Pattern Recognition ▮▮▮ (▮▮▮▮) ▮▮▮–▮▮▮

Contents lists available at ScienceDirect

Pattern Recognition

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

3D face recognition with sparse spherical representations☆

R. Sala Llonch[a], E. Kokiopoulou[b,*], I. Tošić[b], P. Frossard[b]

[a]Hospital Clinic - Universitat de Barcelona, 08028 Barcelona, Spain
[b]Signal Processing Laboratory (LTS4), Ecole Polytechnique Fédérale de Lausanne (EPFL), Lausanne 1015, Switzerland

ARTICLE INFO

Article history:
Received 28 October 2008
Received in revised form 30 June 2009
Accepted 6 July 2009

Keywords:
Sparse representations
Dimensionality reduction
Spherical representations
3D face recognition

ABSTRACT

This paper addresses the problem of 3D face recognition using simultaneous sparse approximations on the sphere. The 3D face point clouds are first aligned with a fully automated registration process. They are then represented as signals on the 2-sphere in order to preserve depth and geometry information. Next, we implement a dimensionality reduction process with simultaneous sparse approximations and subspace projection. It permits to represent each 3D face by only a few spherical functions that are able to capture the salient facial characteristics, and hence to preserve the discriminant facial information. We eventually perform recognition by effective matching in the reduced space, where linear discriminant analysis can be further activated for improved recognition performance. The 3D face recognition algorithm is evaluated on the FRGC v.1.0 data set, where it is shown to outperform classical state-of-the-art solutions that work with depth images.

1. Introduction

Automatic recognition of human faces is an actively researched area, which finds numerous applications such as surveillance, automated screening, authentication or human–computer interaction. The face is an easily collectible, universal and non-intrusive biometric signal [1], which makes it ideal for applications where other biometrics such as fingerprints or iris scanning are not possible.

There has been a considerable progress in the area of 2D face recognition where intensity/color images of human faces are employed. However, these systems are sensitive to illumination, pose variations, occlusions, facial expressions and make-up. On the other hand, recognition systems based on 3D face information have the potential for greater recognition accuracy and are capable of overcoming part of the limitations of 2D face recognition systems [2,3]. The 3D shape of a face, usually given as a 3D point cloud, depends on its anatomical structure. It is independent of its pose, which can be further corrected by rigid rotations in the 3D space [4].

We consider in this paper the problem of 3D face recognition and we design a fully automatic algorithm based on simultaneous sparse expansions on the sphere. We first propose a preprocessing step that registers the 3D point clouds prior to dimensionality reduction. It selects the facial region and registers all the faces by an accurate automatic two-step algorithm based on an average face model (AFM) and on the iterative closest point (ICP) algorithm [4]. Importantly, the proposed registration process does not require any manual intervention, contrarily to most of the existing algorithms. Registered point clouds are then mapped on the 2-sphere[1] where the spherical face functions are created by nearest neighbor interpolation. If the vector (r,θ,φ) denote the spherical coordinates of a 3D point of the face i, the function $s_i(\theta,\varphi)=r$ defines a surface embedded in $\mathbb{R}^3$, which represents a 3D face. Therefore, the mapping from a point cloud to this surface preserves the geometry of the face, up to a small interpolation error. This representation enables the use of spherical signal processing techniques on the sphere, where face signals are considered as combinations of basis functions with diverse shapes, positions and orientations on the sphere.

The spherical face signals then undergo a dimensionality reduction step that represents each face with a reduced set of discriminant features. We build a dictionary of functions on the sphere and we select the discriminant basis functions by simultaneous sparse approximations. The face signals are then projected onto the resulting reduced subspace, in order to generate feature vectors. We finally implement a recognition step where linear discriminant

☆This work has been partly supported by the Swiss National Science Foundation, under Grants NCCR IM2 and 200020-120063.

* Corresponding author at: Seminar for Applied Mathematics, ETH, Raemistrasse 101, 8092 Zürich, Switzerland.
E-mail address: effrosyni.kokiopoulou@sam.math.ethz.ch (E. Kokiopoulou).

[1] A 2-sphere is a 2D spherical surface embedded in the 3D space. Note that a 2-sphere can be embedded in higher dimensional spaces, e.g. in $\mathbb{R}^4$, since $\mathbb{R}^3 \subset \mathbb{R}^4$. However, we consider $\mathbb{R}^3$ in this work, as 3D faces are typically captured as point clouds in $\mathbb{R}^3$.

0031-3203/$ - see front matter © 2009 Elsevier Ltd. All rights reserved.
doi:10.1016/j.patcog.2009.07.005

Please cite this article as: R. Sala Llonch, et al., 3D face recognition with sparse spherical representations, Pattern Recognition (2009), doi:10.1016/j.patcog.2009.07.005

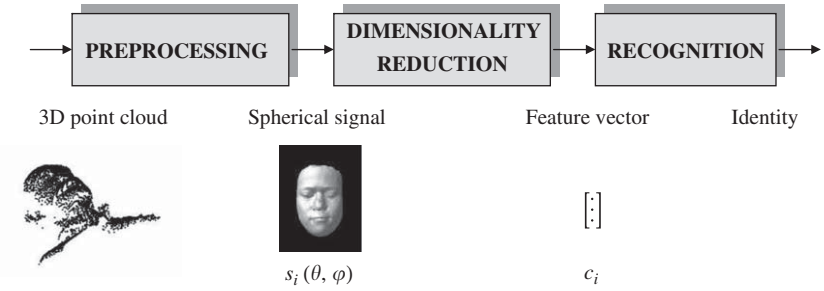

Fig. 1. Block diagram of the 3D face recognition system.

analysis (LDA) is performed on the subspace representation of the faces. The recognition system is illustrated in Fig. 1, where $s_i(\theta, \phi)$ denotes the spherical signal s_i as a function of position (θ, ϕ) on the 2-sphere, and c_i is the corresponding feature vector.

The performance of the 3D face recognition system has been evaluated on the FRGC v.1.0 data set. The proposed algorithm outperforms solutions based on linear discriminant analysis (LDA) and its kernel variants, such as kernel Fisher analysis (KFA) [35] and generalized discriminant analysis (GDA) [36], on depth images. Our fully automatic system provides effective classification performance that shows that 3D face recognition with spherical representations certainly represents a promising solution for person identification.

The paper is organized as follows. We provide an overview of the related work in 3D face recognition in Section 2. Section 3 describes the automatic face registration process that permits to align the 3D points clouds before analysis. The dimensionality reduction step with simultaneous sparse approximations on the sphere is presented in Section 4 and experimental results are provided in Section 5. Section 6 concludes the paper.

2. Related work

3D face recognition has attracted a lot of research efforts in the past few decades due to the advent of new sensing technologies and the high potential of 3D methods for building robust systems with invariance to head pose and illumination variations. We review in this section the most relevant work in 3D face recognition, which can be categorized into methods using point cloud representations, depth images, facial surface features or spherical representations respectively. Further surveys of the state-of-the-art in 3D face recognition can be found in [2,3].

The recognition methods that work directly on 3D point clouds consider the data in their original representation based on spatial and depth information. A priori registration of the point clouds is commonly performed by ICP algorithms [4,6]. The classification is generally based on the Hausdorff distance that permits to measure the similarity between different point clouds [7]. Alternatively, recognition could be performed with "3D eigenfaces" that are constructed directly from the 3D point clouds [8]. Another option is to extract geometrical cues based on eigenvalues and singular values of local covariance matrices defined on the neighborhood of each 3D point [9]. The main drawback of the recognition methods based on 3D point clouds however resides in their high computational complexity that is driven by the large size of the data.

Many recognition systems use depth or range images, where each pixel value represents the distance from the sensor to the facial surface. The 3D face recognition is then formulated as a problem of dimensionality reduction for planar images. Principal component analysis (PCA) [5] and "Eigenfaces" can be used for dimensionality reduction [10]. The basis vectors are however typically holistic and of global support. PCA can be combined with linear discriminant analysis (LDA) to form "Fisherfaces" with enhanced class separability properties [11]. Alternatively, dimensionality reduction can be performed via variants of non-negative matrix factorization (NMF) algorithms [12–14] that produce part-based decompositions of the depth images (see e.g., [37] for a recent evaluation of the NMF performance on 3D face recognition). Part-based decompositions based on non-negative sparse coding [15] have recently been shown to provide improved recognition performance compared to NMF methods in face recognition [16]. Recent methods have proposed to concentrate dimensionality reduction around facial landmarks like the nose tip [17] or in multiple carefully chosen regions [18,19]. Alternatively, one could compute geodesic distances among the selected fiducial points [20]. These methods however require a selection of the fiducial points or areas of interest that is often performed manually and prevents the implementation of fully automatic systems.

Facial surface features have also been proposed for 3D face recognition. The idea of recognizing 3D faces using curvature descriptors has been originally introduced in [21], where features are chosen to represent both curvature and metric size properties of faces. More recently, level sets of the depth function on range images have been used to define sets of facial curves [22]. They are embedded in an appropriately defined shape manifold and compared based on geodesic distances. Facial curve representations provide global information about the whole facial surface, which unfortunately does not permit to take advantage of discriminative local features.

Finally, spherical representations have been used recently for modelling illumination variations [23,24] or both illumination and pose variations in face images [25]. Spherical representations permit to efficiently represent facial surfaces and overcome the limitations of other methods towards occlusions and partial views [27]. To the best of our knowledge, the representation of 3D face point clouds as spherical signals for face recognition has however not been investigated yet. We therefore propose to take benefit of the spherical representations in order to build an effective and automatic 3D face recognition system.

3. Automatic preprocessing of 3D face data

We propose in this section a fully automatic preprocessing method for preparing and aligning 3D face point clouds before feature extraction and recognition. Unlike most of the algorithms in the literature, the preprocessing step does not require any manual intervention. This is an enormous advantage for the design of fully automated face recognition systems. The preprocessing scheme is based on two main tasks, respectively the extraction of the facial region, and the registration of the 3D face. We present these tasks in more details in the rest of the section.

Please cite this article as: R. Sala Llonch, et al., 3D face recognition with sparse spherical representations, Pattern Recognition (2009), doi:10.1016/j.patcog.2009.07.005

3.1. Automatic face extraction

The main purpose of the face extraction step is to remove irrelevant information from the 3D point clouds, such as data that correspond to shoulders, or hair for example. The output of a facial scan typically forms a 3D point cloud $\{X, Y, Z\}$, where X and Y form a uniform Euclidean grid and Z provides the corresponding depth values. The point cloud is also accompanied by a binary matrix A of valid points, which has the same resolution as the grid implied by $X \times Y$. The nonzero pattern of such a sample binary matrix is shown in Fig. 2(a). There is however no guarantee that the points exclusively correspond to face depth information, and face extraction is therefore necessary to ensure that the feature extraction concentrates on capturing discriminative facial information.

The first step in face extraction consists in removing data points on the subject's shoulders. We estimate a vertical projection curve from the point cloud by computing the column sum of the matrix A. Then, we define two lateral thresholds on the left and right inflexion points of the projection curve, and we remove all data points beyond these thresholds, as illustrated in Fig. 2(b). We further remove the data points corresponding to the subject's chest by thresholding of the histogram of depth values. It removes the data points with large depth values that are typically situated behind the data corresponding to frontal face information, as shown in Figs. 2(c) and (d).

We finally have to remove outlier points that remain in regions disconnected from the main facial area, as shown in Fig. 2(e). We therefore perform morphological image processing on the corresponding binary matrix A, where we keep only the largest region that typically corresponds to the facial region, as presented in Fig. 2(f). The automatic face extraction methodology described above worked effectively in the vast majority of facial point clouds in the database, except for very few pathological cases of high measurement noise that were removed from the data set, as is usually done in such cases.

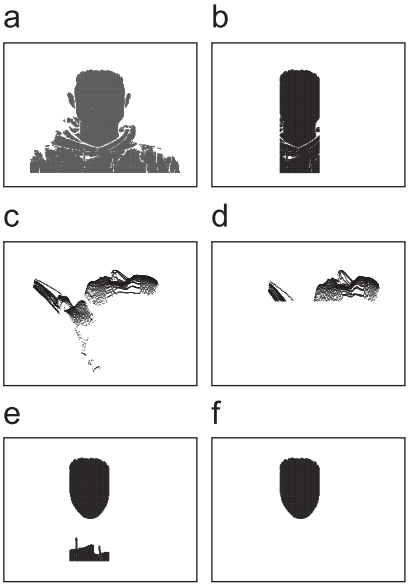

Fig. 2. Main steps in facial region extraction. (a) Binary matrix A; (b) after lateral thresholding; (c) profile view; (d) after depth thresholding (profile view); (e) after depth thresholding; (f) after morphological processing.

3.2. Automatic face registration

After extracting the main facial region from the 3D scans, the face signals have to be registered in order to ensure that all have the same pose before the recognition step. The registration typically applies rigid transformations on the 3D faces in order to align them. We propose a two-step approach for automatic registration, where an average face model (AFM) is first computed and later used for accurate registration.

First, we randomly pick a training face, and we align all the faces approximately to the sample face using the iterative closest point (ICP) algorithm [4]. Given a model and a query point cloud, ICP computes a rigid transformation, consisting of rotations and translations, by minimizing the sum of square errors between the closest model points and query points. After coarse registration with ICP, the face signals are re-sampled on an equiangular grid on the sphere using nearest neighbor interpolation. It permits to construct an AFM, by computing at each grid point the average depth value among all training faces (see Fig. 3). The AFM is subsequently used as reference in order to define an ellipse that contains the main facial region. Since the faces are already registered, this ellipse can be used to crop closely all faces in the training set. The ellipse cropping step removes all the irrelevant information that may be left over from the previous preprocessing steps, as shown in Fig. 4.

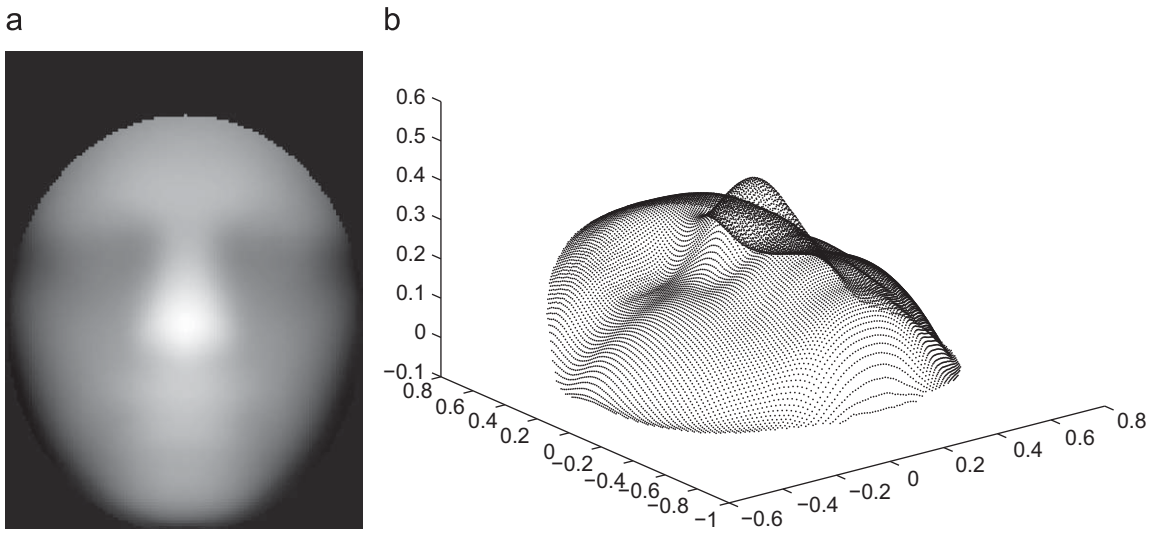

Fig. 3. Average face model given as a depth map or a 3D point cloud. (a) Depth map; (b) point cloud.

Please cite this article as: R. Sala Llonch, et al., 3D face recognition with sparse spherical representations, Pattern Recognition (2009), doi:10.1016/j.patcog.2009.07.005

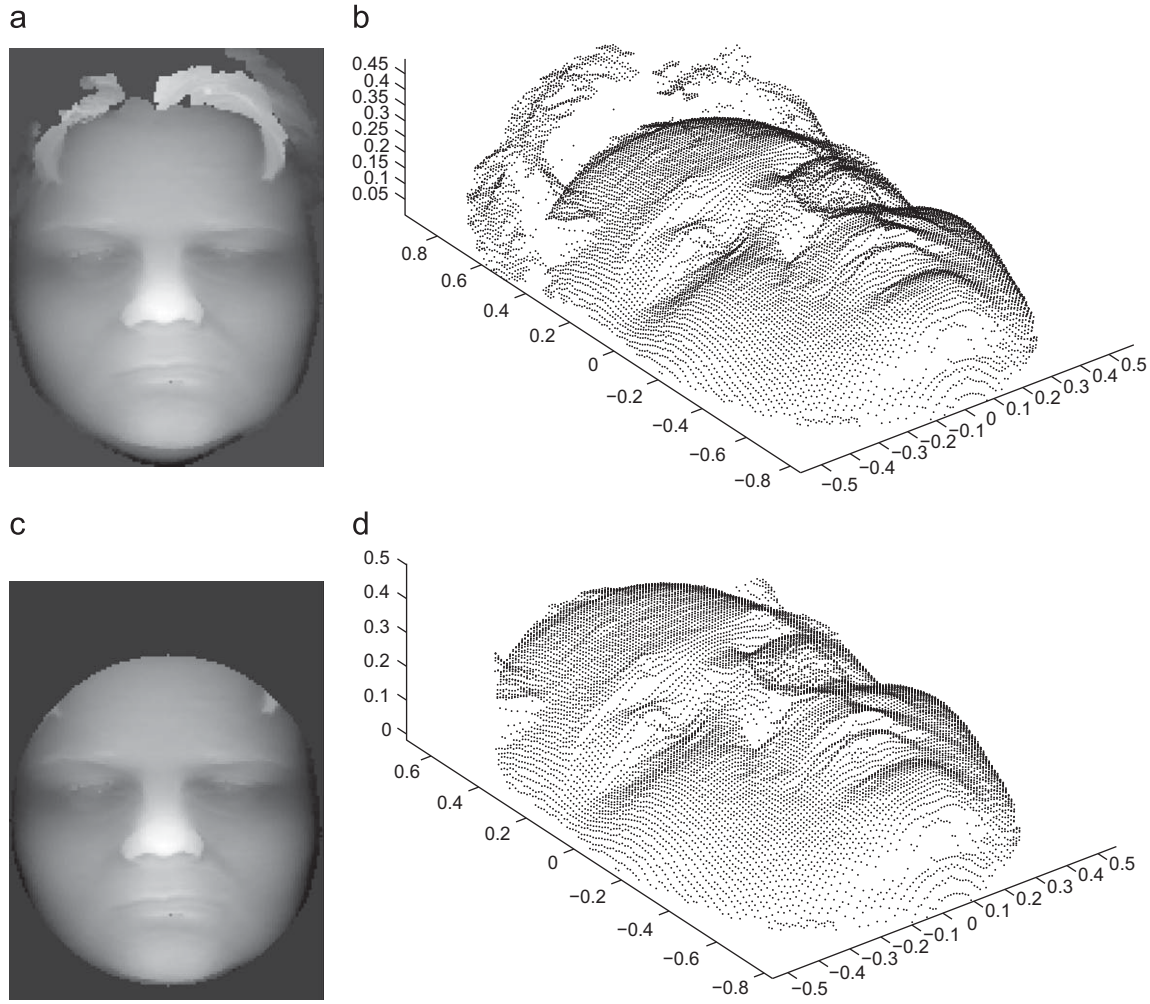

Fig. 4. Illustration of ellipse cropping on depth maps and equivalent 3D point clouds. (a) Before ellipse cropping (depth map); (b) before ellipse cropping (point cloud); (c) after ellipse cropping (depth map); (d) after ellipse cropping (point cloud).

A fine alignment of the faces can now be performed on the signals that have been cleaned from outliers. The accurate alignment is finally obtained by running ICP one more time. The AFM is finally used as a reference face model, and all face signals are finally registered with respect to the AFM.

4. Recognition with sparse spherical representations

4.1. Simultaneous sparse approximations

Efficient face recognition algorithms usually include a dimensionality reduction step, where high dimensional data are represented in a reduced subspace. We propose to use sparse signal representation methods for dimensionality reduction. Such methods have demonstrated good performance in 2D face recognition [28]. They present the advantage of capturing the main signal characteristics in a very small set of meaningful features. These features are defined a priori in a dictionary of functions, called atoms. This presents an interesting advantage compared to classical methods such as PCA, where feature vectors are data-dependent. When the dictionary is composed of spatially localized and oriented anisotropic atoms, sparse approximations of face signals result in a set of salient facial features. Namely, only the atoms that best match facial characteristics are selected. Hence, although the atoms are not data-dependent and are defined a priori by dictionary construction, the atom selection process is data-dependent. We give below a brief overview of sparse approximations, and we show later how we use them for dimensionality reduction on the sphere.

Let s_i, $i=1,\ldots,N$ denote a set of functions in the Hilbert space $\mathscr{H}$. Let further $\mathscr{D}=\{g_\gamma, \gamma\in\Gamma\}$ denote an overcomplete dictionary of unit L_2 norm functions indexed by γ, which spans the space $\mathscr{H}$. A function s_i has a sparse representation in $\mathscr{D}$ if it can be represented in terms of a linear superposition of small set of basis functions $\mathscr{D}_{Ii}=\{g_\gamma\}_{I_i}\subset\mathscr{D}$. In other words, it can be expressed as $s_i=\Phi_{Ii}c_i$, where Φ_{Ii} denotes a matrix whose columns are atoms in $\mathscr{D}_{Ii}$ that form the sparse support of the signal s_i. The vector c_i represents the coefficients of the linear approximation of s_i with atoms in $\mathscr{D}_{Ii}$.

Finding the sparsest representation of a signal in a redundant dictionary $\mathscr{D}$ is in general an NP-hard problem [38]. Greedy algorithms like matching pursuit (MP) [29] have however been shown to provide suboptimal yet efficient solutions in polynomial time. MP selects iteratively the functions from the dictionary that best match the signals s_i. However, we want to form a support set of atoms that is generic with respect to all face signals s_i so that classification can be performed in the corresponding subspace. Finding the sparse support $\mathscr{D}_I$ that is common to all signals s_i can be achieved by the simultaneous MP (SMP) [30] algorithm, which only induces a small increase of complexity compared to MP on a single signal [28]. SMP

Please cite this article as: R. Sala Llonch, et al., 3D face recognition with sparse spherical representations, Pattern Recognition (2009), doi:10.1016/j.patcog.2009.07.005

greedily selects $\mathscr{D}_I$ such that all N functions s_i are simultaneously approximated in the same subspace spanned by atoms in $\mathscr{D}_I$. In what follows, we provide a more detailed description of the algorithm.

Denote by r_i^k the residual of the i th signal at iteration k. Initially, SMP sets the residual signals $r_i^0 = s_i$, $i = 1, \ldots, N$. At each iteration k, the atom in the dictionary $\mathscr{D}$ that best matches all the residual signals is selected. In other words, SMP greedily selects the best matching atom $g_{\gamma_k} \in \mathscr{D}$ by solving a simple optimization problem

$$\gamma_k = \arg\max_{\gamma \in \Gamma} \sum_{i=1}^{N} |\langle r_i^k, g_\gamma \rangle|. \tag{1}$$

The algorithm then updates all the residual signals

$$r_i^{k+1} = r_i^k - \xi_i^k g_{\gamma_k}, \quad i = 1, \ldots, N \tag{2}$$

where $\xi_i^k = \langle r_i^k, g_{\gamma_k} \rangle$. The above step subtracts the contribution of the selected atom g_{γ_k} from all residual signals. The atom selection procedure is repeated iteratively on the updated residuals. The main steps of the SMP algorithm are summarized in Algorithm 1.

Algorithm 1. The SMP algorithm

1: **Input**: $\{s_1, \ldots, s_N\}$: set of signals, K: number of atoms
2: **Output**: Φ_I: selected atoms
3: **Initialization**: $r_i^0 = s_i$, $\Phi_I = [\,]$, $k = 1$
4: **Main iteration**
5: Find index γ_k which solves the optimization problem
$\gamma_k = \arg\max_{\gamma \in \Gamma} \sum_{i=1}^{N} |\langle r_i^k, g_\gamma \rangle|$
6: Augment $\Phi_I = [\Phi_I, g_{\gamma_k}]$
7: Update the residual signals
$\xi_i^k = \langle r_i^k, g_{\gamma_k} \rangle$, $i = 1, \ldots, N$
$r_i^{k+1} = r_i^k - \xi_i^k g_{\gamma_k}$, $i = 1, \ldots, N$.
8: **if** $k = K$ **then**
9: stop
10: **else**
11: increment iteration $k = k + 1$, and go to step (5)
12: **end if**

SMP results in the extraction of K atoms that can be used for representing all signals by linear combination. Each signal can be rewritten as $s_i = \Phi_I c_i$, where Φ_I denotes the matrix whose columns are the atoms in the common sparse support $\mathscr{D}_I \subset \mathscr{D}$. A few iterations are typically sufficient to capture most of the energy of the face signals to be approximated.

4.2. Spherical subspace selection with SMP

We propose to perform dimensionality reduction on the sphere for the classification of 3D faces. We therefore project the 3D point cloud onto the unit sphere S^2, and we then select a subspace that spans these functions on S^2. Since faces are typically star-shaped objects, spherical projection preserves the face geometry information, while reducing the classification complexity by mapping a 3D signal to a 2D spherical signal. Each face, given by a 3D point cloud $\{p_n\} = \{(x_n, y_n, z_n)\}$ is, therefore, represented as a spherical function $s_i(\theta, \varphi) = r$ sampled at points $\{(r_n, \theta_n, \varphi_n)\}$. These points are obtained by transforming Euclidean coordinates from the point cloud to spherical coordinates given by (θ, φ) that represent the elevation and azimuth angles.

Denote as $L^2(S^2)$ the set of all square-integrable functions on S^2 i.e., the set of all functions on S^2 with finite norm. Since we represent 3D faces as functions in $L^2(S^2)$, we can use the SMP to select a subspace of spherical atoms as a dimensionality reduction step. We use a spherical dictionary proposed in [31], where the atoms are created by applying local geometric transforms to a generation function $g(\theta, \varphi)$ defined on the sphere. Local transforms include the atom motion (τ, ν), rotation ψ, and anisotropic scaling by two scales (α, β) in orthogonal directions. Motion and rotation are realized using a rotation in SO(3), which is the group of all rotations about the origin of 3D Euclidean space $\mathbb{R}^3$. Five transform parameters form the atom index $\gamma = (\tau, \nu, \psi, \alpha, \beta) \in \Gamma$, and the redundant dictionary is finally constructed by applying a large set of different transformations γ's to the generating function g. A detailed explanation of the dictionary construction is given in [31]. An example of the generating function is a 2-D Gaussian function in $L^2(S^2)$, given by

$$g(\theta, \varphi) = \exp\left(-\tan^2 \frac{\theta}{2}\right). \tag{3}$$

The function in Eq. (3) represents an isotropic Gaussian function, centered at the North Pole. In Fig. 5 we show a few sample Gaussian atoms that are constructed by applying different local transforms to the generating function in Eq. (3).

Equipped with the spherical dictionary, we can directly apply the SMP presented above in order to find the common support of the spherical faces, where the inner product between two spherical functions $f = f(\theta, \varphi)$ and $g = g(\theta, \varphi)$ is given by

$$\langle f, g \rangle = \int_\theta \int_\varphi f(\theta, \varphi) g(\theta, \varphi) \sin\theta \, d\theta \, d\varphi. \tag{4}$$

In the following, we refer to this special case of SMP for spherical signals using the dictionary defined on the sphere, as *simultaneous spherical matching pursuit* (SSMP). Fig. 6 shows a 3D face in the spherical domain and its approximation with a few Gaussian atoms.

4.3. Recognition on the sphere

The algorithm for recognition of 3D faces on the sphere is finally illustrated in Fig. 7. The first step performs dimensionality reduction, by projecting the spherical signals on the subspace spanned by the selected atoms i.e., span$\{\mathscr{D}_I\}$, as described above. If we denote the set of face signals by $S = [s_1, \ldots, s_N]$, SSMP performs dimensionality reduction by greedily selecting a set of K basis vectors $\mathscr{D}_I = \{g_{\gamma_1}, \ldots, g_{\gamma_K}\}$ from the dictionary $\mathscr{D}$, such that all spherical faces are simultaneously approximated as

$$S \approx \Phi_I \cdot C. \tag{5}$$

The matrix $C \in R^{K \times N}$ holds the coefficient vectors (in its columns) and the matrix $\Phi_I = [g_{\gamma_1}, \ldots, g_{\gamma_K}]$ gathers the basis vectors selected in $\mathscr{D}$.

The matching is performed by comparing the coefficient vectors C, which represent the lower dimensional data samples. The recognition is performed by nearest neighbor classification. We iteratively compute the coefficients c_t of the test face signal s_t on the sub-dictionary $\mathscr{D}_I$. The classification is then performed by computing the l_1 distance between c_t and any coefficient vector c_i corresponding to the training signals

$$d(c_t, c_i) = \sum_{j=1}^{K} |c_t(j) - c_i(j)|. \tag{6}$$

The class of the test signal is finally given by the class of the signal s_i that leads to the smallest distance $d(c_t, c_i)$ between the coefficient vectors. The choice of the l_1 distance metric is mostly empiric as it leads to superior classification performance compared to other metrics.

Although the coefficient vector conveys quite discriminative information about face signals, the class separability of the coefficient vectors in the reduced space could yet be improved by performing

Please cite this article as: R. Sala Llonch, et al., 3D face recognition with sparse spherical representations, Pattern Recognition (2009), doi:10.1016/j.patcog.2009.07.005

Fig. 5. Gaussian atoms on the sphere.

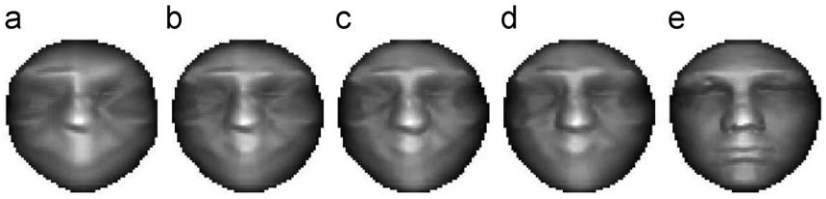

Fig. 6. A spherical face (rightmost panel) and its progressive approximation (panels (a)–(d)) from the common support consisting of Gaussian atoms. (a) 50 atoms; (b) 100 atoms; (c) 150 atoms; (d) 200 atoms; (e) original.

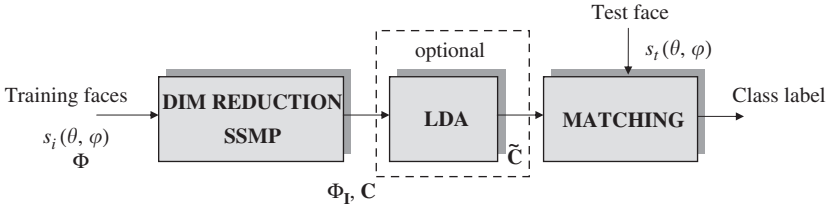

Fig. 7. Block diagram of the recognition process.

an optional linear discriminant analysis (LDA) step before matching. LDA exploits the class labels information of the training samples in order to enhance the discriminant properties of the coefficient vectors. It introduces supervision in the recognition process and permits to build a new set of coefficient vectors $\tilde{C} = CW$ where the weights W are chosen to optimize the ratio of between-class variance and within-class variance for training data [11]. The same classification method can then be used for the coefficients $\tilde{C}$ modified by LDA.

5. Experimental results

5.1. Experimental setup

In this section, we evaluate the performance of the proposed algorithms in both recognition and verification scenarios. For our evaluation, we use the UND (University of Notre Dame) Biometric database [32,33], also known as FRGC v.1.0 database. It contains 953 facial images of 277 subjects, where each subject has between one and eight scans. Each facial scan is provided in the form of a 3D point cloud, along with a corresponding binary matrix of valid points. The number of vertices in a point cloud typically varies between 30,000 and 40,000.

We defined several test configurations for our experimental evaluation. Each configuration is characterized by the number of samples per subject that form the training set. For each configuration T_i, we keep only the subjects from the database that have at least $i+1$ samples, and we use i training samples per class (randomly chosen), while assigning the rest to the test set. The subjects that have only one facial scan cannot be used in the recognition tests. Table 1 summarizes the test configurations and their main characteristics.

SSMP implementation. For the dictionary construction in SSMP-based methods, we have used the 2D Gaussian on the sphere (see Eq. (3)) as the generating function. The atom indexes γ that define the dictionary have to take discrete values in practice. We use here a discretization of the dictionary as in [31], mostly built on empirical choices for atom parameter values. The position parameters, τ and ν are uniformly distributed on the interval $[0, \pi]$, and $[-\pi, \pi)$, respectively, with equal resolution of 128 points. The rotation parameter ψ is uniformly sampled on the interval $[-\pi, \pi)$, with the same resolution as τ and ν. This choice is mostly due to the use of fast computation methods for the correlation on SO(3) in SSMP. In particular, we used the *SpharmonicKit* library,[2] which is part of the *YAW toolbox*.[3] Finally, scaling parameters are distributed in a logarithmic manner, from 1 to half of the resolution of τ and ν, with a granularity of one third of octave. The largest atom covers half of the sphere.

Table 1
Test configurations and their characteristics.

Test configuration	i	Number of subjects	Training set	Test set
T_1	1	200	200	673
T_2	2	166	332	474
T_3	3	121	363	308
T_4	4	86	344	187

An interpolation step is also necessary on the input signal, in order to enable the use of fast computation methods. These require the spherical data to be sampled on an equiangular (θ, φ) grid, defined as

$$G = \left\{(\theta_i, \varphi_j), \theta_i = \frac{(2i+1)\pi}{2N_\theta} \text{ and } \varphi_j = \frac{j2\pi}{N_\varphi}\right\}, \quad (7)$$

[2] http://www.cs.dartmouth.edu/~geelong/sphere/
[3] http://fyma.fyma.ucl.ac.be/projects/yawtb/

Please cite this article as: R. Sala Llonch, et al., 3D face recognition with sparse spherical representations, Pattern Recognition (2009), doi:10.1016/j.patcog.2009.07.005

where $i=0,\ldots,N_\theta-1$ and $j=0,\ldots,N_\varphi-1$. Since 3D face point clouds are projected as scattered data on the sphere, an interpolation step is necessary. For its simplicity we use k-nearest neighbor interpolation, where the value on each spherical grid point (θ_i,φ_j) is computed as an average of its k nearest neighbors. We have used k=4 that resulted in a reasonable compromise between noise reduction and smoothing. The spherical grid resolution is set to $N_\theta=128$, $N_\varphi=128$, which gives the satisfying resolution and smoothness of the spherical faces. Note finally that, for the sake of computational ease, dimensionality reduction with SSMP is performed off-line, using only one training face per subject. The resulting subspace is then used for projecting both training and test samples.

Virtual faces. The size of the training set is important in determining the classification performance. We propose to enrich the training set with *virtual faces* (see e.g., [34] and references therein). These are faces that are artificially generated by slight variations of the original training faces. They are given the corresponding class labels of the training face they originate from, and they are treated as training samples. The use of virtual faces is motivated by two main reasons: (i) they compensate for small registration errors (recall that our registration process is fully automatic and it is expected to contain a few registration errors) and (ii) by augmenting the training set, they may contribute to the performance of sample-based methods (e.g., LDA) that can benefit from large sample sets. Note that the virtual faces do not introduce any new information to the training set, since they are synthetically generated by the original training faces. For computational convenience, we construct them by one-pixel translations in all directions on the sphere. Thus, each training face generates eight virtual faces. Note finally that virtual faces are used only in the SSMP + LDA method.

5.2. Recognition results

We show now experimental results in the recognition scenario. We compare our algorithms with PCA and LDA on depth images that have undergone the same preprocessing step as the data used in the SSMP algorithm. PCA and LDA are well known methods that are typically used as baseline algorithms for comparisons in 3D face recognition. We should note in passing that what we denote by LDA in this paper is essentially PCA followed by LDA.

For the sake of completeness, we include in our comparisons of this section some advanced kernel methods, namely kernel Fisher analysis (KFA) [35] and generalized discriminant analysis (GDA) [36] that are two distinct kernel variants of LDA. In particular, KFA was proposed in order to address the numerical problems of GDA, by including a regularization term in the induced generalized eigenvalue problem in the nonlinear feature space (see [35] for more details). In both GDA and KFA, we use a polynomial kernel $k(x,y)=(x^\top y)^d$, where d is set to its best value, which has been experimentally found to be $d=3$ for both methods. In KFA, the regularization term is set to $\varepsilon=0.001$.

In all LDA-related methods (i.e., SSMP + LDA, PCA + LDA, GDA, KFA), the number of extracted features K cannot be larger than $c-1$. For this reason in our experiments, K is set to the minimum of K and $c-1$, where c is the number of classes (i.e., subjects). Finally, we should mention that virtual faces are used in the SSMP + LDA method in configurations T_1, T_2 and T_3 only, since they correspond to small training sets. Notice finally that LDA, KFA and GDA are not applicable in T_1, since there is only one training sample per class. However, this is not the case with SSMP+LDA thanks to the inclusion of virtual faces.

Fig. 8 shows for each configuration the average classification error rate of all methods, computed over 10 random splits of the data set into training and test sets. The corresponding standard deviations are reported in Table 2, which also includes the performance statistics of the NN classifier with the Euclidean distance between depth images (EUC) and the mean square error between spherical functions (MSE).

Notice the remarkable improvement introduced by the employment of spherical functions for facial representation. This is evident from the fact that the recognition performance of nearest neighbor classification with mean square error (MSE) between spherical signals, outperforms that of Euclidean distances between depth images (EUC). This provides also the main motivation for working on the sphere. Based on this observation, it seems reasonable that our SSMP algorithm outperforms PCA in all configurations.

Observe also that the performance of KFA is superior to that of GDA, as expected. However, they both achieve lower performance than our SSMP + LDA method, which is the best performer. In T_2, SSMP reaches an averaged recognition performance of 76%, while SSMP + LDA reaches 93.2%. The latter goes to 98.5% in T_4 (although no virtual faces have been used for this particular configuration). The maximum recognition rate, 100% is even reached in some of the experiments in T_4.

5.3. Verification results

We consider now the verification scenario, where the test subject claims an identity and the system has to either accept or reject this claim. If the identity is the correct one, then the test subject is called a *client*; otherwise, it is called an *impostor*. In systems that output a confidence score about the test subject, a hard decision (i.e., accept or reject) is typically reached according to a threshold value. We report the verification performances in terms of receiver operating characteristic (ROC) curves, which show the fluctuation of the true positive rate (TPR) versus the false positive rate (FPR) across all values of the threshold. For the computation of the ROC curve we consider every possible pair of subject and claimed identity.

In our experimental setup, we use the dimensions that yields the best performance, which corresponds to 200 atoms in SSMP and 100 dimensions in PCA. In SSMP + LDA we use virtual faces only for configurations T_1 and T_2. Fig. 9 shows the average ROC curves over 10 random experiments for all configurations. Observe again that SSMP consistently outperforms PCA in all configurations and SSMP + LDA stays the best performer.

5.4. Experiments with non-overlapping subject sets

In our experiments so far, faces from all subjects are randomly split into training and test sets. This implies that the same set of subjects are used for both training and testing. However, most person identification systems in practice are highly dynamic, as new subjects are often added in the system. In such a context, one may not want to re-train the system for every new subject, due to complexity and time constraints. Hence, it is rather important for an algorithm to be able to generalize well to unseen subjects.

For this reason, we consider now the case where the subject sets used for training and testing are completely disjoint. In particular, we split the whole UND database of 200 subjects into two groups: (i) subjects that have more than four samples, and (ii) subjects that have from 2 to 4 samples. The first group consists of 531 images of 86 subjects and it is used to train the algorithms. The second group consists of 342 images of 114 subjects and it is used to measure the recognition performances. In this setup, the subjects of the second group are completely unknown to the algorithms.

In the training phase, the subjects of the first group are used to obtain the set of basis vectors. In the testing phase, we split the faces of the second group into gallery and test sets and we project them all in the set of basis vectors that has been obtained in the training phase. For each class (i.e., subject), we assign one face (ran-

Please cite this article as: R. Sala Llonch, et al., 3D face recognition with sparse spherical representations, Pattern Recognition (2009), doi:10.1016/j.patcog.2009.07.005

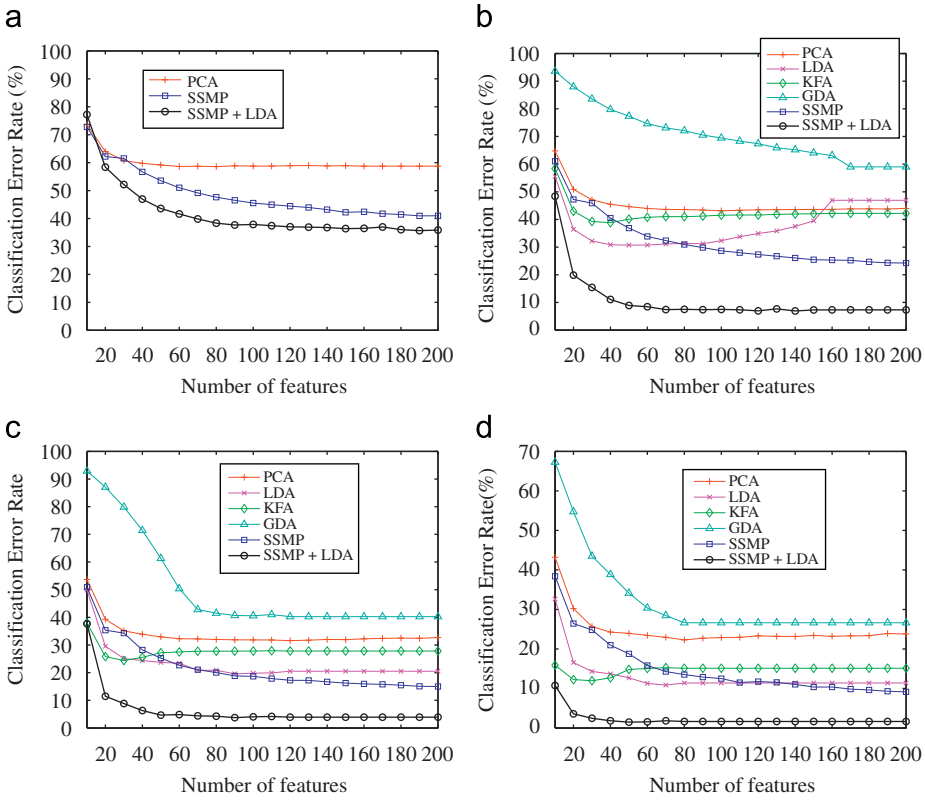

Fig. 8. Rank-1 recognition results: average classification error rate versus the dimension of the subspace. (a) Test configuration T_1; (b) test configuration T_2; (c) test configuration T_3; (d) Test configuration T_4.

Table 2
Summary of recognition performance (in terms of classification error rate (%)) statistics (see also Fig. 8).

K	EUC	MSE	PCA	LDA	KFA	GDA	SSMP	SSMP + LDA
Test configuration T_1								
50	63.2 ± 4.2	46.5 ± 2.7	59.2 ± 1.4	–	–	–	53.6 ± 1.3	43.6 ± 1.7
100			58.8 ± 1.3	–	–	–	45.6 ± 1.6	37.9 ± 1.4
150			58.9 ± 1.2	–	–	–	42.2 ± 1.5	36.4 ± 1.4
200			58.8 ± 1.1	–	–	–	41 ± 1.3	35.9 ± 1.5
Test configuration T_2								
50	46.2 ± 3.7	33.5 ± 2.4	44.6 ± 2.0	30.7 ± 1.9	40.1 ± 5.4	77.3 ± 6.3	36.8 ± 1.6	8.9 ± 1.4
100			43.1 ± 1.8	32.3 ± 2.5	41.5 ± 4.6	69.5 ± 8.5	28.7 ± 1.5	7.5 ± 1.7
150			43.6 ± 2.0	39.5 ± 4.0	42.1 ± 4.3	64.1 ± 8.0	25.4 ± 1.4	7.3 ± 1.5
200			44 ± 2.1	47 ± 2.4	42.2 ± 4.3	59 ± 7.2	24.2 ± 1.5	6.8 ± 1.3
Test configuration T_3								
50	36.4 ± 3.8	19.1 ± 3.1	33 ± 1.8	23.7 ± 1.9	27.2 ± 6.1	61.3 ± 9.5	25.3 ± 1.1	4.6 ± 0.51
100			31.9 ± 2.1	19.8 ± 1.9	27.8 ± 5.7	40.6 ± 8.7	18.7 ± 1.5	4 ± 0.68
150			32 ± 2.3	20.5 ± 2.0	27.8 ± 5.9	40.3 ± 10.3	16.2 ± 1.7	3.9 ± 0.55
200			32.4 ± 2.2	–	–	–	15 ± 1.7	–
Test configuration T_4								
50	27.8 ± 3.4	13.9 ± 3.7	24 ± 2.7	12.7 ± 1.7	14.8 ± 5.3	34.1 ± 3.8	18.8 ± 2.1	1.5 ± 0.91
100			22.3 ± 3.1	11.4 ± 1.4	15.1 ± 5.1	26.6 ± 4.1	12.5 ± 1.6	1.7 ± 1.4
150			23.4 ± 2.9	–	–	–	10.8 ± 1.5	–
200			23.8 ± 3.2	–	–	–	9.1 ± 1.7	–

Results are reported in *mean* ± *std* format.

domly chosen) to the gallery set, and the rest of faces are assigned to the test set. Recognition is done in the reduced space with NN classification.

Fig. 10 shows the average performances over 10 random experiments of all methods using the aforementioned experimental setup. The corresponding standard deviations are shown in Table 3. The recognition rate of the proposed SSMP + LDA solution reaches 84.6%. Thus, the proposed approach seems to generalize well to unseen faces and it certainly represents a promising solution for person identification in highly dynamic environments where new subjects are continuously introduced in the system.

5.5. Discussion

It is worth noting that supervised versions of SSMP could be also used [28]. The idea would be then to select the atoms from

Please cite this article as: R. Sala Llonch, et al., 3D face recognition with sparse spherical representations, Pattern Recognition (2009), doi:10.1016/j.patcog.2009.07.005

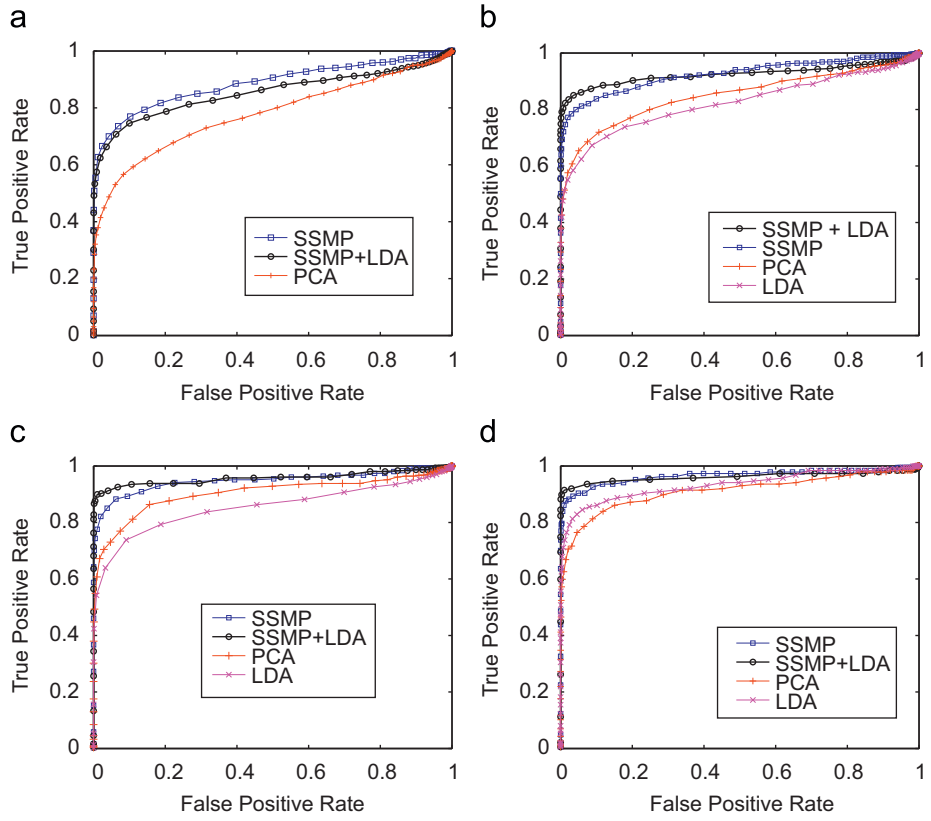

Fig. 9. Verification performance in terms of ROC curves. (a) Test configuration T_1; (b) test configuration T_2; (c) test configuration T_3; (d) test configuration T_4.

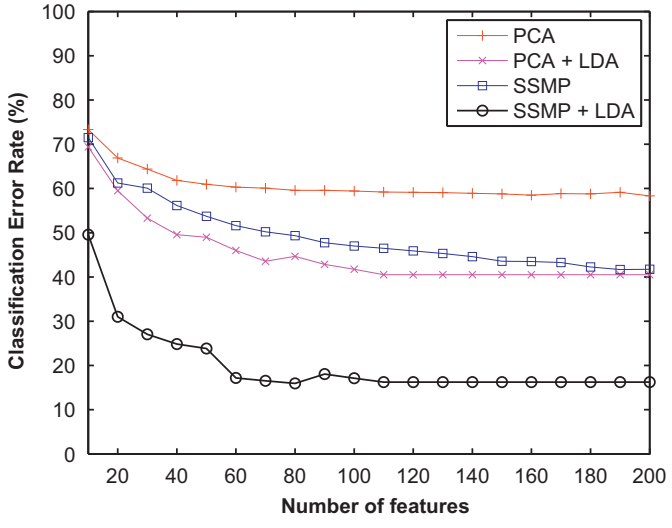

Fig. 10. Rank-1 recognition rate versus number of features used in the non-overlapping recognition experiment.

Table 3
Summary of recognition performance (in terms of classification error rate (%)) statistics in the non-overlapping experiment (see also Fig. 10).

K	PCA	LDA	SSMP	SSMP + LDA
50	61.4 ± 1.7	51.7 ± 1.2	55.5 ± 2.6	25.2 ± 1.2
100	59.5 ± 1.4	43 ± 1.9	48.7 ± 2.4	19.5 ± 1.1
150	59 ± 1.1	40.4 ± 1.0	45.6 ± 2.0	15.4 ± 1.5
200	58.9 ± 1.1	–	44.8 ± 2.1	–

Results are reported in $mean \pm std$ format.

the dictionary according to discriminative criteria. However, in the proposed scheme the supervision information is already taken into account in the LDA postprocessing step, and prior experience has shown that this suffices, when predefined dictionaries are used.

Note also that the importance of each region of the face in terms of recognition performance is certainly not uniform [18]. Although the selection of such regions is typically performed manually and may be sensitive to the testing conditions, one possible approach to take advantage of this observation could be to group the features selected by SSMP into regions by clustering on the sphere, do a classification per region and then fuse the results (e.g., by majority voting). Such an approach however requires a sufficient number of atoms in each area, and the performance of such a region-based classifier has not been convincing.

We should also mention that the proposed dimensionality reduction scheme is generic and simple extensions could be proposed to make the classification more sensitive to some specific areas. For example, the SSMP scheme can easily be adapted to give priorities to regions of high interest such as the nose or the eyes. Such a prioritization can be achieved by giving proper weights to atoms located in different areas, in order to force the dimensionality reduction step to select features in areas that are expected to be more discriminative. This however goes along the lines of supervised versions of SSMP mentioned above with the main difference that discriminative capability in this case is mostly defined in a region-based way.

Please cite this article as: R. Sala Llonch, et al., 3D face recognition with sparse spherical representations, Pattern Recognition (2009), doi:10.1016/j.patcog.2009.07.005

Note finally that one may conceptually relate the framework of sparse representations with the bag-of-features approach (see e.g., [39]) that has been extensively used in the context of content-based image classification. The main idea is to sample a set of local image patches, to evaluate a visual descriptor on each patch and then to convert the distribution of descriptors into a histogram of votes (e.g., by vector quantization against a predefined codebook of centroids). The resulting histogram of votes can be viewed as a global descriptor of the image and subsequently used as a feature vector in a standard classifier. However, such an approach is mostly appropriate for problems with a large within-class variability, such as content-based image classification. On the contrary, the variability of 3D faces is more restricted and this allows for the development of more discriminant methods like ours, which are able to exploit the special structure of the faces.

6. Conclusions

We have proposed a methodology for 3D face recognition based on spherical sparse representations. First, we have introduced a fully automatic process for extraction, preprocessing and registration of facial information in 3D point clouds. Next, we have proposed to convert faces from point clouds to spherical signals. Sparse spherical representation of faces permits effective dimensionality reduction through simultaneous sparse approximations. The dimensionality reduction step preserves the geometry information, which in turn leads to high performance matching in the reduced space. We provide ample experimental evidence that indicates the advantages of the proposed approach over state-of-the-art methods working on depth images.

Acknowledgments

The authors would like to thank Prof. Patrick Flynn for sharing with us the UND Biometrics database. The authors would like also to thank the Associate Editor and the anonymous reviewers for their helpful comments.

References

[1] A. Jain, L. Hong, S. Pankati, Biometric identification, Communications of the ACM 43 (2) (2000) 90–98.
[2] K.W. Bowyer, K. Chang, P. Flynn, A survey of approaches and challenges in 3D and multi-modal 3D + 2D face recognition, Computer Vision and Image Understanding 101 (1) (2006) 1–15.
[3] L. Akarun, B. Gokberk, A.A. Salah, 3D face recognition for biometric applications, in: Proceedings of the European Signal Processing Conference, Antalaya, 2005.
[4] P. Besl, N. McKay, A method for registration of 3D shapes, IEEE Transactions on Pattern Analysis and Machine Intelligence 14 (1992) 239–256.
[5] I.T. Jolliffe, Principal Component Analysis, Springer, New York, 1986.
[6] B. Gökberk, M.O. Irfanoğlu, L. Akarun, 3D shape-based face representation and feature extraction for face recognition, Image and Vision Computing 24 (2006) 857–869.
[7] B. Achermann, H. Bunke, Classifying range images of human faces with Hausdorff distance, in: 15th International Conference on Pattern Recognition, 2000, pp. 809–813.
[8] 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, no. 2, 2004, pp. 884–889.
[9] F.R. Al-Osaimi, M. Bennamoun, A. Mian, Integration of local and global geometrical cues for 3D face recognition, Pattern Recognition 41 (3) (2008) 1030–1040.
[10] M.A. Turk, A.P. Pentland, Face recognition using eigenfaces, Vision and Modeling Group, The Media Laboratory Massachusetts Institute of Technology, 1996.
[11] P.N. Belhumeur, J.P. Hespanha, D.J. Kriegman, Eigenfaces vs. Fisherfaces: recognition using class specific linear projection, IEEE Transactions on Pattern Analysis and Machine Intelligence 19 (7) (1997).
[12] D.D. Lee, H.S. Seung, Algorithms for non-negative matrix factorization, Advances in Neural Information Processing Systems 13 (2001) 556–562.
[13] P. Paatero, U. Tapper, Positive matrix factorization: a non-negative factor model with optimal utilization of error estimates of data values, Environmetrics 5 (1994) 11–126.
[14] P.O. Hoyer, Non-negative matrix factorization with sparseness constraints, Journal of Machine Learning Research 5 (2004) 1457–1469.
[15] P.O. Hoyer, Non-negative sparse coding, in: IEEE Workshop on Neural Networks for Signal Processing, 2002, pp. 557–565.
[16] B.J. Shastri, M.D. Levine, Face recognition using localized features based on non-negative sparse coding, Machine Vision and Applications 18 (2007) 107–122.
[17] S. Jahanbin, H. Choi, A.C. Bovik, K.R. Castleman, Three dimensional face recognition using wavelet decomposition of range images, in: International Conference on Image Processing, 2007, pp. 145–148.
[18] K. Wong, W. Lin, Y. Hu, N. Boston, Optimal linear combination of facial regions for improving identification performance, IEEE Transactions on Systems, Man, and Cybernetics—Part B: Cybernetics 37 (5) (2007).
[19] T.C. Faltemier, K.W. Bowyer, P.J. Flynn, A region ensemble for 3-D face recognition, IEEE Transactions on Information Forensics and Security 3 (1) (2008) 62–73.
[20] S. Gupta, J.K. Aggarwal, M.K. Markey, A.C. Bovik, 3D face recognition founded on the structural diversity of human faces, in: IEEE Conference on Computer Vision and Pattern Recognition, 2007.
[21] G. Gordon, Face recognition based on depth and curvature features, in: SPIE Proceedings on Geometric Methods in Computer Vision, vol. 1570, 1991, pp. 234–247.
[22] C. Samir, A. Srivastava, M. Daoudi, Three-dimensional face recognition using shapes of facial curves, IEEE Transactions on Pattern Analysis and Machine Intelligence 28 (11) (2006).
[23] H. Wang, H. Wei, Y. Wang, Face representation under different illumination conditions, in: IEEE International Conference on Multimedia and Expo, 2003, pp. 285–288.
[24] R. Ramamoorthi, Analytic PCA construction for theoretical analysis of lighting variability in images of a Lambertian object, IEEE Transactions on Pattern Analysis and Machine Intelligence 24 (10) (2002).
[25] Z. Yue, W. Zhao, R. Chellappa, Pose-encoded spherical harmonics for face recognition and synthesis using a single image, EURASIP Journal on Advances in Signal Processing 2008 (2008) 1–18.
[27] M. Hebert, K. Ikeuchi, H. Delingette, A spherical representation for recognition of free-form surfaces, IEEE Transactions on Pattern Analysis and Machine Intelligence 17 (7) (1995).
[28] E. Kokioppoulou, P. Frossard, Semantic coding by supervised dimensionality reduction, IEEE Transactions on Multimedia 10 (5) (2008) 806–818.
[29] S.G. Mallat, Z. Zhang, Matching pursuit with time-frequency dictionaries, IEEE Transactions on Signal Processing 41 (12) (1993) 3397–3415.
[30] J. Tropp, A. Gilbert, M. Strauss, Algorithms for simultaneous sparse approximation. Part I: greedy pursuit, Signal Processing 46 (2006) 572–588 (Special issue "Sparse approximations in signal and image processing").
[31] I. Tosic, P. Frossard, P. Vandergheynst, Progressive coding of 3D objects based on overcomplete decompositions, IEEE Transactions on Circuits and Systems for Video Technology 16 (11) (2006) 1338–1349.
[32] P.J. Flynn, K.W. Bowyer, P.J. Phillips, Assessment of time dependency in face recognition: an initial study, Audio and Video-Based Biometric Person Authentication, 2003, pp. 44–51.
[33] X. Chen, P.J. Flynn, K.W. Bowyer, Visible-light and infrared face recognition, in: ACM Workshop on Multimodal User Authentication, 2003, pp. 48–55.
[34] D. DeCoste, M.C. Burl, Distortion-invariant recognition via jittered queries, in: IEEE International Conference on Computer Vision and Pattern Recognition, 2000.
[35] C. Liu, Capitalize on dimensionality increasing techniques for improving face recognition grand challenge performance, IEEE Transactions on Pattern Analysis and Machine Intelligence 28 (5) (2006).
[36] G. Baudat, F. Anovar, Generalized discriminant analysis using a kernel approach, Neural Computation 12 (10) (2000) 2385–2404.
[37] B. Gökberk, H. Dutağaci, A. Ulas, L. Akarun, B. Sankur, Representation plurality and fusion for 3D face recognition, IEEE Transactions on Systems, Man, and Cybernetics, Part B 38 (2008) 155–173.
[38] G. Davis, S. Mallat, M. Avellaneda, Greedy adaptive approximation, Journal on Constructive Approximation 13 (1) (1997) 57–98.
[39] D. Nistér, H. Stewénius, Scalable recognition with a vocabulary tree, in: IEEE Conference on Computer Vision and Pattern Recognition, 2006.

About the Author—ROSER SALA LLONCH was born in Barcelona, Spain, on July 11th, 1984. She received her M.Sc. degree in Telecommunications Engineering in December 2007 from the Technical University of Catalonia (UPC). From March 2007 to April 2008 she was working within an exchange program in the LTS4 Lab of the EE Institute, at the Swiss Federal Institute of Technology (EPFL), Lausanne, Switzerland under the supervision of Prof. Pascal Frossard, where she first did her M.Sc. thesis and then worked as an intern student. She is currently working towards her Ph.D. degree in Biomedical Science in the University of Barcelona (UB). Her research interests include 2D/3D image processing, medical imaging and computational neuroscience.

Please cite this article as: R. Sala Llonch, et al., 3D face recognition with sparse spherical representations, Pattern Recognition (2009), doi:10.1016/j.patcog.2009.07.005

About the Author—EFFROSYNI KOKIOPOULOU received her Diploma in Engineering in June 2002, from the Computer Engineering and Informatics Department of the University of Patras, Greece. In June 2005, she received a M.Sc. degree in Computer Science from the Computer Science and Engineering Department of the University of Minnesota, USA, under the supervision of Prof. Yousef Saad. In September 2005, she joined the LTS4 Lab of the EE Institute, in the Swiss Federal Institute of Technology (EPFL), Lausanne, Switzerland. She completed her PhD studies in December 2008 under the supervision of Prof. Pascal Frossard. Currently, she is a postdoctoral research associate with the Seminar for Applied Mathematics, ETH, Zurich, Switzerland, working with Prof. Daniel Kressner. Her research interests include pattern recognition, computer vision and numerical linear algebra.

About the Author—IVANA TOSIC received the Dipl.Ing. degree in Telecommunications from the University of Nis, Serbia, in 2003. She is currently pursuing the Ph.D. degree at the Signal Processing Laboratory, Swiss Federal Institute of Technology (EPFL), Lausanne, Switzerland. In 2004, she graduated from the Doctoral School in Information and Communication Sciences at EPFL.
She joined the Signal Processing Laboratory at EPFL in 2004, as a Research and Teaching Assistant. Her research interests include representation and coding of visual information, distributed source coding, nonuniform sampling on the sphere, sparse approximations and dictionary learning.

About the Author—PASCAL FROSSARD (S96,M01,SM04) received the M.S. and Ph.D. degrees, both in Electrical Engineering, from the Swiss Federal Institute of Technology (EPFL), Lausanne, Switzerland, in 1997 and 2000, respectively. Between 2001 and 2003, he was a member of the research staff at the IBM T. J. Watson Research Center, Yorktown Heights, NY, where he worked on media coding and streaming technologies. Since 2003, he has been an assistant professor at EPFL, where he heads the Signal Processing Laboratory (LTS4). His research interests include image representation and coding, nonlinear representations, visual information analysis, joint source and channel coding, multimedia communications, and multimedia content distribution. Dr. Frossard has been the General Chair of IEEE ICME 2002 and Packet Video 2007, and a member of the organizing or technical program committees of numerous conferences. He has been an Associate Editor of the IEEE Transactions on Multimedia (2004–) and of the IEEE Transactions on Circuits and Systems for Video Technology (2006–). He is an elected member of the IEEE Image and Multidimensional Signal Processing Technical Committee (2007–), the IEEE Visual Signal Processing and Communications Technical Committee (2006–), and the IEEE Multimedia Systems and Applications Technical Committee (2005–). He has served as Vice-Chair of the IEEE Multimedia Communications Technical Committee (2004–2006) and as a member of the IEEE Multimedia Signal Processing Technical Committee (2004–2007). He received the Swiss NSF Professorship Award in 2003, and the IBM Faculty Award in 2005.

Please cite this article as: R. Sala Llonch, et al., 3D face recognition with sparse spherical representations, Pattern Recognition (2009), doi:10.1016/j.patcog.2009.07.005