Robust Face Recognition via Sparse Representation

Summary written by Nicholas Hinke
February 22, 2022

Summary: Using automatic face recognition as its case study, this paper proposes a novel method for representing object classes within images in order to efficiently validate and classify them. Termed *Sparse Representation-based Classification* (SRC), the authors sought to tackle two of the predominant problems in image classification specifically with regard to face recognition: 1) the importance of the choice of features extracted, and 2) the typically poor robustness to occlusions and corruptions of other algorithms of the day. Inspired by other works of the time on other popular classifiers [6, 8, 9], the human vision system [10, 15], convex optimization [1, 5, 11, 14], and especially compressed sensing and sparse representations [3, 4, 5, 14], the authors developed an algorithm rooted in a relatively simple idea: represent any image as a linear combination of other images *of the same class*. Moreover, the authors claim that their method outperforms other popular algorithms of the time with respect to the two problems named above [13].

Approach: While the primary contribution of this paper is the SRC algorithm itself, the authors also present detailed mathematical analyses as an argument for the power and versatility of sparse representations. They contend that not only are sufficiently sparse representations very efficient to compute–thanks to the then-recent advances in the field of convex optimization [5]–but also that redundancy and sparsity are the keys to the algorithm's performance regarding error correction and robustness [13].

In contrast to other algorithms of the time [2, 7, 12], the authors claim that their algorithm achieved a notable degree of "feature-invariance", signifying that the choice of features extracted from the face images was unrelated to the success of the algorithm [13]. They claimed that this held true given that the dimension of the utilized feature space was sufficiently large, and that the sparse representation was "correctly computed" using the l^1-norm; this second condition follows from the ability to recover the solution to the non-convex l^0-norm problem via the convex l^1-norm problem given that the solution is sufficiently sparse [3, 4, 5].

More explicitly, the SRC algorithm itself is comprised of the following steps. Before classifying any test images, a dictionary containing multiple image samples of each class is first constructed and stored. Then, given a new test sample y_t, the sparse representation x_t is determined by solving the l^1-norm optimization problem as described above. After finding x_t, class-wise residual errors are then computed for each class by attempting to reconstruct y_t using only the coefficients of x_t that correspond to *one* particular class. Finally, the test image is assigned the identity of the class that minimizes the reconstruction error [13].

In addition to the baseline algorithm as outlined above, the authors describe two natural extensions that further improve its performance. First, they perform validation of test images using what they call the *sparsity concentration index* (SCI). Notably, this validation method does not rely on residuals, but instead utilizes the distribution of sparse coefficients for each class; this resulted in a greater than 10% performance increase when determining whether or not a given test image corresponds to one of the known object classes. Moreover, the SRC algorithm can be easily extended to separate occlusions or noise from a corrupted image by simply adding an error term in a clever way. In fact, SRC vastly outperforms all other popular methods when classifying faces–with recognition rates over 90% for a 70% noisy image or a 40% occluded image!–given that the error is sufficiently sparse [13].

Strengths: Most importantly, the SRC algorithm demonstrated superior invariance to occlusions and noise when compared to any other popular method of the time, as well as the ability to effectively separate clean images from their corruptions. Additionally, the development of the SCI yielded further robustness to invalid test samples than typical residual-based validation methods. Finally, due to the consideration of recent mathematical advancements when designing the algorithm, the authors were able to achieve a level of feature-invariance that had not been seen before, thus allowing one to apply the algorithm to a very wide variety of image features [13].

Weaknesses: Although it results in *zero* time spent training a model, the necessity of storing a dictionary of all test samples makes SRC essentially infeasible for many-class classification problems. This storage cost in addition to the time cost of solving the l^1-norm optimization problem for every new test image makes it extremely challenging for the algorithm to scale efficiently. Finally, as briefly discussed by the authors, the algorithm is not particularly robust to variations in pose [13].

Reflections: As mentioned in the paper, it may prove beneficial during the validation step to somehow combine the SCI with other methods involving residual errors [13]. Additionally, it would be very insightful to examine the performance of this algorithm when used for other image classification tasks aside from face recognition. Finally, it would be interesting to study and update the quoted runtime efficiency of this algorithm using the modern computer hardware and GPU's of today.

References

[1] E. Amaldi and V. Kann. On the approximability of minimizing nonzero variables or unsatisfied relations in linear systems. *Theoretical Computer Science*, 209(1):237–260, 1998.

[2] P. Belhumeur, J. Hespanha, and D. Kriegman. Eigenfaces vs. fisherfaces: recognition using class specific linear projection. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 19(7):711–720, 1997.

[3] E. J. Candès, J. K. Romberg, and T. Tao. Stable signal recovery from incomplete and inaccurate measurements. *Communications on Pure and Applied Mathematics*, 59:1207–1223, 2005.

[4] E. J. Candès and T. Tao. Near-optimal signal recovery from random projections: Universal encoding strategies? *IEEE Transactions on Information Theory*, 52(12):5406–5425, 2006.

[5] D. L. Donoho. For most large underdetermined systems of linear equations the minimal 1-norm solution is also the sparsest solution. *Comm. Pure Appl. Math*, 59:797–829, 2004.

[6] R. O. Duda, P. E. Hart, D. G. Stork, C. R. O. Duda, P. E. Hart, and D. G. Stork. Pattern classification, 2nd ed, 2001.

[7] 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.

[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, 2003.

[9] S. Li and J. Lu. Face recognition using the nearest feature line method. *IEEE transactions on neural networks*, 10(2):439–43, 1999.

[10] P. Sinha, B. Balas, Y. Ostrovsky, and R. Russell. Face recognition by humans: Nineteen results all computer vision researchers should know about. *Proceedings of the IEEE*, 94:1948–1962, 2006.

[11] R. Tibshirani. Regression shrinkage and selection via the lasso. *Journal of the royal statistical society series b-methodological*, 58:267–288, 1996.

[12] M. Turk and A. Pentland. Eigenfaces for recognition. *Journal of Cognitive Neuroscience*, 3(1):71–86, 01 1991.

[13] J. Wright, A. Y. Yang, A. Ganesh, S. S. Sastry, and Y. Ma. Robust face recognition via sparse representation. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 31(2):210–227, 2009.

[14] P. Zhao and B. Yu. On model selection consistency of lasso. *J. Mach. Learn. Res.*, 7:2541–2563, 2006.

[15] W. Zhao, R. Chellappa, P. J. Phillips, and A. Rosenfeld. Face recognition: A literature survey. *ACM Comput. Surv.*, 35:399–458, 2003.