FACE, GENDER AND RACE CLASSIFICATION USING MULTI-REGULARIZED FEATURES LEARNING

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

[1]Research Groups on Intelligent Machines, University of Sfax, BP 1173, Sfax 3038, Tunisia
[2]Department of Mathematics, University of Houston, Houston, TX 77204, USA

ABSTRACT

This paper investigates a new approach for face, gender and race classification, called multi-regularized learning (MRL). This approach combines ideas from the recently proposed algorithms called multi-stage learning (MSL) and multi-task features learning (MTFL). In our approach, we first reduce the dimensionality of the training faces using PCA. Next, for a given a test (probe) face, we use MRL to exploit the relationships among multiple shared stages generated by changing the regularization parameter. Our approach results in convex optimization problem that controls the trade-off between the fidelity to the data (training) and the smoothness of the solution (probe). Our MRL algorithm is compared against different state-of-the-art methods on face recognition (FR), gender classification (GC) and race classification (RC) based on different experimental protocols with AR, LFW, FEI, Lab2 and Indian databases. Results show that our algorithm performs very competitively.

***Index Terms*—** Multi-Regularized Feature Learning, Face Recognition, Gender and Race Classification

1. INTRODUCTION

During the last decade, sparse representations have emerged in pattern classification in many high-performing classification algorithms. In particular, Wright *et al.* have introduced the Sparse Representation Classification (SRC) scheme [1], [5] which casts the recognition problem as a problem of classifying different candidates as multiple linear regressions. 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. A very simple yet very efficient face classification scheme, called Collaborative Representation (CR) based classification with regularized least squares (CRC_RLS) was proposed [2]. Both this method and SRC have shown that the notion of sparsity can play an essential role to improve discrimination among different classes of objects and classification.

Another successful approach is the Regularized Robust Coding (RRC) approach [3] that works by robustly regressing a given signal using regularized regression coefficients. By assuming that the coding residual and the coding coefficient are respectively independent and identically distributed, RRC seeks for a maximum a-posteriori solution of the coding problem. An iteratively reweighted regularized robust coding algorithm was proposed to solve the RRC model efficiently.

The Multi-Task Learning (MTL) originally proposed by Caruana was shown to produce remarkable results in the problem of gender and face recognition [4]. MTL attempts to learn classifiers for multiple tasks jointly and works under the assumption that all tasks should share some common features. Many variants of MTL have appeared in the literature, notably the multi-stage multi-task feature learning (MSMTFL) recently introduced by Gong et al. [5]. Furthermore, a non-convex formulation for multi-task sparse feature learning was defined based on a novel non-convex regularization, called capped-ℓ_1, ℓ_1 regularized model for multi-task feature learning. A similar approach was recently proposed by T. Zhang [6], [7], where a multi-stage convex relaxation scheme is used for solving problems with non-convex objective functions.

In this paper, we propose a new method for face, gender and race classification that we call multi-regularized learning. MRL and can be considered as a refinement of MTFL. The key idea of our MRL approach is the use of a multi-regularization algorithm based on multi-stage convex relaxation to transform non-convex optimization into a convex relaxation.

The rest of the paper is organized as follows. We present our proposed approach in section 2. In section 3, we describe our experimental results and comparison against other state-of-the–art algorithms. We conclude the paper with some remarks about future work.

2. OVERVIEW OF THE PROPOSED APPROACH

The proposed MRL scheme for FR, GC and RC is defined as a cascade of a dimensionality reduction module followed by a classification module. This scheme is implemented using a multi-stage regularization stage as shown in Figure 2.

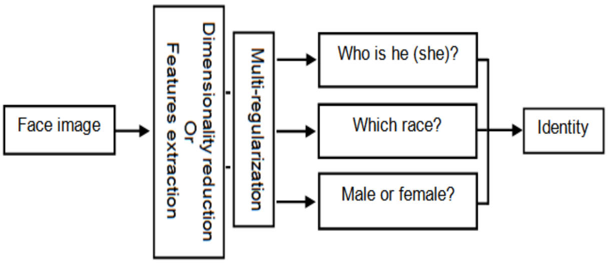

Figure 1. MRL face, gender and race classification schema.

FR, GC and RC are by nature ill-posed problems and a classical way to address the ill-posedness is through regularization theory [8][9]. In practice, rather than looking for an exact solution, it is sufficient to compute an approximate one. Let y be a normalized test face and X is a matrix representing a gallery of faces. In our approach, we consider the classical regularization method known as Lasso [10] that is given by:

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

Here $\lambda > 0$ is an appropriately chosen regularization parameter. Given that noisy image y, $\hat{w}$ is regularized by [7]:

$$\hat{w} = \arg\min_{w} \left[\frac{1}{n} \|Xw - y\|_2^2 + \lambda \sum_{j=1}^{d} g\left(\left|w_j\right|\right) \right] \quad (2)$$

where $g\left(\left|w_j\right|\right)$ is a regularization function. $g(u) = \min(u, \theta)$ is an example of regularizing function which corresponds to the capped-L_1 regularization (θ is a Threshold parameter).

We use convex relaxation to guarantee uniqueness of the solution and efficient computation of the optimization problem. The multi-stage convex relaxation [6] is defined as:

$$\hat{w}^{(\ell)} = \arg\min_{w \in R^d} \left[\frac{1}{n} \|Xw - y\|_2^2 + \sum_{j=1}^{d} \lambda_j^{(\ell-1)} \left|w_j\right| \right] \quad (3)$$

where X is an $n \times d$ matrix, y an $n \times 1$ matrix, $\lambda_j^{(0)} = \lambda$, $\ell = 1, 2, \ldots$ and $j = 1, \ldots, d$.

The proposed MRL makes use of equation (2), (3) under the assumption that $g(u) = \min(u, \theta)$ for the problem of FR. Since optimization problems in equations 1 and 2 are non-convex, the global optimum is difficult to find. Additionally, local minima analysis using the gradient descent method is hard to perform. To address this task, several successful optimization methods such as Lasso, iteratively reweighted least square and recursive least squares have been proposed [3], [11], [12], [13].

The proposed MRL algorithm uses the following simple method to find the optimal solution of the recognition step. First we compute, at every stage:

$$W_x = w_{init} \times X$$

$$W_y = w_{init} \times y$$

then we compute W_s as follows: (4)

$$W_s = W_x \times \lambda^{(l)} \times \sum \min\left(\sum\left(\left|W_y\right|, 2\right), \theta\right) \backslash W_y$$

By doing this we obtain the weight:

$$w = W_x' \times W_s \quad (5)$$

Note that the operator \ is the matrix left division for the linear problem $W_s X = y$, where y is a normalized test face, X is a matrix representing a gallery of faces. The initial value of the weight w_{init} is chosen using the logistic function [14], where the logistic function is:

$$f(x) = \frac{c}{1 + a\exp^{-bx}} \quad (6)$$

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

$$w_{init} = 1/(1 + 1/\exp(-\mu e_{init}{}^2 + \mu\delta)) \quad (7)$$

where μ and δ are positive scalars and e_{init} is the initial residual given by:

$$e_{init} = (y - mean(X))^2 \quad (8)$$

where X is the aligned gallery faces (an $n \times d$ matrix) and y is a normalized test face (an $n \times 1$ matrix). The residual e_{init} and the weight w_i could be updated after optimization. The choice of the parameter $\lambda_j^{(\ell)}$ is given in equation (9) [6]: $\lambda_j^{(\ell)} = \lambda I(\left|w_j^{(\ell)}\right| \le \theta)$ (9)

where

$$j = 1, \ldots, d \text{ and } \lambda = \tau\sigma\sqrt{\ln(d)/n} \quad ; \quad \theta = \mu\lambda \quad (10)$$

with $\tau = 1, 2, 4, 8, 16, \ldots$ And $\mu = 0.5, 1, 2, \ldots$ ($\sigma = 1e-4$).

The dimensionality reduction is made by PCA, note that our recent works use a new features extraction called shearlets network (SN) [22][23][24][25], wavelets network (WN) [26][27][28][29] [30] [31][32][33][34][35] and others features [36][37][38].

The MRL algorithm is summarized below. Let X be the aligned gallery of faces, y the normalized test face, $Iter$ the maximum of iterations chosen and Classnum is the class number (i.e. for GC two classes 1and 2).

Algorithm: MRL

Output: w; $Identity(y)$

1. Compute the residual e_{init}, refer to (8)
2. Compute μ and δ, refer to (11) and (12)
3. Compute w_{init}, refer to (7)
4. Initialize $\lambda_j^{(0)} = \lambda$ and $\theta = \mu\lambda_j^{(0)}$, refer to (□10)
5. For $\ell = 1,2,\ldots$
   For j = 1,…, *Iter*
   - Compute the MRL using formula (4):
     $W_s = MRL(X, y, w_{init}, \lambda_j^{(\ell-1)}, \theta)$
   - Compute w_i using the formula (5)
   - Update the residual e : $e = (Xw_i - y)^2$
   - $e_{init} = e$ (used for w_{init})
   - Update μ and δ
   - Update w_{init}, refer to formula (7)

   End
   - Update $\lambda_j^{(\ell)}$ and θ, refer to □(9) and □(10)
   - $w_{init} = w_i$

   End
   - $y_{rec} = X * w_i$
   - $w = w_i$
6. For k = 1,…, Classnum
   $error(k) = \left\| w^{1/2}(y - X_k w_k) \right\|_2^2$
   End
7. Identity(y) = argmin(error)

μ is given by:

$$\mu = \frac{0.6}{\delta} \tag{11}$$

δ is set to [3]:

$$\delta = \varphi\left(e_{init}\right)_k \tag{12}$$

where $\varphi\left(e_{init}\right)_k$ is the k^{th} largest element of the set $\left\{e_{init}^2(j), j = 1,2,\ldots,n\right\}$. A numerical study was performed on the parameters μ and δ to ensure that properties of the logistic function (i.e., inflection point) are met. Details of this study are not included due to space constraints.

3. EXPERIMENTAL RESULTS

In this section, we show our numerical experiments on benchmark face databases using our MRL algorithm for GC, FR and RC; the results are compared to many standard methods: CRC_RLS [2], SRC [1], SVM, LRC (linear regression classification) [19] and NN. Five face databases, including FEI [15], AR [16], LFW [17], Lab2 [18] and Indian [21], are used to compare the performances of the algorithms.

3.1. Gender classification

3.1.1. AR database

We have selected a non-occluded subset (14 images per subject) of AR [16] consisting of 50 male and 50 female subjects. Images of the first 25 males and 25 females were used for training, and the remaining images for testing. The images were cropped to 60×43. PCA was used to reduce the dimension of each image to 300. Comparing MRL to the CRC_RLS, SRC, SVM, LRC and NN, gives the results in Table 1. MRL accuracy was ranked second after CRC_RLS.

Table 1. GC results using the AR database.

MRL	CRC_RLS	SRC	SVM	LRC	NN
92.83%	**93.70%**	92.30%	92.40%	27.30%	90.70%

3.1.2. FEI database

There are 14 images for each of 200 individuals, for a total of 2800 images. The number of male and female subjects is exactly the same and equal to 100. The first nine images of all subjects are used in the training (1800 images, 900 per gender) and the remaining five images serve as testing images (1000 images, 500 per gender). The images were cropped to 60×43.

Figure 2. One subject from FEI database.

MRL is only compared to the CRC_RLS on different dimensionality. MRL outperforms CRC_RLS with all dimensionality. The results are summarized in Table 2.

Table 2. GC results using the FEI database.

Dim	30	54	120	300
CRC_RLS	88.20%	90.30%	91.40%	93.10%
MRL	**93.70%**	**93.40%**	**94.10%**	**94.00%**

3.2. Face recognition

3.2.1. AR database

In order to have the same experimental protocol as in [1] and [2], a subset on AR database (with only illumination and expression changes) that contains 50 male subjects and 50 female subjects was chosen from the AR dataset [16] in our experiments. For each subject, the seven images from

Session 1 were used for training and the other seven images from Session 2 were used for testing. The images were cropped to 60×43. MRL achieves the best result when the dimensionality is 30, 54 and 120 as shown in Table 3.

Table 3. FR results using the AR database.

Dim	30	54	120	300
NN	62.50%	68.00%	70.10%	71.30%
LRC	66.10%	71.00%	75.40%	76.00%
SVM	66.10%	69.40%	74.50%	75.40%
SRC	73.50%	83.30%	89.50%	93.30%
CRC_RLS	64.25%	80.50%	90.00%	**93.70%**
MRL	**76.29 %**	**84.86%**	**90.43%**	92.57%

3.2.2. Lab2 Database

The Lab2 database [18] contains visible light images and near-infrared images of the subjects. There are 50 subjects. Each subject provides twenty visible light face images (1000 images) and the same number of near-infrared face images. These images were acquired under four different illumination conditions (4 sessions). The face images also have variation in facial expression and pose. 5~15 sample visible light images from the first three sessions for training were chosen. 5 samples were selected from the fourth session for testing purposes. The face images are resized to 32×27.

Figure 3. One subject from visible light Lab2 database [18].

Table 4. FR results using the Lab2 database.

Method	*5*	*10*	*15*
SVM	24.40%	44.80%	65.60%
CRC _RLS	30.00%	47.20%	68.80%
MRL	**33.20%**	**48.00%**	**71.60%**

3.2.3. LFW Database: Uncontrolled Environment

The LFW database [17] contains images of 5,749 different individuals in unconstrained environment. The dataset is composed of 158 subjects from LFW-a. For each subject, 2-to-5 samples were randomly chosen for training and another 2 samples for test. The images were firstly cropped to 121×121 and then resized to 32×32 [20]. The FR rates on the LFW dataset are listed in Table 5. MRL results are noticeably superior as compared to the other methods.

Table 5. FR results using the LFW database.

Method	2	3	4	5
NN	11%	13.20%	14.70%	16.20%
LRC	-	-	-	-
SVM	21.52%	24.68%	30.38%	35.44%
SRC	26.80%	35.90%	41.10%	46.70%
CRC_RLS	26.80%	34.30%	40.40%	45.20%
MRL	**30.38%**	**37.34%**	**41.46%**	**46.84%**

3.3. Race classification

Generally, humans are divided into three races for the task of race classification: yellow, white and black. For the race classification, three races with different manners were selected: Brazilian (race 1), Indian (race 2) and Chinese (race 3) using three databases FEI (white), Indian female face (near black) and Lab2 visible light images (yellow) respectively. Figure 4 gives an example of samples from the three races.

Figure 4. Three races: two samples per race for the gallery, one sample per race for the test.

2748 images were used for training and 1294 images were used for testing, with the following distribution. For race 1, we used nine images of all subjects (1800 images) for training and the remaining five images serve as testing images (1000 images); for race 2, we used nine female images of 22 subjects (198 images) for training and two images for testing (44 images); for race 3, we used 15 visible light images (750 images) for training and five images (250 images) for testing. All the images were cropped and resized to 60×43. MRL results are compared to those of CRC_RLS and RRC [3]. Table 6 shows that MRL gives better results reaching almost 100% for all dimensionalities.

Table 6. Comparison of RC results

Dim	30	54	120
RRC	90.88 %	96.83%	94.20%
CRC_RLS	96.20%	96.91%	97.37%
MRL	**99.40%**	**99.77%**	**99.87%**

4. CONCLUSION

In this paper, we have illustrated the effectiveness of multi-regularized learning (MRL) approach for the task of face gender and race classification. In this approach, we have used a multi-stage regularization learning to share features and make the problem convex. This approach is further refined by a simple weighting strategy in the optimization step. The experimental results on controlled and uncontrolled face databases show that MRL is very competitive against other standard and state-of-the-art face recognition, gender and race classification methods.

AKNOWLEDGMENT

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, to acknowledge, also, Dr. Rafik Borji. D. L acknowledges partial support by NSF DMS 1005799 and DMS 1008900.

5. REFERENCES

[1] J. Wright, A. Y. Yang, A. Ganesh, S. S. Sastry and Y. Ma, “Robust face recognition via sparse representation,” *IEEE PAMI*, 31(2): pp. 210–227, 2009.

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

[3] 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.

[4] R. Caruana, “Multi-task learning,” *Machine Learning*, 28, (1), pp. 41–75, 1997.

[5] P. Gong, J. Ye and C. Zhang, “Multi-Stage Multi-Task Feature Learning,” *Journal of Machine Learning Research*, pp. 2979-3010 (14), 2013.

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

[7] T. Zhang, *Multi-stage convex relaxation for feature selection*, Bernoulli, 2012.

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

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

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

[11] R. Tibshirani, “Regression shrinkage and selection via the LASSO,” *Journal of the Royal Statistical Society B*, vol. 58, no. 1, pp. 267–288, 1996.

[12] I. Daubechies, R. Devore, M. Fornasier and C.S. Gunturk, “Iteratively Re-Weighted least squares minimization for Sparse Recovery,” arXiv: 0807.0575, 2008.

[13] Hayes Monson H, *Statistical Digital Signal Processing and Modeling*, Wiley, 1996.

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

[15] 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.

[16] A. Martinez, “ The ar face database,” CVC Technical Report 24 (1998).

[17] G.B. Huang, M. Ramesh, T. Berg and E. Learned-Miller, “Labeled faces in the wild: A database for studying face recognition in unconstrained environments,” Technical Report 07-49, University of Massachusetts, Amherst (October 2007).

[18] Y. Xu, A. Zhong, J. Yang and D. Zhang, “Bimodal biometrics based on a representation and recognition approach,” *Opt. Eng.* 50, 037202 (Mar 22, 2011); doi:10.1117/1.3554740, 2011.

[19] I. Naseem, R. Togneri, and M. Bennamoun, “Linear regression for face recognition,” *IEEE PAMI*, 32(11), pp. 2106-2112, 2010.

[20] Pengfei Zhu, Lei Zhang, Qinghua Hu and Simon C. K. Shiu, “Multi-scale Patch Based Collaborative Representation for Face Recognition with Margin Distribution Optimization, ” *ECCV* (1), pp. 822-835, 2012.

[21] Indian face database: www.cs.umass.edu / ~vidit / face database.

[22] 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*, pp. 611-620, 2013.

[23] 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, 2014.

[24] M. A. Borgi, D. Labate, M. El’Arbi and C. Ben Amar, “ShearFace: Efficient Extraction of Anisotropic Features for Face Recognition , ” *ICPR*, 2014.

[25] M. A. Borgi, D. Labate, M. El’Arbi and C. Ben Amar, “Sparse Multi-Regularized Shearlet-Network using Convex Relaxation for Face Recognition, ” *ICPR*, 2014.

[26] M. Zaied, S. Said, O. Jemai, and C. Ben Amar, “A novel approach for face recognition based on fast learning algorithm and wavelet network theory,” *IJWMIP International Journal of Wavelets, Multiresolution and Information Processing*, pp 923-945, 2011.

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

[28] 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*, pp. 1279-1319, 2011.

[29] S. Said, B. Ben Amor, C. Ben Amar and Mohamed Daoudi, “Fast and Efficient 3D Face Recognition using Wavelet Networks, ” *ICIP*, pp. 4153-4156, 2009.

[30] O. Jemai, 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*, pp. 111-130, 2011.

[31] M. Zaied, C. Ben Amar, and M. A. Alimi, “Beta Wavelet Networks for Face Recognition, ” *Journal of Decision Systems*, pp. 109-122, 2005.

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

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

[34] M. Othmani, W. Bellil, C. Ben Amar and M. A. Alimi, “A New Structure and Training Procedure for Multi-Mother Wavelet Networks, ” *International Journal of Wavelet Multiresolution and Information Processing, IJWMIP*, Vol. 8, N° 1, pp. 149-175, 2010.

[35] W. Ben Soltana, W. Bellil, C. Ben Amar and Mohamed Adel Alimi, “Multi Library Wavelet Neural Networks for 3D face recognition using 3D facial shape representation, ” *EUSIPCO*, pp. 55-59, 2009.

[36] H. Boughrara, L. Chen, C. Ben Amar and M. Chtourou, “Facial Expression Recognition based on Perceived Facial Images And Local Feature Matching, ” *ICIAP* , pp 591-600, 2013.

[37] H. Boughrara, M. Chtourou and C. Ben Amar, “MLP Neural Network Based Face Recognition System Using Constructive Training Algorithm, ” *ICMCS*, pp. 233-238, 2012.

[38] H. Boughrara, L. Chen, C. Ben Amar and M. Chtourou, “Face Recognition Under Varying Facial Expression Based on Perceived Facial Images and Local Feature Matching, ” *ICITeS*, pp. 1-6, 2012.