Motion Segmentation using Collaborative Low-Rank and Sparse Subspace Clustering

Yao Zhang Min Li
College of Mathematics and Statistics Shenzhen University Shenzhen, PR China
E-mail:limin800@szu.edu.cn

***Abstract*—We propose a method based on the collaborative Low-Rank Representation (LRR) and Sparse Subspace Clustering (SSC) to cluster data drawn from multiple linear subspaces in a high-dimensional space. Given a set of data vectors, Collaborative Low-Rank and Sparse Subspace Clustering (CLRS) want to seek a better representation among the candidates that represent all vectors as affine combination of the bases in a dictionary. Both theoretical and experimental results show that CLRS is a promising method for subspace segmentation.**

***Keywords*-LRR; SSC; CLRS; Subspace Segmentation;**

I. Introduction

Motion segmentation is to divide a video sequence in to its constitant component that is more meaningful and easier to analysis. Up to know, a broad variety of models have been proposed to solve the motion segmentation problem, in which subspace clustering is very efficient [1] [2] [3].

As an example, Figure 1(b) illustrates the set of image points extracted for the first frame of the video shown in Figure 1(a), Figure 1(c) shows the first frame of the point trajectories to be segmented, Figure 1(d) shows the same point trajectories segmented into two subspaces corresponding to the red car and the background respectively.

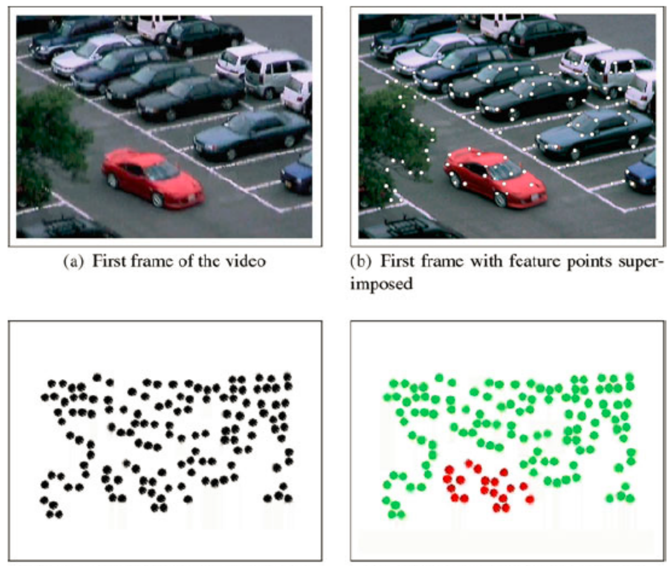
Given a set of data from a union of subspaces, linear model is a attractive choice for us, because it is easy to compute and also effective in many areas with real applications.

Subspace clustering is an unsupervised machine learning method, cluster points in the dataset to their respective subspace.

Existing works on subspace representation and clustering can be roughly divided into four main categories: statistical model, algebraic method, LRR model model [4] [5] and SSC method [4] [6].

Statistical models, such as K-Subspaces and Mixture of Probabilistic Principal Component Analysis model (MPPCA) [7], assume that the data subject to normal distribution. And above all require the numbers of subspaces and the dimensions of each subspace to be known. However, they are sensitive to noise and outliers.

Algebraic methods, such as generalized principal component analysis (GPCA) [8], is able to deal with subspaces



(a) First frame of the video (b) First frame with feature points superimposed

(c) Feature points before segmentation (d) Feature points after segmentation

Figure 1. An example of motion segmentation [4].

of different dimensions. However, GPCA is sensitive to outliers, and its complexity increase exponentially with number of subspaces and their dimensions.

LRR model uses the lowest-rank representation of all data. As a common practice in rank minimization problem, we usually replace the rank function with the nuclear norm. Unfortunately, there is the disadvantage when the data in high-rank space, for example there are 100 subspaces of dimension 5, the low-rank structure is missing.

SSC method uses the sparest representation produced by l_1-minimization to define the affinity matrix of an undirected graph. Then subspace segmentation is performed by spectral clustering algorithms. SSC require that each data point in a union of subspaces only can be expressed as a linear combination of other data points in the same subspace. But this may cause a over spare problem.

The motivation of our study is trying to provide a new way that can combine the advantages of low-rank and sparse. More precisely, this paper is giving the new model CLRS which can avoid the missing of low-rank structure and over spare problem. CLRS seeks a better representation which reflects the features of sparse and low rank in motion segmentation.

The paper is organized as follows: Section II gives the CLRS model, Section III gives the optimization, Section IV gives the experiments, Section V gives the conclusion and future work.

II. Collaborative Low-Rank and Sparse Subspace Clustering

A. Problem Formulation

Let $X = [x_1, x_2, ..., x_N]$ be a matrix, each column of which is a vector $x_i \in R^{D\times 1}$ corresponding to a point of data.

The points of X are drawn from a union of n different linear subspaces $\{S_1, S_2, ..., S_n\}$ of dimensions $\{d_1, d_2, ..., d_n\}$. Then

$$X = [\mathrm{x}_1, \mathrm{x}_2, ..., \mathrm{x}_N] = [\mathrm{X}_1, \mathrm{X}_2, ..., \mathrm{X}_n]\Gamma^T, \tag{1}$$

where $\mathrm{X}_i \in R^{D\times N_i}$ is a matrix containing the $N_i > d_i$ points that lie in subspace S_i, and $\Gamma \in R^{N\times N}$ is an unknown permutation matrix, which sorts the columns of X according to which subspace they are belong to. Our task is to cluster X into different subspace which they are belong to.

When there is noisefree, $\forall x_i \in X$ can be expressed as a linear combination of $\{x_j\}_{j=1}^N$, namely

$$x_i = \sum_{j=1}^{N} x_j z_{j,i}, \tag{2}$$

where $z_{j,i} \in R^{N\times 1}$ is a vector of coefficients corresponding to x_i. If $\sum_{j=1}^{N} z_{j,i} = 1$, we call x_i can be expressed as a affine combination of $\{x_j\}_{j=1}^N$.

When there is noise, a way of modeling errors proposed in [6] is to relax the self-expressiveness constraint to

$$x_i = \sum_{j=1}^{N} x_j z_{j,i} + (x_i - \sum_{j=1}^{N} x_j z_{j,i}) \tag{3}$$

where $x_i - \sum_{j=1}^{N} x_j z_{j,i}$ is the error.

In a matrix form, the equation (3) is equivalent to

$$X = XZ + (X - XZ) \tag{4}$$

where$Z = (z_{i,j})_{N\times N}$, $X - XZ$ is the matrix of errors.

To avoid the low-rank structure is missing and over sparse problem of X. We seek a matrix to represent Z which is called collaborative low-rank and sparse method.

B. Collaborative Low-Rank and Sparse Subspace Clustering

It is well known that, the LRR model [5] is

$$\min_Z rank(Z) + \lambda ||X - XZ||_l \tag{5}$$

where $||\cdot||_l$ is a type of norms, such as $||\cdot||_0, ||\cdot||_2, ||\cdot||_{2,0}$, λ is a parameter which is related to noise.

The SSC model [6] is

$$\min_Z ||Z||_0 + \lambda ||X - XZ||_l,\ s.t.\ diagZ = 0 \tag{6}$$

where $||\cdot||_l$ and λ are the same as above.

In order to reflect the feature of low rank and sparse in motion segmentation, and to avoid the low-rank structure is missing and over sparse problem of X. Moreover the points of data lie in a union of 3-D affine subspaces in some real-world, such as motion segmentation. So our model CLRS is

$$\begin{aligned} \min_Z\ & \lambda_1 rank(Z) + \lambda_2 ||Z||_1 + \lambda_n ||X - XZ||_l \\ & s.t.\ diagZ = 0, 1^T Z = 1^T \end{aligned} \tag{7}$$

where $1 \in R^{N\times 1}, 1 = [1, ..., 1]^T$, λ_1 and λ_2 balance the low-rank and sparsity of Z, λ_n is a parameter that is related to the noise of data. And when $\lambda_2 = 0$, CLRS looks like LRR if without the constraint condition $diagZ = 0$, in the other hand, when$\lambda_1 = 0$, CLRS is SSC model. Here $diagZ = 0$ can remove the trivial solution of Z when λ_1 is to small.

The CLRS optimization problem above is difficult to solve due to the problem is non-convex problem. Fortunately, inspired by the work [9] [10], we can provide a convex optimization:

$$\begin{aligned} \min_Z\ & \lambda_1 ||Z||_* + \lambda_2 ||Z||_1 + \frac{\lambda_3}{2} ||X - XZ||_F^2 \\ & s.t.\ diagZ = 0, 1^T Z = 1^T \end{aligned} \tag{8}$$

In order to compute easily, we set $\frac{\lambda_3}{2}$ in (8) instead of λ_n.

III. Optimization

Now we discuss how to solve and optimize the above problem (8). The alternating direction method of multipliers (ADMM) [4] is an algorithm that solves convex optimization problems by breaking them into smaller pieces, each of which are easier to handle. In this paper, we use ADMM to solve the optimization. We begin with a short review of ADMM, which solves the model in the form of

$$\min_{x,y} f(x) + g(y), s.t. Ax + By = b \tag{9}$$

where $f : R^n \to R$ and $g : R^P \to R$ are two convex function, and $A \in R^{m\times n}, B \in R^{m\times p}$, $b \in R^m$.

The augmented Lagrangian function for problem (9) is

$$L(x,y,\lambda) = f(x) + g(y) + <\lambda, Ax + By - b> + \frac{\mu}{2}||Ax + By - b||_F^2 \tag{10}$$

where $\lambda \in R^m$ is lagrangian multiple term, $< \cdot >$ is inner product, and $\mu > 0$ is a penalty parameter. The ADMM algorithm can solve problem (9) via the following iteration:

$$\begin{aligned} x^{i+1} &= \underset{x}{\operatorname{argmin}} L(x, y^i, \lambda^i) \\ y^{i+1} &= \arg\min_y (x^{i+1}, y, \lambda^i) \\ \lambda^{i+1} &= \lambda^i + \mu(Ax^{i+1} + By^{i+1} - b) \end{aligned} \tag{11}$$

Thus, the problem (8) can be rewritten as

$$\begin{aligned} \min_{Z_1,Z_2,J} \lambda_1||Z_1||_* + \lambda_2||Z_2||_1 + \frac{\lambda_3}{2}||\mathrm{X} - \mathrm{XJ}||_F^2 \\ s.t.\ J = Z_2 - diagZ_2 \\ J = Z_1 \\ 1^T J = 1^T \end{aligned} \tag{12}$$

The augmented Lagrangian function of the problem (12) is

$$\begin{aligned} & L(Z_1, Z_2, J, Y_1, Y_2, Y_3) = \\ & \lambda_1||Z_1||_* + \lambda_2||Z_2||_1 + \frac{\lambda_3}{2}||X - XJ||_F^2 + \\ & < Y_1, J - Z_2 + diagZ_2 > + < Y_2, J - Z_1 > + \\ & < Y_3, 1^T J - 1^T > + \frac{\mu}{2}(||J - Z_2 + diagZ_2||_F^2 + \\ & ||J - Z_1||_F^2 + ||1^T J - 1^T||_F^2) \end{aligned} \tag{13}$$

where $Y_1 \in R^{N\times N}, Y_2 \in R^{N\times N}$, and $Y_3 \in R^{1\times N}$ are the lagrangian multiple terms.

We then apply ADMM method to alternatively solve the problem(13). The resulting algorithm is summarized in Algorithm 1.

Algorithm 1 Solving Optimization (12) by ADMM

Input : data X, parameter λ_1,λ_2,λ_3 and μ
Initialize : $Z_1 = Z_2 = J = Y_1 = Y_2 = 0,\ Y_3 = 0$
While not converged do
1. fix the others and update Z_1 by

$$Z_1 = \arg\min_{Z_1} \frac{\lambda_1}{\mu}||Z_1||_* + \frac{1}{2}||J - Z_1 + \frac{Y_2}{\mu}||_F^2 \tag{14}$$

2. fix the others and update Z_2 by

$$Z_2 = \arg\min_{Z_2} \frac{\lambda_2}{\mu}||Z_2||_1 + \frac{1}{2}||J - Z_2 + \frac{Y_1}{\mu}||_F^2 \tag{15}$$

, next

$$Z_2 = Z_2 - diagZ_2 \tag{16}$$

3. fix the others and update J by

$$\begin{aligned} J = inv(\lambda_3 X^T X + 2\mu I_{N\times N} + \mu 1_{N\times N})(\lambda_3 X^T X - Y_1 \\ - Y_2 - 1_{N\times 1} Y_3 + \mu Z_1 + \mu Z_2 + \mu 1_{N\times N}) \end{aligned} \tag{17}$$

where $inv(A)$, inverse of matrix A; $I_{N\times N} \in R^{N\times N}$ is a identity matrix; $1_{N\times N} \in R^{N\times N}$, all elements of $1_{N\times N}$ are 1s; $1_{N\times 1} \in R^{N\times 1}$, all elements of $1_{N\times 1}$ are 1s.
4. update the multipliers

$$\begin{aligned} Y_1 &= Y_1 + \mu(J - Z_2) \\ Y_2 &= Y_2 + \mu(J - Z_1) \\ Y_3 &= Y_3 + \mu(1^T J - 1^T) \end{aligned} \tag{18}$$

6. end while
Output : Z_1, Z_2

Fortunately, we can prove the following Lemma 1 and Lemma 2 for the subproblems (14) and (15). Thus we give the solutions of these two subproblems that can explicitly expressed in the form of shrinkage.

Lemma 1 [11] Let X be a matrix of rank r, where $X = U\sum V^T$, then the optimal solution of

$$\min_A \frac{1}{2}||X - A||_F^2 + \tau||A||_* \tag{19}$$

is given by $A = D_\tau = US_\tau(\Sigma)V^T$, where D_τ is called the singular-value thresholding operator(SVT), which is defined as

$$S_\tau(x) = sign(x)max(|x - \tau|, 0) = \begin{cases} x - \tau & x > \tau \\ x + \tau & x < -\tau \\ 0 & else \end{cases} \tag{20}$$

Comparing with (14), one can find that $A = Z_1$, $X = J + \frac{Y_2}{\mu}$. Thus Z_1 in (14) can be solved by lemma 1.

Similarly, the Lemma 2 can be described as following form, i.e.

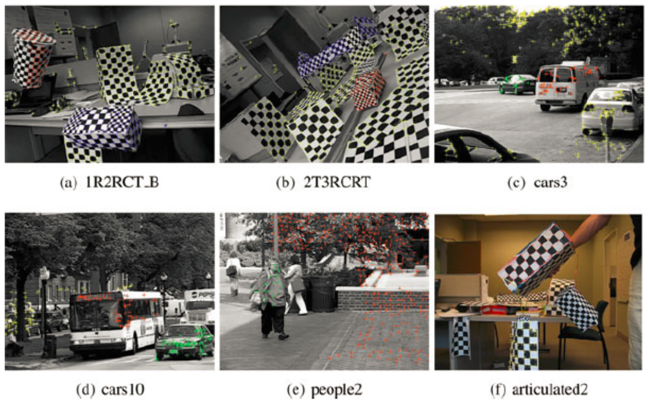


Figure 2. Sample images from some sequences of the Hopkins 155 database with tracked points superimposed [4].

Lemma 2 [12] Let X be a matrix,then the optimal solution of

$$\min_{A} \frac{1}{2}||X - A||_F^2 + \tau||A||_1$$

is given by $A = S_\tau(x)$.

Therefore, Z_2 in (15) also can be soft shrinkage form according to lemma 2.

After solving the CLRS, is computed in the form of $|Z|_{i,j} + |Z^T|_{i,j}$. Finally, we use the spectral clustering algorithms to produce the final segmentation by algorithm 2.

Algorithm 2 Subspace Segmentation

Input : data matrix X, number of subspaces k
1. obtain the representation by Algorithm 1
2. construct an undirected graph by using the representation to define the affinity matrix of the graph
3. use Ncut to segment the vertices of the graph into k clusters

---

IV. Experiments

In this section, we show the ability of motion segmentation by applying the CLRS model (7) to Hopkins 155 dataset.

Hopkins 155: a motion segmentation dataset, there are 155 video sequences with extracted feature points and their tracks across frames, where 120 of the videos have two motions and 35 of the videos have three motions. Figure 2 shows that some sample images from some sequences of Hopkins 155.

We also compare CLRS with local subspace affinity (LSA) [13], spectral curvature clustering (SCC) [14], LRR [5] and SSC [6].

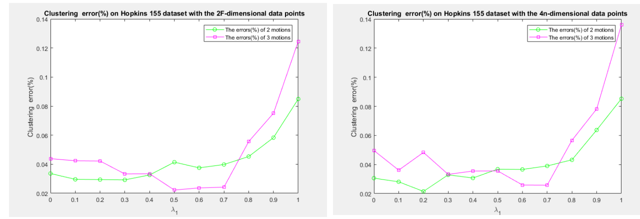


Figure 3. Clustering error(%) on Hopkins 155 dataset when $\lambda_2 = 1 - \lambda_1$.

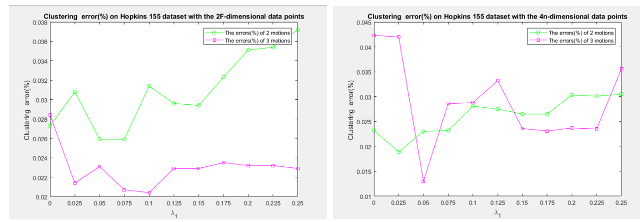


Figure 4. Clustering error(%) on Hopkins 155 dataset when $\lambda_2 = 1 - \lambda_1$.

We apply the same preprocessing steps to described in LSA, SCC, LRR, SSC and CLRS. We use the original 2F-dimensional feature trajectories and we projec the data into a 4n-dimensional subspace using PCA [4] [6] .

In CLRS, $\lambda_3 = \frac{\alpha}{\mu_z}$, α is a constant, μ_z defined as [6],

$$\mu_z = \min_i \max_{j \neq i} |x_i^T x_j|, \forall x_i, x_j \in X \tag{21}$$

where we set $\alpha = 30000$, $\mu = 0.01\lambda_3$, and λ_1, λ_2 should be tuned. In the beginning, set $\lambda_2 = 1 - \lambda_1$. From Figure 3, one can find that the lowest clustering error when $\lambda_1 = 0.6$, which are with the 2F-dimensional data points and 4n-dimensional data points obtained by PCA .

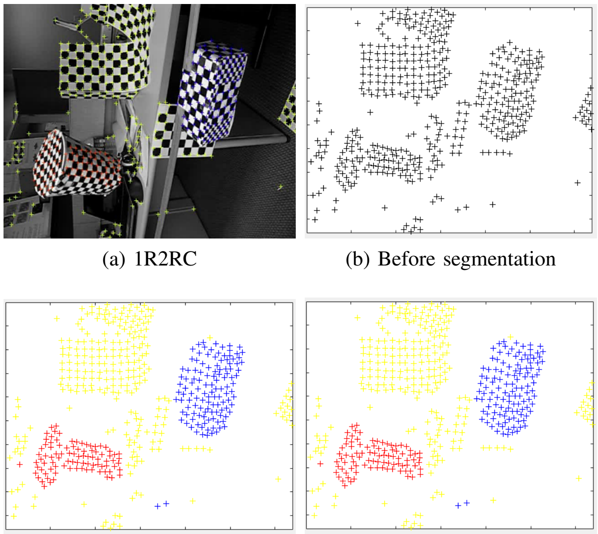
Then fix $\lambda_2 = 1 - \lambda_1 = 0.4$. Through computing, one can find the lowest clustering error may occur when $\lambda_1 = 0.05, 0.075, 0.1$ from Figure 4.

We get optimization parameters $\lambda_1 = 0.05, \lambda_2 = 0.40$ by computing the average misclassification errors of 120 video sequences with 2 motions, 35 video sequences with 3 motions and all video sequences.

The samples result of motion segmentation with CLRS as Figure 5 and Figure 6 .

Table I
Clustering error(%) of LSA,SCC,LRR,SSC and CLRS on the Hopkins 155 with the 2F -dimensional data points.

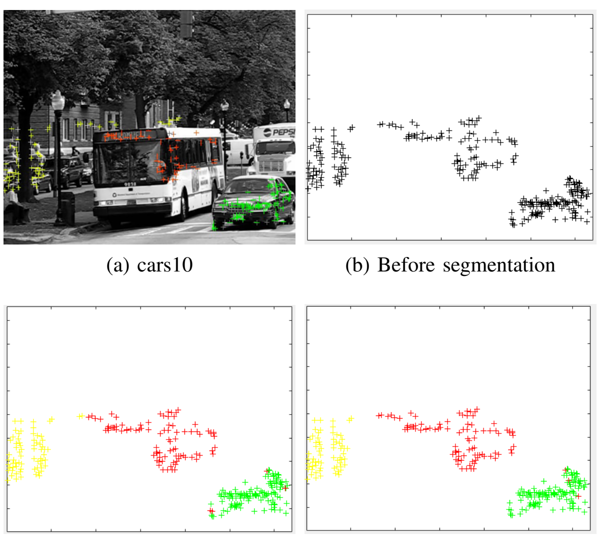
Algorithms	LSA	SCC	LRR	SSC	CLRS
2 Motions					
Mean	4.23	2.89	4.10	1.52	2.59
Median	0.56	0.00	0.22	0.00	0.00
3 Motions					
Mean	7.02	8.25	9.89	4.40	2.31
Median	1.45	0.24	6.22	0.56	0.39
ALL					
Mean	4.86	4.10	5.41	2.18	2.53
Median	0.89	0.00	0.53	0.00	0.00



(a) 1R2RC (b) Before segmentation

(c) After segmentation with 2F (d) After segmentation with 4n

Figure 5. An example of motion segmentation.



(a) cars10 (b) Before segmentation

(c) After segmentation with 2F (d) After segmentation with 4n

Figure 6. An example of motion segmentation.

Table II
Clustering error(%) of LSA,SCC,LRR,SSC and CLRS on the Hopkins 155 with the 4n-dimensional data points.

Algorithms	LSA	SCC	LRR	SSC	CLRS
2 Motions					
Mean	3.61	3.04	4.83	1.83	2.30
Median	0.51	0.00	0.26	0.00	0.00
3 Motions					
Mean	7.65	7.91	9.89	4.40	1.30
Median	1.27	1.14	6.22	0.56	0.41
ALL					
Mean	4.52	4.14	5.98	2.41	2.07
Median	0.57	0.00	0.59	0.00	0.00

Then we tune the parameters of each method and report to their best results.

The average and median misclassification errors are shown in Table I and Table II respectively.

Table I shows that we get a mean misclassification error of 2.31% for sequences with three motions, while the best reported is 4.40% by SSC. Table II also shows that we get a mean misclassification error of 1.30% for sequences with three motions, while the best reported is 4.40% by SSC.

Table II shows that we get a misclassification error of 2.07% for all sequences using CLRS model, while the best reported is 2.41% by SSC. In summary, the CLRS model show the superiority against existing state-of-the-art methods for all the tested task.

V. Conclusion and Future Work

In this paper we proposed a novel method to subspace clustering which is based on collaborative low rank and sparse representation. The new model can avoid effectively the missing of low-rank structure and over sparse problem. Comprehensive experimental comparison with the 3 motions on a dataset of Hopkins 155 demonstrate the efficiency and effectiveness of the CLRS model on motion segmentation. However, the relationship between the parameter λ_1 and λ_2 is unclear. In the future, we believe more efforts should be make for the theoretical explanation of the CLRS model and parameter selection. We will compare CLRS with the newest methods on motion segmentation in the future work.

Acknowledgment

This work was supported by the National Nature Science Foundation of China(61472257,61402290), Shenzhen Basis Research Project(JCYJ20160520161847267). Corresponding author: Min Li.

References

[1] W. Hong, J. Wright, K. Huang, and Y. Ma, "Multiscale hybrid linear models for lossy image representation," *IEEE Trans. Image Process.*, vol. 15, pp. 3655–3671, 2006.

[2] R. Vidal, R. Tron, and R. I. Hartley, "Multiframe motion segmentation with missing data using powerfactorization and GPCA," *International Journal of Computer Vision*, vol. 79, pp. 85–105, 2008.

[3] Y. Han, C. Xu, G. Baciu, and M. Li, "Lightness biased cartoon-and-texture decomposition for textile image segmentation," *Neurocomputing*, vol. 168, pp. 575–587, 2015.

[4] R. Vidal, Y. Ma, and S. Sastry, *Generalized Principal Component Analysis*. New York: Springer New York, 2016.

[5] G. Liu, Z. Lin, S. Yan, J. Sun, Y. Yu, and Y. Ma, "Robust recovery of subspace structures by low-rank representation," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 35, pp. 171–184, 2013.

[6] E. Elhamifar and R. Vidal, "Sparse subspace clustering: Algorithm, theory, and applications," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 35, pp. 2765–2781, 2013.

[7] M.Tipping and C.Bishop, "Mixtures of probabilistic principal component analyzers," *Neural Comput.*, vol. 11, pp. 443–482, 1999.

[8] R. Vidal, Y. Ma, and S. Sastry, "Generalized principal component analysis," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 27, pp. 1945–1959, 2005.

[9] B.Recht, M.Fazel, and P.A.Parrilo, "Guaranteed minimum-rank solutions of linear matrix equations via nuclear norm minimization," *SIAM Review*, vol. 52, pp. 471–501, 2010.

[10] D. Donoho, *For Most Large Underdetermined Systems of Equations, the Minimal L1 -norm Near-solution Approximates the Sparsest Near-solution*. Department of Statistics, Stanford University, 2004.

[11] J. Cai, E. J. Candès, and Z. Shen, "A singular value thresholding algorithm for matrix completion," *SIAM Journal on Optimization*, vol. 20, pp. 1956–1982, 2010.

[12] Z. Lin, M. Chen, and Y. Ma, "The augmented lagrange multiplier method for exact recovery of corrupted low-rank matrices," *arXiv preprint*, vol. 1009.5055, 2010.

[13] J. Yan and M. Pollefeys, *A general framework for motion segmentation: Independent, articulated, rigid, non-rigid, degenerate and non-degenerate*. Springer, 2006.

[14] G. Chen and G. Lerman, "Spectral curvature clustering (scc)," *International Journal of Computer Vision*, vol. 81, pp. 317–330, 2009.