

Time Complexity for Face Recognition under varying Pose, Illumination and Facial Expressions based on Sparse Representation

Steven Lawrence Fernandes
Research Scholar

Dept. of Electronics & Communication Engineering
Karunya University
Coimbatore - 641 114, India

G. Josemin Bala
PhD, Professor & Head

Dept. of Electronics & Communication Engineering
Karunya University
Coimbatore - 641 114, India

ABSTRACT

Sparse representation based face recognition is the most recent technique used, this technique first codes a testing sample as a sparse linear combination of all the training samples, and then classifies the testing sample by evaluating which class leads to the minimum representation error. The l_1 -norm sparsity improves the face recognition accuracy. While most of the research focus has been in increasing the face recognition accuracy, in this paper we analyze the time needed for face recognition under varying Facial expressions, Pose and Illumination. This analysis is done on various public data sets. GRIMACE and ATT data sets provide variations in Facial expressions, SUBJECT data set provides Pose variations and YALEB data set provides 64 illumination conditions. The average time taken is calculated for each of the data set.

Key Words

Sparse representation, Face recognition, Facial expressions, Pose variations, Illumination conditions

1. INTRODUCTION

Face recognition is among the active research topics in pattern recognition and computer vision due to its wide applications in human-computer interaction, automatic photo-tagging, and information security [1]. Linear [2]–[6] and nonlinear [7]–[15] methods, have been widely used in face recognition. The principal component analysis (PCA) [2]–[4] and the linear discriminant analysis (LDA) [5], [6] are two typical examples of linear transform methods. The widely used kernel PCA [7]–[10] and kernel LDA [11]–[13] are two typical examples of nonlinear transform methods. In recent years, sparse representation is a new research direction and hot spot, and its basic idea is most early proposed by Mallat [14]. Sparse representation is studied as a representation method of signals. A purpose of the method is to recover the original signal correctly with as few bases as possible. l_1 -norm minimization is an effective approach to obtain the coefficients which achieve the above-mentioned purpose [15]. The l_1 -norm minimization guarantees that it can recover the original signal correctly in certain circumstances. The sparse representation-based classification (SRC) adopts the sparse representation to face recognition [16]. This method deploys a training data set, namely test images with raster scan order, as bases. The SRC approximates a test sample to linear combination of the bases, and identifies who is the test sample by referring the coefficients. The SRC shows higher recognition rates than traditional classifications including nearest neighbor, nearest subspace and linear support vector machine. In this paper the time needed for face recognition

under varying Facial expressions, Pose, Illumination is analyzed.

This paper is organized as follow: Section 2 describes the methodology for sparse representation based face recognition. Section 3 describes the salient features of the face data sets considered. Section 4 analyzes the experimental results. Finally, Section 5 sums up the conclusion.

2. SPARSE REPRESENTATION BASED FACE RECOGNITION

Let $y \in R^M$ be a given signal, matrix $A \in R^{M \times N}$ ($M < N$) is a set of bases, where M and N denote the dimension of signals and the number of bases respectively. The signal y can be represented as a linear combination of bases, and then it can be written as the following expression:

$$Y = Ax \quad (1)$$

This is an underdetermined system. We can solve this problem with the pseudo-inverse matrix generally. When we hope to represent the signal y with few bases as possible, this problem can be expressed as the following l_0 -norm minimization problem:

$$\hat{x}_0 = \arg \min \|x\|_0 \quad \text{subject to } Ax = y. \quad (2)$$

However, (2) is known as a NP-hard problem; then we cannot obtain a solution within polynomial time. Here, we would like to solve a l_1 -norm minimization problem:

$$\hat{x}_1 = \arg \min \|x\|_1 \quad \text{subject to } Ax = y, \quad (3)$$

Considering an allowable error, (3) can be rewritten as the following:

$$\hat{x}_1 = \arg \min \|x\|_1 \quad \text{subject to } \|A_x - y\|_2^2 \leq \epsilon$$

$$l_0\text{-norm } \|\cdot\|_0 \quad \text{and } l_1\text{-norm } \|\cdot\|_1 \quad (4)$$

Note that l_0 -norm and l_1 -norm are calculated by followings:

$$\|x\|_0 = \sum_{i=1}^n \delta(x_i) \quad \text{subject to } \delta(a) = \begin{cases} 1 & (a \neq 0) \\ 0 & (a = 0) \end{cases} \quad (5)$$

$$\|x\|_1 = \sum_{i=1}^n |x_i| \quad (6)$$

2.1 Accuracy of L1-norm minimization:

We need assumptions on the matrix A to ensure accuracy of L1-norm minimization on spares representation. Let be a minimum δ_k constant, where a solution x has K non-zero coefficients and the matrix A satisfies the following:

$$(1 - \delta)\|x\|_2^2 \leq \|Ax\|_2^2 \leq (1 + \delta)\|x\|_2^2$$

$$\delta_{2K} < \sqrt{2} - 1 \text{ and } K \geq 1, \text{L1-norm} \quad (7)$$

When the natural number K satisfies minimization can recover the original signal correctly. This property is called restricted isometry property. In other words, if the matrix A has above property, the solution of L0-norm minimization corresponds to the solution of L1-norm minimization. If the matrix A is orthogonal, δ_k equals to zero. The constant δ_k expresses how K column vectors selected from the matrix A arbitrary differ from the orthonormality.

2.2 L1-norm minimization as a linear programming problem

Let $t = (t_1, \dots, t_N)^T$ be auxiliary variables, the L1-norm minimization problem (3) can be rewritten as a follow:

$$\min \sum_{i=1}^N t_i \text{ subject to } -t \leq x \leq t, y=Ax \quad (8)$$

This problem is a combination of linear equations and linear in equations, and this is called a linear programming problem. Linear programming problems can be solved by using conventional solutions; for example, the simplex method and the interior method, within polynomial time.

2.3 Algorithm: Sparse representation-based face recognition

Input: A matrix of a training data set $A=[A_1, A_2, \dots, A_k] \in R^{M \times N}$ for k classes, a test sample $y \in R^M$, and a error tolerance $\epsilon > 0$.

Step 1: Normalize columns of A to have unit L2-norm.

Step 2: Make a matrix $B = [A, I]$.

Step 3: Solve a L1-norm minimization problem:

$$\hat{x}_1 = \arg \min \|x\|_1 \text{ subject to } \|Bx - y\|_2^2 \leq \epsilon \quad (9)$$

Step 4: Compute residuals

$$r_i(y) = \left\| y - B\hat{\delta}_i(x_1) \right\|_2 \quad (k=1, 2, \dots, k) \quad (10)$$

Output: A recognition result $\min_i r_i(y)$

3. FACE DATABASE

This section provides the salient features of the four data sets which are GRIMACE, ATT, SUBJECT and YALEB.

3.1 GRIMACE

- Number of individuals: 18
- Image resolution: 180x200 pixels (portrait format)

- Contains images of male and female subjects
- Backgrounds: the background is plain
- Head Scale: small head scale variation
- Head turn, tilt and slant: considerable variation in these attributes
- Position of face in image : some translation
- Image lighting variation: very little
- Expression Variation: major expression variation
- Additional comment: there is no hair style variation as the images were taken in a single session.



Fig 1: Grimace database

3.2 ATT

- Number of individuals: 40
- Image resolution: 92x112 pixels, with 256 grey levels per pixel (portrait format)
- Contains images of male and female subjects
- Backgrounds: the background grey.
- Head Scale: small head scale variation
- Head turn, tilt and slant: considerable variation in these attributes
- Image lighting variation: very little
- Expression Variation: major expression variation
- Additional comment: there is no hair style variation as the images were taken in a single



Fig 2: ATT database

3.3 SUBJECT

- Number of individuals: 90
- Image resolution: 180x200 pixels (portrait format)
- Contains images of male and female subjects
- Backgrounds: the background is grey
- Head Scale: small head scale variation
- Head turn, tilt and slant: images are 5 degree from right profile (defined as +90°) to left profile (defined as -90°) in the pan rotation.
- Position of face in image : some right and some left
- Image lighting variation: no variation in light

- Expression Variation: major expression variation
- Additional comment: there is no hair style variation as the images were taken in a single session.

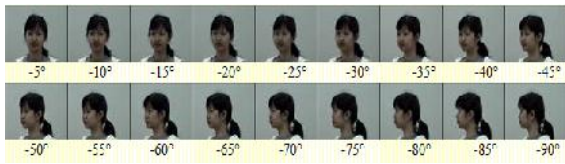


Fig 3: SUBJECT database

3.4 YALEB

- Number of individuals: 39
- Image resolution: 180x200 pixels (portrait format)
- Contains images of male and female subjects
- Backgrounds: the background is grey
- Head Scale: small head scale variation
- Head turn, tilt and slant: images are 5 degree from right profile (defined as +90°) to left profile (defined as -90°) in the pan rotation.
- Position of face in image : some right and some left
- Image lighting variation: 64 illumination condition.
- Expression Variation: minor expression variation
- Each person's image is taken with (9 pose x 64 illumination condition)



Fig 4: YALEB database

4. RESULTS AND ANALYSIS

In order to evaluate the time taken for sparse representation based face recognition a code for the algorithm has been generated using Matlab 7.6 on Intel dual-core processing having 1Gb RAM. This algorithm has been tested using four set of datasets which are GRIMACE, ATT, SUBJECT, YALEB.

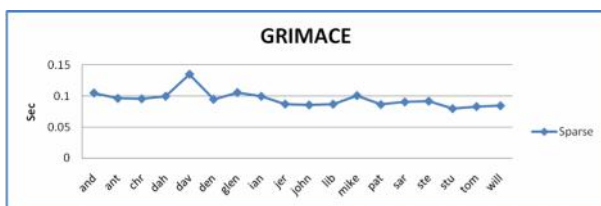


Fig 5: Time taken for Sparse Representation based Face Recognition on GRIMACE database

Figure 5 shows the time taken for Sparse Representation based Face Recognition on GRIMACE database, 0.0947 seconds is the average time on GRIMACE database.

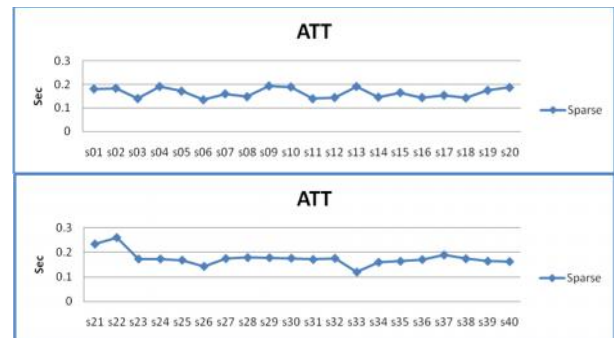


Fig 6: Time taken for Sparse Representation based Face Recognition on ATT database

Figure 6 shows the time taken for Sparse Representation based Face Recognition on ATT database, 0.1698 seconds is the average time on ATT database. Thus the average time by GRIMACE and ATT database is 0.13225 seconds.

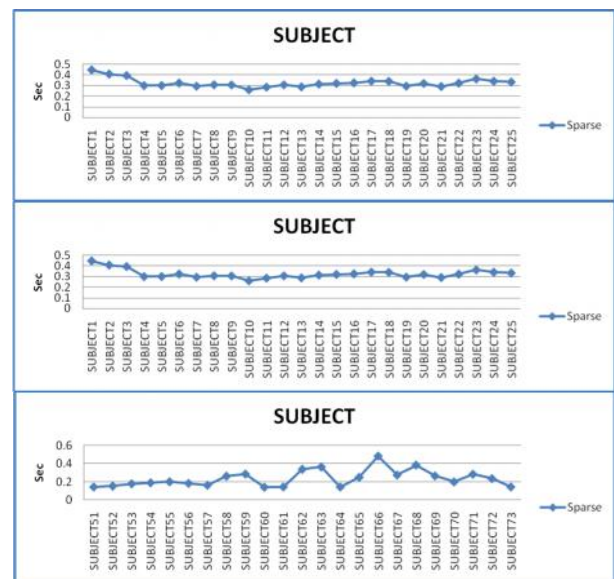


Fig 7: Time taken for Sparse Representation based Face Recognition on SUBJECT database

Figure 7 shows the time taken for Sparse Representation based Face Recognition on SUBJECT database, 0.2858 seconds is the average time on SUBJECT database.

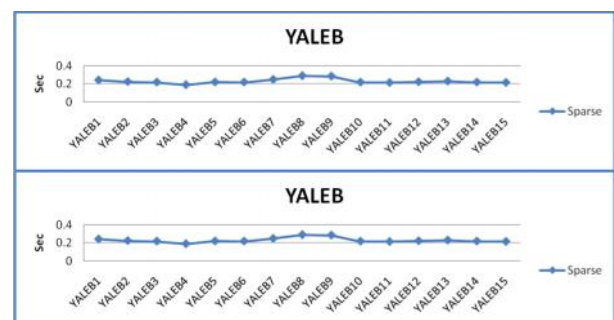


Fig 8: Time taken for Sparse Representation based Face Recognition on YALEB database

Figure 8 shows the time taken for Sparse Representation based Face Recognition on YALEB database, 0.2063 seconds is the average time on YALEB database.

5. CONCLUSION

In this paper we have analyzed the time needed for face recognition under varying Facial expressions, Pose and Illumination. This analysis is done on various public data sets. GRIMACE and ATT data sets provide variations in Facial expressions; the average face recognition time obtained is 0.13225 seconds. SUBJECT data set provides Pose variations; the average face recognition time obtained is 0.2858 seconds. YALEB data set provides 64 illumination conditions; the average face recognition time obtained is 0.2063 seconds. Thus we conclude, sparse representation based face recognition is fastest under varying facial expressions; it still needs improvements during varying pose and illuminations conditions.

6. REFERENCES

- [1] W. Zhao et al., "Face recognition: A literature survey," *ACM Computing Surveys*, 2003. 1
- [2] M. Kirby and L. Sirovich, "Application of the KL phase for the characterization of human faces," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 12, no. 1, pp. 103–108, Jan. 1990.
- [3] Y. Xu, D. Zhang, J. Yang, and J.-Y. Yang, "An approach for directly extracting features from matrix data and its application in face recognition," *Neurocomputing*, vol. 71, nos. 10–12, pp. 1857–1865, 2008.
- [4] J. Yang, D. Zhang, A. F. Frangi, and J.-Y. Yang, "Two-dimensional PCA: A new approach to appearance-based face representation and recognition," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 26, no. 1, pp. 131–137, Jan. 2004.
- [5] Y. Xu and D. Zhang, "Represent and fuse bimodal biometric images at the feature level: Complex-matrix-based fusion scheme," *Opt. Eng.*, vol. 49, no. 3, 2010.
- [6] S. W. Park and M. Savvides, "A multifactor extension of linear discriminant analysis for face recognition under varying pose and illumination," *EURASIP J. Adv. Signal Process.*, vol. 2010, no. 158395, p. 11, 2010.
- [7] M. Debruyne and T. Verdonck, "Robust kernel principal component analysis and classification," *Adv. Data Anal. Classification*, vol. 4, no. 2, pp. 151–167, Sep. 2010.
- [8] Y. Xu, D. Zhang, F. Song, J.-Y. Yang, Z. Jing, and M. Li, "A method for speeding up feature extraction based on KPCA," *Neurocomputing*, vol. 70, nos. 4–6, pp. 1056–1061, Jan. 2007.
- [9] M. E. Tipping, "Sparse kernel principal component analysis," in *Neural Information Processing Systems*, T. K. Leen, T. G. Dietterich, and V. Tresp, Eds. Cambridge: MIT Press, 2000, pp. 633–639.
- [10] B. Schölkopf, A. Smola, and K.-R. Müller, "Kernel principal component analysis," in *Proc. ICANN*, 1997, pp. 583–588.
- [11] B. Schölkopf and A. Smola, *Learning With Kernels*. Cambridge: MIT Press, 2002.
- [12] K.-R. Müller, S. Mika, G. Rätsch, K. Tsuda, and B. Schölkopf, "An introduction to kernel based learning algorithms," *IEEE Trans. Neural Netw.*, vol. 12, no. 2, pp. 181–201, Mar. 2001.
- [13] D. Tao and X. Tang, "Kernel full-space biased discriminant analysis," in *Proc. IEEE ICME*, Jun. 2004, pp. 1287–1290.
- [14] S. Mallat, and Z. Zhang, "Matching Pursuits with Time-frequency Dictionaries," *IEEE Transaction on Signal Process.*, vol. 41 (12): 3397-3415, 1993.
- [15] J. Wright, A. Y. Yang, A. Ganesh, S. Shankar, and Y. Ma, "Robust face recognition via sparse representation," *IEEE Trans. PAMI*, vol. 31, no. 2, pp. 210-227, Feb. 2009.
- [16] J. A. Tropp and S. J. Wright, "Computational methods for sparse solution of linear inverse problem," *Proceedings of the IEEE*, vol. 98, no. 6, pp. 948-958, June 2010.