Robust Regression for Face Recognition

Imran Naseem , Roberto Togneri, Mohammed Bennamoun
The University of Western Australia
imran.naseem@ee.uwa.edu.au, roberto@ee.uwa.edu.au, bennamou@csse.uwa.edu.au

Abstract

In this paper we address the problem of illumination invariant face recognition. Using a fundamental concept that in general, patterns from a single object class lie on a linear subspace [2], we develop a linear model representing a probe image as a linear combination of class-specific galleries. In the presence of noise, the well-conditioned inverse problem is solved using the robust Huber estimation and the decision is ruled in favor of the class with the minimum reconstruction error. The proposed Robust Linear Regression Classification (RLRC) algorithm is extensively evaluated for two standard databases and has shown good performance index compared to the state-of-art robust approaches.

1. Introduction

Face recognition systems critically depend on manifold learning methods. A gray-scale face image of order $a \times b$ can be represented as an ab dimensional vector in the original *image space*. At the feature extraction stage, images are transformed to low dimensional vectors in a *face space*. The main objective is to find a basis function for this transformation, which could distinguishably represent faces in the face space. In the presence of noise, however it is supposed to be an extremely challenging task [1]. It follows from coding theory that iterative measurements are more likely to safely recover information in the presence of noise [13], therefore working in a low dimensional feature space maintaining the aspect of robustness is in fact an ardent problem in object recognition. A number of approaches have been reported in the literature for dimensionality reduction. In the context of robustness, these approaches have been broadly classified in two categories namely *generative/reconstructive* and *discriminative* methods [20]. Reconstructive approaches (such as PCA [22], ICA [25] and NMF [10], [11]) are reported to be robust for the problem related to missing and contaminated pixels, these methods essentially exploit the redundancy in the visual data to produce representations with sufficient reconstruction property. The discriminative approaches (such as LDA [3]), on the other hand, are known to yield better results in "clean" conditions [5] owing to the flexible decision boundaries. Apart from these traditional approaches, it has been shown recently that unorthodox features such as downsampled images and random projections can serve equally well. In fact the choice of the feature space may no longer be so critical [16], [17], [24]. What really matters is the dimensionality of the feature space and the design of the classifier.

In this research we propose a robust classification algorithm for the problem of face recognition in the presence of severe illumination variations. Samples from a specific object class are known to lie on a linear subspace [2], [3]. In our previous work [16], [17] we proposed to develop class specific models of the registered users thereby defining the task of face recognition as a problem of linear regression. In the work presented here, we extend our investigations to the problem of noise contaminated probes, where the inverse problem is solved using a novel application of the robust linear Huber estimation [9], [18] and the class label is decided based on the subspace with the most precise estimation.

The rest of the paper is organized as follows: The fundamental problem of robust estimation is discussed in Section 2 followed by the face recognition problem formulation in Section 3. Section 4 demonstrates the efficacy of the proposed approach for the problem of severely varying illumination. The paper finally concludes in Section 5.

2 The Problem of Robust Estimation

Consider a linear model

$$\mathbf{y} = \mathbf{X}\beta + \mathbf{e} \tag{1}$$

where the dependent or response variable $\mathbf{y} \in \mathbb{R}^{q \times 1}$, the regressor or predictor variable $\mathbf{X} \in \mathbb{R}^{q \times p}$, the vector

1051-4651/10 $26.00 © 2010 IEEE
DOI 10.1109/ICPR.2010.289

of parameters $\beta \in \mathbb{R}^{p\times 1}$ and error term $\mathbf{e} \in \mathbb{R}^{q\times 1}$. The problem of robust estimation is to estimate the vector of parameters $\hat{\beta}$ so as to minimize the residual

$$\mathbf{r} = \mathbf{y} - \hat{\mathbf{y}}; \quad \hat{\mathbf{y}} = \mathbf{X}\hat{\beta} \tag{2}$$

$\hat{\mathbf{y}}$ being the predicted response variable. In classical statistics the error term $\mathbf{e}$ is conventionally taken as a zero mean Gaussian noise [8]. A traditional method to optimize the regression is to minimize the least squares (LS) problem

$$\underbrace{\arg\min}_{\hat{\beta}} \sum_{j=1}^{q} r_j^2(\hat{\beta}) \tag{3}$$

where $r_j(\hat{\beta})$ is the j^{th} component of the residual vector $\mathbf{r}$. Several approaches to robust estimation have been proposed such as R-estimators and L-estimators. However M-estimators have shown superiority due to their generality and high breakdown point [8], [9]. Primarily M-estimators are based on minimizing a function of residuals

$$\hat{\beta} = \underbrace{\arg\min}_{\hat{\beta}\in\mathbb{R}^p} \left\{ F(\hat{\beta}) \equiv \sum_{j=1}^{q} \rho\left(r_j(\hat{\beta})\right) \right\} \tag{4}$$

where $\rho(r)$ is a symmetric function with a unique minimum at zero [9], [18]

$$\rho(r) = \begin{cases} \frac{1}{2\gamma} r^2 & \text{for } |r| \leq \gamma \\ |r| - \frac{1}{2}\gamma & \text{for } |r| > \gamma \end{cases} \tag{5}$$

γ being a tuning constant called the Huber threshold. Many algorithms have been developed for calculating the Huber M-estimate in Equation 4, some of the most efficient are based on Newton's method [14].

3 Robust Linear Regression Classification (RLRC) for Robust Face Recognition

Consider N number of distinguished classes with p_i number of training images from the i^{th} class such that $i = 1, 2, \ldots, N$. Each grayscale training image is of an order $a \times b$ and is represented as $\overset{(m)}{\mathbf{u}_i} \in \mathbb{R}^{a\times b}$, $i = 1, 2, \ldots, N$ and $m = 1, 2, \ldots, p_i$. Each gallery image is downsampled to an order $c \times d$ and transformed to a vector through column concatenation such that $\overset{(m)}{\mathbf{u}_i} \in \mathbb{R}^{a\times b} \rightarrow \overset{(m)}{\mathbf{w}_i} \in \mathbb{R}^{q\times 1}$, where $q = cd$, $cd << ab$. Each image vector is normalized so that the maximum pixel value is 1. Using the concept that patterns from the same class lie on a linear subspace [2], we develop a class specific model $\mathbf{X}_i$ by stacking the q-dimensional image vectors,

$$\mathbf{X}_i = [\overset{(1)}{\mathbf{w}_i} \ \overset{(2)}{\mathbf{w}_i} \ldots\ldots \overset{(p_i)}{\mathbf{w}_i}] \ \in \mathbb{R}^{q\times p_i}, \ \ i = 1, 2, \ldots, N \tag{6}$$

ach vector $\overset{(m)}{\mathbf{w}_i}$, $m = 1, 2, \ldots, p_i$, spans a subspace of $\mathbb{R}^q$ also called the column space of $\mathbf{X}_i$. Therefore at the training level each class i is represented by a vector subspace, $\mathbf{X}_i$, which is also called the *regressor* or *predictor* for class i. Let z be an unlabeled test image and our problem is to classify z as one of the classes $i = 1, 2, \ldots, N$. We transform and normalize the grayscale image z to an image vector $\mathbf{y} \in \mathbb{R}^{q\times 1}$ as discussed for the gallery. If $\mathbf{y}$ belongs to the i^{th} class it should be represented as a linear combination of the training images from the same class (lying in the same subspace) i.e.

$$\mathbf{y} = \mathbf{X}_i\beta_i + \mathbf{e} \ , \ \ i = 1, 2, \ldots, N \tag{7}$$

where $\beta_i \in \mathbb{R}^{p_i\times 1}$. Given that $q \geq p_i$, the system of equations in equation 7 is well-conditioned and β_i is estimated using robust Huber estimation as discussed in Section 2 [9], [18]

$$\hat{\beta}_i = \underbrace{\arg\min}_{\hat{\beta}_i\in\mathbb{R}^{p_i}} \left\{ F(\hat{\beta}_i) \equiv \sum_{j=1}^{q} \rho\left(r_j(\hat{\beta}_i)\right) \right\} \ , \ \ i = 1, 2, \ldots, N \tag{8}$$

where $r_j(\hat{\beta}_i)$ is the j^{th} component of the residual

$$\mathbf{r}(\hat{\beta}_i) = \mathbf{y} - \mathbf{X}_i\hat{\beta}_i \ , \ \ i = 1, 2, \ldots, N \tag{9}$$

The estimated vector of parameters, $\hat{\beta}_i$, along with the predictors $\mathbf{X}_i$ are used to predict the response vector for each class i:

$$\hat{\mathbf{y}}_i = \mathbf{X}_i\hat{\beta}_i, \ \ i = 1, 2, \ldots, N \tag{10}$$

We now calculate the distance measure between the predicted response vector $\hat{\mathbf{y}}_i$, $i = 1, 2, \ldots, N$ and the original response vector $\mathbf{y}$,

$$d_i(\mathbf{y}) = \|\mathbf{y} - \hat{\mathbf{y}}_i\|_2 \ , \ \ i = 1, 2, \ldots, N \tag{11}$$

and rule in favor of the class with minimum distance i.e.

$$\underbrace{\min}_{i} d_i(\mathbf{y}), \ \ i = 1, 2, \ldots, N \tag{12}$$

Table 1. Details of the subsets for Yale Face Database B with respect to light source directions.

Subsets	1	2	3	4	5
Light angle (degrees)	$0-12$	$13-25$	$26-50$	$51-77$	>77
Number of images	70	120	120	140	190

Table 2. Recognition Results for Yale Face Database B

Methods	Subset 3	Subset 4	Subset 5
No Normalization [23]	89.20%	48.60%	22.60%
Histogram Equalization [23]	90.80%	45.80%	58.90%
Linear Subspace [7]	100.00%	85.00%	N/A
Cones-attached [7]	100.00%	91.40%	N/A
Cones-cast [7]	100.00%	100.00%	N/A
Gradient Angle [4]	100.00%	98.60%	N/A
Harmonic Images [26]	99.70%	96.90%	N/A
Illumination Ratio Images [27]	96.70%	81.40%	N/A
Quotient Illumination Relighting [21]	100.00%	90.60%	82.50%
9PL [12]	100.00%	97.20%	N/A
Method in [23]	100.00%	99.82%	98.29%
RLRC	**100.00%**	**100.00%**	**100.00%**

4 Experimental Results

4.1 Yale Face Database B

Yale face database B consists of 10 individuals with 9 poses incorporating 64 different illumination alterations for each pose [7]. The images are divided into 5 subsets with respect to the angle between the light source direction and the camera axis, refer to Table 1. Interested readers may also refer to [7] for further details of the database, all images are downsampled to an order of 50×50.

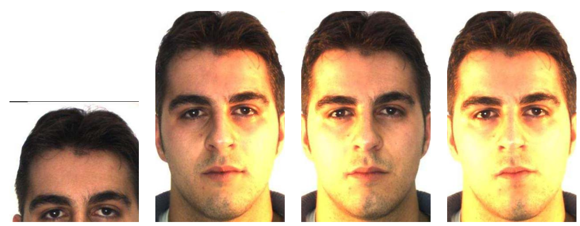
We follow the evaluation protocol as reported in [4], [7], [12], [21], [23], [26], [27]. Training is conducted using subset 1 and the system is validated on the remaining subsets. A detail comparison of the results with some latest approaches is shown in Table 2, all results are as reported in [23]. Note that the error rates have been converted to the recognition success rates. Since subset 3 incorporates moderate luminance variations, most of the state-of-art algorithms report error-free recognition as shown in Table 2. For subset 4 with more adverse illumination variations, the proposed algorithm achieves 100% recognition which is either better than or comparable to all the results reported in the literature. In particular the proposed approach outperforms the Cones-attached, Illumination Ratio Images and Quotient Illumination Relighting methods by 8.60%, 18.60% and 9.40% respectively. It is also found to be fairly comparable to the latest Cone-cast and Gradient Angle approaches. Subset 5 represents the worst case scenario with angle between the light source direction and camera axis being greater than 77°. The proposed RLRC algorithm consistently achieves 100% recognition for the severe alterations comparing favorably with all the reported results in the literature beating the Quotient Illumination Relighting method by more than 17%. Noteworthy is the fact that results for this subset are not available in the literature for most of the contemporary approaches.

Table 3. Results for the AR database.

Method	Left-Light	Right-Light	Both-Lights
1-NN [19]	22.20%	17.80%	3.70%
PCA [19]	7.40%	7.40%	2.20%
LEM [6]	92.90%	91.10%	74.10%
Face-ARG [19]	98.50%	96.30%	91.10%
RLRC	**96.30%**	**94.07%**	**94.07%**

4.2 AR Database

The AR face database contains over 4000 color images incorporating three lighting modes with "left light on", "right light on" and "both lights on" [15], refer to Figure 1.



(a) (b) (c) (d)

Figure 1. Various luminance variations for a typical subject of the AR database.

We follow the experimental setup as proposed in recent works of Face-ARG matching [19] and Line Edge Map (LEM) [6] incorporating 135 subjects of the AR database. The system is trained using neutral lighting conditions i.e. Figure 1 (a) of each subject while Figures 1 (b), (c) and (d) serve as probes, altogether we have 135 gallery images and 405 (135×3) probes. All images are downsampled to an order of 180×180. The results are tabulated in Table 3, noteworthy is the fact that results in [6] are shown for 112 subjects.

The proposed RLRC approach shows a consistent performance across all illumination modes of the AR database. For the cases of "left light on" and "right light on", recognition accuracies of 96.30% and 94.07% are achieved which are fairly comparable to the latest

LEM and Face-ARG approaches as shown in Table 3. For the most challenging problem of illumination with "both lights on" the proposed RLRC approach attains 94.07% recognition which is favorably comparable with the Face-ARG approach and outperforms the LEM approach by a margin of approximately 20%. The conventional methods of PCA and 1-NN reported in [19] are out of discussion as they lag far behind these latest approaches.

5 Conclusion

In this paper we present a novel illumination invariant face recognition algorithm based on the robust Huber estimation approach. Results have been demonstrated on two standard databases, in particular we demonstrate, for the first time, an error-free recognition for the most challenging Subset 5 of the Yale face database B. Apart from the good performance index of the proposed approach, it is interesting to note that recent research has shown competency of the unorthodox features such as downsampled images and random projections [16], [17], [24]. The proposed RLRC approach conforms to this emerging belief. It has been shown that with an appropriate choice of classifier the original image space can produce good results compared to the traditional subspace approaches.

References

[1] A. F. Abate, M. Nappi, D. Riccio, and G. Sabatino. 2D and 3D Face Recognition: A Survey. *Pattern Recognition Letters*, 28(2007):1885–1906, 2007.

[2] R. Barsi and D. Jacobs. Lambertian Reflection and Linear Subspaces. *IEEE Trans. on Pattern Analysis and Machine Intelligence*, 25(3):218–233, 2003.

[3] V. Belhumeur, J. Hespanha, and D. Kriegman. Eigenfaces vs Fisherfaces: Recognition Using Class Specific Linear Projection. *IEEE Tran. PAMI*, 17(7):711–720, July 1997.

[4] H. F. Chen, P. N. Belhumeur, and D. J. Kriegman. In Search of Illumination Invariants. In *IEEE Conf. Computer Vision and Pattern Recognition*, volume 1, pages 13–15, 2000.

[5] R. O. Duda, P. E. Hart, and D. G. Stork. *Pattern Classification*. John Wiley & Sons, Inc., 2000.

[6] Y. Gao and M. K. H. Leung. Face Recognition using Line Edge Map. *IEEE Trans. PAMI*, 24(6):764–779, June 2002.

[7] A. Georghiades, P. Belhumeur, and D. Kriegman. From few to Many: Illumination Cone Models for Face Recognition under Variable Lighting and Pose. *IEEE Trans. PAMI*, 23(6):643–660, 2001.

[8] F. R. Hampel, E. M. Ronchetti, P. J. Rousseeuw, and W. A. Stahel. *Robust Statistics: The Approach Based on Influence Functions*. John Wiley & Sons, 1986, 2005.

[9] P. J. Huber. *Robust Statistics*. New York: John Wiley, 1981.

[10] D. D. Lee and H. S. Seung. Learning the Parts of Objects by Non-negative Matrix Factorization. *Nature*, 401:788–791, 1999.

[11] D. D. Lee and H. S. Seung. Algorithms for Non-negative Matrix Factorization. *Advances in Neural Information Processing Systems*, pages 556–562, 2001.

[12] K. C. Lee, J. Ho, and D. J. Kriegman. Acquiring Linear subspaces for Face Recognition under Variable Lighting. *IEEE Trans. PAMI*, 27(5):684–698, May 2005.

[13] F. Macwilliams and N. Sloane. *The theory of Error-Correcting Codes*. North Holland, 1981.

[14] K. Madsen and H. B. Nielsen. Finite Algorithms for Robust Linear Regression. *BIT Computer Science and Numerical Mathematics*, 30(4):682 – 699, 1990.

[15] A. Martinez and R. Benavente. The AR Face Database. Technical Report 24, CVC, June 1998.

[16] I. Naseem, R. Togneri, and M. Bennamoun. Face Identification using Linear Regression. *IEEE ICIP*, 2009.

[17] I. Naseem, R. Togneri, and M. Bennamoun. Linear Regression for Face Recognition. *IEEE Trans. on PAMI (in press)*, 2009.

[18] H. B. Nielsen. Computing a Minimizer of a Piecewise Quadratic - Implementation. Technical report, Informatics and Mathematical Modelling, Technical University of Denmark, DTU, Sep. 1998.

[19] B.-G. Park, K.-M. Lee, and S.-U. Lee. Face Recognition using Face-ARG matching. *IEEE Trans. on PAMI*, 27(12):1982–1988, 2005.

[20] M. P. Roath and M. Winter. Survey of Appearance-based Methods for Object Recognition. Technical report, Inst. for Computer Graphics and Vision, Graz University, Austria., January 2008.

[21] S. Shan, W. Gao, B. Cao, and D. Zhao. Illumination Normalization for Robust Face Recognition against Varying Lighting Conditions. In *IEEE Workshop on AMFG*, pages 157–164, 2003.

[22] M. Turk and A. Pentland. Eigenfaces for Recognition. *Journal of Cognitive Neurosicence*, 3(1):71–86, 1991.

[23] C. Weilong, M. E. Joo, and S. Wu. Illumination Compensation and Normalization for Robust Face Recognition Using Discrete Cosine Transform in Logarithm Domain. *IEEE Trans. on Systems, Man and Cybernetics*, 36(2):458–464, 2006.

[24] J. Wright, A. Yang, A. Ganesh, S. Sastri, S, and Y. Ma. Robust face recognition via sparse representation. *IEEE Trans. PAMI*, 31(2):210–227, Feb 2009.

[25] P. C. Yuen and J. H. Lai. Face Representation using Independent Component Analysis. *Pattern Recognition*, 35(6):1247–1257, 2002.

[26] L. Zhang and D. Samaras. Face Recognition under Variable Lighting using Harmonic Image Exemplars. In *IEEE Conf. Computer Vision and Pattern Recognition*, volume 1, pages 19–25, 2003.

[27] J. Zhao, Y. Su, D. Wang, and S. Luo. Illumination Ratio Images: Synthesizing and Recognition with Varying Illumination. *Pattern Recognition Letters*, 24:2703–2710, 2003.