Robust Face Recognition via Sparse Representation

Summary written by Mahban Gholijafari
February 10, 2025

Summary: This paper proposes a method based on sparse representation computed by l^1-minimization that can be used for object recognition in general, however, the studies and experiments of this work is done on human frontal face with varying expression and illumination including occlusion and disguise. The presented algorithm addresses critical issues in face recognition such as feature extraction and robustness to occlusion.

Related work: There have been multiple studies to choose a limited subset of features or models from the training data for representing or classifying an input (test) signal [4, 9, 12, 3]. The paper also states that the recent studies are based on the discovery that whenever the optimal representation is sufficiently sparse, it can be efficiently computed by convex optimization [5]. The goal of all these studies was to represent and compress the signals and individual base elements in the dictionary were usually chosen from standard bases like Gabor or even generated from random matrices [2, 1]. The sparsest representation is naturally discriminative which means it selects the subset that most compactly expresses the input signal and rejects all other possible but less compact representations.

Approach: This paper uses the discriminative nature of sparse representation to do classification. Their novelty is to represent test samples in an overcomplete dictionary where base elements are the training samples themselves and it will be a linear combination of just those training samples from the same class. Their approach is to adaptively select the training samples that give the most compact representation and this classifier can be considered a generalization of popular classifiers such as nearest neighbor (NN)[6], nearest subspace (NS)[8] and nearest feature line (NFL) [10] algorithms. Two issues in automatic face recognition motivated this study: the role of feature extraction and the difficulty due to occlusion. Training samples (n_i) are arranged from ith class as columns of a matrix $A_i \doteq [v_{i,1}, v_{i,2}, ..., v_{i,n_i}] \in \mathbb{R}^{m \times n}$ to solve for $y = Ax_0$. A $w \times h$ gray scale image with the vector $v \in \mathbb{R}^m (m = wh)$ will be identified by stacking its columns. Columns of A_i are the training face images of the ith subject. x_0 is a coefficient vector whose entries are zero except those associated with the ith class. A significant difference of this method from NN and NS method is the usage of the entire training set to solve for x rather than one sample or one class at a time methods. If the solution x_0 sought is sparse enough, the solution of the ℓ^0-minimization problem is equal to the solution to the ℓ^1-minimization problem and it can be solved in polynomial time with standard linear programming methods. Noise in this study has been addressed as a convex optimization problem that can be efficiently solved with second-order cone programming. The first algorithm is the Sparse Representation-based Classification (SRC) that computes residuals $r_i(y) = ||y - A\delta_i(\hat{x}_1)||_2$ for $i = 1, ..., k$ for k classes and identifying y as $argmin_i r_i(y)$. δ_i is the characteristic function. To handle Occlusion or corruption, they attempt to recover the sparsest solution by solving the *extended* ℓ^1-minimization problem. The paper still follows algorithm 1 but instead of image matrix A, we will have extended matrix $B = [A, I]$ and instead of we will have $w = [x, e]$ where e is a vector of errors.

Datasets, Experiments and Results: For analyzing algorithm 1, Extended Yale B database [7] containing 2414 images of 38 individuals has been used and dimensions of 30, 56, 120 and 504 has been selected. This algorithms achieves overall recognition rates between 92.1 and 95.6 % for all 120D feature spaces and maximum rate of 98.1 % with 504D randomfaces. For experimental verification, the study tests their SRC algorithm using conventional holistic face features and compares them with randomfaces and downsampled images. Their algorithms also has been compared with NN, NS and linear SVM. The AR database consisting of over 4000 frontal images of 126 individuals has also been used [11] with selected four feature space dimensions of 30, 54, 130 and 540. The recognition rate is between 92% and 94.7%. Partial face features, Random pixel corruption, random block occlusion, and disguise has been tested and SRC correctly recovers the identity of subjects.

Strengths: The best performance of SRC and SVM exceeds the best performance of NN and NS at each individual feature dimensions for both Yale and AR database and their performance does not strongly depend on a good choice of optimal features. The proposed SRC algorithm achieves better recognition rates than NN, NS and SVM in partial face features.

Weaknesses: The proposed SRC algorithm in some cases performs worse than SVM algorithm. This algorithm performs poorly on arbitrary occlusions after the breakdown point. Variation in object pose and misalignment is the weekness of this framework and the algorithms needs a large training samples to represent the distribution of face images.

Reflections: Future work can analyze whether this work can be useful for object detection. Also, the algorithm can be extended to less constrained conditions, especially variations in object pose and misalignment.

References

[1] E. J. Candès et al. Compressive sampling. In *Proceedings of the international congress of mathematicians*, volume 3, pages 1433–1452. Madrid, Spain, 2006.

[2] E. J. Candes and T. Tao. Near-optimal signal recovery from random projections: Universal encoding strategies? *IEEE transactions on information theory*, 52(12):5406–5425, 2006.

[3] T. M. Cover. Geometrical and statistical properties of systems of linear inequalities with applications in pattern recognition. *IEEE transactions on electronic computers*, (3):326–334, 1965.

[4] A. d'Aspremont, L. Ghaoui, M. Jordan, and G. Lanckriet. A direct formulation for sparse pca using semidefinite programming. *Advances in neural information processing systems*, 17, 2004.

[5] D. L. Donoho. For most large underdetermined systems of linear equations the minimal 1-norm solution is also the sparsest solution. *Communications on Pure and Applied Mathematics: A Journal Issued by the Courant Institute of Mathematical Sciences*, 59(6):797–829, 2006.

[6] R. O. Duda, P. E. Hart, and D. G. Stork. Pattern classification second edition john wiley & sons. *New York*, 58:16, 2001.

[7] A. S. Georghiades, P. N. Belhumeur, and D. J. Kriegman. From few to many: Illumination cone models for face recognition under variable lighting and pose. *IEEE transactions on pattern analysis and machine intelligence*, 23(6):643–660, 2001.

[8] J. Ho, M.-H. Yang, J. Lim, K.-C. Lee, and D. Kriegman. Clustering appearances of objects under varying illumination conditions. In *2003 IEEE Computer Society Conference on Computer Vision and Pattern Recognition, 2003. Proceedings.*, volume 1, pages I–I. IEEE, 2003.

[9] K. Huang and S. Aviyente. Sparse representation for signal classification. *Advances in neural information processing systems*, 19, 2006.

[10] S. Z. Li and J. Lu. Face recognition using the nearest feature line method. *IEEE transactions on neural networks*, 10(2):439–443, 1999.

[11] A. Martinez and R. Benavente. The ar face database: Cvc technical report, 24. 1998.

[12] V. Vapnik. *The nature of statistical learning theory*. Springer science & business media, 2013.