Pattern Recognition Letters 78 (2016) 1–7

Contents lists available at ScienceDirect

Pattern Recognition Letters

journal homepage: www.elsevier.com/locate/patrec

3D face recognition using covariance based descriptors☆

Walid Hariri[a,b,*], Hedi Tabia[a], Nadir Farah[b], Abdallah Benouareth[b], David Declercq[a]

[a] ETIS, ENSEA, Univ Cergy-Pontoise, CNRS UMR-8051, 95014 Cergy-Pontoise Cedex, France
[b] Labged Laboratory, Computer science department, Badji Mokhtar Annaba University, B.P.12, Annaba 23000, Algeria

ARTICLE INFO

Article history:
Received 13 September 2015
Available online 9 April 2016

Keywords:
Covariance matrix
Geodesic distances
Face matching

ABSTRACT

In this paper, we propose a new 3D face recognition method based on covariance descriptors. Unlike feature-based vectors, covariance-based descriptors enable the fusion and the encoding of different types of features and modalities into a compact representation. The covariance descriptors are symmetric positive definite matrices which can be viewed as an inner product on the tangent space of (Sym_d^+) the manifold of Symmetric Positive Definite (SPD) matrices. In this article, we study geodesic distances on the Sym_d^+ manifold and use them as metrics for 3D face matching and recognition. We evaluate the performance of the proposed method on the FRGCv2 and the GAVAB databases and demonstrate its superiority compared to other state of the art methods.

1. Introduction

Face recognition is one of the biometric techniques frequently used in security systems. It offers several advantages over other biometric methods as stated in [15]. Face recognition can be used in two different scenarios: verification and identification. In verification, the system compares an input face to the "enrolled" face of a specific user to determine whether they are from the same person (one to one match). In identification, the system compares an input face with all the enrolled users faces in the database to determine if the person is already known under a duplicate or false identity (one to all match).

Although research in automated 2D face recognition has been conducted long time ago, it has only recently caught the attention of the scientific community. Several methods and reviews on face recognition techniques have been published in the literature during the past ten years [1,5,15,41,43]. Besides these efforts, face recognition using 2D images is still not reliable enough, especially in the presence of variations in the pose and/or illumination of subjects. [4] constructed a 3D morphable model with a linear combination of the shape and texture of multiple exemplars. Their model was fitted to a single 2D image to obtain the individual parameters, which were used to characterize the personal features. Creating several synthetic faces could be an efficient solution to overcome the issue of pose invariance in 2D face recognition, but the important characteristics of a 3D model, such as shape details, can not be generated from one single 2D picture. The recent developments in 3D imaging systems have brought an important alternative to overcome unsolved problems in 2D face recognition.

The 3D facial data provide naturally more information on the characteristics of the human face which are likely to improve the performance of recognition systems. The main advantage of the 3D based approaches is that the 3D model retains all the information about the face geometry. The simplest 3D face representation is a 3D polygonal mesh, that consists of a set of points (vertices) connected by edges (polygons). There are many ways to build a 3D mesh, the most widely used being: (i) to combine several 2.5D images, where a 2.5D image is a simplified three-dimensional surface representation which contains one depth (z) value for every point in the (x, y) plane, (ii) to properly tune a 3D morphable model or (iii) to make use of a 3D acquisition system (3D scanner). Similarly to 2D face models, 3D face models experience variations in pose, illumination, facial expressions, and aging. The main challenge is then to propose methods using low dimensional feature representation, with enhanced discriminating capability.

1.1. Related works

3D face recognition techniques can be classified into three categories, depending on the type of features they use: global based, local based and hybrid methods.

The global based methods use the holistic face surface as an input to compute similarity measures between faces. Kamencay et al. [16] proposed a 2D-3D face-matching method based on a principal component analysis (PCA). They use a

☆ This paper has been recommended for acceptance by Dr. J. Yang.

* Corresponding author at: Labged Laboratory, Computer science department, Badji Mokhtar Annaba University, B.P.12, Annaba 23000, Algeria. Tel.: +213 669287644.
E-mail address: hariri@labged.net, walid21548@hotmail.fr (W. Hariri).

http://dx.doi.org/10.1016/j.patrec.2016.03.028
0167-8655/

canonical correlation analysis to learn a mapping between 2D and 3D face images. Another 3D face recognition algorithm is proposed by Xu et al. [40] using 3D eigenfaces. First, they build a 3D mesh model from point cloud, and deduce the 3D eigenfaces from the mesh models using a PCA. Finally, they apply the k-nearest neighbor algorithm to perform recognition. Tsalakanidou et al. [33] proposed a face recognition technique based on the implementation of the principal component analysis algorithm and the extraction of depth and colour eigenfaces.

The local based methods firstly detect landmarks or representative facial regions and then use these landmarks/regions to compute similarity measures between faces. Gupta et al. [12] proposed a method based on anthropometric facial fiducial points. Using the points which are associated with discriminatory anthropometric features, they develop an automatic face recognition algorithm that employs facial 3D Euclidean and geodesic distances and a linear discriminant classifier. Mian et al. [20] presented a feature-based algorithm for the recognition of textured 3D faces. They proposed a keypoint detection algorithm which can repeatably identify locations on a face, and measure similarity between them through their graphs. Perakis et al. [24] proposed a method that treats the partial matching problem, using a 3D landmark detector to detect the pose of the facial scan and a deformable model framework which supports symmetric fitting.

Curve feature based methods belong also to the class of local based methods, and use discriminative surface curves extracted from facial representation as local features. For instance, Drira et al. [10] proposed a geometric framework for analyzing 3D faces. They proposed to represent facial surfaces by radial curves emanating from the nose tips and use elastic shape analysis of these curves to compare faces. Another framework proposed by Heseltine et al. [13] for 3D face matching allows profile and contour based face matching. In their approach, they evaluate profile and contour types, and select subsets of facial curves for face matching. The level curves of a depth function is proposed by Samir et al. [27] to represent 3D facial surfaces using a differential geometric approach which computes geodesic distances between these closed curves on a shape manifold.

The hybrid methods combine global and local features. Among hybrid based approaches, Taghizadegan et al. [32] presented a 3D face recognition method which uses the nose tip as a reference point, and employs two dimensional PCA to obtain features matrix vectors. An Euclidean distance method is employed for classification. Huang et al. [14] presented a method for 3D face recognition using a multi-scale extended Local Binary Patterns as facial descriptors along with a local feature hybrid matching scheme. To combine local and holistic analysis, they used SIFT based matching.

1.2. Contributions of the paper

Recently the image analysis community has shown a growing interest in characterizing image patches with the covariance matrix of local descriptors rather than the descriptors themselves. Covariance methods have been successfully used for object detection and tracking [35], texture [34] and image classification [37]. Motivated by their success in image analysis, in this paper, we propose a 3D face recognition method based on covariance descriptors as an extension of covariance based descriptors presented in [31] for 3D shape retrieval. This article explores the usage of covariance matrices of features as discriminant representation for 3D face recognition problems. Our idea is to represent a 3D face with a set of m landmarks selected from its surface. Each landmark has a region of influence, which we characterize by the covariance of its geometric features instead of directly using the features themselves. These features, each of which captures some properties of the local geometry, can be of different type, dimension or scale. Covariance matrices provide a mean for their aggregation into a compact representation, which is then used for computing distances between 3D faces. Covariance matrices, however, lie on the manifold of Symmetric Positive Definite (SPD) tensors (Sym_d^+). Therefore, matching with covariance matrices requires the computation of geodesic distances on the manifold using proper metrics. Several geodesic distances on the Sym_d^+ manifold have been studied. Once we have chosen the appropriate metric, the next step is to establish covariance matches between 3D faces and compute a global similarity measure. Two different strategies are proposed. The first strategy is to compute optimal match using a Hungarian solution for matching unordered set of covariance matrices. The total cost of matching is used as a measure of dissimilarity between the pair of 3D faces. The second strategy is to compute a mean distance by integrating the chosen metric over the pairs of homologous regions, after spatial registration of the 3D faces. Our experiments conducted on two different 3D face datasets demonstrate the performance of the proposed method compared to state of the art approaches.

The remainder of the paper is organized as follows. In Section 2, the method is detailed. Then, in Section 3 the experiments are presented. Conclusions and future developments end the paper.

2. The proposed 3D face recognition system

Fig. 1 gives an overview of the proposed 3D face recognition method. After the acquisition step, the input face surface is preprocessed. The preprocessing helps improving the quality of the input face which may contain some imperfections as holes, spikes and includes some undesired parts (clothes, neck, ears, hair, etc.) and so on. It consists of applying successively a set of filters (Fig. 2). First, a smoothing filter is applied, which reduces spikes in the mesh surface, followed by a cropping filter which cuts and returns parts of the mesh inside an Euclidean sphere. Finally a filling holes filter is applied, which identifies and fills holes in input meshes. Note that spikes mainly occur in three regions: the eyes, the nose

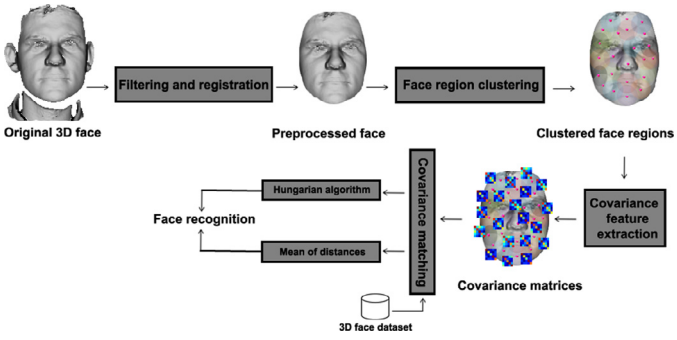

Fig. 1. Overview of the proposed framework.

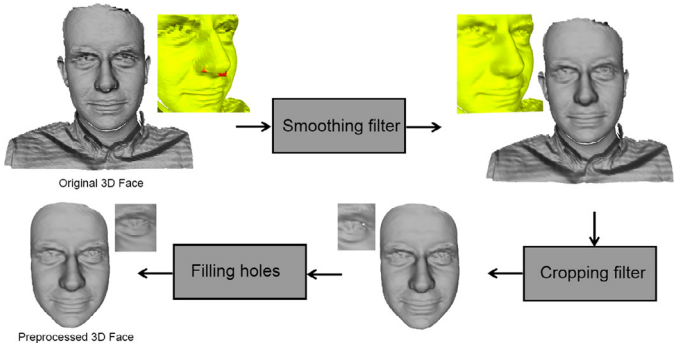

Fig. 2. Automatic 3D face preprocessing.

tip and the teeth. To remove these spikes, we apply a median filter on 3D face vertices. The filter starts by sorting the z coordinate within a neighborhood, finding then the median, and finally replacing the original z coordinate with the value of the median.

After preprocessing, and given a probe face (face to be recognized) F_1 and a gallery face (face in the database) F_0, we uniformly sample m feature points $\{p_1, \ldots, p_m\}$ from the gallery F_0. The m feature points of F_0 are the center of m patches of radius r, and form a paving of the face. We then align F_1 and F_0 by a coarse and fine registration using the Iterative Closest Point (ICP) [3,7,42]. After that, we select, from F_1, $N \leq m$ feature points $\{q_1, \ldots, q_N\}$, which are closest enough to the m points of F_0. In order to do so, we define a distance threshold $\delta = 0.1r$, and for each point p_i, we select its closest point q_i in the probe F_1. The point q_i is considered as a probe feature point only if the Euclidean distance $\|p_i - q_i\| < \delta$. The selected feature points q_j are the centers of the N patches in the probe face, and are used to compute the similarity between F_0 and F_1.

After feature point selection, we compute a covariance matrix for each patch around the feature points. Covariance matrices are not elements of an Euclidean space, they are elements of a Lie group, which has a Riemannian structure. Therefore, matching with covariance matrices requires the computation of geodesic distances on the manifold using a proper metric. In this work, we have compared the performance of our recognition system using several distances for covariance matrices. Finally, for the recognition step, we considered the two following solutions: (i) the Hungarian solution for matching unordered sets of covariance descriptors from two 3D faces, and (ii) the mean of distances between each pair of homologous patches in the two 3D faces.

2.1. Covariance feature extraction

Let $\mathcal{P} = \{P_i, i = 1 \ldots m\}$ be the set of patches extracted from a 3D face. Each patch P_i defines a region around a feature point $p_i = (x_i, y_i, z_i)^t$. For each point p_j in P_i, we compute a feature vector f_j, of dimension d, which encodes the local geometric and spatial properties of the point. In our implementation, we considered the following feature vector:

$$f_j = [x_j, y_j, z_j, k_1, k_2, D_j] \tag{1}$$

Where x_j, y_j and z_j are the three-dimensional coordinates of the point p_j. k_1 and k_2 are respectively the min and max curvatures. D_j is the distance of p_j from the origin defined by $\sqrt{x_j^2 + y_j^2 + z_j^2}$.

We characterize each face patch by the covariance matrix X_i:

$$X_i = \frac{1}{n}\sum_{j=1}^{n}(f_j - \mu)(f_j - \mu)^T \tag{2}$$

Where μ is the mean of the feature vectors $\{f_j\}_{j=1\ldots n}$ computed in the patch P_i, and n is the number of points in P_i. The diagonal entries of X_i represent the variance of each feature and the non-diagonal entries represent their respective co-variations. Using covariance matrices as a region descriptors has several advantages, such as the ability of efficiently combining multiple features into a single descriptor and the invariance with respect to the ordering of points and number of feature vectors used for their computation. The size of covariance matrices does not depend on the size of the region from which they were extracted, but of the size of feature vectors, therefore, they can be computed from variable sized regions. Furthermore, covariance matrices are low dimensional compared to joint feature histograms.

An important aspect to consider is that building covariance-based descriptors requires local features that are correlated to each other otherwise covariance matrices become diagonal and will not provide additional benefits compared to using the individual features. Therefore, the parameters selected in the feature vector f_j need to be carefully selected, and could vary from one database to another. In Section 3, we have performed extensive performance simulations on two databases in order to select the best collection of parameters among the six which are defined in Eq. (1).

2.2. Face comparison

In this section, and for the sake of completeness, we review and discuss the mathematical properties of the space of covariance matrices and the metrics that have been used for computing geodesics and geodesic distances, which in turn have been used as dissimilarity measures for comparing covariance matrices.

2.2.1. The space of covariance matrices

Let $\mathcal{M} = Sym_d^+$ be the space of all $d \times d$ symmetric positive definite (SPD) matrices and thus non-singular covariance matrices. Sym_d^+ is a non-linear Riemannian manifold, i.e. a differentiable manifold in which each tangent space T_X at X has an inner product $\langle \cdot, \cdot \rangle_{X \in \mathcal{M}}$ that smoothly varies from point to point. The inner product induces a norm for the tangent vectors $y \in T_X$ such that $\|y\|^2 = \langle y, y \rangle_X$. The shortest curve connecting two points X and Y on the manifold is called a geodesic. The length $d_g(X, Y)$ of the geodesic between X and Y is a proper metric measuring the dissimilarity between the covariance matrices X and Y. Let $y \in T_X$ and $X \in \mathcal{M}$. There exists a unique geodesic starting at X and shooting in the direction of the tangent vector y. The exponential map $\exp_X : T_X \mapsto \mathcal{M}$ maps elements y on the tangent space T_X to points Y on the manifold $\mathcal{M}$. The length of the geodesic connecting X to Y is given by $d_g(X, \exp_X(y)) = \|y\|_X$.

2.2.2. Distances between covariance matrices

The space of covariance matrices $\mathcal{M} = Sym_d^+$ is a special type of homogeneous space which carries a natural Riemannian structure. More precisely, following the classification given in [21], $\mathcal{M}$ is the Riemannian global symmetric space associated with the Lie algebra. Therefore, we can define a geodesic in the Riemannian space $\mathcal{M}$, or equivalently the space of Hermitian forms, as the shortest curve on $\mathcal{M}$, under the well chosen Riemannian metric or inner product, between two elements of the space $\mathcal{M}$. In this paper, we are interested in computing geodesic lengths, since 3D face recognition only requires a notion of distance between points on the manifold $\mathcal{M}$. In the following, we introduce various distances for dissimilarity measure between two covariance matrices.

A. The affine-invariant and the log-determinant distances:

We first describe the two main distances used in this paper.

- *The affine-invariant distance.* The Riemannian metric of the tangent space T_X at a point X is given as $\langle y, z \rangle_X = \text{trace}(X^{-\frac{1}{2}} y X^{-1} z X^{-\frac{1}{2}})$. The exponential map associated to the Riemannian metric $\exp_X(y) = X^{\frac{1}{2}} \exp(X^{-\frac{1}{2}} y X^{-\frac{1}{2}}) X^{\frac{1}{2}}$ is a global diffeomorphism (a one-to-one, onto, and continuously differentiable mapping in both directions). Thus, its inverse is uniquely defined at every point on the manifold: $\log_X(Y) = X^{\frac{1}{2}} \log(X^{-\frac{1}{2}} Y X^{-\frac{1}{2}}) X^{\frac{1}{2}}$. The symbols exp and log are the ordinary matrix exponential and logarithm operators, while $\exp_X$ and $\log_X$ are manifold-specific operators, which depend on the point $X \in Sym_d^+$. The tangent space of Sym_d^+ is the space of $d \times d$ symmetric matrices and both the manifold and the tangent spaces are of dimension $d(d+1)/2$.

For symmetric matrices, the ordinary matrix exponential and logarithm operators can be computed in the following way. Let $X = UDU^T$ be the eigenvalue decomposition of the symmetric matrix X. The exponential series is defined as: $\exp(X) = \sum_{k=0}^{\infty} \frac{X^k}{k!} =$

$U\exp(D)U^T$, where $\exp(D)$ is the diagonal matrix of the eigenvalue exponentials. Similarly, the logarithm is given by $\log(X) = \sum_{k=1}^{\infty} \frac{-1^{k-1}}{k}(X-I)^k = U\log(D)U^T$. The exponential operator is always defined, whereas the logarithms only exist for symmetric matrices with strictly positive eigenvalues. The affine-invariant distance between two points on Sym_d^+ is given by:

$$d_{ainv}^2(X,Y) = \langle \log_X(Y), \log_X(Y)\rangle_X = \mathrm{trace}\left(\log^2\left(X^{-\frac{1}{2}}YX^{-\frac{1}{2}}\right)\right) \quad (3)$$

- *The log-determinant distance.* From Eq. (3), it is apparent that computing the affine-invariant distance can be inefficient as it requires eigenvalue computations and matrix logarithms, which for large matrices causes significant slowdowns. For an application that need repeatedly to compute distances between numerous pairs of matrices this computational burden can be excessive [8]. Driven by such computational concerns, [6,8,29] introduced a symmetrized log-determinant based matrix divergence. Let $X, Y \in Sym_d^+$, the log-determinant distance between X and Y is defined by:

$$d_{ld}(X,Y) = \sqrt{\log\left(\det\left(\frac{X+Y}{2}\right)\right) - \frac{1}{2}\log(\det(X.Y))} \quad (4)$$

B. Other distances in the Literature:
We describe a few other distances used in the literature.
Let $X, Y \in Sym_d^+$ two SPD matrices.

- *The log-Euclidean distance.* The log-Euclidean distance [2] between X and Y is given by:

$$d_{le}(X,Y) = \|\log(X) - \log(Y)\|_F \quad (5)$$

where $\|\ \|_F$ is the Frobenius norm.

- *The alpha divergence distance.* The α-divergence [9] between X and Y is defined by:

$$d_{\alpha}(X,Y) = \frac{4}{(1-\alpha^2)}\log\frac{\det(\frac{1-\alpha}{2}X + \frac{1+\alpha}{2}Y)}{\det(X)^{\frac{(1-\alpha)}{2}}\det(Y)^{\frac{(1+\alpha)}{2}}} \quad (6)$$

where $\alpha \in (-\infty, +\infty)$.

- *The Kullback–Leibler (KL) divergence.* The KL divergence [39] between X and Y is defined by:

$$d_{KL}(X,Y) = \frac{1}{2}\sqrt{trace(XY^{-1} + YX^{-1} - 2d)} \quad (7)$$

- *The optimal transportation distance.* Optimal transportation distance [36] between X and Y is defined by:

$$d_{ot}(X,Y) = \sqrt{trace(X) + trace(Y) - 2trace((X^{\frac{1}{2}}YX^{\frac{1}{2}})^{\frac{1}{2}})} \quad (8)$$

2.2.3. 3D face matching using SPD matrices

Similar to local features, covariance matrices computed on 3D surfaces can be used as local descriptors for matching two faces. Let us consider a patch center p_i, $i = 1, \ldots, m$ represented by a covariance matrix X_i in a gallery 3D face F_0 and a patch center q_j, $j = 1 \ldots, N$ represented by the covariance matrix Y_j in a probe 3D face F_1. Let $c_{ij} = c(p_i, q_j)$ denotes the cost of matching these two points. This cost is defined as the distance between the two covariance matrices X_i and Y_j.

Given the set of costs c_{ij} between all pairs of points p_i on the gallery face F_0 and q_j on the probe face F_1, we define the total cost of matching the two 3D faces using two different ways:

- *Optimal match.* The total cost of matching is defined by:

$$Cost_1 = \sum_{i=1}^{m} c(p_i, q_{\varphi(i)}), \quad (9)$$

Minimizing $Cost_1$, subject to the constraint that the matching is one-to-one, gives the best permutation $\varphi(i)$. This is an assignment problem, which can be solved using the Hungarian algorithm [17,23]. The input to the assignment problem is a cost matrix with entries c_{ij}. The result is a permutation $\varphi(i)$ such that Eq. (9) is minimized. Finally, once the permutation φ is computed, we use the total cost of matching, defined by Eq. (9), as a measure of dissimilarity between the pair of 3D models.

- *Mean of distances.* An alternative matching cost, simpler than optimal matching consists in computing the mean of distances between each pair of homologous regions. So Eq. (9) becomes:

$$Cost_2 = \frac{1}{N}\sum_{j=1}^{N} c(p_j, q_j), \quad (10)$$

where N is the number of homologous patches.

3. Experiments

We present results from different experiments in which we evaluate the performance of the proposed covariance descriptors. The performance is measured according to the percentage of the correctly recognized faces. We have also studied the impact of the chosen distance and the matching procedure on the recognition performance. The experiments were conducted on two different datasets:

- The *FRGCv2* database [25] which is one of the most comprehensive and popular datasets. It contains 4007 3D face scans of 466 different persons. The scans were acquired in a controlled environment and contain various facial expressions.
- The *GAVAB* dataset [22] contains 549 three-dimensional facial surface images corresponding to 61 individuals (45 male and 16 female). This dataset offers three views per individual in which there are facial expressions (two of them very pronounced). It also includes many variations with respect to the pose of each individual. Fig. 3 shows an example of faces taken from this dataset.

3.1. Experiments on the FRGCv2 dataset

We have first preprocessed the 3D surfaces and selected $m = 40$ feature points on each 3D face in the gallery as described in Section 2. We have then extracted one patch P_i around each point p_i. Each patch has a radius $r = 15\%$ of the radius of the shape's bounding sphere. For each patch, we compute a 5×5 covariance matrices computed from the feature vector $[x, y, z, k_1, k_2]$ (details about the impact of the features, the size of the patch radius r and the number of patches m are given in Section 3.3).

Table 1 presents a comparison of recognition performance on the FRGCv2 database using the different proposed distances (Section 2.2.2) with respect to the two proposed matching methods (Section 2.2.3). When using the Hungarian algorithm, the highest recognition rate is achieved by the log-Euclidian distance 97.9%, followed by the log-determinant distance which achieves

Fig. 3. 3D scans of the same subject from the GAVAB dataset.

Table 1
Recognition rates on FRGCv2 dataset using the different distance metrics presented in Section 2.2.2.

Distance	Hungarian algorithm(%)	Mean of distances(%)
Log-determinant	96.0	**99.2**
Affine-invariant	90.0	99.1
Log-Euclidean	**97.9**	98.7
Alpha divergence	91.0	98.9
KL divergence	70.4	92.9
Optimal transportation	64.1	78.5

Table 2
Comparison with state of the art methods on the FRGCv2 dataset.

Method	Neutral vs. all
[26]	98.4%
[28]	99.0%
[38]	98.3%
[10]	97.0%
[14]	97.6%
[11]	98.1%
Our method	**99.2%**

slightly lower rate 96.0%. When using the mean distance algorithm, the log-determinant distance achieves the highest recognition rate 99.2%. The affine-invariant distance performs 99.1%.

From this experiment, one can notice that when using the geodesic distances (i.e. log-determinant, affine-invariant and log-Euclidean distances), both Hungarian and mean of distances matching techniques behave better than using non-geodesic distances. This demonstrates that the geodesic distances are more discriminative for covariance matrices than the other distances. This is the behavior that one would expect since the non-geodesic distances void one of the benefits of considering the Riemannian structure of the (Sym_d^+) manifold. On the other hand, Table 1 also shows that using the mean of distances matching technique is more suitable for 3D face recognition. This result shows that the spatial relations between covariance matrices are also an important component in the matching process.

Table 2 presents a comparison of our method to several state of the art methods. In this experiment, we evaluate "Neutral versus All" identification experiment, where the first 3D face scan with neutral expression from each subject is used as gallery and the remaining face scans are treated as probes. From Table 2, we can see that our method outperforms the other state of the art methods. This performance can be explained by the fact that covariance matrices provide an elegant way for combining multiple heterogeneous features without normalization or joint probability estimation. This combination significantly boosts the performance of our approach. Note that FRGCv2 contains mostly frontal scans with high quality, so the missing data issues are not treated in this dataset, therefore, many existing methods achieved good performance. In order to evaluate the efficiency of our method against other variations such as pose changes, we have evaluated it on the GAVAB dataset as presented in the next Section.

3.2. Experiments on GAVAB dataset

For the experiment on GAVAB dataset, we use $m = 50$ feature points and 6×6 covariance matrices computed from the feature vector $[x, y, z, k_1, k_2, D]$. Table 3 presents the recognition performance using the different distances presented in Section 2.2.2. The log-determinant and the affine-invariant distances give the highest recognition rates, followed by the alpha divergence, KL divergence, optimal transportation, and the log-Euclidean respectively. From this comparison, we can see that the log-determinant distance gives the highest recognition rate with expressive scans, whereas the affine-invariant distance performs quite well with pose scans. This behavior also demonstrates that the geodesic distances are more efficient for covariance matrices than the other distances. Note that it is also possible to combine the results from each distance, i.e. using vote or training method to further improve the recognition rate.

Table 4 compares the results of our method to results from state of the art methods following the same protocol. We calculated rank-one face recognition rates which show the matching accuracies for different categories of probe faces: including the results with and without expression and pose variations. In this experiment, the first frontal facial scan of each subject was used as gallery while the others were treated as probes, the highest recognition rate achieved by each method is highlighted.

From this comparison, we can see that our method outperforms the majority of the other state of the art approaches in terms of the recognition rate. From Table 4, we can see that for frontal neutral probes, our method provides high recognition rate (100%) similarly as in [10,14,31], note that this rate is obtained by the

Table 3
Recognition rates on the GAVAB dataset using the different distance metrics presented in Section 2.2.2.

Distance	Neutral(%)	Neutral+Expressive(%)	Expressive(%)	Looking down(%)	Looking up(%)	Right profile(%)	Left profile(%)
Log-determinant	**100**	**100**	**100**	97.54	94.26	75.80	81.96
Affine-invariant	100	99.59	99.45	**99.18**	**98.36**	78.69	**83.60**
Alpha divergence	100	99.18	98.90	95.08	93.44	**81.96**	80.32
KL divergence	98.36	97.95	96.72	95.08	91.80	–	–
Optimal transportation	97.54	97.13	95.62	93.44	91.80	–	–
Log-Euclidean	97.54	96.72	95.08	92.34	90.16	–	–

Table 4
Comparison with state of the art methods on the GAVAB dataset.

	Drira et al. [10](%)	Li et al. [18](%)	Tabia et al. [31](%)	Mahoor and Abdel-Mottaleb [19](%)	Huang et al. [14](%)	Our method(%)
Neutral	**100**	96.67	**100**	95	**100**	**100**
Expressive	94.54	93.33	93.30	72	93.99	**100**
Neutral+expressive	95.9	94.68	94.91	78	95.49	**100**
Rotated looking down	**100**	–	–	85.3	96.72	99.18
Rotated looking up	**98.36**	–	–	88.6	96.72	**98.36**
Overall	96.99	–	–	–	–	**97.81**
Right profile	70.49	–	–	–	78.69	**81.96**
Left profile	86.89	–	–	–	**93.44**	83.60

Table 5
Effects of the various geometric features (Section 2.1) on the performance of our face recognition method. Reported results are on both FRGCv2 and GAVAB datasets over all faces.

Features	GAVAB(%)	FRGCv2(%)
$f = [x, y, z]$	95.08	92.7
$f = [k_1, k_2]$	53.16	79.0
$f = [x, y, z, k_1]$	96.72	95.6
$f = [x, y, z, k_2]$	95.08	93.4
$f = [x, y, z, k_1, k_2]$	94.84	**99.2**
$f = [x, y, z, D]$	97.18	92.8
$f = [k_1, k_2, D]$	65.10	78.6
$f = [x, y, z, k_1, k_2, D]$	**97.81**	98.5

log-determinant, the affine-invariant and the alpha divergence distances. For expressive faces, our method with the log-determinant distance provides the highest recognition rate with non-neutral expressions faces (100%) and its performance surpasses all the other methods. The results on (Neutral+Expressive) faces also demonstrate that the proposed method efficiently outperforms the other methods, since we achieve an accuracy of (100%). With looking down faces, our method provides a good recognition rate (99.18%) which is better than the results given by Huang et al. [14] and Mahoor and Abdel-Mottaleb[19] and slightly lower than the result of Drira et al. [10]. Our method also gives the highest recognition rate (98.36%) on looking up faces similarly as in [10], and 97.81% with overall faces. Note that, the performance decreases on left or right sides scanned faces which include many occluded regions, but still outperforms state of the art methods on right side scanned faces. The experimental results on the GAVAB dataset clearly demonstrate that the proposed method can deal with large pose changes and even partial occlusions.

3.3. Effects of the features, the patch size and the number of patches

In this section, we study the performance of the proposed method with respect to the main parameters of the recognition system. First, we studied the impact of the local features that are selected to form the feature vector f (see Eq. (1)). In Table 5, for various choices of feature vectors, we present the performance results on the GAVAB and the FRGCv2 datasets over all faces, using the best performing geodesic distance and the best matching technique, i.e. log-determinant distance and the mean of distances matching algorithm. We can clearly see that the performance of the covariance method highly depends on the chosen features. Although the combination of the six features performs the best in the GAVAB dataset, this experiment shows that the performance of our recognition system does not necessarily improves with the number of selected features. For instance, as shown in Table 5, covarying $[x, y, z]$ features gives slightly better performance than covarying the $[x, y, z, k_1, k_2]$ features. This behavior can be explained by the fact that some feature types are almost orthogonal (i.e. their correlation is low). Thus, their covariance matrix is almost diagonal and therefore not sufficiently discriminative.

We also analyze how the recognition performance of the proposed method varies with respect to the number of sample points. In this experiment, we set the patch radius $r = 15\%$ of the cropped face's bounding sphere and we vary the number of sample points m between 30 to 80. We use the best performing distance and matching technique. Results are summarized in Fig. 4(a). It shows that the performance over all faces becomes stable when the number of sample points is larger than 40 for the FRGCv2 and 50 for the GAVAB dataset. This is predictable since small number of points will result in a coarse representation of the 3D face.

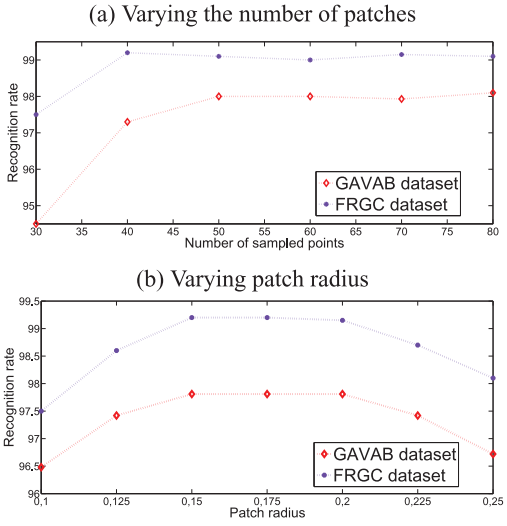

Fig. 4. Effect of the patch radius and the number of patches on the recognition performance of the proposed covariance based method. The reported results are on both FRGCv2 and GAVAB datasets over all faces.

We also analyzed how the recognition performance of the proposed method varies with respect to the patch radius r. For this end, we set the number of sample points $m = 40$ for the FRGCv2 and $m = 50$ for the GAVAB dataset and vary the patch radius between 10% to 25% of the total radius of the cropped face's bounding sphere. Please note that in this setting the patches may overlap. Fig. 4(b) shows that the performance remains stable when r varies between 15% and 20%. The performance starts to drop when choosing values outside this interval. Note that, similar to all local descriptor, this behavior was predictable since very small patches do not capture sufficient geometric properties of the shapes. Large patches on the other hand capture only coarse features, which may not be sufficiently discriminative.

4. Conclusion

In this paper, a new approach for comparing 3D faces using covariance matrices of features has been proposed. We have studied various distances for dissimilarity measure between covariance matrices and proposed two different ways for 3D face matching. Our experiments on two different 3D face recognition datasets, including many variations with respect to the pose and the expression of each individual, show that our method based on covariance matrices with geodesic metrics (i.e. the log-determinant and the affine-invariant distances) outperformed state of the art method. This performance can be explained by the fact that covariance matrices provide an elegant way for combining multiple heterogeneous features without normalization or joint probability estimation. Therefore, analyzing 3D faces with covariance matrices has several advantages compared to individual descriptors. First, covariance matrices enable the fusion of multiple heterogeneous features of arbitrary dimension without normalization, blending weights, or joint probability distribution estimation. Also, spatial relationships can be naturally encoded in the covariance matrices. Moreover, covariance matrices are compact, compared to histogram-based representations, and can be efficiently computed. Finally, although we have experimented with only six types of features, see Section 2.1, our approach is generic and thus other types of features can be added to the framework. An important aspect to consider is that building covariance-based descriptors requires local features that are correlated to each other otherwise covariance matrices become diagonal and will not provide additional

benefits compared to using the individual features instead of their covariance. Although covariance descriptors capture linear dependencies between the various features, dependencies between geometric features could also be non-linear. It would be interesting to investigate whether these non-linear dependencies can be used to further improve the discrimination power of covariance descriptors, e.g. using Brownian or distance covariances [30].

References

[1] A.F. Abate, M. Nappi, D. Riccio, G. Sabatino, 2d and 3d face recognition: A survey, Pattern Recognit. Lett. 28 (14) (2007) 1885–1906.
[2] V. Arsigny, P. Fillard, X. Pennec, N. Ayache, Log-euclidean metrics for fast and simple calculus on diffusion tensors, Magn. Reson. Med. 56 (2) (2006) 411–421.
[3] P.J. Besl, N.D. McKay, Method for registration of 3-d shapes, in: Robotics-DL tentative, International Society for Optics and Photonics, 1992, pp. 586–606.
[4] V. Blanz, T. Vetter, Face recognition based on fitting a 3d morphable model, Pattern Anal. Mach. Intell. IEEE Trans. 25 (9) (2003) 1063–1074.
[5] K.W. Bowyer, K. Chang, P. Flynn, A survey of approaches and challenges in 3d and multi-modal 3d+ 2d face recognition, Comput. Vis. Image Underst. 101 (1) (2006) 1–15.
[6] Z. Chebb, M. Moakher, Means of hermitian positivedefinite matrices based on the log-determinant divergence function, Linear Algebra Appl. 40 (2012).
[7] Y. Chen, G. Medioni, Object modeling by registration of multiple range images, in: Proceedings of the IEEE International Conference on Robotics and Automation., IEEE, 1991, pp. 2724–2729.
[8] A. Cherian, S. Sra, A. Banerjee, N. Papanikolopoulos, Efficient similarity search for covariance matrices via the Jensen-Bregman logdet divergence, in: Proceedings of the IEEE International Conference on Computer Vision (ICCV), IEEE, 2011, pp. 2399–2406.
[9] A. Cichocki, S.-i. Amari, Families of alpha-beta-and gamma-divergences: Flexible and robust measures of similarities, Entropy 12 (6) (2010) 1532–1568.
[10] H. Drira, B. Ben Amor, A. Srivastava, M. Daoudi, R. Slama, 3d face recognition under expressions, occlusions, and pose variations, Pattern Anal. Mach. Intell. IEEE Trans. 35 (9) (2013) 2270–2283.
[11] T.C. Faltemier, K.W. Bowyer, P.J. Flynn, A region ensemble for 3-d face recognition, Inf. Forens. Secur. IEEE Trans. 3 (1) (2008) 62–73.
[12] S. Gupta, M.K. Markey, A.C. Bovik, Anthropometric 3d face recognition, Int. J. Comput. Vis. 90 (3) (2010) 331–349.
[13] T. Heseltine, N. Pears, J. Austin, Three-dimensional face recognition: an eigensurface approach, in: Proceedings of the International Conference on Image Processing, ICIP', vol.2, IEEE, 2004, pp. 1421–1424.
[14] D. Huang, M. Ardabilian, Y. Wang, L. Chen, 3-d face recognition using ELBP-based facial description and local feature hybrid matching, Inf. Forens. Secur. IEEE Trans. 7 (5) (2012) 1551–1565.
[15] R. Jafri, H.R. Arabnia, A survey of face recognition techniques., JIPS 5 (2) (2009) 41–68.
[16] P. Kamencay, R. Hudec, M. Benco, M. Zachariasova, 2d-3d face recognition method based on a modified CCA-PCA algorithm, Int. J. Adv. Robot. Syst. 11 (2014) 36.
[17] H.W. Kuhn, The hungarian method for the assignment problem, Nav. Res. Logist. Q. 2 (1–2) (1955) 83–97.
[18] X. Li, T. Jia, H. Zhang, Expression-insensitive 3d face recognition using sparse representation, in: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, CVPR., IEEE, 2009, pp. 2575–2582.
[19] M.H. Mahoor, M. Abdel-Mottaleb, Face recognition based on 3d ridge images obtained from range data, Pattern Recognit. 42 (3) (2009) 445–451.
[20] A.S. Mian, M. Bennamoun, R. Owens, Keypoint detection and local feature matching for textured 3d face recognition, Int. J. Comput. Vis. 79 (1) (2008) 1–12.
[21] M. Moakher, A differential geometric approach to the geometric mean of symmetric positive-definite matrices, SIAM J. Matrix Anal. Appl. 26 (3) (2005) 735–747.
[22] A. Moreno, A. Sanchez, Gavabdb: a 3d face database, in: Proceedings of the 2nd COST275 Workshop on Biometrics on the Internet, Vigo (Spain), 2004, pp. 75–80.
[23] J. Munkres, Algorithms for the assignment and transportation problems, J. Soc. Ind. Appl. Math. 5 (1) (1957) 32–38.
[24] P. Perakis, G. Passalis, T. Theoharis, G. Toderici, I. Kakadiaris, Partial matching of interpose 3d facial data for face recognition, in: Proceedings of the IEEE 3rd International Conference on Biometrics: Theory, Applications, and Systems, BTAS, IEEE, 2009, pp. 1–8.
[25] P.J. Phillips, P.J. Flynn, T. Scruggs, K.W. Bowyer, J. Chang, K. Hoffman, J. Marques, J. Min, W. Worek, Overview of the face recognition grand challenge, in: Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition, CVPR., vol. 1, IEEE, 2005, pp. 947–954.
[26] C.C. Queirolo, L. Silva, O.R. Bellon, M.P. Segundo, 3d face recognition using simulated annealing and the surface interpenetration measure, Pattern Anal. Mach. Intell. IEEE Trans. 32 (2) (2010) 206–219.
[27] C. Samir, A. Srivastava, M. Daoudi, Three-dimensional face recognition using shapes of facial curves, Pattern Anal. Mach. Intell. IEEE Trans. 28 (11) (2006) 1858–1863.
[28] L. Spreeuwers, Fast and accurate 3d face recognition using registration to an intrinsic coordinate system and fusion of multiple region classifiers, Proc. Int. J. Comput. Vis. 93 (3) (2011) 389–414.
[29] S. Sra, A new metric on the manifold of kernel matrices with application to matrix geometric means, in: Advances in Neural Information Processing Systems, 2012, pp. 144–152.
[30] G.J. Székely, M.L. Rizzo, et al., Brownian distance covariance, The annals of applied statistics 3 (4) (2009) 1236–1265.
[31] H. Tabia, H. Laga, D. Picard, P.-H. Gosselin, Covariance descriptors for 3d shape matching and retrieval, in: Computer Vision and Pattern Recognition (CVPR), 2014 IEEE Conference on, IEEE, 2014, pp. 4185–4192.
[32] Y. Taghizadegan, H. Ghassemian, M. Naser-Moghaddasi, et al., 3d face recognition method using 2DPCA-Euclidean distance classification, ACEEE Int. J. Control Syst. Instrum. 3 (1) (2012).
[33] F. Tsalakanidou, D. Tzovaras, M.G. Strintzis, Use of depth and colour eigenfaces for face recognition, Pattern Recognit. Lett. 24 (9) (2003) 1427–1435.
[34] O. Tuzel, F. Porikli, P. Meer, Region covariance: a fast descriptor for detection and classification, in: ECCV, 2006.
[35] O. Tuzel, F. Porikli, P. Meer, Pedestrian detection via classification on riemannian manifolds, IEEE Trans. Pattern Anal. Mach. Intell. 30 (10) (2008).
[36] C. Villani, Optimal transport: old and new, vol. 338, Springer, 2008.
[37] R. Wang, H. Guo, L.S. Davis, Q. Dai, Covariance discriminative learning: A natural and efficient approach to image set classification, in: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), IEEE, 2012, pp. 2496–2503.
[38] Y. Wang, J. Liu, X. Tang, Robust 3d face recognition by local shape difference boosting, Pattern Anal. Mach. Intell. IEEE Trans. 32 (10) (2010) 1858–1870.
[39] Z. Wang, B.C. Vemuri, An affine invariant tensor dissimilarity measure and its applications to tensor-valued image segmentation, in: Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition, CVPR, vol. 1, IEEE, 2004, pp. I–228.
[40] C. Xu, Y. Wang, T. Tan, L. Quan, A new attempt to face recognition using 3d eigenfaces, in: Proceedings of the Asian Conference on Computer Vision, vol. 2, Citeseer, 2004, pp. 884–889.
[41] X. Zhang, Y. Gao, Face recognition across pose: A review, Pattern Recognit. 42 (11) (2009) 2876–2896.
[42] Z. Zhang, Iterative point matching for registration of free-form curves and surfaces, Int. J. Comput. Vis. 13 (2) (1994) 119–152.
[43] W. Zhao, R. Chellappa, P.J. Phillips, A. Rosenfeld, Face recognition: a literature survey, ACM Comput Surv. (CSUR) 35 (4) (2003) 399–458.