

References

1. Aharon, M., Elad, M., Bruckstein, A.M.: The k-svd: An algorithm for designing of overcomplete dictionaries for sparse representations. *IEEE Transactions on Signal Processing* **54**(11), 4311–4322 (2006)
2. Ahna, J., Oha, J., Choib, S.: Learning principal directions: Integrated-squared-error minimization. *Neurocomputing* **70**, 1372–1381 (2007)
3. Allain, M., Idier, J., Goussard, Y.: On global and local convergence of half-quadratic algorithms. *IEEE Transactions on Image Processing* **15**(5), 1030–1042 (2006)
4. Angst, R., Zach, C., Pollefeys, M.: The generalized trace norm and its application to structure from motion problems. In: *International Conference on Computer Vision*, pp. 2502–2509 (2011)
5. Bajwa, W., Haupt, J., Sayeed, A., Nowak, R.: Compressive wireless sensing. In: *Proceedings of International Conference on Information Processing in Sensor Networks* (2006)
6. Beck, A., Teboulle, M.: A fast iterative shrinkage-thresholding algorithm for linear inverse problems. *SIAM Journal on Imaging Sciences* **2**(1), 183–202 (2009)
7. Belhumeur, P.N., Hespanha, J., Kriegman, D.J.: Eigenfaces vs. fisherfaces: Recognition using class specific linear projection. *IEEE Transactions on Pattern Analysis and Machine Intelligence* **19**(7), 711–720 (1997)
8. Belkin, M., Niyogi, P., Sindhvani, V.: Manifold regularization: A geometric framework for learning from labeled and unlabeled examples. *Journal of Machine Learning Research* pp. 2399–2434 (2006)
9. Bioucas-Dias, J., Figueiredo, M.: A new twist: Two-step iterative shrinkage/thresholding algorithms for image restoration. *IEEE Transactions on Image Processing* **16**(12), 2992–3004 (2007)
10. Bjorck, A.: A direct method for sparse least-squares problems with lower and upper bounds, *numer. Math* **54**, 19–32 (1988)
11. Black, M., Jepson, A.: Eigentracking: Robust matching and tracking of articulated objects using a view-based representation. *International Journal of Computer Vision* **26**(1), 63–84 (1998)
12. Blake, A., Zisserman, A.: *Visual Reconstruction*. MIT Press, Cambridge, MA (1987)
13. Boyd, S., Vandenberghe, L.: *Convex optimization*. Cambridge University Press (2004)
14. Bruckstein, A.M., Elad, M., Zibulevskyy, M.: On the uniqueness of nonnegative sparse solutions to underdetermined systems of equations. *IEEE Transactions on Information Theory* **54**(11), 4813–4820 (2008)
15. Cai, D., He, X., Han, J.: Spectral regression for efficient regularized subspace learning. In: *International Conference on Computer Vision*, pp. 1–7 (2007)

16. Candés, E.J., Romberg, J.: l1-magic: recovery of sparse signals via convex programming. <http://www.acm.caltech.edu/l1magic/> (2005)
17. Candés, E.J., Romberg, J., Tao, T.: Stable signal recovery from incomplete and inaccurate measurements. *Communications on Pure and Applied Math* **59**(8), 1207–1223 (2006)
18. Candés, E.J., Tao, T.: Near optimal signal recovery from random projections: universal encoding strategies. *IEEE Transactions on Information Theory* **52**(12), 5406–5425 (2006)
19. Candés, E.J., Wakin, M., Boyd, S.: Enhancing sparsity by reweighted ℓ_1 minimization. *Journal of Fourier Analysis and Applications* **14**(5), 877–905 (2008)
20. Cetin, M., Karl, W.C.: Feature-enhanced synthetic aperture radar image formation based on nonquadratic regularization. *IEEE Transactions on Image Processing* **10**(4), 623–631 (2001)
21. Champagnat, F., Idier, J.: A connection between half-quadratic criteria and EM algorithms. *IEEE Signal Processing Letters* **11**(9), 709–712 (2004)
22. Charbonnier, P., Blanc-Feraud, L., Aubert, G., Barlaud, M.: Deterministic edge-preserving regularization in computed imaging. *IEEE Transactions on Image Processing* **6**(2), 298–311 (1997)
23. Chartrand, R.: Exact reconstruction of sparse signals via nonconvex minimization. *IEEE Signal Processing Letters* **14**(10), 707–710 (2007)
24. Chartrand, R., Yin, W.: Iteratively reweighted algorithms for compressive sensing. In: *Proceedings of IEEE Conference on Acoustics, Speech and Signal Processing*, pp. 3869–3872 (2008)
25. Chen, S., Donoho, D.L., Saunders, M.: Atomic decomposition by basis pursuit. *SIAM Review* **43**(1), 129–159 (2001)
26. Chen, C., Huang, J., He, L.: Preconditioning for accelerated iteratively reweighted least squares in structured sparsity reconstruction. In: *IEEE Conference on Computer Vision and Pattern Recognition* (2014)
27. Cheng, B., Yang, J., Yan, S., Fu, Y., Huang, T.S.: Learning with ℓ_1 -graph for image analysis. *IEEE Transactions on Image Processing* **4**, 858–866 (2010)
28. Chien, J.T., Wu, C.C.: Discriminant waveletfaces and nearest feature classifiers for face recognition. *IEEE Transactions on Pattern Analysis and Machine Intelligence* **24**(12), 1644–1649 (2002)
29. Combettes, P.L., Pesquet, J.C.: Proximal thresholding algorithm for minimization over orthonormal bases. *SIAM Journal on Optimization* **18**(4), 1531–1576 (2007)
30. Combettes, P.L., Wajs, V.R.: Signal recovery by proximal forward backward splitting. *SIAM Journal on Multiscale Modeling & Simulation* **4**(5), 1168–1200 (2005)
31. Cover, T., Thomas, J.: *Elements of Information Theory*, 2nd edition. New Jersey: John Wiley (2005)
32. Cover, T.M., Thomas, J.A.: *Elements of information theory*, 2nd edition. New York: John Wiley (2005)
33. Daubechies, I., DeFrise, M., De Mol, C.: An iterative thresholding algorithm for linear inverse problems with a sparsity constraint. *Communications in Pure and Applied Mathematics* **57**, 1413–1457 (2006)
34. Daubechies, I., DeVore, R., Fornasier, M., Gunturk, C.: Iteratively re-weighted least squares minimization for sparse recovery. *Communications on Pure and Applied Mathematics* **63**(1), 1–38 (2010)
35. Davis, G., Mallat, S., Avellaneda, M.: Adaptive greedy approximations. *Journal of Constructive Approximation* **13**, 57–98 (1997)
36. Davis, J.V., Kulis, B., Jain, P., Sra, S., Dhillon, I.S.: Information-theoretic metric learning. In: *International Conference on Machine Learning*, pp. 209–216 (2007)
37. De la Torre, F., Black, M.: A framework for robust subspace learning. *International Journal of Computer Vision* **54**(1–3), 117–142 (2003)
38. Ding, C., Zhou, D., He, X., Zha, H.: R1-pca: rotational invariant L1-norm principal component analysis for robust subspace factorization. In: *Proceedings of International Conference on Machine Learning*

39. Donoho, D.L.: Compressed sensing. *IEEE Transactions on Information Theory* **52**(4), 1289–1306 (2006)
40. Donoho, D.L., Stodden, V.: Breakdown point of model selection when the number of variables exceeds the number of observations. In: *Proceedings of the International Joint Conference on Neural Networks*, pp. 16–21 (2006)
41. Donoho, D.L., Tanner, J.: Sparse nonnegative solutions of underdetermined linear equations by linear programming. In: *Proceedings of the National Academy of Sciences*, vol. 102, pp. 9446–9451 (2005)
42. Donoho, D.L., Tanner, J.: Observed universality of phase transitions in high-dimensional geometry, with implications for modern data analysis and signal processing. *Philosophical Transactions of The Royal Society A* **367**(1906), 4273–4293 (2009)
43. Donoho, D.L., Tsai, Y.: Fast solution of l_1 -norm minimization problems when the solution may be sparse. *IEEE Transactions on Information Theory* **54**(11), 4789–4812 (2008)
44. Donoho, D.L., Tsai, Y., Drori, I., Starck, J.C.: Sparse solution of underdetermined systems of linear equations by stagewise orthogonal matching pursuit. *IEEE Transactions on Information Theory* **58**(2), 1094–1121 (2012)
45. Du, L., Li, X., Shen, Y.D.: Robust nonnegative matrix factorization via half-quadratic minimization. In: *International Conference on Data Mining*, pp. 201–210 (2012)
46. Elhamifar, E., Vidal, R.: Sparse subspace clustering: Algorithm, theory, and applications. *Pattern Analysis and Machine Intelligence, IEEE Transactions on* **35**(11), 2765–2781 (2013)
47. Fan, J., Li, R.: Statistical challenges with high dimensionality: Feature selection in knowledge discovery. In: *the Madrid International Congress of Mathematicians*, pp. 595–622
48. Feng, Y., Huang, X., Shi, L., Yang, Y., Suykens, J.A.K.: Learning with the maximum correntropy criterion induced losses for regression. Tech. rep., K.U.Leuven (Leuven, Belgium) (2013)
49. Fidler, S., Skocaj, D., Leonardis, A.: Combining reconstructive and discriminative subspace methods for robust classification and regression by subsampling. *IEEE Transactions on Pattern Analysis and Machine Intelligence* **28**(3), 337–350 (2006)
50. Figueiredo, M., Nowak, R., Wright, S.J.: Gradient projection for sparse reconstruction: application to compressed sensing and other inverse problems. *IEEE Journal of Selected Topics in Signal Processing: Special Issue on Convex Optimization Methods for Signal Processing* **1**(4), 586–597 (2007)
51. Forbes, A.: Classification-algorithm evaluation: Five performance measures based on confusion matrices. *Journal of Clinical Monitoring and Computing* **11**, 189–206 (1995)
52. Fornasier, M.: Theoretical foundations and numerical methods for sparse recovery. *Radon Series on Computational and Applied Mathematics* **9**, 1–121 (2010)
53. Gao, W., Cao, B., Shan, S., Chen, X., Zhou, D., Zhang, X., Zhao, D.: The cas-peal large-scale Chinese face database and baseline evaluations. *IEEE Transactions on System Man, and Cybernetics (Part A)* **38**(1), 149–161 (2008)
54. Geman, D., Reynolds, G.: Constrained restoration and recovery of discontinuities. *IEEE Transactions on Pattern Analysis and Machine Intelligence* **14**, 367–383 (1992)
55. Geman, D., Yang, C.: Nonlinear image recovery with half-quadratic regularization. *IEEE Transactions on Image Processing* **4**(7), 932–946 (1995)
56. Golub, G., Loan, C.V.: *Matrix computations*. 3rd edition. Johns Hopkins, Baltimore (1996)
57. Guyon, I., Elissee, A.: An introduction to variable and feature selection. *Journal of Machine Learning Research* **3**, 1157–1182 (2003)
58. He, R., Ao, M., Xiang, S., Li, S.: nearest feature line: a tangent approximation. In: *Chinese Conference on Pattern Recognition* (2008)
59. He, R., Hu, B.G., Yuan, X.: Robust discriminant analysis based on nonparametric maximum entropy. In: *Asian Conference on Machine Learning* (2009)
60. He, R., Hu, B.G., Yuan, X., Zheng, W.S.: Principal component analysis based on nonparametric maximum entropy. *Neurocomputing* **73**, 1840–1952 (2010)

61. He, R., Hu, B.G., Zheng, W.S., Guo, Y.Q.: Two-stage sparse representation for robust recognition on large-scale database. In: AAAI Conference on Artificial Intelligence, pp. 475–480 (2010)
62. He, R., Sun, Z., Tan, T., Zheng, W.S.: Recovery of corrupted low-rank matrices via half-quadratic based nonconvex minimization. In: IEEE Conference on Computer Vision and Pattern Recognition, pp. 2889–2896 (2011)
63. He, R., Tan, T., Wang, L.: Recovery of corrupted low-rank matrix by implicit regularizers. *IEEE Transactions on Pattern Analysis and Machine Intelligence* **36**(4), 770–783 (2014)
64. He, R., Tan, T., Wang, L., Zheng, W.S.: $\ell_{2,1}$ regularized correntropy for robust feature selection. In: Proceedings of IEEE Conference on Computer Vision and Pattern Recognition, pp. 2504–2511 (2012)
65. He, R., Zheng, W.S., Hu, B.G.: Maximum correntropy criterion for robust face recognition. *IEEE Transactions on Pattern Analysis and Machine Intelligence* **33**(8), 1561–1576 (2011)
66. He, R., Zheng, W.S., Hu, B.G., Kong, X.W.: Nonnegative sparse coding for discriminative semi-supervised learning. In: IEEE Conference on Computer Vision and Pattern Recognition, pp. 2849–2856 (2011)
67. He, R., Zheng, W.S., Hu, B.G., Kong, X.W.: A regularized correntropy framework for robust pattern recognition. *Neural Computation* **23**(8), 2074–2100 (2011)
68. He, R., Zheng, W.S., Hu, B.G., Kong, X.W.: Two-stage nonnegative sparse representation for large-scale face recognition. *IEEE Transactions on Neural Network and Learning System* **34**(1), 35–46 (2013)
69. He, R., Zheng, W.S., Tan, T., Sun, Z.: Half-quadratic based iterative minimization for robust sparse representation. *IEEE Transactions on Pattern Analysis and Machine Intelligence* **36**(2), 261–275 (2014)
70. He, X., Yan, S., Hu, Y., Niyogi, P., Zhang, H.J.: Face recognition using laplacianfaces. *IEEE Transactions on Pattern Analysis and Machine Intelligence* **27**(3), 328–340 (2005)
71. Hellier, P., Barillot, C., Memin, E., Perez, P.: An energy-based framework for dense 3D registration of volumetric brain images. In: Proceedings of IEEE Conference on Computer Vision and Pattern Recognition (2000)
72. Ho, J., Yang, M.H., Lim, J., Lee, K.C., Kriegman, D.: Clustering appearances of objects under varying illumination conditions. In: Proceedings of IEEE Conference on Computer Vision and Pattern Recognition, vol. 1, pp. 11–18 (2003)
73. Hotelling, H.: Analysis of a complex of statistical variables into principal components. *Journal of Educational Psychology* **24**, 417–441 (1933)
74. Hou, L., He, R.: Minimum entropy linear embedding based on gaussian mixture model. In: Asian Conference on Pattern Recognition, pp. 362–366 (2011)
75. Hu, B.G., He, R., Yuan, X.: Information-theoretic measures for objective evaluation of classifications. *Acta Automatica Sinica* **38**(7), 1169–1182 (2012)
76. Hu, B.G., Wang, Y.: Evaluation criteria based on mutual information for classifications including rejected class. *Acta Automatica Sinica* **34**, 1396–1403 (2008)
77. Huber, P.: *Robust statistics*. Wiley (1981)
78. Hyvarinen, A.: Fast and robust fixed-point algorithms for independent component analysis. *IEEE Transactions on Neural Networks* **10**, 626–634 (1999)
79. Hyvärinen, A.: Survey on independent component analysis. *Neural Computing Surveys* **2**, 94–128 (1999)
80. Idier, J.: Convex half-quadratic criteria and interacting auxiliary variables for image restoration. *IEEE Transactions on Image Processing* **10**(7), 1001–1009 (2001)
81. Jenssen, R., D.Erdogmus, Principe, J., Eltoft, T.: Information theoretic angle-based spectral clustering: a theoretical analysis and an algorithm. In: International joint conference on neural networks, pp. 4904–4911 (2006)
82. Jenssen, R., Eltoft, T., Girolami, M., Erdogmus, D.: Kernel maximum entropy data transformation and an enhanced spectral clustering algorithm. In: Neural Information Processing Systems NIPS (2006)

83. Jeonga, K.H., Liu, W.F., Han, S., Hasanbelliu, E., Principe, J.C.: The correntropy mace filter. *Pattern Recognition* **42**(5), 871–885 (2009)
84. Ji, Y., Lin, T., Zha, H.: Mahalanobis distance based non-negative sparse representation for face recognition. In: *Proceedings of International Conference on Machine Learning and Applications*, pp. 41–46 (2009)
85. Jia, H., Martinez, A.M.: Face recognition with occlusions in the training and testing sets. In: *IEEE International Conference on Automatic Face & Gesture Recognition* (2008)
86. Jia, H., Martinez, A.M.: Support vector machines in face recognition with occlusions. In: *Proceedings of IEEE Conference on Computer Vision and Pattern Recognition*, pp. 136–141 (2009)
87. Kapur, J.: *Measures of information and their applications*. John Wiley, New York (1994)
88. Laaksonen, J.: Local subspace classifier. In: *International Conference on Artificial Neural Networks* (1997)
89. Lee, H., Battle, A., Raina, R., Ng, A.Y.: Efficient sparse coding algorithms. In: *Proceedings of Neural Information Processing Systems*, vol. 19, pp. 801–808 (2006)
90. Lei, Z., Chu, R., He, R., Liao, S., Li, S.Z.: Face recognition by discriminant analysis with Gabor tensor representation. In: *Advances in Biometrics* (2010)
91. Leonardis, A., Bischof, H.: Robust recognition using eigenimages. *Computer Vision and Image Understanding* **78**(1), 99–118 (2000)
92. Li, M., Chen, X., Li, X., Ma, B., Vitanyi, M.: The similarity metric. *IEEE Transactions Information Theory* **50**, 3250–3264 (2004)
93. Li, S.Z.: Face recognition based on nearest linear combinations. In: *Proceedings of IEEE Conference on Computer Vision and Pattern Recognition*, pp. 839–844 (1998)
94. Li, S.Z., Lu, J.: Face recognition using the nearest feature line method. *IEEE Transactions Neural Network* **10**(2), 439–443 (1999)
95. Li, W., Swetits, J.J.: The linear ℓ_1 estimator and the Huber m-estimator. *SIAM Journal on Optimization* **8**(2), 457–475 (1998)
96. Li, X.X., Dai, D.Q., Zhang, X.F., Ren, C.X.: Structured sparse error coding for face recognition with occlusion. *IEEE Transactions on Image Processing* **22**(5), 1889–1900 (2013)
97. Lin, D.: An information-theoretic definition of similarity. In: *International Conference on Machine Learning*, pp. 296–304 (1998)
98. Liu, R., Li, S.Z., Yuan, X., He, R.: Online determination of track loss using template inverse matching. In: *International Workshop on Visual Surveillance-VS* (2008)
99. Liu, W.F., Pokharel, P.P., Principe, J.C.: Correntropy: Properties and applications in non-gaussian signal processing. *IEEE Transactions on Signal Processing* **55**(11), 5286–5298 (2007)
100. Luenberger, D.: *Optimization by vector space methods*. Wiley (1969)
101. Mairal, J., Elad, M., Sapiro, G.: Sparse representation for color image restoration. *IEEE Transactions on Image Processing* **17**(1), 53–69 (2008)
102. Mairal, J., Sapiro, G., Elad, M.: Learning multiscale sparse representations for image and video restoration. *SIAM Multiscale Modeling & Simulation* **7**(1), 214–241 (2008)
103. Martinez, A.M., Benavente, R.: The AR face database. Tech. rep., Computer Vision Center (1998)
104. Mathar, R., Schmeink, A.: Saddle-point properties and nash equilibria for channel games. *EURASIP Journal on Advances in Signal Processing* (2009)
105. Mazeta, V., Carteret, C., Brieva, D., Idierc, J., Humbert, B.: Background removal from spectra by designing and minimising a non-quadratic cost function. *Chemometrics and Intelligent Laboratory Systems* **76**(2), 121–133 (2005)
106. Meer, P., Stewart, C., Tyler, D.: Robust computer vision: An interdisciplinary challenge, guest editorial. *Computer Vision and Image Understanding* **78**, 1–7 (2000)
107. Morejon, R.A., Principe, J.C.: Advanced search algorithms for information-theoretic learning with kernel-based estimators. *IEEE Transactions on Neural Networks* **15**(4), 874–884 (2004)
108. Moulin, P., O’Sullivan, J.A.: Information-theoretic analysis of information hiding. *IEEE Transactions on Information Theory* **49**(3), 563–593 (2003)

109. Naseem, I., Togneri, R., Bennamoun, M.: Linear regression for face recognition. *IEEE Transactions on Pattern Analysis and Machine Intelligence* **32**(11), 2106–2112 (2010)
110. Nenadic, Z.: Information discriminant analysis: feature extraction with an information-theoretic objective. *IEEE Transactions on Pattern Analysis and Machine Intelligence* **29**(8), 1394–1407 (2007)
111. Newman, D., Hettich, S., Blake, C., Merz, C.: Uci repository of machine learning databases. <http://www.ics.uci.edu/mllearn/MLRepository.html> (1998)
112. Nikolova, M., NG, M.K.: Analysis of half-quadratic minimization methods for signal and image recovery. *SIAM Journal on Scientific Computing* **27**(3), 937–966 (2005)
113. Nishii, R., Tanaka, S.: Accuracy and inaccuracy assessments in land-cover classification. *IEEE Transactions Geoscience and Remote Sensing* **37**, 491–498 (1999)
114. Niu, G., Dai, B., Yamada, M., Sugiyama, M.: Information-theoretic semi-supervised metric learning via entropy regularization. In: *International Conference on Machine Learning (2012)*
115. Niu, G., Jitkrittum, W., Dai, B., Hachiya, H., Sugiyama, M.: Squared-loss mutual information regularization: A novel information-theoretic approach to semi-supervised learning. In: *International Conference on Machine Learning (2013)*
116. Nowak, R., Figueiredo, M.: Fast wavelet-based image deconvolution using the EM algorithm. In: *Proceedings of Asilomar Conference on Signals, Systems, and Computers*, vol. 1, pp. 371–375 (2001)
117. Parzen, E.: On the estimation of probability density function and the mode. *The Annals of Mathematical Statistics* **33**, 1065–1076 (1962)
118. Patel, V.M., Chellappa, R.: *Sparse Representations and Compressive Sensing for Imaging and Vision*. SpringerBriefs in Electrical and Computer Engineering (2013)
119. Pearson, K.: On lines and planes of closest fit to systems of points in space. *The London, Edinburgh and Dublin Philosophical Magazine and Journal of Science, Sixth Series* **2**, 559–572 (1901)
120. Peng, H., Long, F., Ding, C.: Feature selection based on mutual information: criteria of max-dependency, max-relevance, and min-redundancy. *IEEE Transactions on Pattern Analysis and Machine Intelligence* **27**(8), 1226–1238 (2005)
121. Plumbley, M.: Recovery of sparse representations by polytope faces pursuit. In: *Proceedings of International Conference on Independent Component Analysis and Blind Source Separation*, pp. 206–213 (2006)
122. Pookharel, P.P., Liu, W., Principe, J.C.: A low complexity robust detector in impulsive noise. *Signal Processing* **89**(10), 1902–1909 (2009)
123. Portugal, L.F., Judice, J.J., Vicente, L.N.: A comparison of block pivoting and interior-point algorithms for linear least squares problems with nonnegative variables. *Mathematics of Computation* **63**(208), 625–643 (1994)
124. Principe, J., Xu, D., Zhao, Q., Fisher, J.: Learning from examples with information-theoretic criteria. *Journal of VLSI Signal Processing* **26**, 61–77 (2000)
125. Principe, J.C.: *Information Theoretic Learning: Renyi’s Entropy and Kernel Perspectives*. Springer (2010)
126. Principe, J.C., Xu, D., Fisher, J.W.: Information-theoretic learning. In: S. Haykin, editor, *Unsupervised Adaptive Filtering, Volume 1: Blind-Source Separation*. Wiley (2000)
127. P.Viola, N.Schraudolph, T.Sejnowski: Empirical entropy manipulation for real-world problems. In: *Proceedings of Neural Information Processing Systems*, pp. 851–857 (1995)
128. Rao, S., Liu, W., Principe, J.C., de Medeiros Martins, A.: Information theoretic mean shift algorithm. In: *Machine Learning for Signal Processing (2006)*
129. Renyi, A.: On measures of entropy and information. *Selected Papers of Alfred Renyi* **2**, 565–580 (1976)
130. Reshef, D.N., Reshef, Y.A., Finucane, H.K., Grossman, S.R., McVean, G., Turnbaugh, P.J., Lander, E.S., Mitzenmacher, M. Sabeti, P.C.: Detecting novel associations in large data sets. *Science* **334**, 1518–1524 (2011)
131. Rockfellar, R.: *Convex analysis*. Princeton Press (1970)

132. Roullot, E., Herment, A., Bloch, I., de Cesare, A., Nikolova, M., Mousseaux, E.: Modeling anisotropic undersampling of magnetic resonance angiographies and reconstruction of a high-resolution isotropic volume using half-quadratic regularization techniques. *Signal Processing* **84**(4), 743–762 (2004)
133. Rousseeuw, P.J.: Least median of squares regression. *Journal of the American Statistical Association* **79**(388), 871–880 (1984)
134. Roweis, S.: EM algorithms for PCA and SPCA. In: *Neural Information Processing Systems (NIPS)*, pp. 626–632 (1997)
135. Santamaria, I., Pokharel, P.P., Principe, J.C.: Generalized correlation function: Definition, properties, and application to blind equalization. *IEEE Transactions on Signal Processing* **54**(6), 2187–2197 (2006)
136. Seth, S., Principe, J.C.: Compressed signal reconstruction using the correntropy induced metric. In: *Proceedings of IEEE Conference on Acoustics, Speech and Signal Processing*, pp. 3845–3848 (2008)
137. Shannon, C.: A mathematical theory of communication. *Bell System Technical Journal* **27**, 623–653 (1948)
138. Sharma, A., Paliwal, K.: Fast principal component analysis using fixed-point algorithm. *Pattern Recognition Letters* **28**, 1151–1155 (2007)
139. Shi, Q., Eriksson, A., van den Hengel, A., Shen, C.: Face recognition really a compressive sensing problem. In: *Proceedings of IEEE Conference on Computer Vision and Pattern Recognition*, pp. 553–560 (2011)
140. Shi, Y., Sha, F.: Information-theoretical learning of discriminative clusters for unsupervised domain adaptation. In: *International Conference on Machine Learning* (2012)
141. Siegel, A.F.: Robust regression using repeated medians. *Biometrika* **69**(1), 242–244 (1982)
142. Sun, X.: *Matrix Perturbation Analysis*. Chinese Science Press (2001)
143. Takhar, D., Laska, J., Wakin, M., Duarte, M., Baron, D., Sarvotham, S., Kelly, K., Baraniuk, R.: A new compressive imaging camera architecture using optical-domain compression. In: *Proceedings of Computational Imaging IV at SPIE Electronic Imaging*, pp. 43–52 (2006)
144. Tao, D., Li, X., Wu, X., Maybank, S.: Tensor rank one discriminant analysis - a convergent method for discriminative multilinear subspace selection. *Neurocomputing* **71**, 1866–1882 (2008)
145. Torkkola, K.: Feature extraction by nonparametric mutual information maximization. *Journal of Machine Learning Research* **3**, 1415–1438 (2003)
146. Vapnik, V.: *The Nature of Statistical Learning Theory*. New York: Springer-Verlag (1995)
147. Vincent, P., Bengio, Y.: K-local hyperplane and convex distance nearest neighbor algorithms. In: *Advances in Neural Information Processing Systems*, vol. 14, pp. 985–992 (2001)
148. Vinh, N.X., Epps, J., Bailey, J.: Information theoretic measures for clusterings comparison: Variants, properties, normalization and correction for chance. *Journal of Machine Learning Research* **11**, 2837–2854 (2010)
149. Vo, N., Moran, B., Challa, S.: Nonnegative-least-square classifier for face recognition. In: *Proceedings of International Symposium on Neural Networks: Advances in Neural Networks*, pp. 449–456 (2009)
150. Wakin, M., Laska, J., Duarte, M., Baron, D., Sarvotham, S., Takhar, D., Kelly, K., Baraniuk, R.: An architecture for compressive image. In: *Proceedings of International Conference on Image Processing*, pp. 1273–1276 (2006)
151. Wang, S., Liu, D., Zhang, Z.: Nonconvex Relaxation approaches to robust matrix recovery. In: *International Joint Conferences on Artificial Intelligence* (2013)
152. Weiszfeld, E.: Sur le point pour lequel la somme des distances de n points donnees est minimum. *Mathematical Journal* **43**, 355–386 (1937)
153. Wright, J., Ma, Y.: Dense error correction via ℓ_1 -minimization. *IEEE Transactions on Information Theory* **56**(7), 3540–3560 (2010)
154. Wright, J., Ma, Y., Mairal, J., Sapiro, G., Huang, T.S., Yan, S.: Sparse representation for computer vision and pattern recognition. *Proceedings of IEEE* **98**(6), 1031–1044 (2010)

155. Wright, J., Yang, A.Y., Ganesh, A., Sastry, S.S., Ma, Y.: Robust face recognition via sparse representation. *IEEE Transactions on Pattern Analysis and Machine Intelligence* **31**(2), 210–227 (2009)
156. Wright, S., Nowak, R., Figueiredo, M.: Sparse reconstruction by separable approximation. In: *Proceedings of IEEE Conference on Acoustics, Speech and Signal Processing* (2008)
157. Xing, E.P., Ng, A.Y., Jordan, M.I., Russell, S.: Distance metric learning with application to clustering with side-information. In: *Proceedings of Advances in Neural Information Processing Systems*, vol. 15, pp. 505–512 (2002)
158. Xu, D.: Energy, entropy and information potential for neural computation. Ph.D. thesis, University of Florida (1999)
159. Yan, H., Yuan, X., Yan, S., Yang, J.: Correntropy based feature selection using binary projection. *Pattern Recognition* **44**, 2834–2842 (2011)
160. Yan, S., Xu, D., Zhang, B., Zhang, H., Yang, Q., Lin, S.: Graph embedding and extensions: a general framework for dimensionality reduction. *IEEE Transactions on Pattern Analysis and Machine Intelligence* **29**(1), 40–51 (2007)
161. Yang, A.Y., Sastry, S.S., Ganesh, A., Ma, Y.: Fast ℓ_1 -minimization algorithms and an application in robust face recognition: A review. In: *Proceedings of International Conference on Image Processing* (2010)
162. Yang, M., Zhang, L., Yang, J., Zhang, D.: Robust sparse coding for face recognition. In: *Proceedings of IEEE Conference on Computer Vision and Pattern Recognition*, pp. 625–632 (2011)
163. Yang, S., Zha, H., Zhou, S., Hu, B.G.: Variational graph embedding for globally and locally consistent feature extraction. In: *Europe Conference on Machine Learning (ECML)*, pp. 538–553 (2009)
164. Yin, W., Osher, S., Goldfarb, D., Darbon, J.: Bregman iterative algorithms for ℓ_1 -minimization with applications to compressed sensing. *SIAM Journal on Imaging Sciences* **1**(1), 143–168 (2008)
165. Yuan, X.T., Hu, B.G.: Robust feature extraction via information theoretic learning. In: *Proceedings of International Conference on Machine Learning*, pp. 1193–1200 (2009)
166. Yuan, X.T., Li, S.: Half quadratic analysis for mean shift: with extension to a sequential data mode-seeking method. In: *IEEE International Conference on Computer Vision* (2007)
167. Zhang, L., Yang, M., Feng, X.: Sparse representation or collaborative representation: Which helps face recognition? In: *Proceedings of IEEE International Conference on Computer Vision* (2011)
168. Zhang, T.: Multi-stage convex relaxation for learning with sparse regularization. In: *Proceedings of Neural Information Processing Systems*, pp. 16–21 (2008)
169. Zhang, T.H., Tao, D.C., Li, X.L., Yang, J.: Patch alignment for dimensionality reduction. *IEEE Trans. Knowl. Data Eng.* **21**(9), 1299–1313 (2009)
170. Zhang, Y., Sun, Z., He, R., Tan, T.: Robust low-rank representation via correntropy. In: *Asian Conference on Pattern Recognition* (2013)
171. Zhang, Y., Sun, Z., He, R., Tan, T.: Robust subspace clustering via half-quadratic minimization. In: *International Conference on Computer Vision* (2013)
172. Zhang, Z.: Parameter estimation techniques: A tutorial with application to conic fitting. *Image and Vision Computing* **15**(1), 59–76 (1997)
173. Zou, H.: The adaptive lasso and its oracle properties. *Journal of the American Statistical Association* **101**(476), 1418–1429 (2006)