Shearlet Network-based Sparse Coding Augmented by Facial Texture Features for Face Recognition

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. One open challenge in face recognition (FR) is the single training sample per subject. This paper addresses this problem through a novel approach called Shearlet Network (SN) which takes advantage of the sparse representation (SR) properties of shearlets in biometric applications, specifically, for face coding and recognition. Shearlets are derived from wavelets with composite dilations, a method extending the traditional wavelet approach by allowing for the construction of waveforms defined not only at various scales and locations but also at various orientations. The contributions of this paper are the combination of the power of multi-scale representation with a unique ability to capture geometric information to derive a very efficient representation of facial templates, and the use of a PCA-based approach to design a fusion step by a refined model of belief function based on the Dempster-Shafer rule in the context of confusion matrices. This last step is helpful to improve the processing of facial texture features. We compared our new algorithm (SNPCA) against SN, a wavelet network (WN) implementation and other standard algorithms. Our tests, run on several face databases including FRGC, Extended Yale B database and others, show that this approach yields a very competitive performance as compared to wavelet networks (WN), standard shearlet and PCA-based methods.

Keywords. — Shearlet, Shearlets Network, Sparse Coding, Face Recognition

1   Introduction

The current trend in face recognition (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 [6] 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. Moreover, to deal with the challenges in practical FR system, active shape model and active appearance model [1] were developed for face alignment; LBP [2] and its variants were used to deal with illumination changes; and Eigenimages [3][4] and probabilistic local approach [5] were proposed for FR with occlusion.

Recently, FR via sparse representation based classification (SRC) [7][8] has received more attention as a powerful tool for statistical modeling, sparse representation (or sparse coding) and has been successfully applied to face processing applications. Although SRC performs well when the set of training images is sufficiently large, it is still inadequate for many real world applications where only a single sample per subject is available.

On the other hand, a new generation of multiscale methods has emerged in recent years which go far beyond traditional wavelets and have been shown to have the potential to provide better performing algorithms in a variety of biometric-based approach. The shearlet system is one notable example of these new classes of multiscale systems, which has the ability to capture anisotropic information very efficiently, outperforming traditional wavelets. One particularly appealing feature of shearlets is that they combine a multiscale framework which is particularly effective to capture the geometry of a face, together with a simple mathematical construction which can be associated to a multiresolution analysis and enables fast numerical processing. We recall that some recent work using shearlets for FR has recently appeared in [9][10].

In this work, we present a new method for FR, called *shearlet network (SN),* which is refinement of the classical wavelet network (WN). In this approach, faces will be approximated by a linear combination of weighted sum of shearlets and the weights will be used in the on-line recognition stage to calculate the similarity score between Gallery and probe face. We use PCA-based approach for a fusion step with SN to provide more depth to the facial texture appearance of the face; this fusion is achieved via a model of belief function which will be explained below.

The rest of this paper is organized as follows. In Sec. 2, we briefly describe some background on shearlets. Sec. 3 presents the proposed face coding and recognition method. In Sec. 4, the experimental results of the proposed algorithm are demonstrated and compared with other algorithms. Finally, Sec. 5 concludes this paper.

2   Background on Shearlets

The shearlet representation has emerged in recent years as one of the most effective frameworks for the analysis and processing of multidimensional data [11]. The shearlet approach is derived from the theory of wavelets with composite dilatation, a method providing a general framework for the construction of waveforms defined not only at various scales and locations, as traditional wavelets, but also at various orientations and with different scaling factors in each coordinate. As shown in several publications (e.g., [12][13][14]) shearlets are particularly effective in a number of image processing application, such as denoising and feature extraction, where it is important

to capture the geometric information efficiently. As a generalization of the traditional wavelet approach, the *continuous shearlet transform* [15] is defined as the mapping:

$$SH_\psi(a,s,t) = \left\langle f, \psi_{a,s,t} \right\rangle, a > 0, s \in \mathbb{R}, t \in \mathbb{R}^2 \tag{1}$$

where, $\psi_{ast}(x) = \left| \det M_{as} \right|^{-\frac{1}{2}} \psi(M_{as}^{-1}(x-t))$, and $M_{as} = \begin{pmatrix} a & s \\ 0 & \sqrt{a} \end{pmatrix}$. The shearlet transform

is a function of three variables: the *scale a*, the *shear s* and the *translation t*. By choosing the generator function appropriately, one can construct a collection of functions $\psi_{ast}(x)$, called shearlets, which are well-localized waveforms at various scales, orientations and locations. One of the main properties of the Continuous Shearlet Transform is its ability to describe very precisely the geometry of the singularities of a 2-dimensional function f. By sampling the Continuous Shearlet Transform $SH_\psi(a,s,t)$ on an appropriate discrete set we obtain a discrete 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:

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

Figure 1 illustrates the two-level shearlet decomposition of a face image from Yale database, where the first-level and and the second-level decomposition generates 8 subbands corresponding to the different directional bands.
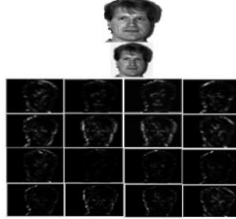
Fig. 1. The top image is the original Face image. The image below the top image contains the approximation shearlet coefficients. Images of the detail shearlet coefficients are shown below.

3 The Proposed Approach

The proposed approach is organized as follows. The Gallery faces are approximated by a shearlet network (SN) to produce a compact biometric signature. It is this signature, constituted by shearlets and their weights, which will be used to match a Probe with all faces in the Gallery. In the recognition stage, the Test (Probe) face is projected on the shearlet network (SN) of the Gallery face and new weights specific to this face are produced. The family of shearlets remains then unchanged (that of the

Gallery face). Finally, a similarity score is produced by comparing the weights of the Gallery face with the weights of the test face. In another section of the pipeline, PCA is used for FR [28] in order to generate the eigenfaces from Gallery dataset which will be operated for the projection step of Probe dataset. Those two matchers, SN and PCA, will generate two confusion matrixes which are used to fill two belief mass matrixes with belief functions. Finally, we combined those matrixes using a conjunction fusion rule. The pipeline of all these stages is illustrated in Figure 2.
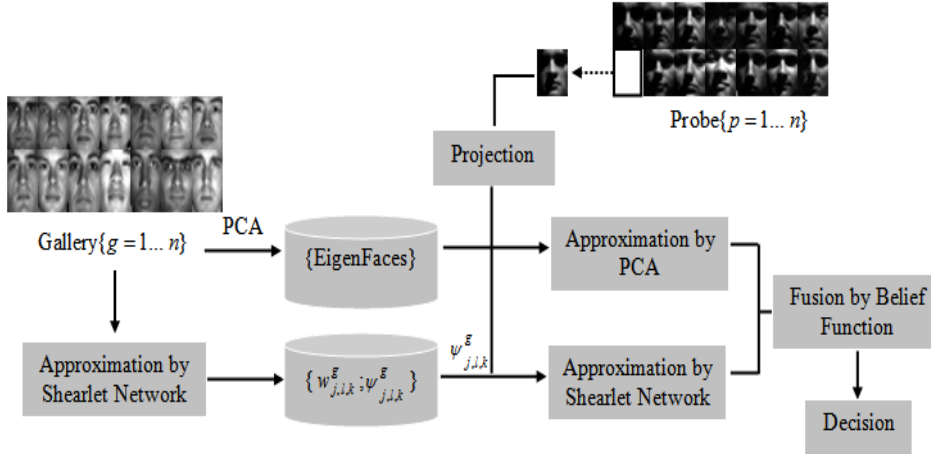

Fig. 2. Framework of *2D* face recognition by Shearlet Networks and PCA.

3.1 Approximation by Shearlet Network (SN)

A SN is a combination of the RBF neural network and the shearlet decomposition. The SN algorithm approximates a 2D face image f using a linear combination of shearlet functions in the network that are multiplied by corresponding weights according to:

$$\tilde{f} = \sum_{j,l,k} w_{j,l,k} \psi_{j,l,k} \tag{3}$$

where f is the face image, $\tilde{f}$ is the face image approximation, $w_{j,l,k}$ are the weights and $\psi_{j,l,k}$ are the shearlet functions. An important part in the design of this method is the weights optimization. For an image f, an energy (error) function is specified which is minimized through a learning process with respect to the desired SN parameters.

$$E = \min_{i} \left\| f_i - \sum_{j,l,k} w_{j,l,k} \psi_{j,l,k} \right\| \tag{4}$$

The mother shearlet used in our work to construct the family $\left\{\psi_{j,l,k}\right\}$ is the second derivate of the beta function. Details regarding beta function can be found in [16].

The algorithm used for the FR by SN is based on the dyadic sampling necessary for building the shearlet candidates library to use in the network and the frames theory for the weights calculations [17]. By sampling the continuous shearlet transform $SH_{\psi}(a,s,t)$ on an appropriate discretization lattice for the scaling, shear, and translation parameters (a,s,t), one obtains a discrete transform which is associated to a Parseval (tight) frame for $L^2(\mathbb{R}^2)$. Indeed, we obtain a discrete system of shearlets $\left\{\psi_{j,l,k}\right\}$, for $j,l \in \mathbb{Z}, k \in \mathbb{Z}^2$ which, for an appropriate generator [11], satisfies the reproducing formula:

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

In the optimization stage, a shearlet coefficient from the library is processed through the hidden layer of the network and used to update the weights. The low frequency shearlets associated with the coarse approximation level of the image to be analyzed are processed first; the high frequency shearlets are used next to refine the approximated image. In order to know if a shearlet (n) will be an activation function of a new "neuron", we must verify if it is a linear combination of the other (n-1) shearlets of the network. In other words, it must be linearly independent with respect to the other (n-1) shearlets. The calculation of the weights connection in every stage is obtained by projecting the image to be analyzed on a family of shearlets. Note however that the beta shearlet (shearlets whose generator is a beta function) are not necessarily orthogonal. Hence, for a given family of shearlets it is not possible to calculate the weights by direct projection of the function f. At every stage of the optimization process we need to know the dual family of the shearlets forming our shearlet network. The dual shearlet family is calculated by the formula:

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

It is clear that the quality of reconstructed image increases with the number of shearlets used as illustrated in Figure 3.
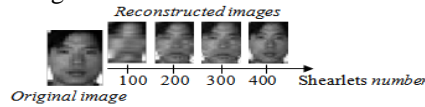
Fig. 3. Example of a *2D* face and the corresponding reconstructed images by beta shearlet networks.

3.2 Fusion by Belief Function

A step to normalize similarity matrices is essential, before combination, to generate meaningful distances from the two matchers: SN and PCA. In this paper, we use the Min-Max normalization, which has the following formula:

$$d' = \frac{d - d_{MIN}}{d_{MAX} - d_{MIN}} \tag{7}$$

where d_{MAX}, d_{MIN}, d, d' denote respectively the maximum, the minimum, the original and the normalized distances. In the second step, we use a Bayesian assignment [18] to compute the belief mass functions values from the confusion matrix. Finally, a conjunctive combination rule [19], defined by the following equation (8) is used to fuse confusion matrices.

$$m_{SN \cap PCA}(A) = m_{SN} \cap m_{PCA}(A) = \sum_{B \cap C = A} m_{SN}(B)\, m_{PCA}(C), \quad \forall A \subseteq \Omega \tag{8}$$

With m_{SN} and m_{PCA} the two Belief mass functions associated to SN and PCA.

4 Experimental Results

We run our experiments using standard benchmark face databases to evaluate the performance of the proposed approach. We used the Extended Cohn-Kanade (CK+) [20], Georgia Tech [21], FEI [22], Extended YALE B [23] and FRGC v1 [24] face databases. All the images are copped except the Extended Cohn-Kanade (CK+) database and resized to 27×32. In this paper, we chose to select **randomly** the face image both for Gallery and Probe dataset. A description of the databases used in Table 1 and examples of subjects (Gallery and Probe) in Figure 4.

Database	CK+	Georgia Tech	FEI	Yale B	FRGC v1
Database Size	593	750	2800	16128	4003
Subject number	123	50	200	38	152

Table 1. Description of the different databases used.
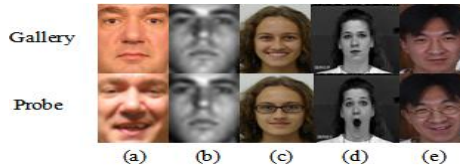

Fig. 4. A subject from Gallery and Probe with different databases. (a) FRGC. (b) Yale B. (c) FEI. (d) Extended CK+. (e) Georgia Tech.

Most work in the literature uses multiple training samples per subject. However, an emerging tendency in FR is to use **Single Training Sample per Subject (STSS)** [25][26] [27], which is a more challenging problem. By applying such strategy to the images from the databases indicated above, we obtain a similarity matrix of 38x38 comparisons for Extended YALE B, 123 x123 comparisons for Extended Cohn Kanade (CK+), 152 x152 comparisons for FRGC v1, 200 x200 comparisons for FEI and 50 x50 comparisons for Georgia Tech, which significantly reduce the computational complexity of the algorithm compared to traditional multiple training samples per subject. Hence, the similarity values located in the diagonal of the matrix are intra-class (the same person) and the others are inter-class (different persons) or imposter access. In this paper, we considered the problems of face coding, recognition and authentication. Results are provided in next subsections.

4.1 Face Coding (FC) Evaluation:

We use some standard quality measurement tool, such as the Peak Signal to Noise Ratio (PSNR) and the Normalized Cross Correlation (NCC). We compared the two approaches, e.g, SN vs. WN, and to the standard shearlet representation. We test with 3 face images for the same person from YALE database. The SN (table 2) is found to outperform the two   other   methods that we considered.

Method	Test1		Test2		Test3	
	PSNR	NCC	PSNR	NCC	PSNR	NCC
Shearlet	27.1227	0.9913	26.1698	0.9853	26.1582	0.9861
WN	36.9600	0.9967	33.4986	0.9942	32.5192	0.9931
SN	*37.7788*	*0.9973*	*36.2613*	*0.9970*	*35.8314*	*0.9962*

Table 2. PSNR and NCC of different methods. (**higher is better**)

In order to have the same experimental protocol and conditions for FC we chose a shearlets number (SN) equal to wavelets number (WN) equal to filters number of shearlet-based approach. We have also plotted the reconstruction error versus the number of employed wavelets (WN) and shearlets of (SN). Details for these results are not reported here for reasons of space.

4.2 Face Recognition and Authentication using STSS

The following experiments illustrate our results regarding authentication and recognition tasks. In these experiments, similarity matrices and mass belief functions matrices are produced, rank-one recognition rates (Rank-1) are computed and evaluation Receiver Operator Characteristic (ROC) curves are plotted. Table 3 summarize the rank-one recognition rates (Rank-1). SNPCA, the proposed approach, presents a Rank-1 which outperforms SN, WN, PCA approaches with whole the face databases.

Method	Extended CK+	FEI	Georgia Tech	Yale B	FRGC v1
PCA	80.49	59.00	6.00	51.50	12.50
WN	91.87	78.50	16.00	86.84	23.03
SN	95.94	79.50	26.00	97.37	34.87
SNPCA	**96.75**	**85.50**	**30.00**	**97.37**	**43.50**

Table 3. Rank-1 with different subsets database (%).

The authentication experiments are illustrated in Figure 5, one image for each one of the 5 face database considered. Each figure shows an ROC curve, where the x-axis represents the False Accept rate. The figures show that SNPCA provides consistently the best performance.
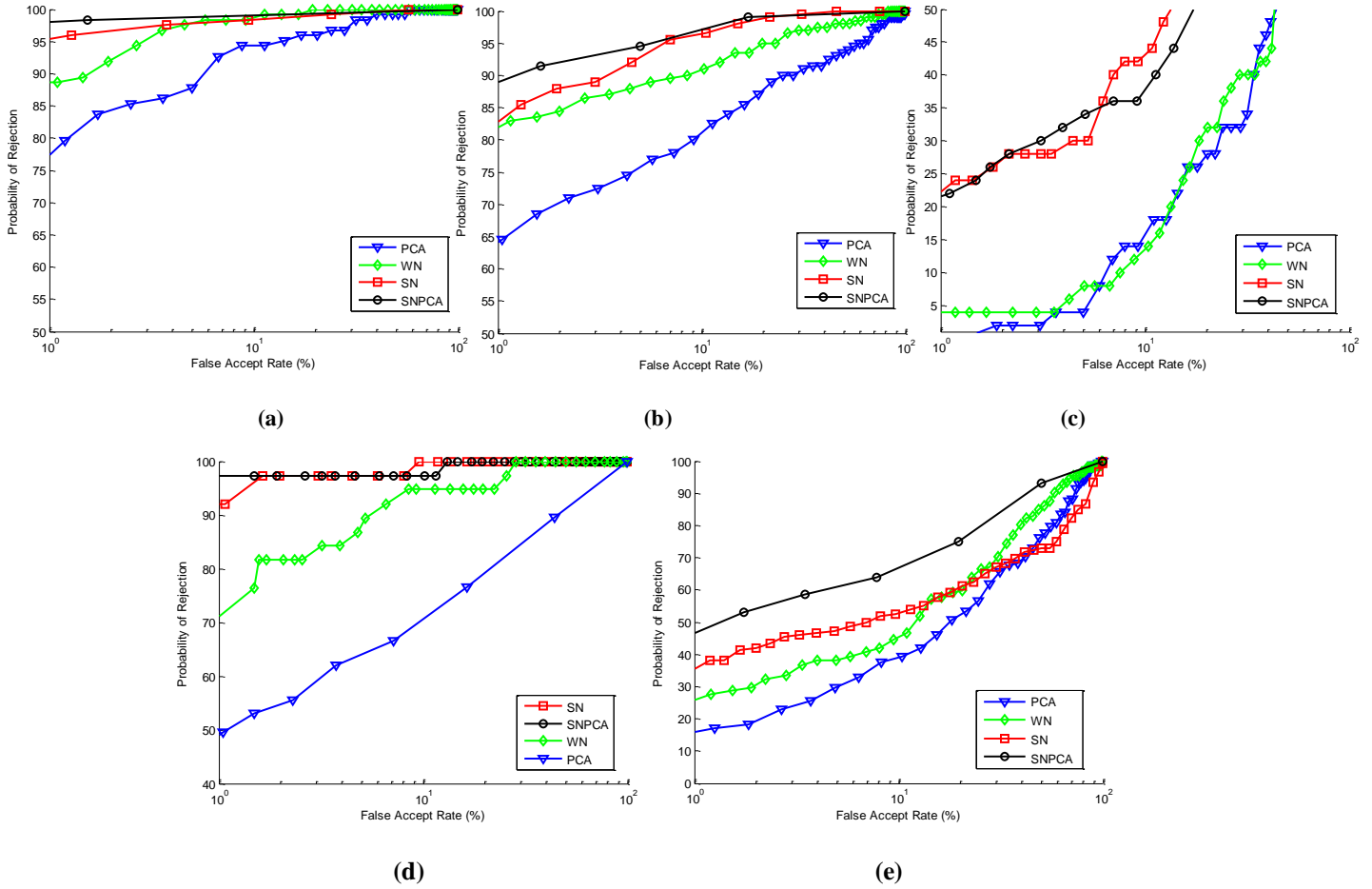
Fig.5. ROC curves with (a) Extended Cohn-Kanade (CK+) database, (b) FEI database, (c) Georgia Tech database , (d) Extended Yale B database, (e) FRGC v1 database.

From Table 2, we notice that SN gives better results in terms of PSNR and NC as compared to WN and shearlet coding. This is due to the superior ability of SN to capture the geometry of the data. Also for FR applications the SN performs better than PCA and WN as shown by Table 3 and Figure 5. Moreover, the combination of PCA and SN (SNPCA) using belief functions improves considerably the recognition rates. This shows that features (texture feature for PCA and anisotropic feature for SN) from matchers based on different approaches can be effectively combined.

5   Conclusion

This paper presents a novel sparse coding (SC) model for robust face recognition (FR) called Shearlet Network (SN). This method was performed via fusion step by PCA-based approach to provide more depth to facial texture feature, using a refined model of belief function based on the Dempster-Shafer rule. One important advantage of SNPCA is its robustness to various types of challenging scenarios of FR (pose, expression and illumination). This paper focussed on the challenging problem of a single training sample per subject. In order to illustrate and validate our approach we used the Extended Cohn-Kanade (CK+), Georgia Tech, FEI, Extended YALE B and FRGC v1 face databases. Our results show that our approach is very competitive in terms of recognition rate with respect to other standard and state-of-the-art methods.

6   References

1. A. Lanitis, C.J. Taylor, and T.F. Cootes, "Automatic Interpretation and Coding of Face Images Using Flexible Models," *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol. 19, no. 7, pp. 743-756, 1997.
2. T. Ahonen, A. Hadid, and M. Pietikainen, "Face description with local binary patterns: Application to face recognition," *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol. 28, no. 12, pp.2037–2041, 2006.
3. A. Leonardis and H. Bischof, "Robust recognition using eigenimages," *Computer Vision and Image Understanding*, vol. 78, no. 1, pp. 99-118, 2000.
4. S. Chen, T. Shan, and B.C. Lovell, "Robust face recognition in rotated eigenspaces," *Proc. Int'l Conf. Image and Vision Computing New Zealand*, 2007.
5. A.M. Martinez, "Recognizing Imprecisely localized, partially occluded, and expression variant faces from a single sample per class," *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol. 24, no. 6, pp. 748-763, 2002.
6. Lu, J., Plataniotis Kostantinos, N., Venetsanopoulos Anastasios, "Face recognition using LDA-based algorithms". *IEEE Trans. Neural Networks* 14 (1), 195–200. 2003.
7. M. Yang, L. Zhang, J. Yang and D. Zhang, "Robust sparse coding for face recognition," *Proc. IEEE Conf. Computer Vision and Pattern Recognition*, 2011.
8. M. Yang, L. Zhang, J. Yang and D. Zhang, "Regularized Robust Coding for Face Recognition." *IEEE Trans. on Image Processing*, 2013.
9. W. R. Schwartz and R. D. da Silva and L. S. Davis and H. Pedrini. "A Novel Feature Descriptor Based on the Shearlet Transform". *IEEE International Conference on Image Processing*, pp. 1033-1036, 2011.

10. W. R. Schwartz and H. Guo and J. Choi and L. S. Davis. "Face Identification Using Large Feature Sets". *IEEE Transactions on Image Processing*, vol. 21, n. 4, pp. 2245-2255, 2012.

11. P. S. Negi and D. Labate, "3D discrete shearlet transform and video processing*", IEEE Trans. Image Process*. **21**(6) pp. 2944-2954, 2012.

12. G. R. Easley, and D. Labate, "Critically sampled wavelets with composite dilations", *IEEE Trans. Image Process*., **21** (2) pp. 550-561, 2012.

13. G. R. Easley, D. Labate and V. Patel, "Directional multiscale processing of images using wavelets with composite dilations", to appear *in J. Math. Imag. Vision*, 2013.

14. 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.

15. G. Kutyniok and D. Labate, "Resolution of the wavefront set using continuous shearlets", *Trans. Amer. Math. Soc.* **361** pp. 2719-2754, 2009.

16. 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.

17. M. Zaied, O. Jemai and C. Ben Amar, "Training of the Beta wavelet networks by the frames theory: Application to face recognition", *the international Workshops on Image Processing Theory, Tools and Applications*, Tunisia, November 2008.

18. D. Mercier, G. Cron, T. Denoeux and M. Masson. "Fusion of multi-level decision systems using the Transferable Belief Model", *Proc. FUSION, Philadelphia, PA, USA*, July 2005.

19. T. Denoeux. "Conjunctive and Disjunctive Combination of Belief Functions Induced by Non Distinct Bodies of Evidence". *Artificial Intelligence*, Vol. 172, pages 234–264, 2008.

20. Lucey, P., Cohn, J. F., Kanade, T., Saragih, J., Ambadar, Z., & Matthews, I. "The Extended Cohn-Kande Dataset (CK+): A complete facial expression dataset for action unit and emotion-specified expression", *Proc. CVPR,Workshop H CBA*, 2010.

21. Georgia Tech Face Database, ftp://ftp.ee.gatech.edu/pub/users/hayes/facedb/.

22. 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

23. A. S. Georghiades and P. N. Belhumeur, "From Few to Many: Illumination Cone Models for Face Recognition under Variable Lighting and Pose", *IEEE Trans. on. PAMI*, vol.23, no.6, pp.643-660, June 2001

24. 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", *Proc. CVPR05*, no. 1, pp. 947-954, June 2005

25. Meng Yang, Lei Zhang, and David Zhang, "Efficient Misalignment Robust Representation Representation for Real-Time Face Recognition Classification," *Proc. ECCV* 2012.

26. P. Zhu, L. Zhang, Q. Hu, Simon and C.K. Shiu, "Multi-scale Patch based Collaborative Representation for Face Recognition with Margin Distribution Optimization," *Proc. ECCV* 2012.

27. Y. Su, S.G. Shan, X.L. Chen, and W. Gao, "Adaptive Generic Learning for Face Recognition from a Single Sample per Person," *Proc. CVPR*, 2010.

28. A. Mian, M. Bennamoun, and R. Owens, "2D and 3D Multimodal Hybrid Face Recognition," *Proc. ECCV*, pp. 344–355, 2006.