Subspace Clustering Under Multiplicative Noise Corruption

Baohua Li(✉) and Wei Wu

School of Mathematical Sciences, Dalian University of Technology, Dalian, China
libaoh@mail.dlut.edu.cn

Abstract. Traditional subspace clustering models generally adopt the hypothesis of additive noise, which, however, dose not always hold. When it comes to multiplicative noise corruption, these models usually have poor performance. Therefore, we propose a novel model for robust subspace clustering with multiplicative noise corruption to alleviate this problem, which is the key contribution of this work. The proposed model is evaluated on the Extend Yale B and MNIST datasets and the experimental results show that our method achieves favorable performance against the state-of-the-art methods.

Keywords: Subspace clustering · Multiplicative noise

1 Introduction

In computer vision society, some missions can be modeled as drawing high-dimensional samples from the union of multiple low-dimensional linear subspaces, that is subspace clustering. By subspace clustering, the missions such as motion trajectories segmentation [1,2], face images clustering [3], and video segmentation [4] can be reasonably interpreted and whose underlying structure of data also be well uncovered. Subspace clustering has received significant attention in recent years and many elaborate models have been developed. Thereinto, the reconstruction based clustering methods obtain good performance by employing the data *self-expressiveness* strategy which assumes that each data point in a union of subspace can be reconstructed by a linear combination of other points in the dataset. As for reconstruction based clustering, many methods have been proposed. Literature [5–8] proposed the sparse representation based algorithm to seek a sparse coefficient matrix in some sense. While, literatures [9–12] intend to find the low rank solution of the clustering models by imposing the nuclear norm constraint. Using the ridge regression technique, authors in [13] proposed the least square regression method for subspace clustering under enforced block diagonal condition and grouping effect guarantee. To balance the sparse and density, literatures [14,15] combine a quadratic data-fidelity term with trace lasso and elastic net regularization terms respectively to model the subspace clustering problem. In [16–18], the non-Gaussian noise corruption and

Y. Sun et al. (Eds.): IScIDE 2017, LNCS 10559, pp. 461–470, 2017.
DOI: 10.1007/978-3-319-67777-4_41

impulsive noise corruption are considered. In addition, more general noise pollution is involved in [19] by mixture Gaussian regression. The mentioned methods perform subspace clustering issues in two stages: (1) using the estimated coefficient matrix to construct the affinity matrix which characterizes the degree of affinity among the different data points; (2) Applying the Normalized Cuts [30] to the affinity matrix to find the final clustering result. Therefore, the core of successful subspace clustering lies in how to find a suitable coefficient matrix.

Most reconstruction based subspace clustering methods are based on the assumption of additive noise, in which the captured data $A \in R^{m\times n}$ can be uniform shown as

$$A = AX + Z \tag{1}$$

where $X \in R^{n\times n}$ is the desired coefficient matrix, and $Z \in R^{m\times n}$ is the unknown noise term.

However, coherent imaging systems such as Doppler imaging, synthetic aperture radar (SAR) imaging, synthetic aperture sonar (SAS) imaging and ultrasound imaging, which play an important in medical, military, aerospace and other fields, are usually interfered by multiplicative noise. The multiplicative noise scenarios have been carefully studied by many researchers and obtained exciting performance on image restoration [20–24] etc. Usually, the degraded image F under multiplicative noise corruption is given by

$$G = (RF) \odot \varGamma \tag{2}$$

where $\odot$ denotes the Hadamard product which works in element-wise multiplication manner. R is the blur matrix, $\varGamma$ denotes the noise, and G is the observed image. Here G, H, F and $\varGamma$ have the agreed dimension. Especially, when R is the identity matrix, the abovementioned method (2) fails in the multiplicative noise removal problem which motivates us to consider the subspace clustering problem under multiplicative noise scenario. We contrast the Eq. (1) with (2), and proceed to provide the mathematical description of Eq. (1) with multiplicative noise flavour as

$$A = B \odot Z \tag{3}$$

where the elements of Z is not zero, A is the observed data, and B is the clean data. Note that, our ideal is to estimate a clean data matrix and then find its reconstruction coefficient. Therefore, we use B instead of AX in (3), where X is the coefficient matrix. If the elements of Z are all equal to 1, the data is clean. In this case, (3) boils down the noise free subspace clustering problem. Based on (3), a novel subspace clustering model is proposed by us.

The rest of this paper is organized as follows. Section 2 presents the necessary notations and preliminaries and based on which our proposed model is developed. Following that, we design an ADMM [26,27] based algorithm to solve the proposed model in Sect. 3. Experimental results are presented in Sect. 4. Finally, we make a conclusion in Sect. 5.

2 Notations and Preliminaries

We use capital and lowercase letters to represent matrixes and scalars respectively. Two operators will be utilized in the following pages. For a matrix C, we denote $vec\,(C)$ to return a column vector which contains all the elements of C. The operation $vec\,(C)$ is the same as the Matlab command *vec*. We denote $diag\,(C)$ to return a diagonal matrix, which is equal to the Matlab command $diag\,(vec\,(C))$. Let $\parallel C \parallel_F$, $\parallel C \parallel_1$, and $\parallel C \parallel_\star$ denote the matrix Frobenius norm, ℓ_1 norm and nuclear norm of a matrix C respectively.

Given the corrupted data matrix $A = (A_1, \cdots, A_n)$ whose columns are data points drawn from k subspaces, different from the existing subspace clustering methods, we solve the following model

$$\min_{H,X,B} \frac{1}{2} \parallel H - \alpha E \parallel_F^2 + \lambda_1 \parallel B - A \odot H \parallel_1 + \frac{1}{2} \parallel BX - B \parallel_F^2 + \lambda_2 \parallel X \parallel_* \quad (4)$$

to conduct the subspace clustering problem in multiplicative noise scenario. Where H is the matrix whose entries are the reciprocal of noise, α is set to be the mean of H, and E is a matrix in which the entries are equal to 1. $\lambda_1 > 0$ and $\lambda_2 > 0$ are regularization parameters to balance the objective function. According the characteristic of subspace clustering data, we remove the total variation regularization term.

In addition, the common subspace clustering models directly employ the observed data, which are serious corrupted. It weakens the correlation of the data points within the same subspace, and leads to the poor performance. Instead of utilizing the observed data, our proposed model (4) resorts to an estimated data matrix which is obtained by preprocessing the observed data. Since the proposed model (4) involves the variance of the inverse of the noise, we need to manually input the mean of the inverse of noise before solving the object function. We list some frequently used multiplicative noises [21,24] etc. and the mean value of their inverse as bellow:

– The probability density function of Gamma noise is:

$$p\,(x) = \frac{x^{k-1}}{\Gamma\,(k)\,\theta^k} \exp\left(\frac{-x}{\theta}\right)$$

with $\theta > 0$ and $k > 1$. The mean value of $\frac{1}{x}$ is derived by

$$\begin{aligned}\mathbb{E}\left(\frac{1}{x}\right) &= \frac{\Gamma\,(k-2)}{\theta^2 \Gamma\,(k)} \int_0^{+\infty} \frac{x^{k-2}}{\Gamma\,(k-2)\,\theta^{k-2}} \exp\left(\frac{-x}{\theta}\right) dx \\ &= \theta \frac{\Gamma\,(k-2)}{\theta^2 \Gamma\,(k)} (k-2) \\ &= \frac{1}{\theta\,(k-1)}\end{aligned}$$

– The probability density function of Rayleigh noise is:

$$p(x) = \frac{x}{\sigma^2} \exp\left(-\frac{x^2}{2\sigma^2}\right)$$

The mean value of $\frac{1}{x}$ is derived by:

$$\begin{aligned}\mathbb{E}\left(\frac{1}{x}\right) &= \int_0^{+\infty} \frac{1}{x}\frac{x}{\sigma^2} \exp\left(\frac{-x^2}{2\sigma^2}\right) dx \\ &= \frac{\sqrt{2}}{\sigma} \int_0^{+\infty} \exp\left(-\left(\frac{x}{2\sigma}\right)^2\right) d\frac{x}{\sqrt{2}\sigma} \\ &= \sqrt{\frac{\pi}{2\sigma^2}}\end{aligned}$$

– Besides in additive noise situation, the Gaussian noise is also frequently used in multiplicative noise case. The probability density function of Gaussian noise is:

$$p(x) = \frac{1}{\sigma\sqrt{2\pi}} exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right)$$

Unlike the first two density functions, the mean value of $\frac{1}{x}$ in case of Gaussian

$$\mathbb{E}\left(\frac{1}{x}\right) = \frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^{+\infty} \frac{1}{x} \exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right) dx$$

is not showed as a expression. Thanks to the numerical approximation [25] for above integral, which makes the approximate value of the inverse of Gaussian noise feasible.

3 The Numerical Algorithm

We are now in the position of numerical method section. Model (4) is reformulated as following constrained optimization problem:

$$\begin{aligned} \min_{H,X,B} \frac{1}{2} \parallel H - \alpha E \parallel_F^2 + \lambda_1 \parallel Z \parallel_1 + \frac{1}{2} \parallel BW - B \parallel_F^2 + \lambda_2 \parallel X \parallel_* \\ s.t.\ B - A \odot H = Z \\ W = X \end{aligned} \tag{5}$$

by introducing the auxiliary variables Z and W. The augmented Lagrangian function of (5) is:

$$\begin{aligned} &\mathcal{L}(H, X, B, W, Z, \Lambda_1, \Lambda_2) = \frac{1}{2} \parallel H - \alpha E \parallel_F^2 + \frac{1}{2} \parallel BW - B \parallel_F^2 + \lambda_1 \parallel Z \parallel_1 \\ &+ \lambda_2 \parallel X \parallel_* + tr\left(\Lambda_1^\top (B - A \odot H - Z)\right) + tr\left(\Lambda_2^\top (W - X)\right) \\ &+ \frac{\beta}{2}\left(\parallel B - A \odot H - Z \parallel_F^2 + \parallel W - X \parallel_F^2\right) \end{aligned} \tag{6}$$

Here, Λ_1, Λ_1 are multipliers, and β is the penalty parameter. Operation $tr(\cdot)$ denotes the trace of a square matrix. Our optimization contains two major steps: optimizing variables and optimizing multipliers. When updating a certain variable or multiplier, we always assume the rest items are fixed.

Ignoring the irrelevant terms, we update B via solving following optimization problem:

$$\min_{B} \frac{1}{2} \parallel BW - B \parallel_F^2 + \frac{\beta}{2} \parallel B - A \odot H - Z + \frac{\Lambda_1}{\beta} \parallel_F^2 \tag{7}$$

Similar to the problem (7), we find W, Z, X and H as follows: Update W

$$\min_{W} \frac{1}{2} \parallel BW - B \parallel_F^2 + \frac{\beta}{2} \parallel W - X + \frac{\Lambda_2}{\beta} \parallel_F^2 \tag{8}$$

Update Z

$$\min_{Z} \frac{1}{2} \parallel B - A \odot H - Z + \frac{\Lambda_2}{\beta} \parallel_F^2 + \parallel Z \parallel_1 \tag{9}$$

Update X

$$\min_{X} \frac{1}{2} \parallel W - X + \frac{\Lambda_2}{\beta} \parallel_F^2 + \frac{\lambda_2}{\beta} \parallel X \parallel_* \tag{10}$$

Updata H

$$\min_{H} \frac{1}{2} \parallel H - \alpha E \parallel_F^2 + \frac{\beta}{2} \parallel B - A \odot H - Z + \frac{\Lambda_2}{\beta} \parallel_F^2 \tag{11}$$

In order to decouple the $A \odot H$ and make the subproblem (11) easy to be carried out, we reformulate (11) as

$$\begin{aligned} &\min_{vec(H)} \frac{1}{2} \parallel vec(H) - \alpha vec(E) \parallel_2^2 \\ &\quad + \frac{\beta}{2} \parallel vec(B) - diag(A)\, vec(H) - vec(Z) + \frac{vec(\Lambda_2)}{\beta} \parallel_F^2 \end{aligned} \tag{12}$$

Our overall ADMM based algorithm is summarized in Algorithm 1. Since the update of the multipliers Λ_1 and Λ_2 does not need to solve optimization function, we just list their update processes in Algorithm 1. In addition, the shrinkage operator $\mathcal{S}_\tau[x]$ will be exploited in our algorithm to simplify the expressions of optimization problems:

$$\mathcal{S}_\tau[x] = \begin{cases} x - \tau, \ if \ \ x > \tau \\ x + \tau, \ if \ \ x < -\tau \\ 0, \ otherwise \end{cases} \tag{13}$$

Moreover, the discussion of convergence guarantee of Algorithm 1 is similar to [28,29], we thus omit it. After calculating the reconstruction coefficient X, we construct the affinity matrix by $\frac{(|X|+|X^\top|)}{2}$. Then, the Normalized Cuts [30] is applied to perform the subspace clustering. The clustering accuracy is defined by:

$$Accuracy = 1 - \frac{missclustered\ points}{total\ data\ points} \times 100\%$$

Algorithm 1: Finding the solution by ADMM

Input: The dataset, implicative noise, regularization parameter λ_1, λ_2 and β, the mean α of inverse noise, the number of subspaces k, and tolerate error ε

Given the initial estimation of W, Z, X, H, Λ_1, Λ_2

Repeat:

1. Update B via solving (7)

$B^{(n+1)} = \beta\left(A \odot H^{(n)} - Z^{(n)} + \frac{\Lambda_1^{(n)}}{\beta}\right)\left(\left(W^{(n)} - I\right)\left(W^{(n)} - I\right)^\top + \beta I\right)^{-1}$

2. Update W via solving (8)

$W^{(n+1)} = \left(B^{(n+1)\top} B^{(n+1)} + \beta I\right)^{-1}\left(B^{(n+1)\top} B^{(n+1)} + \beta X^{(n)} - \Lambda_2^{(n)}\right)$

3. Update Z via solving (9)

$Z^{(n+1)} = \mathcal{S}_{\frac{\lambda_2}{\beta}}[B^{(n+1)} - A \odot H^{(n)} + \frac{\Lambda_2^{(n)}}{\beta}]$

4. Update X via solving (10)

$(U, \Sigma, V) = svd\left(W^{(n+1)} + \frac{\Lambda_2^{(n)}}{\beta}\right)$

$X^{(n+1)} = U\mathcal{S}_{\frac{\lambda_2}{\beta}}[\Sigma]V^\top$

5. Update H via solving (12)

$vec\,(H)^{(n+1)} = \left(I + \beta\,(diag\,(A))^2\right)^{-1}\left(\alpha vec\,(E) + \beta vec\left(B^{(n+1)} - Z^{(n+1)} + \frac{\Lambda_2^{(n)}}{\beta}\right)\right)$

we finish updating H by transforming $vec\,(H)^{(n+1)}$ in to matrix $H^{(n+1)}$

6. Update the multipliers Λ_1 and Λ_2 as follows

$\Lambda_1^{(n+1)} = \Lambda_1^{(n)} + \beta\left(B^{(n+1)} - A \odot H^{(n+1)} - Z^{(n+1)}\right)$

and

$\Lambda_2^{(n+1)} = \Lambda_2^{(n)} + \beta\left(W^{(n+1)} - X^{(n+1)}\right)$

7. Until

$\| X^{(n+1)} - X^{(n)} \|_F \leq \varepsilon$

$\| W^{(n+1)} - W^{(n)} \|_F \leq \varepsilon$

$\| Z^{(n+1)} - Z^{(n)} \|_F \leq \varepsilon$

$\| H^{(n+1)} - H^{(n)} \|_F \leq \varepsilon$

End loop. Output X

4 Experiments

In this section we will report the performance of our new method on Extended Yale B database and the MNIST database of hand-written digits under different commonly used multiplicative noise corruption respectively. Our method is evaluated with four state-of-the-art algorithms, that are SSC [6], LRR [11], LSR [13] and CASS [14]. The experimental results are list in tables. The parameters for each clustering algorithm are tuned to reach the best performance. The clustering accuracies of our method are shown in red color.

The extend Yale B database contains 38 objects in total and each object contains 64 images. Here, we take it's first 5, 8 and 10 classes data and corrupt them with frequently used multiplicative noises, then proceed to group these data. The data are reduced into a 30, 48, 60-dimension new data respectively before solving the object function. Figure 1 illustrates the clean data, the corrupted data and the restored data. Note that, we take its first 1 class 64 images without reducing dimensions, and show the restoring result. The restored images are closer to the clean ones, which avoids poor performance caused by serious corruption of noises. The results of clustering accuracies prove the superiority of employing restored estimated data matrix in model (4). Tables 1, 2 and 3 state the clustering result. It demonstrates that our method outperforms the SSC, LRR, LSR, and CASS. For the 5 objects face clustering issue, all the mentioned methods reach satisfied accuracies under the serous multiplicative noise corruption. As for the 8 objects and 10 objects situations, our method makes a remarkable improvement which dues to both the noise removing step and the grouping effect of nuclear norm regularization term of our model. We can see that LRR method, which adopts nuclear norm regularization term too, however, doesn't perform very well. It further proves that exploiting estimated data matrix through noise removing step is reasonable. It is also observed that the performance of each method decreases from the Tables 1, 2 and 3 as the numbers of object get larger. For the reason that the trace lasso norm is adaptive to the correlation of data, CASS performs better than SSC, LRR and LSR. The SSC enjoys the property of variable selection [31] while LSR and LRR possess the grouping effect. These elegant properties guarantee the performance of above-mentioned methods. However, the serious multiplicative noise may break the correlation of data and depress the elegant properties. Therefore, the performance of CASS, SSC, LRR, LSR are limited. Our model adopt noise removed data to estimate the coefficient matrix, which is different from the previous methods. By employing the learned data matrix which is less affected by the noise, we obtain the competitive clustering accuracies.

The MNIST database of hand-written digitals is another challenging data set in subspace clustering society. It contains 10 objects which are corresponding to 0–9 of size 28×28 pixels. We select a subset of it in which each object of 0–9 contains 50 samples. Before grouping the chosen data by all methods, we corrupt them by the three common multiplicative noises, and use PCA to project them into 12-dimensional data inspired by [32]. We evaluate the models according to different noises and report the clustering accuracies in the Table 4. As we can see, in all cases, our method performances better than the others. According to the experimental results, we know that the clustering accuracies heavily depend on the noises removal step in the multiplicative noises corruption.

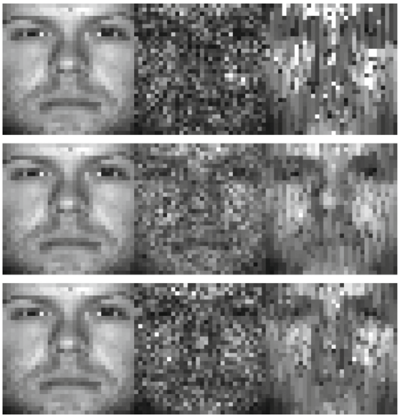

Fig. 1. (a) Clean images, (b) The noisy images corrupted by multiplicative Rayleigh, Gamma and Gaussian noise from the top down. (c) The corresponding noise removed results.

Table 1. The clustering accuracies about 5 objects

	5 objects		
	Gamma	Gaussian	Rayleigh
Our method	64.06%	61.75%	59.07%
SSC	52.64%	53.08%	50.43%
LRR	52.73%	51.77%	47.33%
LSR	50.35%	48.04%	45.17%
CASS	56.97%	55.04	52.72%

Table 2. The clustering accuracies about 8 objects

	8 objects		
	Gamma	Gaussian	Rayleigh
Our method	53.05%	51.83%	49.31%
SSC	45.69%	43.83%	40.76%
LRR	46.74%	44.01%	42.33%
LSR	41.63%	39.78%	40.02%
CASS	48.76%	46.55%	43.75%

Table 3. The clustering accuracies about 10 objects

	10 objects		
	Gamma	Gaussian	Rayleigh
Our method	41.96%	40.32%	39.31%
SSC	37.56%	34.85%	33.61%
LRR	34.07%	31.86%	32.63%
LSR	33.73%	29.06%	28.61%
CASS	38.77%	36.01%	33.88%

Table 4. The clustering accuracies of MINST data

	MINST data		
	Gamma	Gaussian	Rayleigh
Our method	39.59%	39.04%	38.88%
SSC	35.06%	33.67%	30.96%
LRR	34.66%	31.07%	28.99%
LSR	32.86%	29.43%	27.96%
CASS	36.11%	35.01%	31.96%

5 Conclusion

In this work, we propose a novel reconstruction based subspace clustering model with multiplicative noises corruption. To the best of our knowledge, it is the first work to address the subspace clustering issue under the multiplicative noises scenario. An effective numerical algorithm is designed based on ADMM. At last, the experimental results on the Extended Yale B and MNIST datasets state the effectiveness of our method.

References

1. Costeira, J.P., Kanade, T.: A multibody factorization method for independently moving objects. Int. J. Comput. Vis. **29**(3), 159–179 (1998)
2. Rao, S., Tron, R., Vidal, R., Ma, Y.: Motion segmentation in the presence of outlying, incomplete, or corrupted trajectories. IEEE Trans. Pattern Anal. Mach. Intell. **32**(10), 1832–1845 (2010)
3. Basri, R., Jacobs, D.W.: Lambertian reflectance and linear subspaces. IEEE Trans. Pattern Anal. Mach. Intell. **25**(2), 218–233 (2003)
4. Xiao, S., Tan, M., Xu, D.: Weighted block-sparse low rank representation for face clustering in videos. In: Fleet, D., Pajdla, T., Schiele, B., Tuytelaars, T. (eds.) ECCV 2014. LNCS, vol. 8694, pp. 123–138. Springer, Cham (2014). doi:10.1007/978-3-319-10599-4_9
5. Elhamifar, E., Vidal, R.: Sparse subspace clustering. In: Proceedings of the IEEE Computer Vision and Pattern Recognition, pp. 2790–2797 (2009)
6. Elhamifar, E., Vidal, R.: Sparse subspace clustering: algorithm, theory, and applications. IEEE Trans. Pattern Anal. Mach. Intell. **35**(11), 2765–2781 (2013)
7. Li, C.-G., Vidal, R.: Structured sparse subspace clustering: a unified optimization framework. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pp. 277–286 (2015)
8. Yin, M., Guo, Y., Gao, J., He, Z., Xie, S.: Kernel sparse subspace clustering on symmetric positive definite manifolds. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pp. 5157–5164 (2016)
9. Liu, G., Lin, Z., Yu, Y.: Robust subspace segmentation by low-rank representation. In: Proceedings of the 27th International Conference on Machine Learning, pp. 663–670 (2010)
10. Favaro, P., Vidal, R., Ravichandran, A.: A closed form solution to robust subspace estimation and clustering. In: Proceedings of the IEEE Computer Vision and Pattern Recognition, pp. 1801–1807 (2011)
11. Liu, G., Lin, Z., Yan, S., Sun, J., Yu, Y., Ma, Y.: Robust recovery of subspace structures by low-rank representation. IEEE Trans. Pattern Anal. Mach. Intell. **35**(1), 171–184 (2013)
12. Vidal, R., Favaro, P.: Low rank subspace clustering (LRSC). Pattern Recogn. Lett. **43**, 47–61 (2014)
13. Lu, C.-Y., Min, H., Zhao, Z.-Q., Zhu, L., Huang, D.-S., Yan, S.: Robust and efficient subspace segmentation via least squares regression. In: Fitzgibbon, A., Lazebnik, S., Perona, P., Sato, Y., Schmid, C. (eds.) ECCV 2012. LNCS, vol. 7578, pp. 347–360. Springer, Heidelberg (2012). doi:10.1007/978-3-642-33786-4_26

14. Lu, C., Feng, J., Lin, Z., Yan, S.: Correlation adaptive subspace segmentation by trace lasso. In: Proceedings of the IEEE International Conference on Computer Vision, pp. 1345–1352 (2013)
15. You, C., Li, C.-G., Robinson, D.P., Vidal, R.: Oracle based active set algorithm for scalable elastic net subspace clustering. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pp. 3928–3937 (2016)
16. Lu, C., Tang, J., Lin, M., Lin, L., Yan, S., Lin, Z.: Correntropy induced L2 graph for robust subspace clustering. In: Proceedings of the IEEE International Conference on Computer Vision, pp. 1801–1808 (2013)
17. Zhang, Y., Sun, Z., He, R., Tan, T.: Robust subspace clustering via half-quadratic minimization. In: Proceedings of the IEEE International Conference on Computer Vision, pp. 3096–3103 (2013)
18. He, R., Zhang, Y., Sun, Z., Yin, Q.: Robust subspace clustering with complex noise. IEEE Trans. Image Process. **24**(11), 4001–4013 (2015)
19. Li, B., Zhang, Y., Lin, Z., Lu, H.: Subspace clustering by mixture of Gaussian regression. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pp. 2094–2102 (2015)
20. Huang, Y.-M., Lu, D.-Y., Zeng, T.: Two-step approach for the restoration of images corrupted by multiplicative noise. SIAM J. Sci. Comput. **35**(6), 2856–2873 (2013)
21. Zhao, X.-L., Wang, F., Ng, M.K.: A new convex optimization model for multiplicative noise and blur removal. SIAM J. Imaging Sci. **7**(1), 456–475 (2014)
22. Aubert, G., Aujol, J.-F.: A variational approach to removing multiplicative noise. SIAM J. Appl. Math. **68**(4), 925–946 (2008)
23. Rudin, L., Lions, P.-L., Osher, S.: Multiplicative denoising and deblurring: theory and algorithms. In: Osher, S., Paragios, N. (eds.) Geometric Level Set Methods in Imaging, Vision, and Graphics, pp. 103–119. Springer, Heidelberg (2003). doi:10.1007/0-387-21810-6_6
24. Wang, F., Zhao, X.-L., Ng, M.K.: Multiplicative noise and blur removal by framelet decomposition and l_1-based l-curve method. IEEE Trans. Image Process. **25**(9), 4222–4232 (2016)
25. Shampine, L.F.: Vectorized adaptive quadrature in matlab. J. Comput. Appl. Math. **211**(2), 131–140 (2008)
26. Boyd, S., Parikh, N., Chu, E., Peleato, B., Eckstein, J.: Distributed optimization and statistical learning via the alternating direction method of multipliers. Found. Trends® Mach. Learn. **3**(1), 1–122 (2011)
27. Ng, M.K., Wang, F., Yuan, X.: Inexact alternating direction methods for image recovery. SIAM J. Sci. Comput. **33**(4), 1643–1668 (2011)
28. Mairal, J., Bach, F., Ponce, J., Sapiro, G.: Online dictionary learning for sparse coding. In: International Conference on Machine learning, pp. 689–696 (2009)
29. Nguyen, H., Yang, W., Sheng, B., Sun, C.: Discriminative low-rank dictionary learning for face recognition. Neurocomputing **173**, 541–551 (2016)
30. Shi, J., Malik, J.: Normalized cuts and image segmentation. IEEE Trans. Pattern Anal. Mach. Intell. **22**(8), 888–905 (2000)
31. Tibshirani, R.: Regression shrinkage and selection via the lasso. J. Roy. Stat. Soc. Ser. B (Methodol.) **73**(3), 267–288 (1996)
32. Hastie, T., Simard, P.Y.: Metrics and models for handwritten character recognition. Stat. Sci. **13**(1), 54–65 (1998)