



HAL
open science

Towards a methodology for retrieving suspects using 3D facial descriptors

Naoufel Werghi, Hassen Drira

► **To cite this version:**

Naoufel Werghi, Hassen Drira. Towards a methodology for retrieving suspects using 3D facial descriptors. RFMI, 2016, Sidi Bou Said, Tunisia. hal-01703251

HAL Id: hal-01703251

<https://hal.science/hal-01703251>

Submitted on 7 Feb 2018

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

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

Towards a methodology for retrieving suspects using 3D facial descriptors

Naoufel Wergchi¹ and Hassen Drira²

¹ Khalifa University, Department of Electrical and Computer Engineering, UAE,

² Institut Mines-Tlcom/Tlcom Lille, Centre de Recherche en Informatique, Signal et Automatique de Lille (UMR CNRS 9189), France

Abstract. In this paper, We propose a first investigation towards a methodology for exploiting 3D facial images for suspect retrieval in the context of crime investigation. In this field, the standard method is to construct a facial composite, based on witness description, by an artist or via a software, then search a match for it in legal databases. An alternative or complementary scheme would be to define a system of 3D facial attributes that can fit human verbal face description and use them to annotate face databases. Such framework would allow a more efficient search of legal face database and more effective suspect shortlisting. In this contribution, we describe some first steps towards that goal, whereby we define some novel 3D face attributes, we analyze their capacity for face categorization through a hierarchical clustering analysis. Then we present some experiments, using a cohort of 107 subjects, assessing the extent to which faces annotation based on some of these attributes meets its human-based counterpart. Both the clustering analysis and the experiments results reveal encouraging indicators for this novel proposed scheme.

Keywords: 3D Face, clustering, face recognition.

1 Introduction

In criminology and police investigation, facial sketches (called also facial composite) are commonly used in searching and identifying crime preprato in the absence of the suspect(s) photos [14]. In addition to identification, facial composite can be used as additional evidence to assist investigation at checking points and to defuse warning of vulnerable population against serial offenders. The sketch of the face used in criminal investigations can be divided into two categories: a) Legal sketches: these sketches are drawn by forensic artists referring to the description provided by a witness. Judicial Sketches have been used in criminal investigations since the 19th century[16]; b) Composite sketches: the sketches of faces are rather built using software allowing an operator to select and combine different elements of the face. Composite sketches are increasingly used. It is now estimated that 80% of law enforcement agencies use software to create facial sketches of suspects [16].

The current procedure of suspect identification based on witness description as it is currently adopted by authorities does not yet seem to profit from all the available resources. In particular face database maintained by legal authorities and which are

continuously fed from network of cameras deployed at access control points and public places. Performance-wise, the current procedures suffer from several shortcomings. Legal sketches production is subjective and depends on the artist skills. Facial composite software, while offer comprehensive construction functionalities, they often produce a mismatched outcomes. Moreover, both categories use 2D face reconstruction, which does not accurately reflect the actual 3D shape features of the subject. Recently some methods proposed to match the face sketch to mugshots (photos of person taken after being arrested) [20, 15] and composite sketches to mugshots [22, 12]. In both of these two schemes, witness description goes through a human interpretation stage, namely the expert artist for face sketch, and the software operator for the composite sketch. Both face sketch and composite sketch are therefore subjected to reconstruction error. Time required to generate the sketch can be problematic for cases requiring immediate investigation.

More recently, a face retrieval approach trend was pioneered by Klare et al. [8]. In this approach a set of textual description of the suspect face are used to interrogate a face database and retrieve a set of potential suspect(s).

The primary investigation, conducted with 2D images, showed that such scheme achieve retrieval performance comparable to the sketch-based part, and has the potential of improving further the accuracy through fusion.

In the this work we proposes investigating a 3D facial image approach of this scheme and capitalizing on the intrinsic advantages the 3D facial images especially with the regard of the shape information. Indeed, a large number of pertinent facial attributes emanate from the facial shape, that one can notice when contemplating a face. These include global attributes (e.g. overall face shap) and local attributes(nose and eye shapes). This approach has higher potential for retrieving facial trait and features that are not preserved in 2D images because of the loss of geometry by projection. The practical deployment of this approach in surveillance and investigations scenarios involving large date sets require an automatic annotation of these last. In this scope, the paper proposes first steps towards this objective.

2 Shape-based 3D Face Description

2.1 Face Kernel

The face kernel is a concept inspired from the "starshapeness" framework [10] in which the kernel (Kern) of a surface is the space (e.g. the set of points) from which the interior of whole surface is visible. It was firstly proposed by Werghi [17] for the purpose of spherical mapping and alignment of facial surfaces. Here we suggest the face kernel as a mean for describing global properties of the facial surface. This intuition behind this suggestion is that the face kernel reflects the convexity of a surface. For instance the kernel of convex surface, (plane or sphere) is the whole space encompassed by that surface. Its counterpart for a non-convex surface will be much more reduced depending on the amount of self-occlusion inferred by protrusions and cavities in this surface. Fig 1 depicts some surface kernels illustrating such difference.

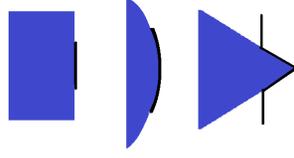


Fig. 1. Examples of kernel for a planar patch, ellipsoid patch and prismatic convex surface. kernels are cut from the left side because of limited space.

We suggest that the size of the face kernel has the potential of reflecting the facial landscape features in terms of the extent of protrusion and concavities that it does exhibit.

Face kernel construction For a triangular mesh manifold surface $\mathcal{S}(V, F)$, where V and F refer to the vertices and the facets, respectively, we can demonstrate that the kernel of the surface \mathcal{S} as follows [17]

$$Kern(\mathcal{S}) = \bigcap_{i=1}^n \mathcal{H}_i \quad (1)$$

Where n is the number of facets in the mesh surface \mathcal{S} , and \mathcal{H}_i is the negative half spaces associated to the plane containing the triangular facet f_i . For an oriented plane, the negative half space is the set of points that fall beneath that plane, as opposite to positive half space that include points which are above that plane. The above definition allows an iterative construction of the surface kernel in a space-carving fashion by initializing it to the whole space then successively discarding from it the positive half space associated to the facet f_i , as presented in the following algorithm

Kernel construction

Generate a 3-D Grid of points \mathcal{G} encompassing the surface \mathcal{S}_n .

$Ker(\mathcal{S}_n) \leftarrow \mathcal{G}$

For each facet t_i

Find, in $Ker(\mathcal{S}_n)$, the set of points Y_i that lie in the half-space \mathcal{H}_i

$Ker(\mathcal{S}_n) \leftarrow Y_i$

End For

End Kernel construction

Fig. 2(a-c) depicts different stages of the kernel construction of a facial surface.

Practically, there is a need to check the integrity of the normal across all the facial mesh surface (to avoid wrongly flipped facet normals) and to apply an optimal smoothing and mesh-regularization of the facial surface to avoid the kernel being affected by mesh artifacts as we will see in the experiments.

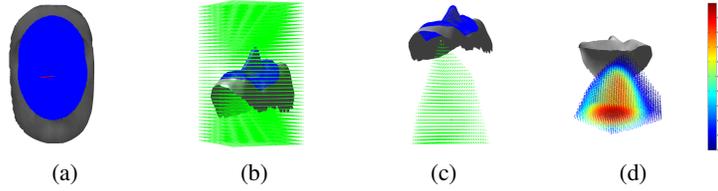


Fig. 2. Kernel construction: (a) Spherical cropping of the facial surface, (b) initial kernel composed a dense set of points encompassing the facial surface. (c) the final kernel. (d) The "Goodness of Visibility computed at each point in the kernel.

The Goodness of Visibility A complementary aspect to the kernel concept is what we call the "goodness of visibility" of the surface, which we define according to the rule of thumb: A surface is best viewed when the line of sight reaches it perpendicularly. While the interior of a surface is visible from any point in the kernel, some points allow a better view than others. For example, for a sphere surface, for which the kernel is its whole interior, the sphere's center is the point having the best view since any ray fired from this point towards the surface, is colinear with the normal at the intersection point.

For a given point in the kernel, We define the "goodness of visibility" by

$$\mathcal{V} = \frac{1}{K} \int_{\mathcal{S}} \varrho_s ds \quad (2)$$

where $\varrho_s ds$ is scalar product of the unit vector defining the orientation of the ray fired from the kernel point towards the facial surface and the local normal at the interception point. K is a normalizing factor. Fig 2.d shows the Goodness of visibility colormapped at each point of the a face kernel. Notice that points that are kernel borders, particularly the closed to the surface, have a less visibility. In contrast with those located in the central zone and at a larger setback distance from the facial surface. These observations fit with the human intuition that a best view of given surface is the one which is centrality and symmetry with respect to the surface. Also while it was not an intention of this research we believe, that the concept of the surface kernel (1) and the goodness of visibility derived from it (2) is a novel and an original criterion, expected to be strong competitor to other standard best viewpoint criteria proposed in the literature [4]

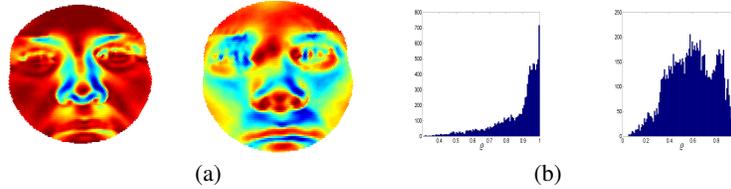


Fig. 3. (a): Examples showing mappings of ϱ on the cropped facial surface obtained for the best (maximum) and worst (minimum) \mathcal{V} . (b): The corresponding ϱ distributions.

2.2 Nasal Profile

In this section we investigate three nasal profiles for human recognition. The first curve is the geodesic path between eyes corners. The second is the geodesic path between nose corners, whereas the third curve (the vertical profile curve) is the geodesic between the mid-eye point and the point at the middle of upper lip edge.

Examples of extracted curves are illustrated in Fig. 4.

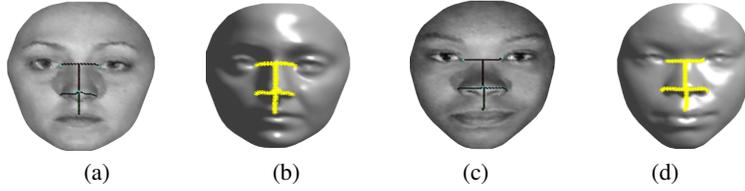


Fig. 4. Illustration of nasal curves: the curves represent the geodesic paths between several facial landmarks.

Background on Shape Analysis of Profile Curves Let $\beta : I \rightarrow \mathbb{R}^2$, represent a parameterized curve representing a nasal profile, where $I = [0, 1]$. To analyze the shape of β , we shall represent it mathematically using the *square-root velocity function* (SRVF) [19], denoted by $q(t)$, according to: $q(t) = \frac{\dot{\beta}(t)}{\sqrt{\|\dot{\beta}(t)\|}}$; $q(t)$ is a special function of β that simplifies computations under elastic metric.

Actually, under \mathbb{L}^2 -metric, the re-parametrization group acts by isometries on the manifold of q functions, which is not the case for the original curve β . Let's define the preshape space of such curves:

$\mathcal{C} = \{q : I \rightarrow \mathbb{R}^2 \mid \|q\| = 1\} \subset \mathbb{L}^2(I, \mathbb{R}^2)$, where $\|\cdot\|$ implies the \mathbb{L}^2 norm. With the \mathbb{L}^2 metric on its tangent spaces, \mathcal{C} becomes a Riemannian manifold. Also, since the elements of \mathcal{C} have a unit \mathbb{L}^2 norm, \mathcal{C} is a hypersphere in the Hilbert space $\mathbb{L}^2(I, \mathbb{R}^2)$. The geodesic path between any two points $q_1, q_2 \in \mathcal{C}$ is given by the great circle, $\psi : [0, 1] \rightarrow \mathcal{C}$, where

$$\psi(\tau) = \frac{1}{\sin(\theta)} (\sin((1-\tau)\theta)q_1 + \sin(\tau\theta)q_2), \quad (3)$$

and the geodesic length is $\theta = d_c(q_1, q_2) = \cos^{-1}(\langle q_1, q_2 \rangle)$.

In order to study *shapes* of curves, one identifies all rotations and re-parameterizations of a curve as an equivalence class. Define the equivalent class of q as: $[q] = \text{closure}\{\sqrt{\dot{\gamma}(t)}O.q(\gamma(t)), \gamma \in I\}$, where $O \in SO(3)$ is a rotation matrix in \mathbb{R}^3 . The set of such equivalence classes, denoted by $\mathcal{S} \doteq \{[q] \mid q \in \mathcal{C}\}$ is called the *shape space* of open curves in \mathbb{R}^2 . As described in [19], \mathcal{S} inherits a Riemannian metric from the larger space \mathcal{C} due to the quotient structure. To obtain geodesics and geodesic distances between elements of \mathcal{S} , one needs to solve the optimization problem:

$$(O^*, \gamma^*) = \underset{\gamma \in I, O \in SO(3)}{\text{argmin}} d_c(q_1, \sqrt{\dot{\gamma}}O.(q_2 \circ \gamma)). \quad (4)$$

Let $q_2^*(t) = \sqrt{\gamma^*(t) \cdot O^* \cdot q_2(\gamma^*(t))}$ be the optimal element of $[q_2]$, associated with the optimal re-parameterization γ^* of the second curve and the optimal rotation O^* , then the geodesic distance between $[q_1]$ and $[q_2]$ in \mathcal{S} is $d_s([q_1], [q_2]) \doteq d_c(q_1, q_2^*)$ and the geodesic is given by Eqn. 3, with q_2 replaced by q_2^* . This representation was previously investigated for biometric [5, 7, 1, 6] and soft-biometric applications [21, 2] based on the face shape.

3 Experiments

We conducted a series of experiments aiming at 1) Analyzing the distribution of the kernel size and the Goodness of Visibility to investigate the presence of potential semantic partition; And 2) Searching for some evidence that can support the concordance of these face descriptors with the human perception when categorizing face based on some morphological traits. In the experiments we used a dataset of 105 scans, from the Bosphorus database [18] corresponding to the set of subjects scanned in neutral pose including male and female instance. This data was first-reprocessed to uniform the mesh, and to remove artifacts using Laplacian smoothing.

3.1 Clustering Analysis

In the first experiments we conducted a series of Hierarchical clustering on the proposed facial attributes. Different Hierarchical clustering methods Can be investigated [13]. Most of the works of Hierarchical clustering of facial images were related to subject recognition [3, 9]. Recently Grant and Flynn investigate Hierarchical clustering beyond subject identification, as to prove the existence of cluster by gender race, and illumination condition. Little or nothing has been done on on whole 3D facial images to the best of our knowledge. The goal of this analysis is to explore the extent to which our attributes can form the basis of semantic partition and a meaningful categorization of the facial shapes, and therefore can be adopted to face annotation. We adopted an agglomerative hierarchical clustering using the standard average methods. Other variants such as the single, complete and Ward [13] could be used as well.

Fig 5.a shows the dendrogram of the nasal profiles based classification. We notice that the dendrogram exhibits two main distinctive clusters. The examination of these samples Fig 5.b reveals clear different aspects in nasal profiles.

Fig 6.(a) shows the dendrograms of the kernel size. We notice that the dendrogram exhibits three distinctive and fairly balanced clusters. On the right are three representative samples from the extrema leaves in the tree. The examination of these samples (Fig6.(b)) reveals a clear dissimilarities aspect between the two groups. Indeed, we can notice that that first group show even shape with moderate variation. In the opposite the second group exhibits ample protrusion (nose) and intrusion (eye sockets) marking salient features of the face. We notice in particular the the second sample shows a lateral nose deformation. Such feature reduces further the visibility of the surface.

The dendrogram of the goodness of visibility is depicted in Fig. 7. Here also we notice three distinctive clusters. As for the kernel size, the three samples of the extrema tree clusters show clear contrast. The first group exhibit rather smooth and even-shaped

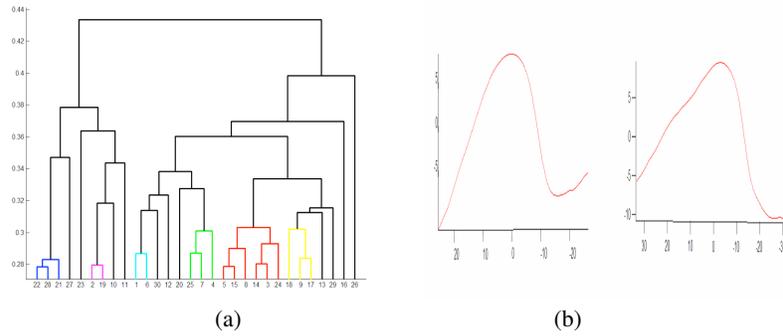


Fig. 5. (a): Nasal profile based dendrogram. (b) Representative samples from two extrema clusters of nasal profiles.

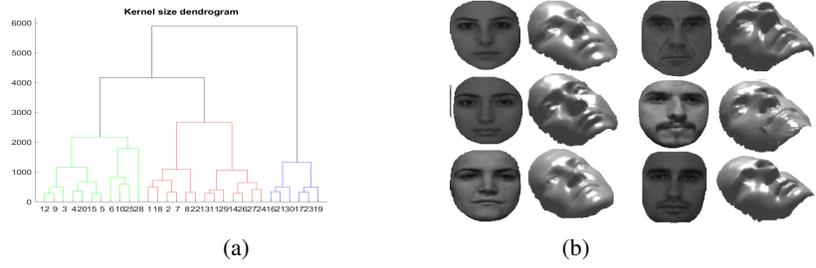


Fig. 6. (a):kernel size dendrogram. (b) Representative samples from two extrema clusters

face, again with an overall flattens aspect. The other group is characterized by a blatant saliency appearance exhibiting significant eye socket intrusion and nose-mouth protrusion with an overall acute shape.

To have an idea on the range of variation of the kernel size and the Goodness of visibility we plotted the values of these two descriptors in ascending order (see Fig 8) for the 105 subjects. From the plots we can notice a range amplitude between the minimum and the maximum values of around 3 and 2 for the kernel size and the Goodness of visibility respectively. We can also notice the clear contrast in the facial shape between the group of the three samples corresponding to the three extrema values in each category.

3.2 Human judgment matching

This experiment aimed at assessing the extent to which face categorization based on the proposed facial attributes, namely, the kernel size and the goodness of visibility, can match human perception. The experiment was set as follows: A cohort of thirty participants composed of undergrad and postgrad students including equal portions of males and females was selected. The group does not include students that are familiar with the databases faces, (e.g. through research projects) as this might affect the perceptual process [11]. Each participant watches a brief video of about 8 seconds showing

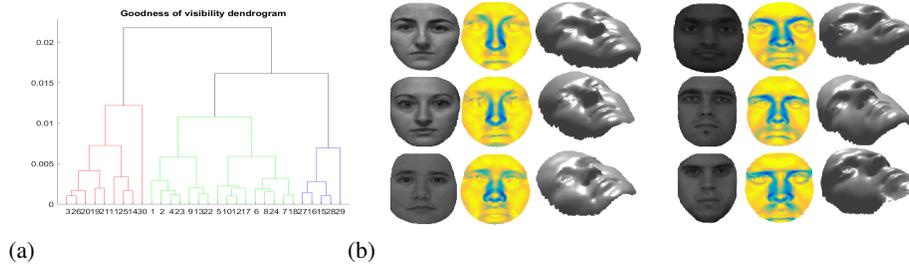


Fig. 7. (a): Goodness of visibility dendrogram. (b) Representative samples from two extrema clusters with their corresponding q mappings.

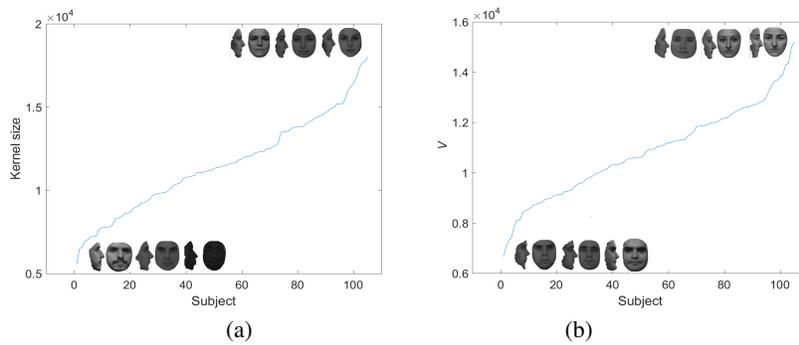


Fig. 8. (a): plots of the ranked kernel size(a) and the Goodness of Visibility (b) for the 100 subjects. Shown also the subjects having the maximum and minimal values for each descriptor.

the 3D face model rotating left to right then right. Afterwards he is asked to select a choice among three options (a: Too little, b Somewhat c: Too much) to a question in the following form: To what extent the face looks having a *Face_description* appearance, where *Face_description* is a brief description of the targeted face profile. Here based on the findings of Section 3.1, we defined two different profiles, namely: *Profile_1*: Wide, flat, unmarked face; And *Profile_2*: Marked face exhibiting protruding nose, intruding eyes. This procedure is repeated for all the 105 models in the dataset with a pause of 3 minutes after each judgment.

Scores collected from the participants are averaged for correlation with scores obtained from the kernel size and the goodness of visibility criterion. For that purpose, we mapped score obtained with these two attributes into three sets representing three segments of their related ranges, and labeled with the three aforementioned options. However, rather than using crisp sets, we considered a mapping to three fuzzy sets as shown in Fig. 9.a This is motivated by the appropriateness of fuzzy rating accommodating comparative judgment and confidence ambiguity characterizing facial description by human [11].

Scores related to the Goodness of Visibility show a slightly better match. The rate of matched scores are reported in Fig. 9.(b) and Fig. 9.(c) for the kernel size and the

goodness of visibility respectively. First we notice that matches with the third set (“Too much”) have the largest and reasonable rate (above 80%). Less matches are obtained for the first set (“Too little”), whereas the middle set shows a relatively low matches. Considered relatively to each other, the matching scores give some indication that both face descriptors concord well with human perception for assigning to (and with a less degree rejection from) the aforementioned profiles. This there is some evidence that significant values of these descriptors can be utilized for labeling subject with these profiles.

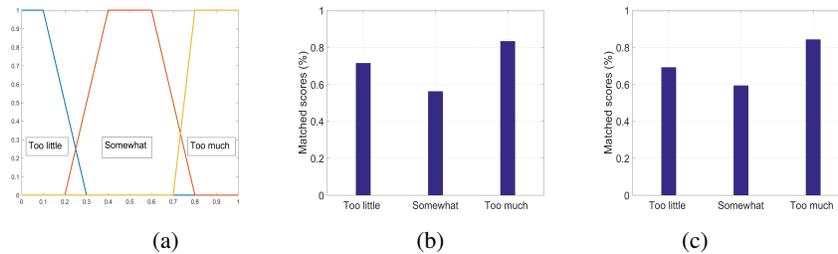


Fig. 9. (a):Fuzzy sets associated with the three judgment options. (b) matched score rate for the kernel size, (c) matched score for the Goodness of Visibility.

4 Conclusion and Discussion

In this paper, we have presented a novel approach for using 3D facial image for retrieving suspects based on witness description. We proposed two global facial descriptors for categorizing facial morphology. the clustering analysis, we performed, seem providing an encouraging indication about the plausibility of these tools for a semantic subject partition that can be verbally described, and thus having a potential to be utilized for annotation. The experiment assessing the extent to which human perception can meet face categorization based on the proposal global descriptors revealed positive trend in this regard, and confirms further the utility of 3D images,

While it is true that the dataset we used is not exhaustive and does not encompass the full spectrum of face morphology (little presence of far east ethnicity), the approach we proposed remain, in our opinion, valid for a more diverse set. To accommodate this diversity, there is a need to consider, in addition to other global attributes, local face attributes reflecting pertinent traits, such as nose and mouth shapes and size of the eyes. The next step in our work is to integrate the nasal morphology and work out related descriptors that can manage the wide spectrum of facial nose. This work developed in [7] can provide appropriate guidance.

References

1. Boulbaba Ben Amor, Hassen Drira, Lahoucine Ballihi, Anuj Srivastava, and Mohamed Daoudi. An experimental illustration of 3d facial shape analysis under facial expressions. *Annales des Télécommunications*, 64(5-6):369–379, 2009.
2. Boulbaba Ben Amor, Hassen Drira, Stefano Berretti, Mohamed Daoudi, and Anuj Srivastava. 4-d facial expression recognition by learning geometric deformations. *IEEE Trans. Cybernetics*, 44(12):2443–2457, 2014.
3. P. Antonopoulos, N. Nikolaidis, , and I. Pitas. Hierarchical face clustering using sift image features. In *Proc. IEEE Symposium on Computational Intelligence in Image and Signal Processing*, pages 325–329, 2007.
4. A.Secord, J.Lu, A.Finkelstein, M.Singh, and A.Nealen. Perceptual models of viewpoint preference. *ACM Transactions on Graphics*, 30:1–13, 2011.
5. Hassen Drira, Boulbaba Ben Amor, Mohamed Daoudi, and Anuj Srivastava. Pose and expression-invariant 3d face recognition using elastic radial curves. In *British machine vision conference*, pages 1–11, 2010.
6. Hassen Drira, Boulbaba Ben Amor, Anuj Srivastava, Mohamed Daoudi, and Rim Slama. 3d face recognition under expressions, occlusions, and pose variations. *IEEE Trans. Pattern Anal. Mach. Intell.*, 35(9):2270–2283, 2013.
7. Hassen Drira, Ben Amor Boulbaba, Srivastava Anuj, and Daoudi Mohamed. A riemannian analysis of 3d nose shapes for partial human biometrics. In *International Conference on Computer Vision*, pages 2050–2057, 2009.
8. B.F. Klare et al. Suspect identification based on descriptive facial attributes. In *Proc. IEEE/IAPR International Joint Conference on Biometrics*, pages 1–8, 2014.
9. W. Fan and D.Y. Yeung. Face recognition with image sets using hierarchically extracted exemplars from appearance manifold. In *Proc. IEEE 7th International Conference on Automatic Face and Gesture Recognition*, pages 177–192, 2007.
10. F.A.Torzanos. The points of local nonconvexity of starshaped objects sets. *Pacific Journal of Mathematics*, 11:25–35, 1982.
11. C. Frowd. Craniofacial identification. In *Craniofacial Identification*, C. Wilkinson and C. Rynn Eds, pages 42–56, 2012.
12. H. Han, B. Klare, and K. Bonnen. Matching composite sketches to face photos: A component based approach. 8:191204, 2013.
13. A. Jain and R. C. Dubes. Algorithms for clustering data. *Prentice- Hall, Inc., Upper Saddle River*, 1988.
14. A. Jain, B. Klare, and U. Park. Face matching and retrieval in forensics applications. *IEEE Multimedia*, 19:20–28, 2012.
15. B. Klare, Z. Li, and A. Jain. Matching forensic sketches to mug shot photos. *IEEE Trans. Pattern Analysis and Machine Intelligence*, 33:639646, 2011.
16. D. Mcquiston, L. Topp, and R. Malpass. Use of facial composite systems in us law enforcement agencies. *Psychology, Crime and Law*, 12:505–517, 2006.
17. N.Werghi. The 3d facial kernel: Application to facial surface spherical mapping and alignment. In *Proc. IEEE Conf. Systems, Men and Cybernetics*, pages 1777–1784, 2010.
18. A. Savran, N. Alyüz, H. Dibeklioğlu, O. Çeliktutan, B. Gökberk, B. Sankur, and L. Akarun. Bosphorus database for 3D face analysis. In *Proc. First COST 2101 Workshop on Biometrics and Identity Management*, May 2008.
19. Anuj Srivastava, Eric Klassen, Shantanu H. Joshi, and Ian H. Jermyn. Shape analysis of elastic curves in euclidean spaces. *IEEE Trans. Pattern Anal. Mach. Intell.*, 33(7):1415–1428, 2011.

20. X. Wang and X. Tang. Face photo-sketch synthesis and recognition. *IEEE Trans. on Pattern Analysis and Machine Intelligence*, 31:1955–1967, 2009.
21. Baiqiang Xia, Boulbaba Ben Amor, Hassen Drira, Mohamed Daoudi, and Lahoucine Ballihi. Combining face averageness and symmetry for 3d-based gender classification. *Pattern Recognition*, 48(3):746–758, 2015.
22. P. Yuen and C. Man. Human face image searching system using sketches. *IEEE Trans. SMC, Part A: Systems and Humans*, 37:493504, 2007.