REGULARIZED SHEARLET NETWORK FOR FACE RECOGNITION USING SINGLE SAMPLE PER PERSON

Mohamed Anouar Borgi[1], Demetrio Labate[2], Maher El'Arbi[1], Chokri Ben Amar[1]

[1]Research Groups on Intelligent Machines, University of Sfax, Sfax 3038, Tunisia
[2]Department of Mathematics, University of Houston, Houston, TX 77204, USA
{anoir.borgi@ieee.org； dlabate@math.uh.edu ; maher.elarbi@gmail.com ; chokri.benamar@ieee.org }

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 includes a module based on regularization theory (RSN) 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-arts method using several facial databases, such as AR, FERET, FRGC, FEI, CK. Our tests show that the RSN approach is very competitive and outperforms several standard face recognition methods.

Index Terms— 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 [1], Fisherfaces [2], SVM [3] and Metaface [4] have been proposed in the past two decades.
One of the standard statistical methods for FR is the method of subset selection (L_0 regularization) [19], which consists in computing the following estimator:

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

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 a different insight into the classical face recognition 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 could result in the minimal reconstruction error. However, it has been indicated in [6] 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 [7] [8], which is able to 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 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, is based on the use of the shearlet representation, a multiscale framework which combines the power of classical multiresolution analysis with high directional sensitivity. The shearlet approach was successfully introduced during the last decade to overcome the limitations of classical wavelets. Indeed, despite their extensive use in image processing, traditional wavelets have a limited ability to deal with directional information. By contrast, shearlets are especially designed to capture directional and anisotropic features with high efficiency, are optimally sparse for the representation of 2D/3D images containing edges and have fast numerical implementations [9]. 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 rest of this paper is organized as follows. In Sec. 2, we briefly describe the necessary background on shearlets. Sec. 3 presents the proposed 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, Sec. 5 concludes this paper.

2. THE SHEARLET TRANSFORM

The shearlet transform, introduced by one of the authors and their collaborator 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 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 that 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.

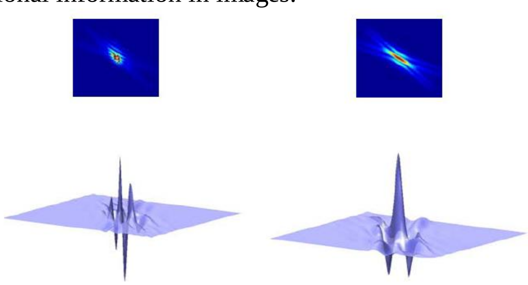

Figure 1. Directional filters of shearlet.

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 functions of the form:

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

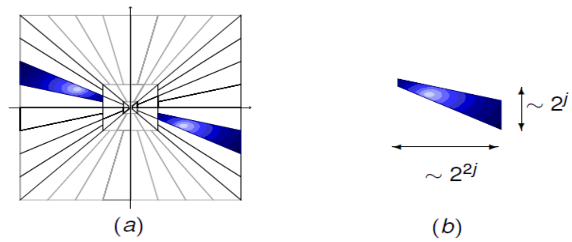

Figure 2. (a) Spatial-frequency plane of the shearlets, (b) Frequency support.

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

3. REGULARIZED SHEARLET NETWORK (RSN)

Our proposed Regularized Shearlet Network scheme for FR is defined as a cascade of a feature extraction module followed by a recognition module. We will implement this scheme by the use of regularization theory, where the extraction of directional features is controlled by the Shearlet Network (SN), as shown in Figure 3.

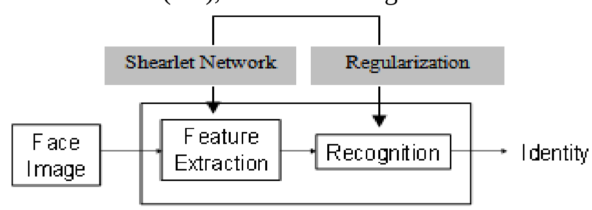

Figure 3. Augmented face recognition schema.

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 which can be solved using regularization theory [15, 16, 17]. In practice, rather than looking for an exact solution, we will settle for an approximate one. A very popular and effective approximation method 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].

3.1. 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. One main feature of our approach is that this signature, consisting of the shearlet coefficients 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 shearlet coefficients remains then unchanged (this is the Gallery face).

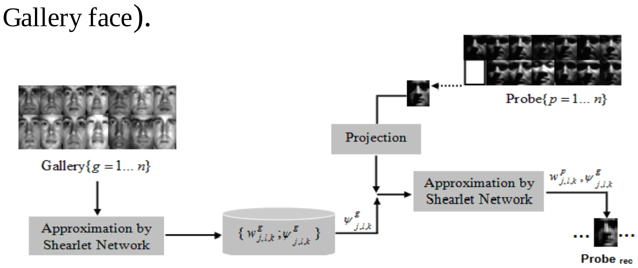

Figure 3. Overview of SN modeling.

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

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

We will use this formula to define the Shearlet Network approach, similar to the wavelet network [33], as a combination of the RBF neural network and the shearlet decomposition. In the optimization stage, a shearlet coefficient from the library is processed through the hidden layer of the network or used to update the weights. The calculation of the weights connection in every stage is obtained by projecting the signal to be analyzed on a family of shearlets. For this, 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 shearlet generator used to construct the family $\psi_{j,l,k}$ is the second derived of the Beta function [31]. 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 not a 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 $\tilde{\psi}^{i}_{j,l,k} = \psi_{j,l,k}$.

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 the learning stops, otherwise another shearlet is selected and we return to **2.**

3.2. RSN algorithm

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 $f : y = f / norm \quad f, 2$

- 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} in [0 .. 1] (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]:

     $w_i = Lasso(X, y, X^t, y^t, w_{init}, w\_thre)$
   - $w_{init} = w_i$

   EndFor

   $y_{rec} = X * w_i ; \; w = w_i$

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

   $$Identity(y) = \arg\min_k \left\| w^{1/2} (y - X_k \hat{w}_k) \right\|_2$$

4. EXPERIMENTAL RESULTS

An emerging tendency in FR is to use **Single Training Sample per Subject (STSS)** [21]. In our experiments, we applied STSS, using standard benchmark face databases to evaluate the performance of the proposed approach. We 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 have **randomly** selected the face image both for Gallery and Probe dataset. We compare 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].

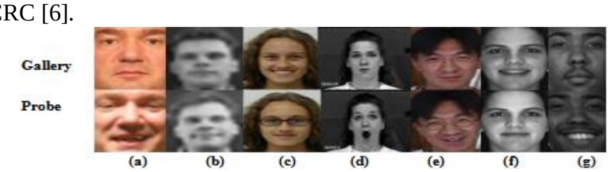

Figure 4. A subject from Gallery and Probe with different face databases. (a) FRGC. (b) ORL. (c) FEI. (d) CK+. (e) GT. (f) AR. (g) FERET.

	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 1. Face recognition rates with FRGC v1, ORL, CK+.

	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**

Table 2. Face recognition rates with FEI, GT, AR.

	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**

Table 3. Face recognition rates with FERET.

Our approach gives the best recognition rates with FRGC, GT, FERET (100 images), ORL (with others approaches); the 2nd best results (RRC is the 1st) with the others database. It demonstrates the effectiveness of RSN based FR.

4.1 Running time comparison

The running time of all methods (NN, SVM_OAA, SVM_DAG, BHDT, MetaFace, RKR, RRC, CRC, our RSN), is evaluated using STSS based FR experiments with three of the databases for reasons of space. We have used 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.

	Database		
Method	FRGC v1	ORL	AR
NN	-	0.7703	-
SVM_OAA	0.6415	0.0133	0.1680
SVM_DAG[28]	0.0610	0.0113	0.0433
BHDT [29]	**0.0109**	0.0019	**0.0055**
MetaFace [4]	0.5042	0.6500	0.3153
RKR [30]	0.0160	0.0160	0.0150
RRC [8]	0.0867	0.0102	0.0405
CRC [6]	**0.0027**	**7.7e-04**	**0.0038**
RSN (our)	0.0784	**0.0094**	0.0419

Table 4.The average running time (seconds).

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, which promotes sparsity and is especially able to extract directional information. In our approach, these features are used to control the trade-off between fidelity to the gallery and 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.

ACKNOWLDGEMENTS

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. DL acknowledges support of grants NSF-DMS 1008900, 100579.

6. 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*, 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 FR,” *ICIAP*, 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.