Multimed Tools Appl (2015) 74:11281–11295
DOI 10.1007/s11042-014-2228-3

Regularized directional feature learning for face recognition

Mohamed Anouar Borgi · Maher El'Arbi · Demetrio Labate · Chokri Ben Amar

Received: 18 February 2014 / Revised: 24 June 2014 / Accepted: 7 August 2014 /
Published online: 3 September 2014

Abstract This paper presents an improved approach to face recognition, called Regularized Shearlet Network (RSN), which takes advantage of the sparse representation properties of shearlets in biometric applications. One of the novelties of our approach is that directional and anisotropic geometric features are efficiently extracted and used for the recognition step. In addition, our approach is augmented by regularization theory (RSN) in order to control the trade-off between the fidelity to the data (gallery) and the smoothness of the solution (probe). In this work, we address the challenging problem of the single training sample per subject (STSS). We compare our new algorithm against different state-of-the-art methods.

Keywords Shearlet · Regularized shearlets network · Face recognition

1 Introduction

Face recognition (FR) is a classical problem in computer vision and pattern recognition and many methods, such as Eigenfaces [42], Fisherfaces [3], SVM [19] and Metaface [45] have been proposed in the past two decades. One of the standard statistical methods for

M. A. Borgi (✉) · M. El'Arbi · C. Ben Amar
Research Group on Intelligent Machines, University of Sfax, ENIS, BP 1173, Sfax 3038 Tunisia
e-mail: anoir.borgi@ieee.org

M. El'Arbi
e-mail: maher.elarbi@gmail.com

C. Ben Amar
e-mail: chokri.benamar@ieee.org

D. Labate
Department of Mathematics, University of Houston, Houston, TX 77204, USA
e-mail: dlabate@math.uh.edu

FR is subset selection (L_0 regularization) [51], which consists in computing the following estimator

$$\hat{w}_{L_0} = argmin_w \| Xw - y\|_2^2 \; subject\ to\ \|w\|_0 \leq \delta \tag{1}$$

where δ is a tuning parameter, **y** is a normalized test face, **X** is a matrix representing a gallery of faces and w is the weight which control the trade-off between the fidelity to the gallery faces and the smoothness of the test face.

This statistical approach has received renewed interest in recent years due to the notion of sparse representations, which offers the possibility of recasting the face recognition problem. For example, the recently proposed Sparse Representation Classification (SRC) scheme [44] casts the recognition problem as one of classifying among multiple linear regression and uses sparse representations computed via L_1 minimization for efficient feature extraction. By coding a query image as a sparse linear combination of all the training samples, SRC classifies the query image by evaluating which class could result in the minimal reconstruction error. However, it has been indicated in [52] that SRC actually owes its success to its utilization of collaborative representation on the query image rather than the l1-norm sparsity constraint on coding coefficient. Besides SRC, another powerful method recently proposed is the Regularized Robust Coding (RRC) approach [46, 48], which could robustly regress a given signal with regularized regression coefficients. By assuming that the coding residual and the coding coefficient are respectively independent and identically distributed, the RRC seeks for a maximum a posterior solution of the coding problem. An iteratively re-weighted regularized robust coding algorithm was proposed to solve the RRC model efficiently.

In this paper, we propose a method called RSN, which combines sparsity and regularization theory. Sparsity, in particular, will be based on the use of the shearlet representation, one of the systems introduced during the last decade to go beyond classical wavelets. Indeed, despite their extensive use in image processing, traditional wavelets are known to have a limited ability to deal with directional information. By contrast, shearlets provide a simple multiscale framework which is especially effective to capture directional and anisotropic features with high efficiency, are optimally sparse for the representation of 2D/3D images and have fast numerical implementations [24]. As part of this work, we will assess the performance of the Regularized Shearlets Network approach for FR and compare it against competitive algorithms. The main contributions of this paper include:

- A new feature-extraction approach for efficient FR based on directional features.
- A regularization optimization, using those features, based on Lasso method.

The rest of this paper is organized as follows. In Section 2, a description of the related works regarding regularized sparse coding is presented. In section 3, we briefly describe the necessary background on shearlets. Section 4 presents the proposed Regularized Shearlet Network algorithm. In Section 5, we present several numerical experiments to demonstrate the efficacy of the proposed algorithm and compare it against competing algorithms. Finally, Section 6 concludes this paper.

2 Related works

The current trend in FR emphasizes biometrics which can be collected on the move, so that there is significant interest in more sophisticated and robust methods to go beyond current state-of-the-art FR methods. One of the most successful approaches to template-based face representation and recognition is based on Principal Component Analysis (PCA). However, PCA approximates texture only, while the geometrical information of the face is not properly captured. In addition to PCA, many other linear projection methods have been considered in face recognition applications. The LDA (Linear Discriminant Analysis) has been proposed in [31] as an alternative to PCA. This method provides discrimination among the classes, while the PCA deals with the input data in their entirety without paying much attention for the underlying structure.

The Discrete Cosine Transform (DCT) is also one of the most popular linear projection techniques for feature extraction like principal components analysis (PCA) and linear discriminant analysis (LDA) but recently a Discrimination power analysis (DPA) has been proposed as a statistical analysis combining discrimination concept with DCTCs properties. Unfortunately there is not a uniform and effective criterion to optimize the shape and size of premasking window on which the effect of DPA excessively relies. Proper premasking is an auxiliary process to select the feature coefficients that have more discrimination power (DP). Dynamic weighted DPA (DWDPA) is proposed in [27, 28] to enhance the DP of the selected DCTCs without premasking window, in other words, it does not need to optimize the shape and size of pre-masking window. The experimental results on ORL, Yale and PolyU databases show that DWDPA outperforms DPA obviously.

Moreover, to deal with the challenges in practical FR system, active shape model and active appearance model [25] were developed for face alignment; LBP [1] and its variants were used to deal with illumination changes; and Eigenimages and probabilistic local approach [33] were proposed for FR with occlusion.

The recognition of a test face image is usually accomplished by classifying the features extracted from this image. The most popular classifier for FR may be the nearest neighbor (NN) classifier due to its simplicity and efficiency. In order to overcome NN's limitation that only one training sample is used to represent the test face image, the authors in [30] proposed the nearest feature line (NFL) classifier, which uses two training samples for each class to represent the test face. In [30] contributors proposed the nearest feature plane (NSP) classifier, which uses three samples to represent the test image. Later on, classifiers using more training samples for face representation were proposed, such as the local subspace classifier (LSC) [23] and the nearest subspace (NS) classifiers [11, 26, 29, 35], which represent the query sample by all the training samples of each class. Though NFL, NSP, LSC and NS achieve better performance than NN, all these methods with holistic face features are not robust to face occlusion.

3 The shearlet transform

The shearlet transform, introduced by one of the authors and their collaborator in [24], is a genuinely multidimensional version of the traditional wavelet transform, and is especially designed to represent data containing anisotropic and directional features with high efficiency. As a result, this approach provides optimally sparse approximations for images with edges, outperforming traditional wavelets. Thanks to their properties, shearlets have been successfully employed in a number of image processing application including denoising, edge detection and feature extraction [12, 14, 49]. Formally, the Continuous Shearlet Transform [22] is defined as the mapping:

$$SH_{\psi}(\alpha, s, t) = \langle f, \Psi_{\alpha,s,t} \rangle, \alpha > 0, s \in R, t \in R^2 \quad (2)$$

where $\Psi_{\alpha,s,t}(x) = \mid det M_{\alpha,s} \mid^{-\frac{1}{2}} \Psi(M_{\alpha,s}^{-1}(x-t))$ and $M_{\alpha,s} = \begin{pmatrix} \alpha & s \\ 0 & \sqrt{\alpha} \end{pmatrix}$. Observe that each matrix $M_{\alpha,s}$ can be factorized as $B_s A_\alpha$, where $B_s = \begin{pmatrix} 1 & -s \\ 0 & 1 \end{pmatrix}$ is a shear matrix and $A_\alpha = \begin{pmatrix} \alpha & 0 \\ 0 & \sqrt{\alpha} \end{pmatrix}$ is an anisotropic dilation matrix.

Thus, the shearlet transform is a function of three variables: the scale α, the shear **s** and the translation **t**. One of the main properties of the Continuous Shearlet Transform is its ability to detect very precisely the geometry of the singularities of a 2-dimensional function **f**. This property is going far beyond the properties of the wavelet transform and explains why shearlets are so effective at capturing edges and other directional information in images.

By sampling the Continuous Shearlet Transform $SH_\psi(\alpha, s, t)$ on an appropriate discrete set we obtain a discrete transform. Specifically, $M_{\alpha,s}$ is "discretized" as $M_{j,l} = B_l A^j$, where **B**= $\begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$, **A**= $\begin{pmatrix} 4 & 0 \\ 0 & 2 \end{pmatrix}$ are the shear matrix and the anisotropic dilation matrix, respectively. Hence, the discrete shearlets are the functions of the form:

$$\psi j, l, k(x) = 2^{\frac{3j}{2}} \psi(B_l A^j x - k),\ j \geq 0, -2^j \leq l \leq 2^j - 1, k \in Z^2 \tag{3}$$

Frequency support of shearlets is a pair of trapezoids, symmetric with respect to the origin. This is illustrated in Fig. 1.

By choosing the generator function appropriately, the discrete shearlets form a tight frame of well-localized waveforms defined at various scales, orientations and locations.

Shearlets are a variant of wavelets with composite dilatation originally introduced in [16, 17] offering a particularly general framework which allows one to derive a variety of powerful data representation schemes; many constructions are obtained within this framework and recently the authors in [6] introduced Gabor shearlets, a variant of shearlet systems, which combine elements from Gabor and wavelet frames in their construction; another interesting construction is hyperbolic shearlets [13].

In the spatial domain, a 2D Gabor filter is a Gaussian kernel function modulated by a sinusoidal plane wave, as illustrated in Fig. 2.

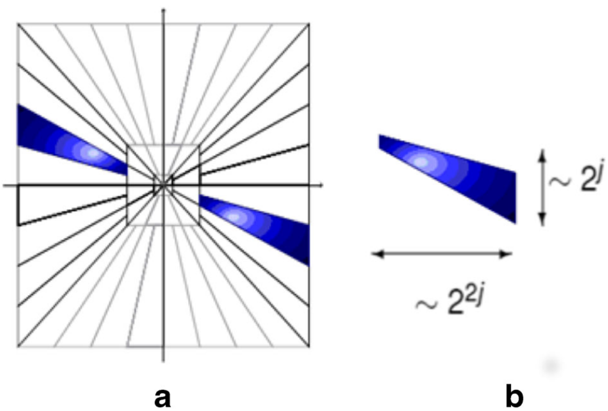

Fig. 1 **a** Shearlet tiling of the frequency plane. **b** Frequency support

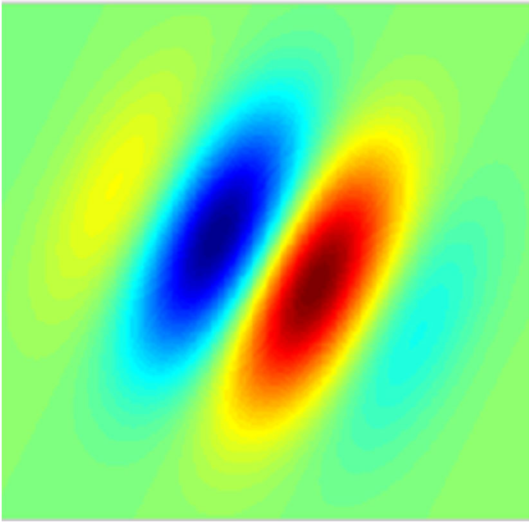

Fig. 2 A two-dimensional Gabor filter

4 The proposed approach

The proposed approach, RSN, for FR is defined as a cascade of a feature extraction module followed by a recognition (recognition or verification) module. We will perform this schema by the use of regularization theory to control the solution (Probe or Test) and its closeness to the data (Gallery), where the extraction of directional features is controlled by the Shearlet Network (SN) as shown in Fig. 3.

Analytically, the FR problem can be casted as a regression problem of approximating a multivariate function from sparse data. This is an ill-posed problem and a classical way to solve it is though regularization theory [4, 5, 41]. In practice, rather than looking for an exact solution, one can only find an approximate one. The most popular approximation method is the L_1 regularization method which is often referred to as Lasso [40] and is given by:

$$\hat{w}_{L_1} = argmin_w \left[\frac{1}{n} \| Xw - y\|_2^2 + \lambda \|w\|_1 \right] \tag{4}$$

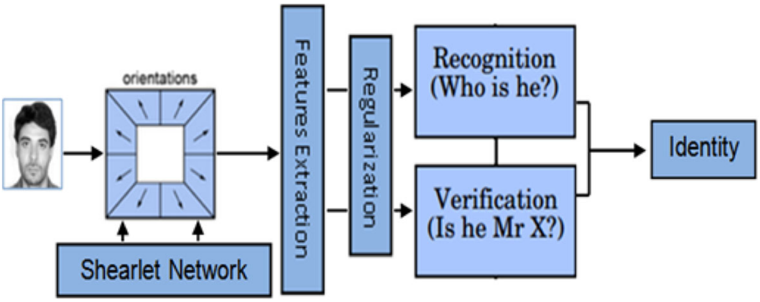

Fig. 3 Augmented face recognition schema

where $\lambda > 0$ is an appropriately chosen regularization parameter, **y** is a normalized test face and **X** is an **n*d** matrix representing a gallery of faces and w is the weight which will be explained in the paragraph RSN algorithm. The global optimum of Eq. 4 can be easily computed using standard convex programming techniques. It is known that, in practice, L_1 regularization often leads to sparse solutions, although they are often sub-optimal. The theoretical performance of this method has been analyzed recently [7, 8, 50].

4.1 SN for modeling and features extraction

Our proposed RSN approach is initialized by training an SN [9] to models the faces. The Gallery faces are approximated by a shearlet network to produce a compact biometric signature. One main feature of this approach is that this signature, constituted by the shearlets and their weights, will be used to match a Probe with all faces in the Gallery. The test (Probe) face is projected on the shearlet network of the Gallery face and new weights specific to this face are produced. The family of shearlets remains then unchanged (this is the Gallery face) as illustrated in Fig. 4.

Recall that the shearlets form a tight frame, meaning that, for any image in the space of square integrable functions we have the reproducing formula:

$$f = \sum_{j,k \in R, k \in R^2} \langle f, \psi_{j,l,k} \rangle \psi_{j,l,k} \tag{5}$$

We will use this formula to define the Shearlet Network approach, similar to the wavelet network [2], as a combination of the RBF neural network and the shearlet decomposition. In the optimization stage, the calculation of the weights connection in every stage is obtained by projecting the signal to be analyzed on a family of shearlets. We need the dual family of the shearlets forming our shearlet network, which is calculated by the formula:

$$\tilde{\psi}^t_{j,l,k} = \sum_{m=1}^{N} (\Psi_{i,m})^{-1} \psi^m_{j,l,k} \; with \; \Psi_{i,m} = \langle \psi^i_{j,l,k}, \psi^m_{j,l,k} \rangle \tag{6}$$

In our approach, the mother shearlet used to construct the family $\psi_{j,l,k}$ is the second derived of the Beta function [18]. Note that the number of shearlets may be chosen by the

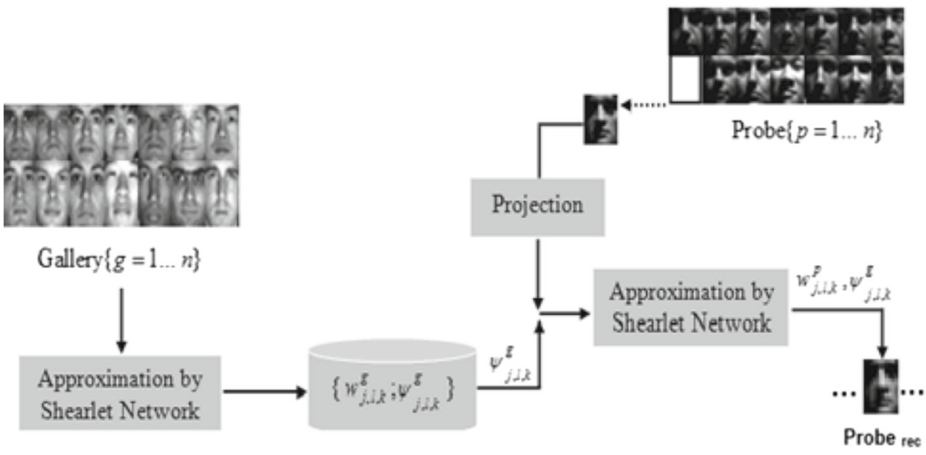

Fig. 4 Overview of SN architecture

user. We construct a library of shearlets with different scales, shears and the translations which form a shearlet frame and finally calculate the weights by direct projection of the image on the dual shearlet; details are reported in the following algorithm:

Algorithm 1: Training SN

Input: image f
Output: reconstructed image f_{rec}

1. Select a shearlet $\psi_{j,l,k}$ as activation function of the shearlet network:
   - Choose the mother shearlet.
   - Build a library formed by the shearlets which form a shearlet frame.
   - Set as a stop learning condition based on the difference of input and the output network and iterate the following steps:
2. If not tight then Calculate the dual basis $\tilde{\psi}^t_{j,l,k}$ formed by the shearlets of the network and the new selected shearlet according (6) else $\psi_{j,l,k} = \tilde{\psi}^t_{j,l,k}$.
3. Calculate the weights by direct projection of the image on the dual shearlet $w_i = < f, \tilde{\psi}^t_{j,l,k} >$.
4. Calculate the output of the network f_{rec}.
5. If the number of shearlets is reached the learning stops, otherwise another shearlet is selected and we return to 2.

4.2 RSN algorithm

The initial value of the weight in Eq. 4 w_{init} is chosen using the logistic function [50]. In fact these functions, like exponential functions, grow quickly at first, but eventually grow more slowly and then level off. The formula for the logistic function is as follow:

$$f(x) = \frac{c}{1 + aexp^{-bx}} \tag{7}$$

This involves three positive parameters a, b and c. One best choice of w_{init} is:

$$w_{init} = \frac{1}{1 + \frac{1}{exp(-\mu e^2_{init} + \mu\delta)}} \tag{8}$$

where μ and δ are positive scalars.

In our experiments we choose μ as:

$$\mu = \frac{0.6}{\delta} \tag{9}$$

As in [46] we set δ as:

$$\delta = \varphi(e_{init})_l \tag{10}$$

where $\varphi(e_{init})_k$ is the k^{th} largest element of the set $e^2_{init}(j) j = 1, 2, ..., n$ and e_{init} is the initial residual given by:

$$e_{init} = (\boldsymbol{y} - mean(\boldsymbol{X}))^2 \tag{11}$$

X is the aligned gallery faces (an n*d matrix) and **y** a normalized test face (an n*1 matrix). Note that, after optimization, we can update the residual e using the following formula:

$$e = (\boldsymbol{X} w_i - \boldsymbol{y})^2 \tag{12}$$

Where w_i is the solution given by the Lasso optimization.

How to classify? The query sample **y** is classified to the class which gives the minimal:

$$error(y) = \|w^{\frac{1}{2}}(\boldsymbol{y} - X_k w_k)\|_2^2 \tag{13}$$

Below we present the algorithm of RSN, where **X** represents the reconstructed gallery faces after extraction of the features by training SN, **y** is the reconstruct test face with the features extracted after projection of the real test face on the frame of shearlets produced by the gallery faces.

Algorithm 2: RSN

Input:

- y: normalized test face f:y=f / norm (f,2)
- X: aligned gallery faces X=X / $\sqrt{\sum x * x}$
- Iter: max of iteration;w_{thre} ; $\lambda \in]0..1[$

Output:w,identity(y)

1. Choose w_{init} (refer to [40])
2. Diagonalizable X ; $X^t = X' * X$; $y^t = X' * y$;
3. For j = 1 .. Iter
   - Calculate $\hat{w} = argmin_{w \in Rp}[\frac{1}{n}\| \; X_w - y\|_2^2 + \lambda\|w\|_1]$ Use Lasso [41]: $w_i = lasso(X, y, X^t, y^t, w_{init}, w_{thre})$
   - $w_{init} = w_i$
     End
   - $y_{rec} = X * w_i$; $w = w_i$
4. For k = 1.. Classnum error(k)=$\|w^{1/2}(y - X_k w_k)\|_2^2$
   End
   If we consider here a class then the identity is:
   Identity(y)=$argmin_h(error)$

Classnum is the class's number of **X** ; where **classnum** $\geq$ d; if **classnum=d** then we consider the case of STSS.

---

5 Experimental results

An emerging tendency in FR is to use STSS [32]. In our experiments, we applied STSS, using standard benchmark face databases to evaluate the performance of the proposed approach.

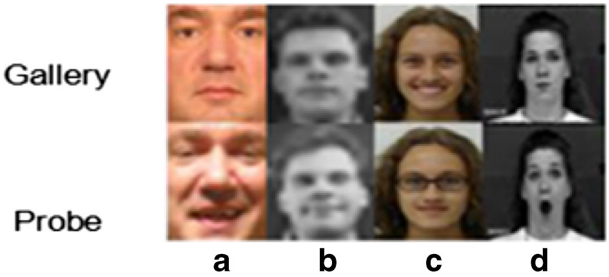

Fig. 5 A subject from Gallery and Probe with different face databases **a** FRGC **b** ORL **c** FEI **d** CK+ **e** GT **f** AR **g** FERET

5.1 Datasets

We used the Extended Cohn-Kanade (CK+) [15] (123 images), Georgia Tech (GT) [39] (50 images), FEI [34] (200 images), AR [37] (100 images some of them with occlusions like in figure), FRGC v1 [36] (152 images), FERET [21] (with different dimension 100, 150 and 200 images) and ORL (40 images) face databases. All the images are resized to 27×32. The pre-processing of the different images are released by commercial face alignment software [43]. In this paper, we chose to select randomly the face image both for Gallery and Probe database. Samples from the used databases are illustrated in Figs. 5 and 6.

In this paper, we chose to select randomly the face image both for Gallery and Probe database. We compare our approach with NN (nearest neighbor) SVM-OAA (one against all), SVM-DAG (Directed Acyclic Graph) [21], BHDT [10], MetaFace [45], RKR [47], RRC [48], CRC [52].

5.2 Experimental protocol

An emerging tendency in FR is to use Single Training Sample per Subject (STSS) which is a challenging problem.

By applying to the images from the databases indicated above, we obtain a similarity matrix of 123×123 comparisons for Extended Cohn Ka-nade (CK+), 152×152 comparisons for FRGC v1, 200 × 200 comparisons for FEI, 50 × 50 comparisons for Georgia Tech, 100 × 100 comparisons for AR, 40 × 40 comparisons for ORL and three different matrix for FERET for three tests (100 × 100 comparisons, 150 × 150 comparisons and 200 × 200 comparisons) which significantly reduce the computational complexity of the algorithm compared to traditional multiple training samples per subject. Hence, the similarity values

Fig. 6 Subjects from AR database with occlusions

Table 1 Recognition accuracy on the FRGC V1, ORL, CK+, FEI, GT and AR databases

Method	FRGC v1	ORL	CK+	FEI	GT	AR
NN	–	0.6994	–	–	–	0.4810
SVM-OAA	0.5921	0.8750	0.9837	0.9600	0.2800	0.8800
SVM-DAG	0.6053	0.8750	0.9837	0.9600	0.2800	0.8200
BHDT	0.2697	0.7500	0.9187	0.6250	0.2000	0.6371
MetaFace	0.6842	0.8750	0.9837	0.9700	0.2800	0.8528
RKR	0.6316	0.8250	0.9837	0.9750	0.2400	0.9286
RRC	0.7105	0.8500	1	0.9800	0.2800	0.9571
CRC	0.6316	0.8500	0.9837	0.9750	0.2800	0.8900
RSN (our)	0.7171	0.8750	0.9919	0.9750	0.3800	0.9500

located in the diagonal of the matrix are intra-class (the same person) and the others are inter-class (different persons) or imposter access.

5.3 Recognition accuracy

The recognition accuracy (RA) is defined as:

$$RA = (1 - EER) \tag{14}$$

where ERR is the Equal Error Rate.

The recognition accuracy of the proposed method is evaluated on test databases and compared to state of the art methods. Obtained results are given in Tables 1 and 2.

From the table we can notice that RSN and RRC lead to much improvement in FR rate compared with the other methods with FRGC v1, ORL and CK+ databases. We can see also using FEI, GT and AR, that RSN and RRC work much better than other methods. RRC gives a 0.98 of recognition while RSN achieves 0.9750 with others methods using FEI database. RSN gives the best accuracy, 0.38, using GT database; while with AR database,

Table 2 Recognition accuracy on the FERET databases

Method	100	150	200
NN	–	–	–
SVM-OAA	0.7700	0.7200	0.6850
SVM-DAG	0.7700	0.7333	0.7150
BHDT	0.5000	0.4200	0.3350
MetaFace	0.8900	0.8933	0.8950
RKR	0.8900	0.8533	0.8500
RRC	0.8800	0.8800	0.9050
CRC	0.8700	0.8400	0.8750
RSN (our)	0.9000	0.8733	0.8950

Table 3 The average running time in seconds on FRGC V1, OR, CK+, FEI, GT and AR databases

Method	FRGC v1	ORL	CK+	FEI	GT	AR
NN	–	0.7703	–	–	–	–
SVM-OAA	0.6415	0.0133	0.1146	0.1516	0.0212	0.1680
SVM-DAG	0.0610	0.0113	0.0473	0.0786	0.0138	0.0433
BHDT	0.0109	0.0019	0.0046	0.0057	0.0022	0.0055
MetaFace	0.5042	0.6500	0.5238	1.0325	1.0684	0.3153
RKR	0.0160	0.0160	1.2e-004	7.5e-005	0	0.0150
RRC	0.0867	0.0102	0.1443	0.1758	0.1178	0.0405
CRC	0.0027	7.7e-04	0.0017	0.0031	0.0012	0.0038
RSN (our)	0.0784	0.0094	0.1954	0.2341	0.1600	0.0419

Table 4 The average running time in seconds on FERET databases

Method	100	150	200
NN	–	–	–
SVM-OAA	0.1692	0.6996	0.5001
SVM-DAG	0.0397	0.0794	0.1074
BHDT	0.0053	0.0121	0.0120
MetaFace	0.4781	0.6991	0.9191
RKR	1.5e-004	1.1e-004	1.6e-004
RRC	0.1366	0.1564	0.1751
CRC	0.0014	0.0076	0.0037
RSN (our)	0.2486	0.2505	0.2519

RSN achieves the second best value; In fact RSN work better when the chosen faces contains a different head pose like the faces with GT database. RSN gives the best recognition using 100 images from FERET database, using 150 and 200 images RSN is competitive compared to the others methods.

5.4 Running time comparison

The average running time of all methods, is evaluated using STSS based FR experiments. We use Matlab version 7.0.1 environment with Intel core 2 duo 2.10 GHz CPU and with 2.87Go RAM. All the methods are implemented using the codes provided by the authors using STSS. In fact, in practical applications, training is usually an offline stage while recognition is usually an online step so that the recognition time is usually much more critical than the training time. Results are summarized in Tables 3 and 4.

RKR and CRC gives the best results.

6 Conclusion

The objective of this paper is to present a new method for face recognition called Regularized Shearlet Network. This approach has the ability to capture face features very efficiently thanks to the use of the shearlet representation, which promotes sparsity and is especially able to extract directional and anisotropic features. In our approach, these features are used to control the trade-off between the fidelity to the gallery and the smoothness of the probe faces in context of regularization theory. The experimental results using single training sample per subject on several face databases show that our new approach is very competitive against several state-of-the-art methods for face recognition.

Acknowledgment The authors would like to acknowledge the financial support of this work by grants from General Direction of scientific Research (DGRST), Tunisia, under the ARUB program. D. Labate acknowledges partial support by NSF DMS 1005799 and DMS 1008900.

References

1. Ahonen T, Hadid A, Pietikainen M (2006) Face description with local binary patterns: Application to face recognition. IEEE Trans Pattern Anal Mach Intell 28(12):2037–2041
2. Ben Amar C, Zaied M, Alimi MA (2005) Beta wavelets. Synthesis and application to lossy image compression. J Adv Eng Softw Elsevier Ed 36(7):459–474
3. Belhumeur PN, Hespanha JP, Kriengman DJ (1997) Eigenfaces vs. fisherfaces: recognition using class specific linear projection. IEEE PAMI 19(7):711–720
4. Bertero M (1986) Regularization methods for linear inverse problems. In: Talenti CG (ed) Inverse problems. Springer, Berlin
5. Bertero M, Poggio T, Torre V (1988) Ill-posed problems in early vision. Proc IEEE 76:869–889
6. Bodmann BG, Kutyniok G, Zhuang X Gabor Shearlets. Appl Comput Harmon Anal. to appear
7. Borgi MA, Labate D, El'arbi M, Ben Amar C (2013) Shearlet network-based sparse coding augmented by facial texture features for face recognition. ICIAP 2:611–620
8. Borgi MA, Labate D, El'arbi M, Ben Amar C (2014) Regularized Shearlet network for face recognition using single sample per person. In: IEEE international conference on acoustics, speech and signal processing, ICASSP2014. Firenze, Italy, pp 514–518
9. Borgi MA, El'arbi M, Ben Amar C (2013) Wavelet Network and Geometric Features Fusion Using Belief Functions for 3D Face Recognition. CAIP 2:307–314
10. Cevikalp H (2010) New clustering algorithms for the support vector machine based hierarchical classification. Pattern Recognit Lett 31(11):1285–1291
11. Chien JT, Wu CC (2002) Discriminant waveletfaces and nearest feature classifiers for face recognition. IEEE Trans Pattern Anal Mach Intell 24(12):1644–1649
12. Easley GR, Labate D (2012) Critically sampled wavelets with composite dilations. IEEE Trans Image Process 21(2):550–561
13. Easley GR, Labate D, Patel V (2012) Hyerbolic shearlets. ICIP, Orlando
14. Easley GR, Labate D, Patel V (2014) Directional multiscale processing of images using wavelets with composite dilations. J Math Imag Vision 48(1):13–34
15. Georgia Tech Face Database, ftp://ftp.ee.gatech.edu/pub/users/hayes/facedb/
16. Guo K, Lim W-Q, Labate D, Weiss G, Wilson E (2004) Wavelets with composite dilations? Electron Res Announc Amer Math Soc 10:78–87
17. Guo K, Labate D, Lim W, Weiss G, Wilson E (2006) The theory of wavelets with composite dilations. In: Heil C (ed) Harmonic analysis and applications. Birkauser, pp 231–249
18. Hastie T, Tibshirani R, Friedman J (2003) The elements of statistical learning. Springer Series in Statistics
19. Heisele B, Ho P, Poggio T (2001) Face recognition with support vector machine: Global versus component-based approach. ICCV 2:688–694
20. Hiriart-Urruty J, Lemarechal C (1996) Convex analysis and minimization algorithms. Springer-Verlag
21. Hsu CW, Lin CJ (2002) A comparison of methods for multiclass support vector machines. IEEE Trans Neural Networks 13(N. 2):415–425
22. Kutyniok G, Labate D (2009) Resolution of the wavefront set using continuous shearlets. Trans Amer Math Soc 361:2719–2754
23. Laaksonen J (1997) Local subspace classifier. Proc Int Conf Artif Neural Netw

24. Labate D, Lim W-Q, Kutyniok G, Weiss G (2005) Sparse multidimensional representation using shearlets. Wavelets XI (San Diego, CA). SPIE, Bellingham, pp 254–262
25. Lanitis A, Taylor CJ, Cootes TF (1997) Automatic interpretation and coding of face images using flexible models. IEEE Trans Pattern Anal Mach Intell 19(7):743–756
26. Lee K, Ho J, Kriegman D (2005) Acquiring linear subspaces for face recognition under variable lighting. IEEE Trans Pattern Anal Mach Intell 27(5):684–698
27. Leng L, Zhang J, Khurram Khan M, Chen X, Alghathbar K (2010) Dynamic weighted discrimination power analysis: A novel approach for face and palmprint recognition in DCT domain? Int J Phys Sci 5(17):2543–2554
28. Leng L, Zhang J, Xu J, Khurram Khan M, Alghathbar K (2010) Dynamic weighted discrimination power analysis in DCT domain for face and palmprint recognition. In: International Conference on Information and Communication Technology Convergence (ICTC), pp 467–471
29. Li SZ (1998) Face recognition based on nearest linear combinations. Proc IEEE Int Conf Comput Vis Pattern Recognit
30. Li SZ, Lu J (1999) Face recognition using nearest feature line method. IEEE Trans Neural Netw 10(2):439–443
31. Lu J, Plataniotis Kostantinos N, Anastasios V (2003) Face recognition using LDA-based algorithms. IEEE Trans Neural Networks 14(1):195–200
32. Lucey P, Cohn JF, Kanade T, Saragih J, Ambadar Z, Matthews I (2010) The Extended Cohn-Kande Dataset (CK+): A complete facial expression dataset for action unit and emotion-specified expression. CVPR, Workshop H CBA, pp 94–101
33. Martinez AM (2002) Recognizing Imprecisely localized, partially occluded, and expression variant faces from a single sample per class. IEEE Trans Pattern Anal Mach Intell 24(6):748–763
34. Martinez A, benavente R (1998) The AR face database. Technical Report 24, CVC
35. Naseem I, Togneri R, Bennamoun M (2010) Linear regression for face recognition. IEEE Trans Pattern Anal Mach Intell 32(11):2106–2112
36. Phillips PJ, Moon H, Rizvi SA, Rauss PJ (2000) The FERET evaluation methodology for face-recognition algorithms. IEEE Trans Pattern Anal Mach Intell 22:1090–1104
37. Phillips PJ, Flynn PJ, Scruggs T, Bowyer KW, Chang J, Hoffman K, Marques J, Min J, Worek W (2005) Overview of the face recognition grand challenge. CVPR 1:947–954
38. Su Y, Shan SG, Chen XL, Gao W (2010) Adaptive generic learning for face recognition from a single sample per person. CVPR, pp 2699–2706
39. Thomaz CE, Giraldi GA (2010) A new ranking method for Principal Components Analysis and its application to face image analysis. Image Vis Comput 28(6):902–913
40. Tibshirani R (1996) Regression shrinkage and selection via the lasso. J R Stat Soc Ser B (Stat Methodol) 58:267–288
41. Tikhonov AN, Arsenin VY (1977) Solutions of Ill-posed problems. W. H. Winston, Washington
42. Turk M, Pentland A (1991) Eigenfaces for recognition. J Cogn Neurosci 3(1):71–86
43. Wolf L, Hassner T, Taigman Y (2009) Similarity scores based on background samples,? Computer Vision? ACCV, pp 88–97
44. Wright J, Yang AY, Ganesh A, Sastry SS, Ma Y (2009) Robust face recognition via sparse representation. IEEE Trans Pattern Anal Mach Intell 31(2):210–227
45. Yang M, Zhang L, Zhang D (2010) Metaface learning for sparse representation based face recognition. ICIP, Hong Kong, pp 1601–1604
46. Yang M, Zhang L, Yang J, Zhang D (2011) Robust sparse coding for face recognition. CVPR, pp 625–632
47. Yang M, Zhang L, Zhang D, Wang Shenlong (2012) Relaxed collaborative representation for pattern classification. CVPR, pp 2224–2231
48. Yang M, Zhang L, Yang J, Zhang D (2013) Regularized robust coding for face recognition. IEEE Trans Image Process 22(5):1753–1766
49. Yi S, Labate D, Easley GR, Krim H (2009) A Shearlet approach to edge analysis and detection. IEEE Trans Image Process 18(5):929–941
50. Zhang T (2010) Analysis of multi-stage convex relaxation for sparse regularization. J Mach Learn Res 11:1081–1107
51. Zhang T (2012) Multi-stage convex relaxation for feature selection. Bernoulli
52. Zhang L, Yang M, Feng X (2011) Sparse representation or collaborative representation: which helps face recognition. ICCV, pp 471–478

Mohamed Anouar Borgi received the B.S. degree in computer science from University of Monastir, Tunisia, in 2001, Msc1 and the Msc. degree in computer science from the University of Sfax, Tunisia, in 2007 and 2010. He is Ph.D student in Regim Laboratory, University of sfax. His research interests include machine learning, pattern recognition and computer vision.

Maher El'Arbi was born in Kerkennah Sfax (Tunisia) in September 1976. He received the maitrise Diploma from the University Of Management and Economics of Sfax (FSEGS) in 2000, the DEA in computer science from the High Institute of Management of Tunis (ISGT) and the Phd degree in Engineering of Computer Systems from the National Engineering School of Sfax in 2010. Nowadays, he is an assistant in the department of Multimedia of the High Institute Of Computer and Multimedia of Gabs Tunisia and a research member in the Research Group of Intelligent Machine (REGIM) Tunisia. His research interests include digital watermarking and data hiding, multimedia authentication and image processing.

Demetrio Labate received the Ph.D. degree in electrical engineering from the Politecnico di Torino, Italy, in 1995, and the Ph.D. degree in mathematics from the Georgia Institute of Technology, Atlanta, in 2000. He was a Chauvenet lecturer at Washington University in St.Louis and an Assistant Professor at NC State University. Since 2009, he is Associate Professor at the Department of Mathematics of the University of Houston. His research interests include harmonic analysis, wavelet theory, time-frequency analysis and sparse representations. Prof. Labate received the NSF CAREER Award in Applied Mathematics in 2008.

Chokri Ben Ammar received the B.S. degree in Electrical Engineering from the National Engineering School of Sfax (ENIS) in 1989, the M.S. and PhD degrees in Computer Engineering from the National Institute of Applied Sciences in Lyon, France, in 1990 and 1994, respectively. He spent one year at the University of "Haute Savoie" (France) as a teaching assistant and researcher before joining the higher School of Sciences and Techniques of Tunis (ESSTT) as Assistant Professor in 1995. In 1999, he joined the Sfax University (USS), where he is currently a professor in the Department of Electrical Engineering of the National Engineering School of Sfax (ENIS).

His research interests include Computer Vision and Image and video analysis. These research activities are centered on two main areas: Social Web and the use of Wavelets and Wavelet networks for data Classification and approximation, Pattern Recognition and image and video coding, indexing and watermarking.