Robust Face Recognition via Sparse Representation

Summary written by Sahana Srihari
February 24, 2020

Summary: The paper deals with a method of robust face recognition based on sparse representation of images and how this plays a crucial role in feature extraction and invariance to occlusion.The main goal of the paper is to show how compressed sensing allows picking non-crucial and random features which provides good classification and also how their algorithm performs well in the case of adversities such as illumination, expression and occlusion.They have experimentally demonstrated and validated their results.

Related work: Convex optimization leading to optimal representation can be noted in [6] which provides the basis for solving for a sparse representation.The classifiers are based on a generalization of NN,NS and NFL as seen in [7][9][10].Details on Sparse representation and l^1-minimization can be drawn from [6][4][5]. The minimization for *l1* is based on [2][3]. For classification comparison, methods were tested with the regular Fisherfaces[1],Eigenfaces[11]and Laplacianfaces [8]

Approach: Based on the notion that a *sufficiently sparse* representation can be obtained through convex optimisation and the test image is represented as an over-complete dictionary of the training samples. The methods discussed in this paper can broadly be divided into 3 parts - Section 1 deals with l^1-minimization- utilision to compute the sparse representations for classification. Section 2 explains the cases of feature extraction and robustness to occlusion. Section 3 contains the experimental set up for validation.

Section 1: Based on the assumption that training sample lie on a linear subspace and hence a test sample lies in the linear span of the training sample, represented as $y = Ax_0$ where x_0 is the gives the class of the test sample. Through their geometric interpretation, they were able to find that the l^1-minimization recovers x_0 fully containing only 1 non-zero value. The algorithm pipeline: (i)***A*** is constructed from all training samples of all classes and test sample ***y*** is considered the inputs to the system. (ii) Each class values composing the ***A*** is l^2-normalised for elimination of the system being under-determined, but this leads to a dense solution to the problem. (iii) l^1-normalization through solving the system for $\hat{x_0} = \underset{x}{\mathrm{argmin}}||x||_1$ subject to $||Ax-y||_2 < \epsilon$ where ϵ is introduced for dealing with small dense noise. This step ensures the recovery of a sparse representation of the image. (iv)Classification is attempted by how well the training samples can reproduce ***y*** with the coefficients. It is computed through the notion of residuals $r_i(y) = ||y - A\delta_i(\hat{x_1})||_2$ where δ_i is a function used for selecting the coefficients of a particular class *i*. The aim is to minimize the residual. (v) Classification is accomplished as identity(***y***) = $\underset{i}{\mathrm{argmin}} r_1(\boldsymbol{y})$

Section 2: To tackle the issue of large dimensional data,transformation matrix ***R*** is utilized to reduce to a lower dimensional feature space and the linear system can be represented now as $\tilde{y}$ = $RA\boldsymbol{x_0}$ and solved under l^1-minimization. As compared the classical face features, *Randomfaces* are efficient as the transformation is independent of the training set. This sets ground to the criteria for extracting precise features is relaxed. To deal with errors due to occlusion, the linear system is remodelled as y = $y_0 + e_0$; e_0 models the corrupted pixels, represented as y = $B\mathbf{w}$ and solved via *extended* l^1-minimization guided by spatial locality.

Section 3: Experimental results were verified using extended YaleB and AR datasets. Feature extraction and classification was compared with respect to other classifiers and later the robustness to occlusion was also tested. SRC performed better than NN, NS and SVM in terms of recognition rates. Both SRC and SVM outperform the remaining two algorithms for the 2 datasets. SRC showed good recognition rates for partial face features and also for partial pixel corruptions. SRC showed good performance relative to the others in the case of over 50% corruption and for over 30% random block occlusions.

Strengths: The paper proves that through sparse enough representations, there is less weight on finding precise feature extraction methods and instead the approach adopted is a simple linear system transformation capable of being optimised through l^1-minimization. The method also efficiently deals with some of the most prominent issues regarding face recognition - illumination, expression and occlusion and enough experimental backing is provided to show the algorithm being superior than the rest.

Weaknesses: The recognition system focuses only on the frontal view of faces, it does not consider the change in pose. The experiment is focused on faces with a simple background but fails to evaluate the performance of recognising faces in a natural scene setting. Given that it performs well in general occlusions and pixel corruption, there is not enough evidence of good performance with respect to the computational time for the recognition system.

Reflections: The authors have approached the problem with a unique perspective compared to the usual Fisherfaces or Eigenfaces approach for face recognition. It addresses a very important issue the need for high performance feature extraction. Although there is scope of improvement with regards to dealing with a more complex problem, this provides a good platform to build on for future work.

References

[1] P. N. Belhumeur, J. P. Hespanha, and D. J. Kriegman. Eigenfaces vs. fisherfaces: Recognition using class specific linear projection. *IEEE Transactions on pattern analysis and machine intelligence*, 19(7):711–720, 1997.

[2] S. Boyd, S. P. Boyd, and L. Vandenberghe. *Convex optimization*. Cambridge university press, 2004.

[3] E. Candes and J. Romberg. l1-magic: Recovery of sparse signals via convex programming. *URL: www. acm. caltech. edu/l1magic/downloads/l1magic. pdf*, 4:14, 2005.

[4] E. J. Candes, J. K. Romberg, and T. Tao. Stable signal recovery from incomplete and inaccurate measurements. *Communications on Pure and Applied Mathematics: A Journal Issued by the Courant Institute of Mathematical Sciences*, 59(8):1207–1223, 2006.

[5] 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.

[6] D. Donoho. For most large underdetermined systems of equations, the minimal l1-norm near-solution approximates the sparsest near-solution. *Communications on Pure and Applied Mathematics*, 59:907 – 934, 07 2006.

[7] R. O. Duda, P. E. Hart, and D. G. Stork. Pattern classification 2nd ed. *John Willey & Sons Inc*, 2001.

[8] X. He, S. Yan, Y. Hu, P. Niyogi, and H.-J. Zhang. Face recognition using laplacianfaces. *IEEE transactions on pattern analysis and machine intelligence*, 27(3):328–340, 2005.

[9] 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.

[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] M. Turk and A. Pentland. Eigenfaces for recognition. *Journal of cognitive neuroscience*, 3(1):71–86, 1991.