on the FRGC v2.0 dataset, and the identification rate is 97.8%.

Table 3 summarizes the contributions which adopted facial curves for face analysis.

3.3 Facial model and thresholding

Some techniques employed for overcoming the challenges caused by facial occlusions consist in *adaptively selecting non-occluded facial regions* [17,35], which, combined together, are used for FR instead of using only the nasal region or the entire facial surface. For instance, if the left side of the face is occluded by a hand, the non-occluded regions such as right eye, mouth, and nose, are automatically considered. This way, the regions where occlusions are present are discarded, whereas all the free regions are considered and used for adding information and improving the FR rate. The idea comes from the observation that 3D FR systems which consider the entire face for recognition purposes perform sufficiently well only for neutral faces. However, the accuracy decreases if some parts of the face contain surface deformations like facial expressions or occlusions. Considering the face as a composition of different and independent components aids to solve this problem and leads to propose different region-based systems and techniques [55–58], primarily applied for expression variations management. Alyuz et al. in 2008 [17,35]adopted this technique for facial occlusion management and, with the testing phase led in [17], clearly showed that nose and eye/forehead regions, less affected by surface deformations, are not sufficient by themselves for recognition. Whereas combining all the regions supports the enhancement of the RR.

In order to determine the validity of each region, different strategies are designed and tested. In particular, Alyuz et al. in 2008 [17] employed *thresholds*, calculated on a non-occluded neutral training set, to determine whether a given region is occlusion-free or not. The experimental results on the Bosphorus database demonstrated that the proposed system improves the performances, obtaining for mouth occlusion the best identification rate(97.87%) among the others due to its ability to automatically detect and remove the chin part. In these experiments, the Bosphorus v.2 set is used and 47 subjects are considered. Each one has four occluded scans (eye/cheek, mouth, glasses and hair occlusions), regrouping totally 170 scans in the probe set.

The research group endorsed the strategy in 2012 [35], when they proposed an occlusion-resistant three-dimensional face registration method. Again, the non-occluded facial regions, here named patches, are selected using an *average face model*, generated on a non-occluded neutral training set. The FRGCv2 is used for the construction of the average face and patch models, and for the determination of threshold values used for validity check over template matching scores. Hence, the non-occluded facial regions such as eyes and mouth are automatically detected to form an adaptive face model for registration, whereas the regions where occlusions are present are discarded. Then, the combined regions are used for alignment estimation instead of using only the nasal

Table 2 Applications of radial curves for facial occlusions discovery

Method	Research group	Year	Ref.	Application	Dimension	DB	Reliability	Explanation
Radial curves	Drira et al.	2013	[28]	FR	3D	Bosphorus	89.25% RR	By applying a quality filter on the radial facial curves, curves which are partially covered or deformed by the occluded surface are discarded
	Yu et al.	2016	[33,48]	FR	3D	Bosphorus	95.3% RR	A new attributed string matching algorithm is tested to effectively eliminate the occluded parts
	Bhave et al.	2016	[50]	FR	3D	Bosphorus	83.7% RR	An occlusion detection and removal step that is based on recursive-ICP is developed to handle occlusions

Table 3 Applications of facial curves for facial occlusions discovery

Method	Research group	Year	Ref.	Application	Dimension	DB	Reliability	Explanation
Facial curves	Li et al.	2012	[37]	FR	3D	FRGC v.2	96% Verification rate and 97.8% RR	Facial curves can express the deformation of the region which contains the facial curve, and used for detecting occluded facial areas

region. The UMB-DB face database is used to evaluate the registration system and to undertake the recognition experiments. The gallery set contains first neutral scan of each of the 142 subjects; instead the remaining 590 occluded scans constitute the challenging occlusion subset. In the 86.94% of cases, patches are selected correctly. The errors appear mostly in the mouth area.

In the same year, Alyuz et al. [59] proposed the same approach for handling facial occlusions and proposed another occlusion-resistant three-dimensional (3D) system, where occluded areas are determined automatically via a generic *face model*. In detail, in this approach, occlusion detection is handled by thresholding the absolute difference computed between an average face model and the input face. This way, the facial areas occluded by exterior objects are located, for restoring or ignoring missing parts. The Bosphorus database was used for FR experiments. Neutral images are used to construct the gallery set, which includes 1–4 scans per subject and has a total of 299 scans. The acquisitions that contain occlusion variations form the probe set, which consists of 381

images of four different types of occlusions. This approach achieves 94.23% rank-1 identification accuracy.

In 2013, [24] the group proposed to consider the 3D surface as a combination of several non-occluded independent regions, constructed from the combination of several non-overlapping patches. To locate facial areas occluded by exterior objects, the research group used what they considered the most straightforward approach for occlusion detection, which is to analyse the difference between an input image and a mean face, as in the previous works. If an exterior object covers part of the facial surface, the difference for this specific area will be more evident. Therefore, occluded areas can be detected by thresholding the difference map obtained by computing the absolute difference between the average face and the input face. The FRGC v.2 neutral subset is employed as a separate training set for the construction of the average face and the determination of threshold values. Then, the proposed system is evaluated on two main 3D face databases that contain realistic occlusions: the Bosphorus and the UMB-DB databases. For the Bophorus, the first neutral scan of each subject is used to construct the gallery

set (105 scans), whereas the 381images with occlusion variations form the probe set. For the UMB-DB, the gallery set contains 142 neutral scans, one for each subject. Instead, the probe set consists of the 590 occluded scans. The testing phase shows that, in presence of occlusions, it is beneficial to act independently on separate regional classifiers, obtaining an improved overall performance. As a matter of fact, if the facial area partially occluded by external objects is considered, the incorrect information regarding the covered regions will cause the global recognition approaches to fail or decrease its performances. This approach achieves 92.91% rank-1 identification accuracy with the Bosphorus database, and 68.47% with UMB-DB, which contains more challenging occlusions, such as eyeglasses, hands, hats, scarves, and other objects, whose location and amount vary greatly.

Along similar lines, Bagchi et al. [26] in 2014 compared each occluded image with the *mean face* and applied a thresholding approach to detect occluding objects. Their main contribution is that they have tried to generalize the thresholding technique for detecting real-world occlusions. The procedure for occlusion detection and initial threshold setting starts with calculating the difference map between the mean image and the occluded image. Then, for each column, the maximum depth values are calculated and stored. It can be clearly seen that in each column the maximum values are the occluded portions of the image. The threshold T, applied in the condition for invalidating occluded pixels, is obtained by taking the maximum intensity value of the difference map image sampled at each column of the difference map image. The experimental results are obtained on the occluded facial images from the Bosphorus 3D face database. The gallery consists of 92 neutral scans, whereas 368 occluded scans form the training set. This scheme attains a recognition accuracy of 91.30%.

Ganguly et al. [30] in 2015 proposed a novel threshold based technique for occlusion detection and localization from a given 3D face image. Firstly, the 2.5D or range face image is created from the 3D input face. In the range image, instead of pixel intensity value, the depth value ('Z') of each point is represented. Then, facial surface is investigated to detect the obtruded sections that, if the face is occluded, will be more than one (one is the nose area). The bit-plane slicing [60,61] is the crucial technique used to identify occlusions. From the bit-plane sliced images, it is investigated that in 7th bit-plane slice there is always an extra 'component' or 'obtrude section' compared to same bit-plane of non-occluded range face image. Hence, the outward components can be detected using multilevel threshold method [62] to depth values of range images. If occlusions are present, at the threshold level 7, two components are adequately identified, whereas neutral images, at same threshold level, have only one single obtrude section. The method has been tested on the Bosphorus database, and experiment results show that

this threshold-based technique can detect the occluded faces with accuracy 82.85%. The detection rate, in accordance with [63], is defined as the fraction of total occluded images which have been correctly detected as occluded. The success rate of automatic occlusion detection is compared with manually occlusion detection. Maximizing this rate, along with the spatial location of the occlusion, is the goal of the occlusion detection.

Another method has been investigated by Ganguly et al. [30] for occlusion detection. They noticed that obtrude sections have high depth density than any other facial surface. Besides, non-occluded faces only have nose region as an outward part, whereas occluded faces have more than one noticeable depth density. Hence, in this methodology, two blocks of size of 3×3 and 5×5 are rolled over 2.5 D face images in a row-major order. Thus, the depth values from 2.5D image under these windows are counted. With this mechanism, the outward facial regions are identified, and occluded faces are detected. In particular, the face images with more than one outward region have been detected as occluded faces, without selecting any threshold value. This second approach, namely block-based, can successfully detect the same occluded objects with a success rate of 91.43%.

Finally, Zohra et al. [20] in 2016 adopted a threshold-based approach for localizing the occluded regions in the face images, acquired from the Kinect. Kinect sensors are low-cost 3D acquisition devices with high acquisition speed. They can capture 3D images in a wide range of lighting conditions. These devices return low-quality red, green, blue and depth (RGB-D) data and usually generate very noisy and low resolution depth information, but still usable for FR purposes. Moreover, these speed devices reduce the amount of user cooperation in an application scenario and favour real-time FR. However, nowadays, there are still very limited researches on occlusion detection using the depth information provided by Kinect RGB-D sensors, even if it is possible to effectively localize the occluded region. As a matter of fact, if an occlusion is present, it appears with higher intensity values in the depth image, as the intensity values indicate the estimated distance from the camera [64] and an occlusion has a lower distance from the camera than other facial parts. Based on this hypothesis, Zhora et al. proposed to extract the high-intensity values from the occluded depth images with a threshold-based approach where the threshold value T is set empirically. The proposed method is validated using the EURECOM Kinect Face Dataset [21].In particular, for the method implementation, 160 frontal and 160 occluded faces are used as training data, whereas the testing folder contains images of 40 front and 40 occluded faces from the database. Experimental results show that the proposed method for identifying the front and occluded faces provides an average classification rate of 97.50%.

However, even *automatic detection of nose regions* on 3D face images is relevant for 3D face registration, facial landmarking, and FR. For this reason, in 2014, a nose detection approach based on template matching of depth imaged has been designed by Liu et al. [32] for an automatic detection of nose regions in 3D face images, even in situations such as occlusions and missing facial parts. The nose region plays a significant role in 3D face registration because nose is both rigid and with a predefined shape. It is also the most protrusive feature at the frontal view of 3D face images. However, in spite of the recent research works which use symmetrical and curvature information, the nose detection on 3D face images is still a challenging issue, especially due to facial occlusions. The reason is that occlusions and missing 3D data may bring regions with high curvatures, as the nose region is, decreasing the detection accuracy of curvature-based approaches [59,65–67]. The second problem is that large external object may destroy the symmetrical property of 3D faces so that the methods based on the symmetrical assumption of face images [68–70] do not perform well in these situations. In this work, to overcome these issues and to well perform even in situations of missing facial parts, three nose templates have been constructed which correspond to the whole nose, the left and right half of nose, respectively. The nose templates are constructed based on the *Average Nose Model (*ANM) and obtained as part of an *Average Face Model* (AFM) [71,72]. The AFM is obtained by averaging all depth images in the training dataset, which is composed by a set of frontal and neutral expression faces selected from the Bosphorus 3D face database. Once the AFM has been constructed, the ANM of the whole nose is obtained by cropping the nose region with a rectangle. Correspondingly, the ANMs of the left and right half of nose are obtained by separating the whole nose model into its left and right parts. Then, the ANM templates are matched, using the normalized cross correlation (NCC) as matching criterion [73], calculated between the input image and the templates using the fast algorithm proposed by [74], to obtain the nose candidates. Finally, the optimum nose region is detected among these candidates by using a sorting criterion. The proposed method obtains an overall 98.75% detection accuracy on the Bosphorus database, reaching an higher 99.48% of detection accuracy, if only occluded faces are considered.

Table 4 summarizes the contributions adopting these techniques, which empirically set thresholds and facial models or templates.

3.4 Other thresholding techniques

In 2012, Liu et al. [31] proposed an efficient *variant of the ICP algorithm* for 3D FR robust to facial occlusion changes. The ICP algorithm has been proved to be an effective manner for shape registration by iteratively minimizing the mean

square error (MSE) between all points of one query face and a face model; it has been successfully applied for 3D FR purposes [75,76]. However, the original ICP algorithm dramatically decreases its performances if a large area of outliers is present, as happened with facial occlusions [17]. To solve this issue, Liu et al. embedded a *rejection strategy* into the ICP method for discarding the occluding parts, by rejecting the worst *n*% of points, which have given the maximum point-to-point square error. The proposed method was tested on the Bosphorus database (105/381), and more unreliable facial parts are eliminated with the increasing of *n*% threshold. Hence, this parameter may be set appropriately, to eliminate facial occlusions and preserve the fundamental identity information for recognition. The proposed method is validated in face verification and identification, and the experimental results clearly demonstrate its effectiveness and efficiency. The overall identification rate obtained is 97.9%, reaching the 100% of accuracy with the eye occlusion. The worst accuracy is obtained with the hair occlusions with 96.97%, whereas the best verification rate is 85.78%.

In 2016, Bellil et al. [27] proposed a 3D method to identify and remove occlusions for 3D face identification and recognition. Firstly, a *discreet wavelet transform* (DCT) is applied. Then, occluded regions detection is performed by considering that occlusions can be defined as local face deformations. Occlusions are detected by comparing the coefficients of the input facial test wavelet coefficients and wavelet coefficients of generic face model formed by the mean data base faces, previously calculated. Hence, occluded regions are refined *by removing pixels above a certain threshold*. To test the proposed method, UMB-DB and BOSPHO-RUS database are used. The experimental results on these databases demonstrate that the 90.43% of facial occlusions have been successfully detected and the proposed approach improves RR to 93.57%. In particular, the RR is higher (99.46%) for eyeglasses occlusions, whereas the lowest rate is obtained by newspaper occlusion (79.22%), since a biggest surface in the face is covered.

Table 5 summarizes the relevant contributions involving these methodologies.

3.5 Multimodal Algorithms

Some FR strategies have led to a new paradigm using *mixed 2D-3D FR systems* where 3D data is used in the training phase but either 2D or 3D information can be used in the recognition depending on the scenario [77]. This type of 3D FR approach is called *multimodal algorithm* and has proved to be very robust in pose variation scenarios. Tsapatsoulis et al. [78] in 1998, Chang et al. [79] in2005, and Samani et al. [80] in 2006 carried out experimentations in this direction, showing that, with frontal view face images and depth maps,

Table 4 Applications of thresholds and face models for facial occlusion localization

Method	Research group	Year	Ref.	Application	Dimension	DB	Reliability	Explanation
Face model	Alyuz et al.	2012	[35]	Face registration	3D	UMB-DB	65.08% Registration accuracy	An average face model is generated and, in order to determine the validity of each patch, thresholding is applied, discarding regions where occlusions are present
			[59]	FR	3D	Bosphorus	94.23% RR	Occlusion detection is handled by thresholding the difference map obtained by computing the absolute difference between an average face model and the input face
Thresholds		2008	[17]	Face registration and recognition	3D	Bosphorus	93.69% RR	A threshold enables to determine whether a given surface is occlusion-free or not
Thresholds		2013	[24]	Face registration and recognition	3D	Bosphorus and UMB-DB	92.91% Recognition accuracy	By thresholding the absolute difference of the input image from a mean face, occluded areas can be detected
Thresholds	Bagchi et al.	2014	[26]	FR	3D	Bosphorus	91.30% Recognition accuracy	The main contribution is that the thresholding technique has been generalized for detection of real-world occlusions

Table 4 continued

Method	Research group	Year	Ref.	Application	Dimension	DB	Reliability	Explanation
Thresholds	Ganguly et al.	2015	[30]	Occlusion detection	2.5D	Bosphorus	82.85% Occlusion detection accuracy	Firstly, the 2.5D or range face images are created from input 3D face images. Then for detecting the occluded faces, a threshold-based approach is used
Block							91.43% Occlusion detection accuracy	Firstly, the 2.5D or range face images are created from input 3D face images, and a block-based approach is undertaken to distinguish the obtrude sections
Thresholds	Zohra et al.	2016	[20]	Occlusion detection	Depth information from Kinect RGB-D cameras	EURECOM Kinect Face Dataset	97.50% Classification rate	For localizing occluded regions, a threshold-based approach is proposed, as objects close to the camera have higher pixel intensity than the object further from the camera
Template matching	Liu et al.	2014	[32]	Nose region detection	3D	Bosphorus	99.48% Nose detection accuracy	Three nose templates have been constructed

Table 5 Contributions proposing other approaches to face facial occlusions detection

Method	Research group	Year	Ref.	Application	Dimension	DB	Reliability	Explanation
n%	Liu et al.	2012	[31]	3D FR	3D	Bosphorus	85.8% Verification rate and 97.9% FR rate	A rejection strategy is embedded into the ICP algorithm for discarding the occluding parts
Thresholds	Bellil et al.	2016	[27]	3D face identification and recognition	3D	Bosphorus and UMB-DB	90.43% Occlusion detection accuracy and 93.57% RR	Occlusions are detected by analysing and thresholding the wavelet coefficients

the accuracy for the multimodal algorithms outperforms the RRs of either 2D or 3D modalities alone.

In 2009, Colombo et al. [23] proposed a new algorithm for 3D face detection which takes advantages from both 2D and 3D data. It maintains the main multimodal structure of the previous one [81], adding the capability to detect faces even in presence of self-occlusions, occlusions generated by any kind of objects, and missing data, and to indicate the regions covered by the occluding objects. The input of the algorithm is a single *range image* (2.5D), where the occluding object is something with a different luminance/colour pattern from a typical face image, i.e. a face template. Then, the image is analysed in order to find occluding objects, and, if present, occluding objects are eliminated from the image invalidating the corresponding pixels. As training set, 100 neutral acquisitions from the BU3DFE database have been used for computing the mean face template, whereas the algorithm has been tested using the non-occluded UND. Therefore, the database has been processed with an artificial occlusions generator producing realistic acquisitions that emulate unconstrained scenarios. A rate of 89.8% of correct detections has been reached. In order to verify the results obtained with artificial occlusions, a small set of acquisitions containing real occlusions has been also captured. This test set (IVL test set) is composed of 102 acquisitions containing 104 faces (in two acquisitions there were two subjects at the same time). In this case, the 90.4% of faces have been correctly detected.

In 2014, Alyuz et al. [25] proposed two different multimodal occlusion detection approaches. In the first case, each pixel is checked for validity as a facial point. To decide whether each pixel is from the facial surface or the occluding object, the difference between an image and the average face model, computed using Gaussian mixture models (GMMs), is calculated and compared with a threshold value. The second detection technique takes advantage of neighbouring pixel relations. In detail, the facial surface is represented as a graph and each pixel corresponds to a node in the graph. Additionally, there are two terminal nodes, namely the source and the sink nodes, which respectively represent the object

(occlusion) and the background (face). The face or occlusion pixels will be connected only to one of these terminals. Then, the occlusion detection problem becomes a binary image segmentation problem, where 'face' and 'occlusion' pixels form two distinct sets of surface pixels, taking into account the spatial relation of pixels, and is solved using graph cut techniques [82,83]. Both of the occlusion detection techniques require statistical modelling of facial pixels, and the models are trained on the FRGC v.2 dataset. Separated from the training set, the Bosphorus and UMB-DB are used as the evaluation sets. Therefore, for the Bosphorus database, 105 and 381 images are included in gallery and probe sets, respectively. Instead, for the UMB-DB, there are 142 gallery images and 590 probe images database. Using the first method, an overall 87.66% detection accuracy is reached with the Bosphorus database as training set, whereas a lowest rate is obtained with the UMB-DB (66.27%), since it includes highly challenging scans. On the contrary, the second method leads to better performances. In detail, using the Bosphorus and UMB-DB databases, 89.24% and 67.63% detection accuracies are reached, respectively.

In the same year, also Srinivasan et al. [36] investigated the multimodal techniques and developed an innovative occlusion detection strategy for the recognition of three-dimensional faces partially occluded by unforeseen objects. Firstly, the three-dimensional faces are encoded as range images. Then, an efficient method is used for detection of occlusions, which specifies the missing information in the occluded face. The results are verified using UMB-DB database, with separate testing and training subsets. The overall FR rate is 95.49%. Moreover, the system handles occlusions like eyeglass and hat efficiently, reaching the 99.03% and 98.52%FR rate, respectively.

Table 6 summarizes the multimodal techniques developed by the three research groups.

3.6 Similarity values

Finally, despite the increasing amount of related literature, *automatic landmarking* is still an open research branch, as

Table 6 Contributions about multimodal algorithms for facial occlusions detection

Method	Research group	Year	Ref.	Application	Dimension	DB	Reliability	Explanation
Face template	Colombo et al.	2009	[23]	3D face detection	Range image	UND + Artificial occ.	89.8% Face detection rate	The image is analysed in order to find occluding objects, as the occluding object has a different luminance/colour pattern from the face template
Average face model	Alyuz et al.	2014	[25]	Occlusion detection	2D	Bosphorus and UMB-DB	87.66% Occlusion detection accuracy	The difference between an image and the average face model is calculated
Neighbouring pixel relations						Bosphorus and UMB-DB	89.24% Occlusion detection accuracy	The occlusion detection problem is treated as a binary image segmentation problem
Average face model	Srinivasan et al.	2014	[36]	Occlusion detection	Range image	UMB-DB	95.49% RR	Firstly, the 2.5D or range face images are created from input 3D face images. Then, the occluded regions are located by thresholding a facial model

current techniques lack both accuracy and robustness, particularly in presence of self-occlusions and occlusion by accessories.

For this reason, Zhao et al. [34] in 2011 proposed a framework for reliable landmark localization on 3-D facial data under challenging conditions, an occlusion classifier, and a fitting algorithm. To undertake occlusion detection, information from the depth map is extracted, as the presence of occlusion significantly changes local shape. Then, the local *similarity values* are computed for all points in a local region and a histogram of local similarity values is build. Finally, the distance between histograms is computed using the Euclidean metric. The identification of the partially occluded faces and the occlusion type classification is performed using a simple a k-nearest neighbour (k-NN) classifier. This information about occlusions is also integrated into the function used in the fitting process to handle landmarking on partially occluded 3-D facial scans. The technique effectiveness is supported by experiments on the Bosphorus data set containing partial occlusions. During the testing, partial occlusions caused by glasses, by a hand near the mouth region or near the ocular region, are consider in addition to non-occluded 3-D scans. On the contrary, partial occlusions by hair are excluded as they usually do not occur in the landmark regions. The experiment was carried out on 292 scans, and an occlusion detection accuracy of 93.8% was achieved. Moreover, 71.4% of the landmarks were localized with a 10mm precision, and 97% with a 20mm precision.

Table 7 summarize the contributions adopting these techniques.

4 Occlusion management

After the occlusion detection, two different approaches dealing with missing data can be undertaken, by either removing or restoring the occluded parts.

In 2006, Colombo et al. [84] presented an approach to restore the face, called *Gappy Principal Component Analysis* (GPCA), which relied on the estimation of the occluded parts [85]. Gappy PCA was proposed as a Principal Component Analysis (PCA) variant to handle data with missing components. With Gappy PCA, it is possible to reconstruct original signal when the signal contains missing values. However, in order to estimate the unknown facial data by the Gappy PCA method, locations of the missing components are required and a general model of facial data has to be constructed using a training set of non-occluded images, to approximate the incomplete data. The same GPCA approach of restoring images presenting invalid regions of pixels, due to occlusion detection or holes generated by acquisition artefacts, was improved by the authors in 2009 [23].

A different approach was undertaken by Alyuz et al. [17] in 2008. They, instead of restoring the missing parts, proceeded to the automatic elimination of the detected occluded regions. According to them, automatically removing occluded regions improves the correct classification rate. This idea is supported by the research group also in their following paper [35] of 2012, where they proposed to discard regions where occlusion artefacts are present. Moreover, the occluded areas are removed automatically, achieving even better identification rates (67.32%) than by using manually removed occlusions (65.08%).

Another contemporary contribution of the same group [59] proposed to apply GPCA after the removal of non-facial parts introduced by occlusions. The purpose was to restore the whole face from occlusion-free facial surfaces. Differently from the group of Colombo [23,84], the reconstructed data were here used only for the missing components and the original data were kept for the non-occluded facial regions. The experimental results showed that restoring missing regions actually decreased the identification rates. In detail, the identification accuracies obtained without occlusion restoration is 83.99%, higher than the 81.63% of accuracy recorded with restoration. The explication provided by the research group is that the restoration of missing regions does not necessarily involve an increase in person-specific surface characteristics, as the occlusion-free regions are the areas which contain the most discriminatory information, not the restored-estimated facial areas. Hence, the restoration via a PCA-like model does not add discriminatory information useful for recognition purposes.

In 2013, Alyuz et al. [24] continued the experiments in both directions, applying occlusion removal and restoration. For removal, the global masked projection has been used. Instead, the partial Gappy PCA has been chosen for restoration; this means that the occluded parts are first removed from the surface. Then, to complete the facial surface, the missing components are estimated using eigenvectors computed by PCA. Again, the results indicate that, instead of restoring occluded areas, it is beneficial to remove the occluded parts and employ only the non-occluded surface regions. As a matter of fact, for the Bosphorus database, the 87.40% of accuracy is reached, only removing the occluded parts and obtaining a performance increase of 2–4% (83.46% of accuracy is obtained, applying the facial reconstruction). Instead, for the UMB-DB, an enhanced performance (about 6–9%) is achieved, passing from 60.34 to 69.15%. A similar approach was followed in 2014, by atomically removing facial occlusions.

Also Li et al. in 2012 [86] discarded occluded regions covered by hair or moustache. Firstly, they identified faces in which deformations were presented. Then, they undertake

Table 7 Contributions proposing other approaches to face facial occlusions detection

Method	Research group	Year	Ref.	Application	Dimension	DB	Reliability	Explanation
Depth map and a k-NN classifier	Zhao et al.	2011	[31]	Landmark localization	3D	Bosphorus	93.8% Classification accuracy	Depth map and a k-NN classifier

the recognition process by placing more emphasis on those regions which are least affected by the deformation.

Yu et al. [48] in 2016 proposed to address the recognition problems with partial occlusions, by eliminating the occluded parts.

There are also several approaches that rely on the estimation of the occluded parts. Drira et al. in 2013 [87] tried to predict and complete the partial facial curves using a statistical shape model. In detail, this method uses the *PCA algorithm* on the training data, to form a training shape, used to estimate and reconstruct the full curves. Results showed that the process of restoring occluded parts, by estimating the missing curves that include important shape data, significantly increased the accuracy of recognition. In detail, the RR is 78.63% when the occluded parts are removed and the FR algorithm is applied using the remaining parts only. Whereas, the RR improves when the restoration step and the estimation of missing parts and curves are performed, reaching the 89.3% of accuracy. However, the performances are worse in the case of occlusion by hair, more difficult to restore. In this case, the removal of curves passing through occlusion is better (90.4% RR) than restoring them (85.2% RR). In 2014 [29], the group adopted the same approach to restore the missing parts due to external objects occlusion. A restoration step, which uses statistical estimation on curves, specifically by using PCA on tangent, is again proposed to handle the missing data and to complete the partially observed curves.

In the same year, also Bagchi et al. [26] proposed a restoration step of facial occlusions. Firstly, the occluded regions are automatically detected by thresholding the depth map values of the 3D image. Then the restoration is undertaken by Principal Component Analysis (PCA). The rank-1 recognition rate is 66%,

Also Srinivasan et al. [36] in 2014 preferred a restoration method to eliminate facial occlusions. The restoration strategy uses *GPCA*, independent from the method used to detect occlusions and usable also to restore faces subject to noise and missing pixels due to acquisition inaccuracies. The restoration module recovers the whole face by exploiting the information provided by the non-occluded parts of the face and renders a restored facial image. The occluding extraneous objects may be glasses, hats, scarves and may differ greatly in shape or size, introducing a high level of variability in appearance. Thanks to the restoration module carried

out, the features of the face not available due to occlusion are retained, improving the recognition efficiency.

In 2014, a new occlusion restoration strategy is follow by Bellil et al. [27], who restored 3D occluded regions with a Multi Library Wavelet Neural Network [88].

In 2014, Gawali et al. [89] performed an occlusion removal step, based on the *recursive ICP algorithm*. Then, they pursued restoration with a curve, which uses statistical estimation on shape manifolds of curves to handle missing data. Specifically, the partially observed curves are completed by using the curves model produced through the PCA technique.

Srinivasan et al. [36] in 2014 eliminated occlusions and rendered a restored facial image. To achieve this, GPCA face restoration method was used. The restoration module recovers the whole face by exploiting the information provided by the non-occluded part of the face. Moreover, their 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. Hence, thanks to the restoration module carried out, the features of the face not available due to occlusion are retained.

This strategy of reconstructing missing part is considered valuable also by Ganguly et al. [30] in 2015. They proposed two novel techniques for occlusion detection and localization and affirm to be working on the reconstruction of the facial surface, to recreate the depth values in localized occluded sections.

Finally, in 2016, Bahve et al. [50] introduced in their paper a restoration step that used statistical estimation on curves to handle missing data. Specifically, the partially observed curves are completed using the curves model constructed by using PCA.

Table 8 summarizes the decisions taken by the each research group about how manage the occlusions localized, i.e. removing o restoring them.

5 Discussion

The last sessions outlined the current methodologies to detect, localize, remove and/or restore facial occlusions. Generally speaking, we may sum up that, in presence of facial occlusion, the first step consists in *localizing occluding objects*. For this task, most of the different approaches, proposed by various authors, used thresholds and average facial

Table 8 Choice of removal or restore occlusions made by each research group

Research group and references	Removal	Restoration
Alyuz et al. [17,24,25,35,59]	✓	
Bagchi et al. [26]		✓
Bahve et al. [50]		✓
Bellil et al. [27]		✓
Colombo et al. [84]		✓
Ganguly et al. [30]		✓
Gawali et al. [89]		✓
Li et al. [86]	✓	
Srinivasan et al. [36]		✓
Yu et al. [48]	✓	

models. These methods usually use a different non-occluded database for the training part, necessary for the threshold setting and the face model creation. Also a new paradigm, which consists in using either 2D or 3D information for FR purposes with partial occlusions, has been introduced. Less work has been done using this multimodal approach but it has proved to be robust, even outperforming in some contexts the RRs of both 2D and 3D modalities alone. Finally, other techniques employed facial and radial curves to efficiently discover occluded parts, as recent significant progresses have been done in the analysis of curves shapes.

Then, once the occlusion has been detected, two possible choices can be made. Some authors preferred to follow a *removal* step, whereas others affirm that a *restoration* step is necessary to improve FR performances. The PCA and GPCA algorithms are the most used techniques for face restoration, in the presence of facial occlusion.

Like stated earlier, none of these methods is designed to handle all the possible occlusion challenges, but each author focuses on a subset of typical occlusions like hand occlusions or occlusions caused by glasses, a scarves, an hat or facial hair.

The analysis of the main techniques presented in the previous sections is organized in Table 9. It catalogues the types of methods used for managing and detecting occluded parts, the databases selected the number and origin of training and testing sets used, the occlusions considered. The table offers entire collection of literature analysed in this survey.

Given that occluding objects partially cover the face view and appear as a part of the facial surface, in these situations the FR approaches decrease their performances. However, the strategies undertaken by these research groups, to isolate the non-occluded patches and restoring the missing part, have produced good results and high FR accuracy is guaranteed even in these challenging situations.

The most studied and tested occlusions are the occlusions caused by free hands in front of face and eyeglasses, but also other occlusions have been considered, to develop sounder algorithms able to deal with any type of occlusions.

Table 10 summarizes the results (RRs of the respective FR methods) obtained by the different authors. In the FR scenario, the objective is to answer to the question 'Who is the probe person?'. Each face in the probe set is compared against the whole gallery set, returning the most similar face. The overall accuracy is reported as well as the results obtained for the various types of occlusions. Recognition rates change according to the type of the occluded object and in particular its size and position.

These results indicate that significant progress has been made in the area of face recognition in presence of facial occlusion. However, the proposed study also shows that, even if occlusion problems have gained increasing interest in the research community, there is still room for further improvement. Firstly, greater attention has to be paid to the multiple sources of occlusion that can affect face recognition in natural environments. As a matter of fact, most of the works focus on the same occlusions caused by sunglasses, hand, and hair, whereas only few researchers have extended their testing to other sources of occlusions such as hat, scarf, scissor, and newspaper.

Hence, as a future work, it is of interest to extend the different approaches in order to address face recognition under general occlusions, including beards, caps, extreme facial make-ups, but also semi-transparent occlusion, caused by the reflection of windscreen of cars or waters, which still are unsolved challenges in practical scenarios, despite their importance for many security management and access control purposes. There exists a large variety of occlusions in real-world scenarios which take into consideration such new types of occlusions may be helpful to improve face recognition in different applications and in the surveillance world.

Moreover, referring to real-world scenarios, the presence of facial occlusions can be together with the presence of other variations (i.e. light source change, facial expressions, pose variations, and image degradations). Hence, even if up to date almost all occlusion studies focus on the occlusion problem itself, future works should investigate on how to handle the occlusion problem in these more challenging conditions, in order to achieve robust face recognition in non-controlled environments. For this purpose, databases which combined different factors are needed. They should consist of face scans affected both by facial occlusion and pose variation, as well as occlusion and expression.

Another strong limitation of these approaches is that they often need manual or semi-automated setting during different phases of the FR procedure. For example, fiducial landmarks needed for estimating the pose and detecting occluded areas are often manual or semi-automated annotated for scans with

Table 9 Framework outlining the literature contributions handling partial occlusions in FR systems

References	TrainingSets	TestingSets	Occl. Considered	Occl. Detec.	Occl. Removed	Occl. Restored
Alyuz et al. [17]	-	Bosphorus (47 neutral scans for the gallery, 170 occluded scans for the probe set)	Mouth occlusions by hand, occlusions of the face with hair and eyeglasses, occlusion of the left eye and forehead regions by hands	Yes A threshold enables to determine whether a given surface is occlusion-free or not	Yes Only non-occluded regions are considered	No
Alyuz et al. [35]	FRGCv2 (2365 neutral scans)	UMB-DB (441 neutral scans for the gallery, 590 occluded scans for the probe set)	Occlusions caused by hair, eyeglasses, hands, hats, scarves, and other objects	Yes Using a threshold applied on template, the validity of each patch is determined	Yes Only non-occluded regions are considered	No
Alyuz et al. [59]	FRGCv2 (2365 neutral scans)	Bosphorus (299 neutral scans for the gallery, 381 occluded scans for the probe set)	Mouth occlusions by hand, occlusions of the face with hair and eyeglasses, occlusion of the left eye and forehead regions by hands	Yes Throughout a generic face model-based occlusion detection method	Yes	No
Alyuz et al. [24]	FRGCv2 (2365 neutral scans)	Bosphorus (105 neutral scans for the gallery, 381 occluded scans for the probe set) and the UMB-DB (142 neutral scans for the gallery, 590 occluded scans for the probe set)	Occlusions caused by hair, eyeglasses, hands, hats, scarves, and other objects	Yes Throughout an occlusion detection method based on difference from an average face model	Yes	No

Table 9 continued

References	TrainingSets	TestingSets	Occl. Considered	Occl. Detec.	Occl. Removed	Occl. Restored
Bagchi et al. [26]	Bosphorus	Bosphorus (92 neutral scans for the gallery, 368 occluded scans for the probe set)	Mouth occlusions by hand, occlusions of the face with hair and eyeglasses, occlusion of the left eye and forehead regions by hands	Yes By comparing each occluded image with the mean face, occluding objects are detected	Yes By invalidating the occluded pixels	Yes Throughout the PCA method
Bellil et al. [27]	-	Bosphorus (381 occluded scans for the probe set) and UND-DB (76 occluded scans for the probe set)	Occlusions due to cap, glasses, free hands, newspaper, hand with phone, scarf, scissors and hat	Yes Applying the discreet wavelet transform	Yes By removing the coefficients whose distance is greater than a threshold	Yes, Applying a Gappy Wavelet Neural Network (GWNN)
Bhave et al. [50]	-	UND and the University of Houston databases	Occlusions due to glasses and hair	With an occlusion detection step based on recursive ICP Yes	With an occlusion removal step based on recursive ICP Yes	Using PCA to complete the partially observed redial curves Yes,
Colombo et al. [84]	UMB (132 images which do not present occlusions)	UMB (76 acquisitions, included 52 occluded faces)	Occlusion due to glasses, scarves, caps, by the hands of the subject, or by other extraneous objects	Yes The deformations are detected from a comparison of the input image with a generic face model	Yes	Yes The approach used is the Gappy Principal Component Analysis

Table 9 continued

References	TrainingSets	TestingSets	Occl. Considered	Occl. Detec.	Occl. Removed	Occl. Restored
Colombo et al. [23]	BU3DFE (100 scans)	UND (951 scans)	Self-occlusions, occlusions generated by any kind of objects, missing data	Yes The deformations are detected from a comparison of the input image with a generic face model	Yes	Yes The approach used is the Gappy Principal Component Analysis
Drira et al. [28]	–	Bosphorus (105 neutral scans for the gallery, 381/360 occluded scans for the probe set)	Mouth occlusions by hand, occlusions of the face with hair and eyeglasses, occlusion of the left eye and forehead regions by hands	Yes By comparing the given scan with a template scan and using a threshold technique.	Yes An occlusion removal step based on recursive-ICP is developed to handle occlusions.	Yes Throughout the PCA method, the missing part are reconstructed and only the curves that pass the quality filter are used for FR purposes
Drira et al. [29]	–	Bosphorus (105 neutral scans for the gallery, 381/360 occluded scans for the probe set)	Mouth occlusions by hand, occlusions of the face with hair and eyeglasses, occlusion of the left eye and forehead regions by hands	Yes An occlusion detection step based on recursive-ICP is developed to handle occlusions	Yes An occlusion removal step based on recursive-ICP is developed to handle occlusions	Yes Throughout the PCA method, the missing part is reconstructed and only the curves that pass the quality filter are used for FR purposes
Ganguly et al. [30]	Bosphorus (30 range face images)	Bosphorus (105 neutral scans for the gallery, 381 occluded scans for the probe set)	Mouth occlusions by hand, occlusions of the face with hair and eyeglasses, occlusion of the left eye and forehead regions by hands	Yes Throughout a threshold-based approach applied on the 7th bit-plane slice	Yes	Yes
Gawali et al. [89]	–	GavabDB, and Bosphorus	Mouth occlusions by hand, occlusions of the face with hair and eyeglasses, occlusion of the left eye and forehead regions by hands	Yes Based on the recursive ICP algorithm	Yes Based on the recursive ICP algorithm	Yes The partially observed curves are completed by using the curves model produce through the PCA technique

Table 9 continued

References	TrainingSets	TestingSets	Occl. Considered	Occl. Detec.	Occl. Removed	Occl. Restored
Li et al. [37]	FRGCv2 and ZJU3DFED databases	FRGCv2	Hair and moustache	Yes, Analysing the deformation of the facial curves	No	No, The occluded regions are simply ignored during the matching face
Liu et al. [31]	–	Bosphorus (105 neutral scans for the gallery, 381 occluded scans for the probe set)	Mouth occlusions by hand, occlusions of the face with hair and eyeglasses, occlusion of the left eye and forehead regions by hands	Yes, The occlusions are the face outliers of the ICP algorithm	Yes, N% points are eliminated as unreliable	No
Yu et al. [48],[33]	–	Bosphorus (105 neutral scans for the gallery, 360 occluded scans for the probe set)	Mouth occlusions by hand, occlusions of the face with hair and eyeglasses, occlusion of the left eye and forehead regions by hands	Yes, Throughout a new string matching algorithm	Yes, The partial matching mechanism eliminates occluded parts	No, Only the most discriminative parts are used during the matching process
Zhao et al. [34]		Bosphorus (105 neutral scans for the gallery, 381 occluded scans for the probe set)	Mouth occlusions by hand, occlusions of the face with eyeglasses, occlusion of the left eye and forehead regions by hands	Yes, Occlusions are detecting analysing the depth map, as the presence of occlusion significantly changes local shape	No	No
Zohra et al. [20]		Eurecom Kinect Face Dataset	Occlusion by paper, occlusion by hand, and occlusion by glasses	Yes, A support vector machine (SVM)[6], computed on the feature histogram obtained from the face images using LBP operator, detect if the input image contains some kind of occlusions	Yes, A threshold-based approach extract high-intensity values from the occluded depth images	No

Table 10 RRs obtained in presence of each type of occlusion

Type of Occlusions	P/G	Overall (%)	Eyeglass (%)	Hat (%)	Hair (%)	Eye (%)	Mouth (%)	Scarf (%)	Scissors (%)	Newspaper (%)
Alyuz et al. [17]	170/47	94.12	97.87	–	89.66	93.62	93.62	–	–	–
Alyuz et al. [35]	590/142	67.63	88.00	81.42	66.67	63.64	63.64	41.72	–	–
Bellil et al. [27]	–	92.25	98.91	94.38	–	85.64	85.64	86.99	**93.58**	**87.01**
Drira et al. [28]	360/105	89.2	94.2	–	85.2	**98.9**	78.5	–	–	–
Liu et al. [31]	381/105	**97.9**	98.08		96.97	**100**	**98.1**			
Srinivasan et al. [36]	–	95.49	**99.03**	98.52	89.62	–	95.14	**95.39**	–	–
Yu et al. [48]	360/105	96.3	94.2	–	**98.0**	97.2	96.1	–	–	–

P/G denotes the size of probe and gallery sets, respectively. For each type of occlusion, the best result obtained is written in bold characters

occlusions and large pose variation [28,29,65,90], and rarely fully automated [66]. In other cases, the processing itself has some manual components, such as the manual segmentation of the facial area into several meaningful parts [24,35]. Future works may address more importance to this issue and develop fully automated techniques, so that also the overall performance of the system can be increased.

It is also worth noting that the annotation of a large number of occluded faces with ground points is a costly and time-intensive operation, which increases the time consumption of already time-hungry algorithms. As a matter of fact, the computational cost is an important issue [77], as each step such as pre-processing, registration, occlusion detection, features extraction, increases time consumption, which is directly proportional to the number of faces treated in the gallery and testing set. Hence, the calculation time becomes an important parameter to consider and therefore in future works should be included in the results. Drira et al. [24] reports the computational time needed to compare two faces and Alyuz et al. [47] compares the time consumption of each step, making evident how there is still room to further acceleration of the computation process. Moreover, the exploitation of parallel techniques for independent tasks may improve the performance and computational efficiency in the coming works.

As a future work, the use of Kinect as well as the progressive adaptation of the algorithms developed for high-quality 3D face data to low-quality scans may bring to reliable, robust and more practical face recognition systems. As a matter of fact, processing using Kinect, which is less time-consuming, can bring to systems able to work in real time and to be potentially commercialized for real-world products. Moreover, both depth and RGB information from Kinect can be simultaneously exploited, to find an optimal strategy for robustly performing face recognition in presence of all types of partial occlusions. For example the occlusion issue may be solved by marking as occlusions all points having a colour which is not likely to be found on a face.

Mixed 2D–3D face recognition philosophy has been already partially used for facing occlusion problem [23,25,

36]. However, this idea to adapt the techniques originally born in 2D to 3D may be extended also to other statistical approaches that have shown good results.

In 2015, Li et al. [40] have adopted the well-known local key-point descriptor matching scheme 2D SIFT to 3D, demonstrating that the proposed approach is very robust also for 3D face recognition under expression changes, occlusions, and pose variations, without using any occlusion detection and restoration scheme. The authors reported a rank-one recognition rate of 100, 100, 100, and 95.52% for the case of eye occlusion by hand, mouth occlusion by hand, glass occlusion, and hair occlusion, totally achieving a recognition rate of 99.21% under the experimental setting of 105 neutral gallery vs. 381 occlusion probes over the Bosphorus database. Hence, this local technique should be further investigated.

Moreover, recently, in the 2D domain, deep learning methods have shown to well perform face recognition task and have become the dominant approach for a wide variety of computer vision tasks. However, deep learning lacks in efficiency, requires complex model, and heavily depends on big data which is computationally expensive, thus not always suitable for real-time applications [91]. Another limitation of this approach, which makes it very difficult for current usage in 3D, is that it requires a wide range of training scans. This lack of an adequate amount of diversified training and testing data to support deep model training should be overcome in the future with the introduction of new databases or the enrichment of existing ones, with a higher number of subjects and types of occlusions. In the meantime, Kim et al. [92], who in 2017 try to undertake this path, overcome this problem by training a CNN using principally 2D facial images, and adding a relatively small number of 3D facial scans. The proposed method was tested on the Bosphorus, reaching 99.2% rank-1 accuracy. More experiments in this field should be undertaken.

Summing up, significant progress have been made in this area of 3D face recognition under occlusions, but the state of the art is also far from having a robust solution able to treat

any type of occlusion, thus it deserves thorough investigations.

6 Conclusion

This survey deals with various 3D methods which tried to overcome the problem of partial occlusions. Different approaches for facial occlusion localization are detailed, like threshold and face models techniques, radial and facial curves usage and the newest multimodal approaches. Among them, many contributions adopted threshold-based methods where the absolute difference is computed between an average face model and the input face, to determine whether a given surface is occlusion-free or not. Less works focused on multimodal-based approach. Then, either by automatically detecting occluded regions or removing/restoring the missing data, it is possible to improve the correct classification rate. PCA e GPCA are the most spread technique used for face restoration.

Regarding occlusions, the most studied and tested are the occlusions caused by free hands in front of face and eyeglasses, but also other occlusions have been taken into consideration, to develop sounder algorithms for dealing with any type of occlusions.

Further improvements, which can be made when dealing with partial occlusion, may be to incorporate other types of occlusions, such as occlusions caused by scarves and foulard, hats, and other accessories, or occlusions of different facial parts. Reducing the computational time and memory usage are other two objectives. The main motivation behind this is the urgent need to have robust and effective applications, which can be used in real-world applications like citizen identification, criminal identification in surveillance, access control in important sites like power plants, military installations, airports, and sensitive offices.

The aims of this survey were to group together and organize all the works related to 3D partial occlusion management. Many researchers have worked on FR systems under controlled environments but fewer contributions have investigated uncontrolled conditions like occlusions.

References

1. Saini, R., Rana, N.: Comparison of various biometric methods. Int. J. Adv. Sci. Technol. (IJAST) **2**(1), 24–30 (2014)
2. Pato, J., Millett, L.: Biometric Recognition: Challenges and Opportunities. National Academies Press, Washington, D.C. (2010)
3. Jain, A., Ross, A., Prabhakar, S.: An introduction to biometric recognition. IEEE Trans. Circuits Syst. Video Technol. **14**(1), 4–20 (2004)
4. Zhao, W., Chellappa, R., Phillips, P., Rosenfeld, A.: Face recognition: a literature survey. ACM Comput. Surv. **35**, 399–458 (2003)
5. Zhao, W., Chellappa, R.: 3D Model Enhanced Face Recognition. In: Proceedings 2000 International Conference on Image Processing 3, 50–53 (2000)
6. Bowyer, K.W., Chang, K.I., Flynn, P.J.: A survey of approaches to three-dimensional face recognition. Int. Conf. Pattern Recognit. **1**, 358–361 (2004)
7. Kakadiaris, I.A., Toderici, G., Evangelopoulos, G., Passalis, G., Chu, D., Zhao, K., Shah, S.K., Theoharis, T.: 3D–2D face recognition with pose and illumination normalization. Comput. Vis. Image Underst. **154**, 137–151 (2017)
8. Ayyavoo, T., Suseela, J.J.: Illumination pre-processing method for face recognition using 2D DWT and CLAHE. IET Biom. (2017)
9. Bowyer, K., Chang, K., Flynn, P.: A survey of approaches and challenges in 3D and multi-modal 3D+2D face recognition. Comput. Vis. Image Underst. **101**(1), 1–15 (2006)
10. Pal, R.: Innovative Research in Attention Modelling and Computer Vision Applications. IGI Global, Hershey, PA (2015)
11. Ekenel, H., Stiefelhagen, R.: Why Is Facial Occlusion a Challenging Problem? Lecture Notes in Computer Science **5558** (2009)
12. Li, H., Huang, D., Morvan, J.M., Chen, L., Wang, Y.: Expression-robust 3D face recognition via weighted sparse representation of multi-scale and multi-component local normal patterns. Neurocomputing **133**, 179–193 (2014)
13. Chellappa, R., Wilson, C., Sirohey, S.: Human and machine recognition of faces: a survey. Proc. IEEE **83**(5), 705–741 (1995)
14. Samal, A., Iyengar, P.A.: Automatic recognition and analysis of human faces and facial expressions: a survey. Pattern Recogn. **25**, 65–7 (1992)
15. Azeem, A., Raza, M., Murtaza, M.: A survey: face recognition techniques under partial occlusion. Int. Arab J. Inf. Technol. **11**, 1–10 (2014)
16. Savran, A., Alyüz, N., Dibeklioğlu, H., Çeliktutan, O., Gökberk, B., Sankur, B., Akarun, L.: Bosphorus Database for 3D Face Analysis. Lecture Notes in Computer Science **5372** (2008)
17. Alyuz, N., Gokberk, B., Akarun, L.: A 3D face recognition system for expression and occlusion invariance. Biometrics: Theory, Applications and Systems **29** (2008)
18. Colombo, A., Cusano, C., Schettini, R.: UMB-DB: A database of partially occluded 3D faces. Computer Vision Workshops (ICCV Workshops) (2011)
19. Moreno, A.B., Sanchez, A.: GavabDB: a 3D face database. Workshop on Biometrics on the Internet, 77–85 (2004)
20. Zohra, F.T., Rahman, W., Gavrilova, M.: Occlusion Detection and Localization from Kinect Depth Images. In: International Conference on Cyberworlds (2016)
21. Min, R., Kose, N., Dugelay, J.: KinectFaceDB: a kinect database for face recognition. IEEE Trans. Syst. Man Cybern. **44**(11), 1534–1548 (2014)
22. Phillips, P., Flynn, P., Scruggs, T., Bowyer, K., Chang, J., Hoffman, K., Marques, J., Min, J., Worek, W.: Overview of the face recognition grand challenge. IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit. **1**, 947–954 (2005)
23. Colombo, A., Cusano, C., Schettin, R.: Gappy PCA classification for occlusion tolerant 3D face detection. J. Math. Imaging Vision **35**, 193 (2009)
24. Alyuz, N., Gokberk, B., Akarun, L.: 3-D face recognition under occlusion using masked projection. IEEE Trans. Inf. Forensics Secur. **8**, 789–802 (2013)
25. Alyuz, N., Gokberk, B., Akarun, L.: Detection of realistic facial occlusions for robust 3D face recognition. In: 22nd International Conference on Pattern Recognition (2014)
26. Bagchi, P., Bhattacharjee, D., Nasipuri, M.: Robust 3D face recognition in presence of pose and partial occlusions or missing parts. Int. J. Found. Comput. Sci. Technol. (IJFCST) **4**(4), 21–35 (2014)

27. Bellil, W., Brahim, H., Ben Amar, C.: Gappy wavelet neural network for 3D occluded faces: detection and recognition. Moltimed. Tools Appl. **75**, 36–380 (2016)
28. Drira, H., Ben Amor, B., Srivastava, A., Daoudi, M., Slama, R.: 3D face recognition under expressions, occlusions and pose variations. IEEE Trans. Pattern Anal. Mach. Intell. **35**, 2270–2283 (2013)
29. Drira, H., Ben Amor, B., Srivastava, A., Daoudi, M., Slama, R.: 3D face recognition using geodesic facial curves to handle expression, occlusions and pose variations. Int. J. Comput. Sci. Inf. Technol. (IJCSIT) **53**, 4284–4287 (2014)
30. Ganguly, S., Bhattacharjee, D., Nasipuri, M.: Depth based occlusion detection and localization from 3D face image. Int. J. Image Graph. Signal Process. **5**, 20–31 (2015)
31. Liu, P., Wang, Y., Huang, D., Zhaoxiang, Z.: Recognizing occluded 3D faces usign an efficient ICP variant. In: IEEE International Conference on Multimedia and Expo (2012)
32. Liu, R., Hu, R., Yu, H.: Nose detection on 3D face images by depth-based template matching. In: The 2014 7th international congress on image and signal processing (2014)
33. Yu, X., Gao, Y., Zhou, J.: Boosting Radial String for 3D Face Recognition with Expressions and Occlusions. In: International Conference on Digital Image Computing: Tecniques and Applications (DICTA) (2016)
34. Zhao, X., Dellandrea, E., Chen, L.: Accurate Landmarking of Three-dimensional facial data in the presence of facial expressions and occlusions using a three-dimensional statistical facial feature model. IEEE Trans. Syst. Man Cybern. **41**(5), 1417–1428 (2011)
35. Alyuz, N., Gokberk, B., Akarun, L.: Adaptive registration for occlusion robust 3D face recognition. In: European Conference on Computer Vision **7585** (2012)
36. Srinivasan, A., Balamurugan, V.: Occlusion detection and image restoration in 3D face image. In: IEEE Region Conference TEN-CON (2014)
37. Li, X., Da, F.: Efficient 3D face recognition handling facial expression and hair occlusion. Image Vis. Comput. **30**(9), 668–679 (2012)
38. Yin, L., Wei, X., Sun, Y., Wang, J., Rosato, M.J.: A 3D facial expression database for facial behavior research. In: FGR '06: Proceedings of the 7th International Conference on Automatic Face and Gesture Recognition, pp. 211–216 (2006)
39. Min, R., Hadid, A., Dugelay, J.L.: Efficient detection of occlusion prior to robust face recognition. Sci. World J. (2014)
40. Li, H., Huang, D., Morvan, J.M., Wang, Y., Chen, L.: Towards 3D face recognition in the real: a registration-free approach using fine-grained matching of 3D keypoint descriptors. Int. J. Comput. Vision **113**(2), 128–142 (2015)
41. Srivastava, A., Klassen, E., Joshi, S.H., Jermyn, I.H.: Shape analysis of elastic curves in euclidean spaces. IEEE Trans. Pattern Anal. Mach. Intell. **33**(7), 1415–1428 (2011)
42. Younes, L., Michor, P.W., Shah, J., Mumford, D.: A metric on shape space with explicit geodesics. Rendiconti Lincei—Matematica e Applicazioni **19**(1), 25–57 (2008)
43. Samir, C., Srivastava, A., Daoudi, M.: Three-dimensional face recognition using shapes of facial curves. IEEE Trans. Pattern Anal. Mach. Intell. **28**, 1858–1863 (2006)
44. Samir, C., Srivastava, A., Daoudi, M., Klassen, E.: An intrinsic framework for analysis of facial surfaces. Int. J. Comput. Vision **82**(1), 80–95 (2009)
45. Samir, C., Srivastava, A., Daoudi, M.: Automatic 3D face recognition using shapes of facial curves. In: IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP) (2006)
46. Raed, M., Zou, B., Hiyam, H.: Nose tip detection in 3D face image based on maximum intensity algorithm. Int. J. Multimed. Ubiquitous Eng. **10**(5), 373–382 (2015)
47. Drira, H., Ben Amor, B., Daoudi, M., Srivastava, A.: Pose and expression invariant 3D face recognition using elastic radial curves. In: British Machine Vision Conference, pp. 1–11 (2010)
48. Yu, X., Gao, Y., Zhou, J.: 3D face recognition under partial occlusions using radial strings. In: IEEE International Conference on Image Processing (ICIP) (2016)
49. Gao, Y., Leung, M.: Human face profile recognition using attributed string. Pattern Recogn. **35**, 353–360 (2002)
50. Bhave, D., Choudhary, R., Gavali, R., Gholap, P.: 3D face recognition under expressions, occlusions, and pose variations. Int. J. Adv. Res. Comput. Commun. Eng. **5**(3), 2270–2283 (2016)
51. Pan, G., Zheng, L., Wu, v.: Robust metric and alignment for profile-based face recognition: an experimental comparison. In: Proceedings of Workshop Applications of Computer Vision, pp. 1–6 (2005)
52. Veltkamp, R.C., Haar, F.B.: A 3D face matching framework for facial curves. Graph. Models **71**(2), 77–91 (2009)
53. Li, C., Barreto, A., Zhai, J., Chin, C.: Exploring face recognition by combining 3D profiles and contours. In: IEEE Proceedings SoutheastCon 576–579 (2005)
54. 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)
55. Kakadiaris, I.A., Passalis, G., Toderici, G., Murtuza, M.N., Lu, Y., Karampatziakis, N., Theoharis, T.: Three-dimensional face recognition in the presence of facial expressions: an annotated deformable model approach. IEEE Trans. Pattern Anal. Mach. Intell. **29**(4), 640–649 (2007)
56. 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**, 1695–1700 (2006)
57. Faltemier, T., Bowyer, K.W., Flynn, P. J.: 3D face recognition with region committee voting. In: Third International Symposium on 3D Data Processing, Visualization, and Transmission 318–325 (2006)
58. Mian, A.S., Bennamoun, M., Owens, R.: An efficient multimodal 2D–3D hybrid approach to automatic face recognition. IEEE Trans. Pattern Anal. Mach. Intell. **29**(11), 1927–1943 (2007)
59. Alyuz, N., Gokberk, B., Spreeuwers, L., Veldhuis, R., Akarun, L.: Robust 3D face recognition in the presence of realistic occlusions. In: 5th IAPR international conference on biometrics (ICB) (2012)
60. Gonzalez, R.C., Woods, R.E.: Digital Image Processing (2007)
61. Srinivas, T., Mohan, P., Shankar, R., Reddy, C.S., Naganjaneyulu, P.V.: Face recognition using PCA and bit-plane slicing. Lect. Notes Electr. Eng. **150**, 515–523 (2013)
62. Otsu, N.: In a threshold selection method from gray-level histograms. IEEE Trans. Syst. Man Cybern. **9**(1), 62–66 (1979)
63. Nirmala Priya, G., Wahida Banu, R.S.D.: Occlusion invariant face recognition using mean based weight matrix and support vector machine. Sadhana **39**(2), 303–315 (2014)
64. Saxena, A., Chung, S.H.: 3D depth reconsrtuction from a single still image. Int. J. Comput. Vision **76**(1), 53–69 (2008)
65. Lu, X., Jain, A.: Automatic feature extraction for multiview 3D face recognition. In: Proceedings of the 7th International Conference on Automatic Face and Gesture Recognition, pp. 585–590 (2006)
66. Passalis, G., Perakis, P., Theoharis, T.: Using facial symmetry to handle pose variations in real-world 3d face recognition. IEEE Trans. Pattern Anal. Mach. Intell. **33**(10), 1938–1951 (2011)
67. Passalis, G., Perakis, P., Kakadiaris, I.: 3D facial landmark detection under large yaw and expression variations. IEEE Trans. Pattern Anal. Mach. Intell. **35**, 1552–1564 (2013)
68. Peng, X., Bennamoun, M., Mian, A.: A training-free nose tip detection method from face range images. Pattern Recogn. **44**, 544–558 (2011)
69. Faltemier, T., Bowyer, K., Flynn, P.: Rotated profile signatures for robust 3d feature detection. In: Proceedings of the 8th IEEE International Conference on Automatic Face and Gesture Recognition, pp. 1–7 (2008)
70. Spreeuwers, L.: Fast and accurate 3d face recognition. Int. J. Comput. Visi. IJCV11 **93**, 389–414 (2011)

71. Gokberk, B., Irfanoglu, M., Akarun, L.: 3D shape-based face representation and feature extraction for face recognition. J. Image Vis. Comput. **24**, 857–869 (2006)

72. Salah, A., Alyuz, N., Akarun, L.: Registration of three-dimensional face scans with average face model. J. Electron. Imaging **17**, 1–14 (2008)

73. Blanken, H., De Vries, A.P., Blok, H.E., Feng, L.: Multimedia Retrieval, pp. 163–165. Springer-Verlag, Berlin (2007)

74. Briechle, K., Hanebeck, U.D.: Template matching using fast normalized cross correlation. Proc. SPIE Opt. Pattern Matching XII **4387**, 95–102 (2001)

75. Maurer, T., Guigonis, D., Malov, I., Presenti, B., Tsaregorodtsev, A., West, D., Medioni, G.: Preformance of Geometrix ActiveID 3D Face Recognition engine on the FRGC data. IEEE CVPR (2005)

76. Chang, K.I., Bowyer, K.W., Flynn, P.J.: Multiple nose egion matching for 3D face recogntion under varying facial expression. IEEE Trans. Pattern Anal. Mach. Intell. **28**, 1695–1700 (2006)

77. Calvo, A.R., Ruiz, F.T., Rurainsky, J., Eisert, P.: 2D-3D Mixed Face Recognition Schemes. In: Recent Advances in Face Recognition, pp. 125–149 (2008)

78. Tsapatsoulis, N., Doulamis, N., Doulamis, A., Kollias, S.: Face extraction from non-uniform background and recognition in compressed domain. Proc. IEEE Int. Conf. Acoust. Speech Signal **5**, 2701–2704 (1998)

79. Chang, K.I., Bowyer, K.W., Flynn, P.J.: An evaluation of multimodal 2D+3D face biometrics. IEEE Trans Pattern Anal. Mach. Intell. **27**, 619–624 (2005)

80. Samani, A., Winkler, J., Niranjan, M.: Automatic face recognition using stereo images. In: IEEE International Conference on Acoustics, Speech and Signal Processing (2006)

81. Colombo, A., Cusano, C., Schettini, R.: 3D face detection using curvature analysis. Pattern Recogn. **39**(3), 444–455 (2006)

82. Boykov, Y.Y., Jolly, M.P.: Interactive graph cuts for optimal boundary & region segmentation of objects in nd images. Int. Conf. Comput. Vis. **1**, 105–112 (2001)

83. Boykov, Y., Funka-Lea, G.: Graph cuts and efficient ND image segmentation. Int. J. Comput. Vision **70**(2), 109–131 (2006)

84. Colombo, A., Cusano, C., Schettini, R.: Detection and restoration of occlusions for 3D face recognition. In: IEEE International Conference on Multimedia and Expo, pp. 1541–1544 (2006)

85. Everson, R., Sirovich, L.: Karhunen–Loeve procedure for gappy data. J. Opt. Soc. Am. A **12**(8), 1657–1664 (1995)

86. Xiaoli Li, F.D.: Efficient 3D face recognition handling facial expression and hair occlusion. Image Vis. Comput. **30**, 668–679 (2012)

87. Drira, H., Ben Amor, B., Srivastava, A., Daoudi, M., Slama, R.: 3D Face recognition under expressions, occlusions and pose variations. In: IEEE Transactions on Pattern Analysis and Machine Intelligence, Institute of Electrical and Electronics Engineers, pp. 2270–2283 (2013)

88. Othmani, M., Bellil, W., Amar, C., Alimi, M.: A New structure and training procedure for multimother wavelet networks. IJWMIP **8**(1), 149–175 (2010)

89. Gawali, S., Deshmukh, R.R.: 3D face recognition using geodesic facial curves to handle expression, occlusion and pose variations. Int. J. Comput. Sci. Inf. Technol. **5**(3), 4284–4287 (2014)

90. Colombo, A., Cusano, C., Schettini, R.: Three-dimensional occlusion detection and restoration of partially occluded faces. J. Math. Imaging Vis. **40**(1), 105–119 (2010)

91. U. D. R. C. i. S. Processing, FG2018 special session: is deep learning always the best solution for face recognition? https://udrc.eng.ed.ac.uk/about/news/20170725/fg2018-special-session-deep-learning-always-best-solution-face-recognition

92. Kim, D., Hernandez, M., Choi, J., Medioni, G.: Deep 3D face identification. Computer Vision and Pattern Recognition (2017)

Nicole Dagnes graduated in Computer Engineering in 2016 at Polytechnic of Turin with a thesis on 3D Face Recognition. She is currently undertaking her PhD on 3D face recognition, under an international Italian-French co-tutored agreement signed by the Polytechnic of Turin and the University of Technology of Compiègne (Sorbonne Universités).

Enrico Vezzetti received the Master's Degree in Management Engineering from the Polytechnic of Turin in 1997 and the PhD in Manufacturing Engineering from the University of Parma in 2001. In 1998, he was in the 3D Printing Laboratory of the Massachusetts Institute of Technology, USA. He is currently Full Professor of Technical Drawing and Design Tools for Industrial Engineering at the Department of Production and Management Engineering of the Polytechnic of Turin. He is author of more than 80 technical papers on product design and development with specific attention on reverse engineering, rapid prototyping and collaborative design topics. His research interests include computer-aided design, computer-aided inspection, 3D scanners and virtual reality.

Federica Marcolin is an Assistant Professor at the Department of Production and Management Engineering of the Polytechnic of Turin. After a Master's Degree in Mathematics in Engineering in 2010, she received her PhD in Production Systems & Industrial Design on January 2015 with a dissertation on Differential Geometry applied to 3D face analysis. In 2014, she was at the i·Bug group of Imperial College London to work on the same topics and now is dealing with 3D face analysis, landmarking, face recognition, face expression recognition.

Stefano Tornincasa was born in Napoli on April 2, 1952. He received his University Degree in Mechanical Engineering in 1977 and was a researcher at the Engineering Faculty of the Polytechnic of Turin from 1983 to 1992. In 1992, he was Associate Professor in Industrial Plants Design and in 1999 he was also Professor of Technical and Industrial Design Numerical in new engineering motor vehicle courses. From 2002 he is full Professor of the ING-IND/15 group at the Polytechnic of Turin.