ShearFace: Efficient Extraction of Anisotropic Features for Face Recognition

Mohamed Anouar Borgi[1], Demetrio Labate[2]
[1]Research Groups on Intelligent Machines,
University of Sfax,
Sfax 3038, Tunisia
{anoir.borgi@ieee.org; dlabate@math.uh.edu}

Maher El Arbi[1], Chokri Ben Amar[1]
[2] Department of Mathematics,
University of Houston, Houston,
TX 77204, USA
{maher.elarbi@gmail.com; chokri.benamar@ieee.org}

Abstract— **This paper presents an improved approach to face recognition, called Regularized Shearlet Network (RSN), that takes advantage of the sparse representation properties of shearlets in biometric applications. The main novelty of our approach is the efficient extraction of geometric features based on the properties of the shearlet decomposition, a multiscale directional method which is especially designed to capture directional and anisotropic information in multidimensional data. To further improve the performance of our face recognition algorithm, we include a regularization step to control the trade-off between the fidelity to the data (gallery) and smoothness of the solution (probe). In this work, we focus on the challenging problem of the single training sample per subject (STSS). We compare our new algorithm against different state-of-the-arts method using several facial databases including AR, FERET, FRGC, FEI and CK. Our tests show that our RSN algorithm is very competitive and outperforms several state-of-the-art face recognition methods.**

Keywords— **Shearlet, Regularized Shearlets Network, Face Recognition.**

I. Introduction

Face recognition (FR) is a classical problem in computer vision and pattern recognition and many methods, such as Eigenfaces [1], Fisherfaces [2], SVM [3] and Metaface [4] have been proposed during the past two decades.

One of the standard statistical methods for FR is *subset selection* (L_0 regularization) [19], which consists in computing the following estimator:

$$\hat{w}_{L_0} = \arg\min_{w \in Rp} \|Xw - y\|_2^2 \text{ subject to } \|w\|_0 \le \delta \qquad (1)$$

where δ is a tuning parameter, y is a normalized test face and X is a matrix representing a gallery of faces. 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 as a minimization problem. For example, the recently proposed Sparse Representation Classification (SRC) scheme [5] 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 would result in the minimal reconstruction error. However, it was shown in [6] that SRC actually owes its success to the use 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 [7] [8] that robustly regresses 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 reweighted regularized robust coding algorithm was proposed to solve the RRC model efficiently

In this paper, we propose a method called Regularized Shearlets Network (RSN), which combines sparsity and regularization theory. Sparsity, in particular, will be based on the use of the shearlet representation, an innovative multiscale framework which combines the classical multiresolution analysis with high directional sensitivity and provides optimally sparse approximations for a large class of images. 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 are especially effective to capture directional and anisotropic features with high efficiency. Furthermore, they have a well understood mathematical theory and fast numerical implementations [9]. Regularization theory is another important component of our approach, that allows us to control the trade-off between fidelity to the data and smoothness of the solution.

The rest of this paper is organized as follows. In Sec. 2, we briefly describe the necessary background on shearlets. In Sec. 3, we describe our Regularized Shearlet Network algorithm. In Sec. 4, we present several numerical experiments to demonstrate the efficacy of the proposed algorithm and compare it against competing algorithms. Finally, we make some concluding remarks in Sec. 5.

II. The Shearlet Transform

The shearlet transform, introduced by one of the authors and his collaborators in [10], is a genuinely multidimensional version of the traditional wavelet transform, and is especially designed to represent data containing anisotropic and directional features with very 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 [11][12][13]. Formally, the *Continuous Shearlet Transform* [14] is defined as the mapping:

$$SH_{\psi}(a,s,t) = \left\langle f, \psi_{a,s,t} \right\rangle, a > 0, s \in \mathbb{R}, t \in \mathbb{R}^2 \qquad (2)$$

where, $\psi_{ast}(x) = \left|\det M_{as}\right|^{-\frac{1}{2}} \psi(M_{as}^{-1}(x-t))$, and $M_{as} = \begin{pmatrix} a & s \\ 0 & \sqrt{a} \end{pmatrix}$. Observe each matrix M_{as} can be factorized as $B_s A_a$, where $B_s = \begin{pmatrix} 1 & -s \\ 0 & 1 \end{pmatrix}$ is a *shear matrix* and $A_a = \begin{pmatrix} a & 0 \\ 0 & \sqrt{a} \end{pmatrix}$ is an anisotropic dilation matrix. Thus, the shearlet transform is a function of three variables: the *scale a*, 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 $SH_{\psi}(a,s,t)$ on an appropriate discrete set, we obtain the corresponding Discrete Shearlet Transform. Specifically, M_{as} is discretized as $M_{jl} = B_l A^j$, 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 waveforms:

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

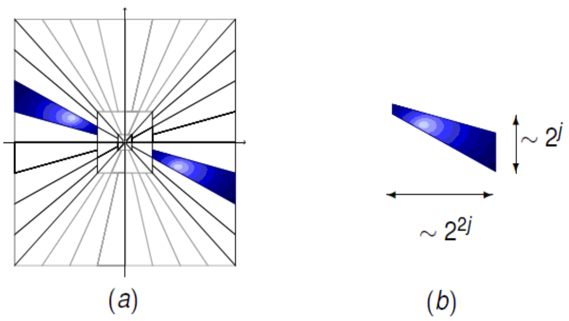

Fig. 1. (a) Spatial-frequency plane of the shearlets, (b) Frequency support.

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

III. THE PROPOSED APPROACH

Our novel approach for FR that we call Regularized Shearlet Network (RSN) is defined as a cascade of a feature extraction module followed by a recognition (or verification) module. We handle the extraction of the features using the Shearlet Network (SN). Thanks to the properties of shearlets, this step is very efficient to capture the essential geometry of the image. We implement the recognition step by the use of regularization theory which allows us to satisfy both fidelity to the solution (Probe or Test) and closeness to the data (Gallery) [36]. The structure of our algorithm is shown in Figure 3.

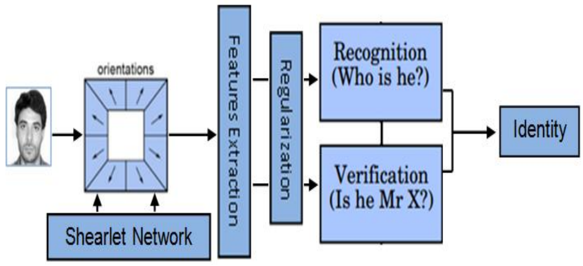

Fig. 2. Augmented face recognition schema.

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 [15, 16, 17]. In practice, rather than looking for the exact solution, we settle for an approximate one which satisfies some type of *regularity*. One of the most popular and effective approximation methods is the L_1 regularization method which is often referred to as Lasso [32] and is given by:

$$\hat{w}_{L_1} = \arg\min_{w \in Rp} \left[\frac{1}{n} \left\| Xw - y \right\|_2^2 + \lambda \left\| w \right\|_1 \right] \qquad (4)$$

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.

The global optimum of (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 suboptimal. The theoretical performance of this method has been analyzed recently [18][19].

A. SN for Modeling and Features Extraction

Our proposed RSN approach is initialized by training a shearlet network (SN) [20] to models the faces. The Gallery faces are approximated by a shearlet network to produce a compact biometric signature as wavelet network [38]. One main feature of this approach is that this signature, constituted 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).

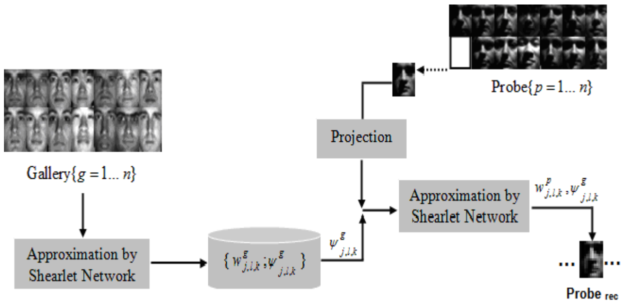

Fig. 3. Overview of SN Architecture.

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:

$$f = \sum_{j,l\in\mathbb{Z},k\in\mathbb{Z}^2} \left\langle f, \psi_{j,l,k} \right\rangle \psi_{j,l,k} \quad (5)$$

We will use this formula to define the Shearlet Network approach, similar to the wavelet network [33] [39] [40] [41], 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:

$$\tilde{\psi}^{i}_{j,l,k} = \sum_{m=1}^{N} (\Psi_{i,m})^{-1} \psi^{m}_{j,l,k} \text{ with } \Psi_{i,m} = \left\langle \psi^{i}_{j,l,k}, \psi^{m}_{j,l,k} \right\rangle \quad (6)$$

In our approach, the mother shearlet that we use to construct the family $\psi_{j,l,k}$ is the second derived of the Beta function [31] [37] which has the advantage of being well localized. Note that the number of shearlets may be chosen by the user.

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:

a. Choose the mother shearlet.
b. Build a library formed by the shearlets which form a shearlet frame.
c. Set as a stop learning condition based on the difference of input and the output network and iterate the following steps:

2. If frame is not tight: Calculate the dual basis $\tilde{\psi}^{i}_{j,l,k}$ formed by the shearlets of the network and the new selected shearlet according (6); else $\psi_{j,l,k} = \tilde{\psi}^{i}_{j,l,k}$.

3. Calculate the weights by direct projection of the image on the dual shearlet $w_i = < f, \tilde{\psi}^{i}_{j,l,k} >$.

4. Calculate the output of the network f_{rec}.

5. If the number of shearlets is reached then learning stops; otherwise another shearlet is selected and we return to **2.**

B. RSN Algorithm

Below we present the algorithm of RSN, where X represents the reconstructed gallery faces after extraction of the features by training SN and 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 [35])
2. Diagonalizable X ; $X^t = X' * X; y^t = X' * y$
3. For j = 1 … Iter
   - Calculate $\hat{w} = \arg\min_{w\in Rp} \left[\frac{1}{n} \|Xw - y\|_2^2 + \lambda \|w\|_1 \right]$ Use Lasso [34]:
     $w_i = Lasso(X, y, X^t, y^t, w_{init}, w\_thre)$
   - $w_{init} = w_i$

   End

   $y_{rec} = X * w_i$; $w = w_i$
4. For k = 1,…, Classnum

   $error(k) = \left\| w^{1/2} (y - X_k w_k) \right\|_2^2$

   End

   If we consider here a class h then the identity is:

   $Identity(y) = \arg\min_h (error)$

In the algorithm above: Classnum is the classes' number of X , where $Classnum \geq d$; if $Classnum = d$ then we consider the case of Single Training Sample per Subject (STSS).

IV. EXPERIMENTAL RESULTS

In this paper, we focus on the problem called Single Training Sample per Subject (STSS) that is receiving considerable attention in FR [21]. For our experiments, we have used several standard benchmark face databases to evaluate the performance of our approach.

A. Datasets

We have used the Extended Cohn-Kanade (CK+) [22] (123 images), Georgia Tech (GT) [23] (50 images), FEI [24] (200 images), AR [25] (100 images), FRGC v1 [26] (152 images), FERET [27] (with different dimension 100, 150 and 200 images) and ORL (40 images) face databases. All the images are resized to 27×32.

In this paper, we chose to select **randomly** the face image both for Gallery and Probe dataset.

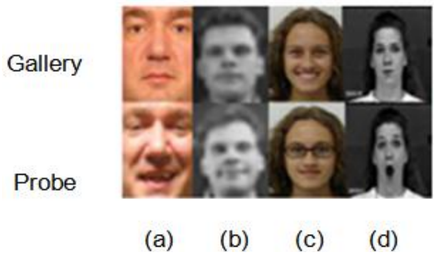

Fig. 4. A subject from Gallery and Probe with different face databases. (a) FRGC. (b) ORL. (c) FEI. (d) CK+.

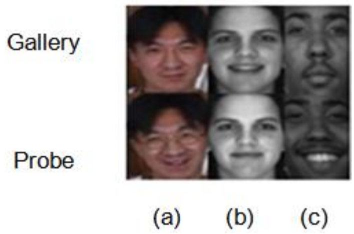

Fig. 5. A subject from Gallery and Probe with different face databases. (a) GT. (b) AR. (c) FERET.

We have compared our approach with NN (nearest neighbor) SVM_OAA (one against all), SVM_DAG (Directed Acyclic Graph) [28], BHDT [29], MetaFace [4], RKR [30], RRC [8], CRC [6].

B. Recognition accuracy

Table I shows that RSN (our method) and RRC are the best performing methods in terms of FR rate when compared with many other classical and state-of-the-art methods using the FRGC v1, ORL and CK+ databases.

TABLE I. RECOGNITION ACCURACY ON THE FRGC V1, ORL & CK+ DATABASES.

	Database		
Method	***FRGC v1***	***ORL***	***CK+***
NN	-	0.6994	-
SVM_OAA	0.5921	0.8750	0.9837
SVM_DAG[28]	0.6053	0.8750	0.9837
BHDT [29]	0.2697	0.7500	0.9187
MetaFace [4]	0.6842	0.8750	0.9837
RKR [30]	0.6316	0.8250	0.9837
RRC [8]	**0.7105**	0.8500	**1**
CRC [6]	0.6316	0.8500	0.9837
RSN (our)	**0.7171**	**0.8750**	**0.9919**

Table II shows that, also using the FEI, GT and AR databases, the RSN and RRC methods are the top performers.

TABLE II. RECOGNITION ACCURACY ON THE FEI, GT & AR DATABASES.

	Database		
Method	***FEI***	***GT***	***AR***
NN	-	-	0.4810
SVM_OAA	0.9600	0.2800	0.8800
SVM_DAG[28]	0.9600	0.2800	0.8200
BHDT [29]	0.6250	0.2000	0.6371
MetaFace [4]	0.9700	0.2800	0.8528
RKR [30]	0.9750	0.2400	0.9286
RRC [8]	**0.9800**	0.2800	**0.9571**
CRC [6]	0.9750	0.2800	0.8900
RSN (our)	**0.9750**	**0.3800**	**0.9500**

In Table III, we test FR using the FERET database, with 100, 150 and 200 images. Also in this case, our method is among the top performers.

TABLE III. RECOGNITION ACCURACY ON THE FERET DATABASE.

	FERET Database		
Method	***100***	***150***	***200***
NN	-	-	-
SVM_OAA	0.7700	0.7200	0.6850
SVM_DAG[28]	0.7700	0.7333	0.7150
BHDT [29]	0.5000	0.4200	0.3350
MetaFace [4]	0.8900	**0.8933**	0.8950
RKR [30]	0.8900	0.8533	0.8500
RRC [8]	0.8800	0.8800	**0.9050**
CRC [6]	0.8700	0.8400	0.8750
RSN (our)	**0.9000**	0.8733	**0.8950**

C. Runing Time Comparison

For a fair comparison, we have measured the average running time of all methods. For all our experiments, we have used Matlab version 7.0.1 environment with Intel core 2 duo 2.10 GHz CPU and with 2.87Go RAM. For all methods cited from the literature, we have applied the implemented codes as provided by the authors in the case of STSS. The tables below report the average running times for the various methods considered. Note that the algorithms RKR [30] and CRC [6] are overall the least computationally intensive. Our approach requires a computational time comparable to CRC in many case, even though the performance in terms of running times depends on the database considered.

TABLE IV. THE AVERAGE RUNING TIME (SECONDS) ON FRGC V1, ORL & CK+ DATABASES.

	Database		
Method	***FRGC v1***	***ORL***	***CK+***
NN	-	0.7703	-
SVM_OAA	0.6415	0.0133	0.1146
SVM_DAG[28]	0.0610	0.0113	0.0473
BHDT [29]	0.0109	0.0019	0.0046
MetaFace [4]	0.5042	0.6500	0.5238
RKR [30]	0.0160	0.0160	1.2e-004
RRC [8]	0.0867	0.0102	0.1443
CRC [6]	0.0027	7.7e-04	0.0017
RSN (our)	0.0784	0.0094	0.1954

TABLE V. THE AVERAGE RUNING TIME (SECONDS) ON FEI, GT & AR DATABASES.

	Database		
Method	***FEI***	***GT***	***AR***
NN	-	-	-
SVM_OAA	0.1516	0.0212	0.1680
SVM_DAG[28]	0.0786	0.0138	0.0433
BHDT [29]	0.0057	0.0022	0.0055
MetaFace [4]	1.0325	1.0684	0.3153
RKR [30]	7.5e-005	0	0.0150
RRC [8]	0.1758	0.1178	0.0405
CRC [6]	0.0031	0.0012	0.0038
RSN (our)	0.2341	0.1600	0.0419

TABLE VI. T HE AVERAGE RUNING TIME (SECONDS) ON FERET DATABASE.

	FERET Database		
Method	***100***	***150***	***200***
NN	-	-	-
SVM_OAA	0.1692	0.6996	0.5001
SVM_DAG[28]	0.0397	0.0794	0.1074
BHDT [29]	0.0053	0.0121	0.0120
MetaFace [4]	0.4781	0.6991	0.9191
RKR [30]	1.5e-004	1.1e-004	1.6e-004
RRC [8]	0.1366	0.1564	0.1751
CRC [6]	0.0014	0.0076	0.0037
RSN (our)	0.2486	0.2505	0.2519

5. 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, a method which promotes sparsity and is especially able to extract geometric features with high accuracy. In our approach, these features are fed into a shearlet network and processed through a regularization stage to control the trade-off between fidelity to the gallery and smoothness of the probe faces. The experimental results for FR on the problem of Single Training Sample per Subject run on several face databases show that our new approach is very competitive when compared against several state-of-the-art methods.

ACKNOWLEDGMENT

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

REFERENCES

[1] M. Turk, and A. Pentland, “Eigenfaces for recognition,” J.Cognitive Neuroscience, 3(1), pp. 71–86, 1991.

[2] P.N. Belhumeur, J.P. Hespanha, and D.J. Kriengman, “Eigenfaces vs. fisherfaces: recognition using class specific linear projection,” IEEE PAMI, 19(7), pp. 711–720, 1997.

[3] B. Heisele, P. Ho, and T. Poggio, “Face recognition with support vector machine: Global versus component-based approach,”, ICCV, vol.2, pp. 688 – 694, 2001.

[4] M. Yang, L. Zhang and D. Zhang. “Metaface learning for sparse representation based face recognition,” ICIP, pp. 1601-1604, Hong Kong, 2010.

[5] J. Wright, A.Y. Yang, A. Ganesh, S.S. Sastry, and Y. Ma, “Robust face recognition via sparse representation,” IEEE Trans. Pattern Analysis and Machine Intelligence, vol. 31, no. 2, pp. 210–227, 2009.

[6] L. Zhang, M. Yang, and X. Feng, “Sparse Representation or Collaborative Representation: Which Helps Face Recognition?,” ICCV, pp. 471-478, 2011.

[7] M. Yang, L. Zhang, J. Yang and D. Zhang, “Robust sparse coding for face recognition,” CVPR, pp. 625-632, 2011.

[8] M. Yang, L. Zhang, J. Yang and D. Zhang, “Regularized Robust Coding for Face Recognition,” IEEE Trans. on Image Processing, 22(5), pp. 1753-1766, 2013.

[9] K. Guo and D. Labate, “Optimally sparse multidimensional representation using shearlets,” SIAM J Math. Anal., vol. 39, pp. 298-318, 2007.

[10] D. Labate, W.-Q. Lim, G. Kutyniok and G. Weiss, “Sparse multidimensional representation using shearlets,” Wavelets XI (San Diego, CA), SPIE, Bellingham, WA, pp. 254–262, 2005.

[11] G.R. Easley, and D. Labate, “Critically sampled wavelets with composite dilations,” IEEE Trans. Image Process., 21 (2), pp. 550-561, 2012.

[12] G.R. Easley, D. Labate and V. Patel, “Directional multiscale processing of images using wavelets with composite dilations,” J. Math. Imag. Vision, 2013.

[13] S. Yi, D. Labate, G.R. Easley, and H. Krim, “A Shearlet approach to Edge Analysis and Detection,” IEEE Trans. Image Process, 18 (5), pp. 929–941, 2009.

[14] G. Kutyniok and D. Labate, “Resolution of the wavefront set using continuous shearlets,” Trans. Amer. Math. Soc. 361, pp. 2719-2754, 2009.

[15] M. Bertero, “Regularization methods for linear inverse problems,” Inverse Problems, ed. C.G. Talenti, Springer, Berlin, 1986.

[16] M. Bertero, T. Poggio and V. Torre, “Ill-posed problems in early vision,” Proc. IEEE, 76 , pp. 869–889, 1988.

[17] A.N. Tikhonov and V.Y. Arsenin, “Solutions of Ill-posed Problems,” W.H. Winston, Washington, DC, 1977.

[18] T. Zhang, “Analysis of multi-stage convex relaxation for sparse regularization,” Journal of Machine Learning Research, 11, pp. 1081–1107, 2010.

[19] T. Zhang, “Multi-stage convex relaxation for feature selection,’’ Bernoulli, 2012.

[20] M.A. Borgi, D. Labate, M. El'arbi, and C. Ben Amar, “Shearlet Network-based Sparse Coding Augmented by Facial Texture Features for Face Recognition,” ICIAP (2), pp. 611-620, 2013.

[21] Y. Su, S.G. Shan, X.L. Chen, and W. Gao, “Adaptive Generic Learning for Face Recognition from a Single Sample per Person,” CVPR, pp. 2699 – 2706, 2010.

[22] P. Lucey, J. F. Cohn, T. Kanade, J. Saragih, Z. Ambadar, and I. Matthews, “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, 2010.

[23] A. Martinez and R. benavente, “The AR face database,” Technical Report 24, CVC, 1998.

[24] C. E. Thomaz and G. A. Giraldi. “A new ranking method for Principal Components Analysis and its application to face image analysis,” Image and Vision Computing, vol. 28, no. 6, pp. 902-913, June 2010.

[25] A. Martinez and R. benavente, “The AR face database,” Technical Report 24, CVC, 1998.

[26] P. J. Phillips, P. J. Flynn, T. Scruggs, K. W. Bowyer, J. Chang, K. Hoffman, J. Marques, J. Min, and W. Worek, “Overview of the face recognition grand challenge,” CVPR, no. 1, pp. 947-954, June 2005.

[27] P.J. Phillips, H. Moon, S.A. Rizvi, and P.J. Rauss, “The FERET Evaluation Methodology for Face-Recognition Algorithms,” IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 22, pp. 1090-1104, 2000.

[28] C.W Hsu, and C.J Lin, “A Comparison of Methods for Multiclass Support Vector Machines,” IEEE Transactions on neural networks, vol. 13, N°. 2, pp. 415 – 425, March 2002.

[29] H. Cevikalp, “New clustering algorithms for the support vector machine based hierarchical classification,” Pattern Recognition Letters, 31(11), pp. 1285-1291, 2010.

[30] M. Yang, L. Zhang, D. Zhang, and Shenlong Wang, “Relaxed collaborative Representation for Pattern Classification,” CVPR, pp. 2224-2231, 2012.

[31] C. Ben Amar, M. Zaied and M.A. Alimi, “Beta Wavelets. Synthesis and application to lossy image compression,” Journal of Advances in Engineering Software, Elsevier edition, 36(7), pp. 459-474, July 2005.

[32] R. Tibshirani, “Regression shrinkage and selection via the lasso,” Journal of the Royal Statistical Society, Series B (Statistical Methodology), 58, pp. 267–288, 1996.

[33] M.A. Borgi, M. El'arbi and C. Ben Amar, “Wavelet Network and Geometric Features Fusion Using Belief Functions for 3D Face Recognition,” CAIP (2), pp. 307-314, 2013.

[34] J. Hiriart-Urruty and C. Lemarechal, “Convex analysis and minimization algorithms,” Springer-Verlag, 1996.

[35] T. Hastie, R. Tibshirani, and J. Friedman, “The Elements of Statistical Learning,” Springer Series in Statistics, 2003.

[36] M.A. Borgi, D. Labate, M. El'arbi, and C. Ben Amar, “Regularized Shearlet Network for face Recognition using Single Sample Per Person,” ICASSP, pp. 514-518, Firenze Italy, May 4-9, 2014. (accepted)

[37] M. Zaied, C. Ben Amar, and M. A. Alimi, “Beta Wavelet Networks for Face Recognition, Journal of Decision Systems,” Lavoisier 2005 Edition, Vol. 14, N°. 1-2/2005, pp. 109-122, 2005.

[38] S. Said, B. Ben Amor, M. Zaied, C. Ben Amar and M. Daoudi, “Fast and efficient 3D face recognition using wavelet networks,” ICIP. pp. 4153-4156, 2009.

[39] O. Jemai, M. Zaied, C. Ben Amar, and M. A. Alimi, “FBWN: an architecture of Fast Beta Wavelet Networks for Image Classification,” 2010 IEEE World Congress on Computational Intelligence, IJCNN 2010, International Joint Conference on Neural Networks, Barcelona, pp. 1953-1960, July 18-23, 2010.

[40] O. Jemai, M. M. Zaied, C. Ben Amar, and M. A. Alimi, “Pyramidal Hybrid Approach: Wavelet Network with OLS Algorithm Based-Image Classification,” International Journal of Wavelets, Multiresolution and Information Processing, IJWMIP, Vol. 9, No. 1, pp. 111-130, 2011.

[41] O. Jemai, M. Zaied, C. Ben Amar, and M. A. Alimi, “Faster Learning Algorithm of Wavelet Network Based on Fast Wavelet Transform,” International Journal of Pattern Recognition and Artificial Intelligence, IJPRAI, Vol. 25, N° 8, pp. 1279-1319, 2011.