Stochastic Sparse Subspace Clustering

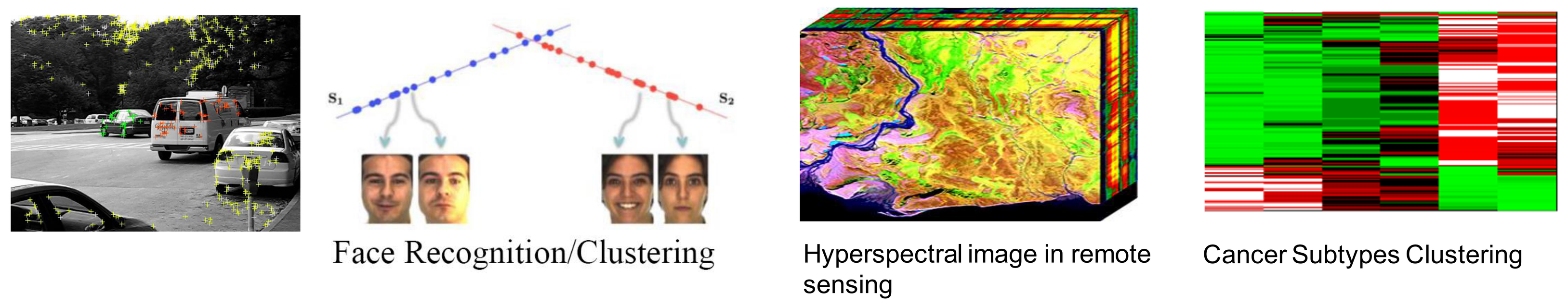
Ying Chen (谌瑛)[1], Chun-Guang Li (李春光)[1], Chong You (由翀)[2]



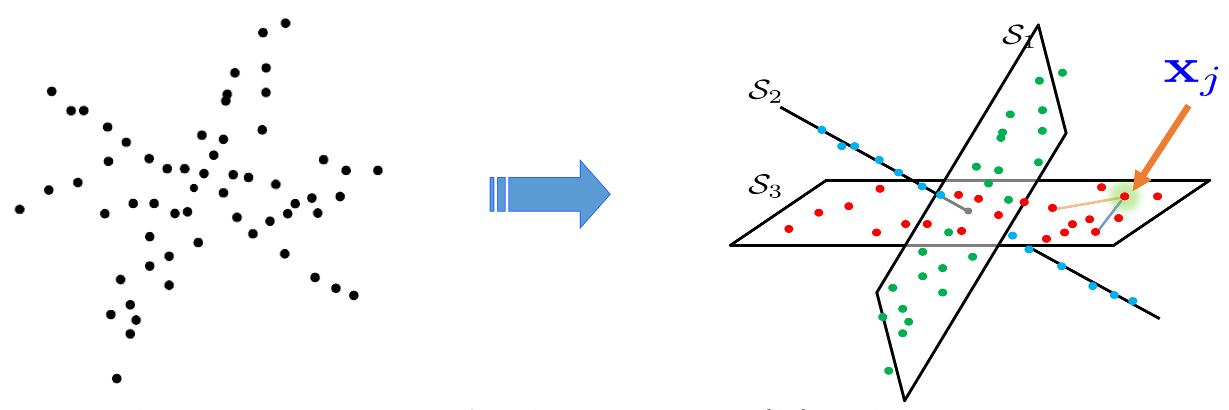
[1] 北京邮电大学 [2] 加州大学 伯克利分校

Introduction

➤ **High dimensional dataset often consists of multiple low-dimensional subspaces**, i.e. **a union of subspaces**

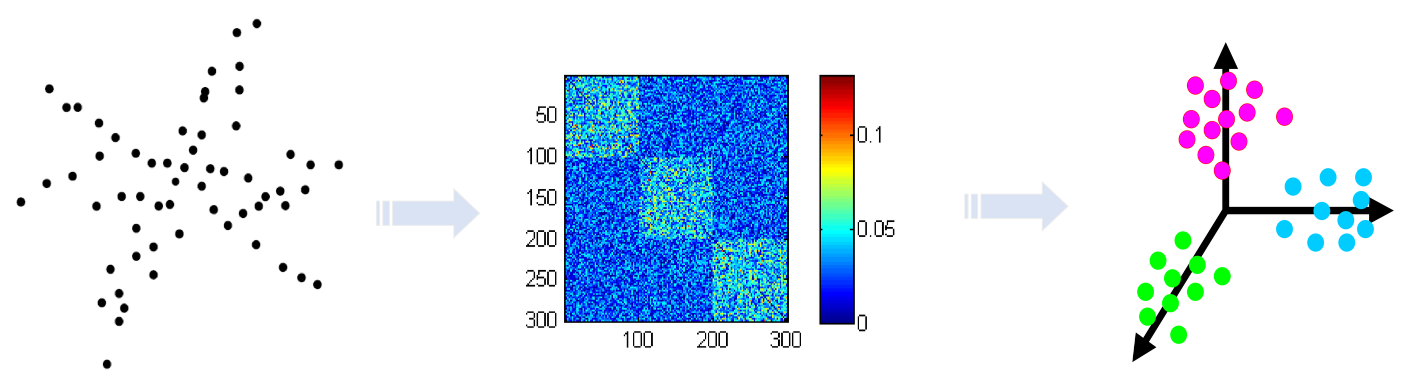


➤ **Subspace clustering: to segment the data into each subspace**



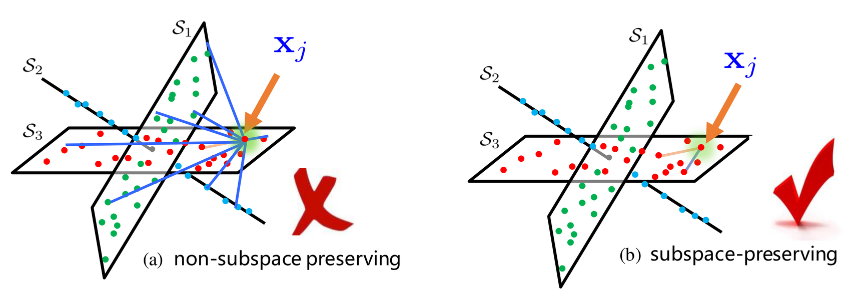
(a) data points lying in a union of subspaces (b) Subspace segmentation

➤ **State-of-the-art subspace clustering methods follow a two-step approach:**



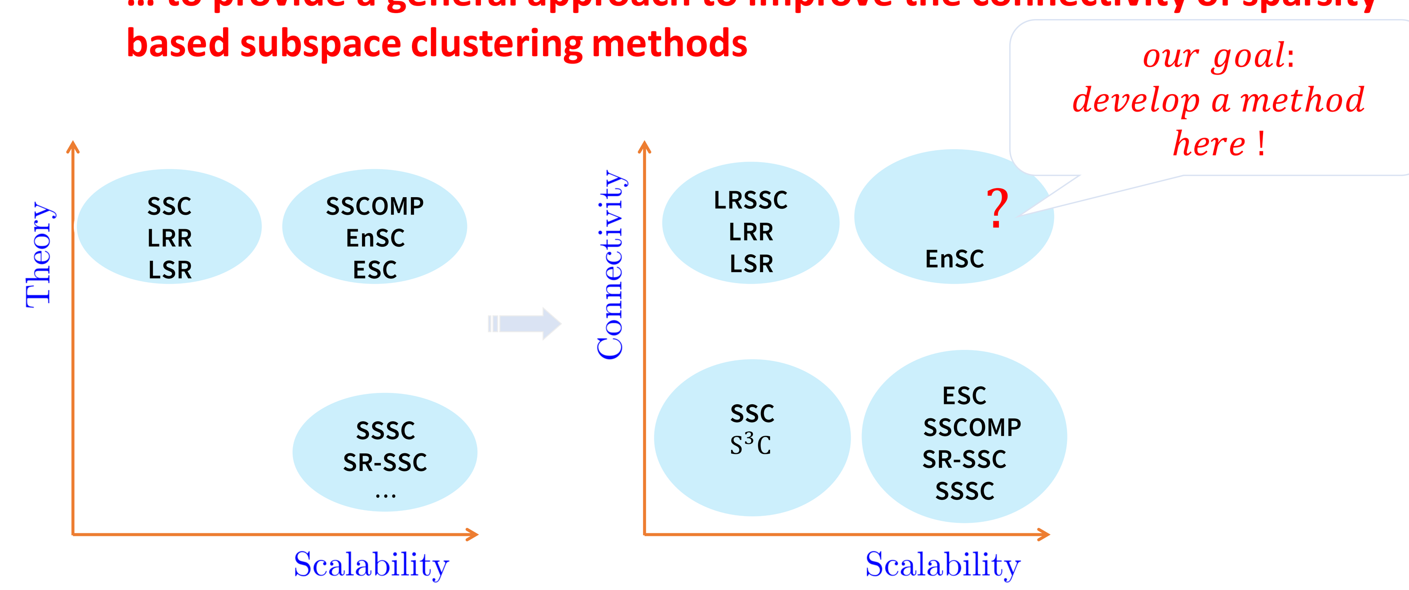
- **Step 1: Self-Expression Model:**

$$x_j = c_1 x_1 + c_2 x_2 + \cdots + c_{j-1} x_{j-1} + 0 \cdot x_j + c_{j+1} x_{j+1} + \cdots + c_N x_N$$



- **Step 2: Spectral Clustering**
  - ✓ **Conditions to yield correct clustering:**
    1. nonzero entries in affinity matrix are correct (i.e. subspace preserving)
    2. affinity graph well-connected for each subspace (i.e. good connectivity)

➤ **Our Goal** … to provide a general approach to improve the connectivity of sparsity-based subspace clustering methods



- SSCOMP is very promising except for its suffering from the connectivity issue → **How to improve the connectivity of SSCOMP?**

➤ **Our Idea and Proposal** Dropout + Self Expression Model

$$x_j = c_1 x_1 + c_2 x_2 + \cdots + c_{j-1} x_{j-1} + 0 \cdot x_j + c_{j+1} x_{j+1} + \cdots + c_N x_N$$

✓ **Introduce dropout into self-expression model to drop the columns uniformly at random**

✓ **We prove that: dropout → an implicit squared ℓ_2 norm, i.e.**

Theorem 1: Let $\{\xi_i\}_{i=1}^N$ be i.i.d. Bernoulli random variables, we have that:

$$\mathbb{E}_\xi \|x_j - \sum_i \xi_i c_{ij} x_i\|_2^2 = \|x_j - \sum_i c_{ij} x_i\|_2^2 + \frac{\delta}{1-\delta} \sum_i \|x_i\|_2^2 c_{ij}^2$$
$$= \|x_j - \sum_i c_{ij} x_i\|_2^2 + \frac{\delta}{1-\delta} \|c_j\|_2^2 \quad \text{if } \|x_i\|_2 = 1$$

➤ SSCOMP (You et al. CVPR16)

$$\min_{c_j} \|x_j - X c_j\|_2^2, \quad \text{s.t.} \quad \|c_j\|_0 \le s, \quad c_{jj} = 0,$$

where $\|\cdot\|_0$ is the ℓ_0 pseudo-norm and s is a parameter that controls the sparsity

➤ Dropout meets SSCOMP:

$$\mathbb{E}_\xi \|x_j - \sum_i \xi_i c_{ij} x_i\|_2^2 \Rightarrow \frac{1}{T} \sum_{t=1}^T \|x_j - \sum_i \xi_i^{(t)} c_{ij} x_i\|_2^2 \quad \text{(sample mean)}$$

$$\min_{c_j} \frac{1}{T} \sum_{t=1}^T \|x_j - \sum_i \xi_i^{(t)} c_{ij} x_i\|_2^2 \quad \text{s.t.} \quad \|c_j\|_0 \le s, \quad c_{jj} = 0,$$

→ Stochastic Sparse Subspace Clustering

Algorithm

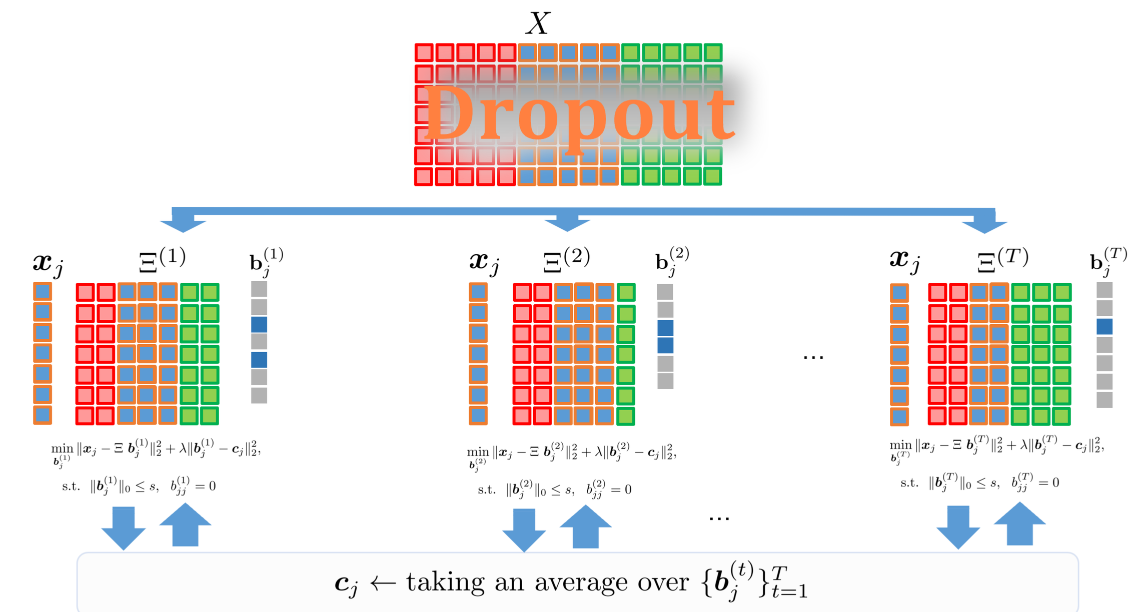
➤ We develop a **Consensus Orthogonal Matching Pursuit** algorithm to solve problem:

$$\min_{c_j, \{b_j^{(t)}\}} \frac{1}{T} \sum_{t=1}^T \|x_j - \sum_i \xi_i^{(t)} b_{ij}^{(t)} x_i\|_2^2 + \lambda \|b_j^{(t)} - c_j\|_2^2$$
$$\text{s.t.} \quad \|b_j^{(t)}\|_0 \le s, \quad b_{jj}^{(t)} = 0, \quad t = 1, \cdots, T, \qquad (3)$$

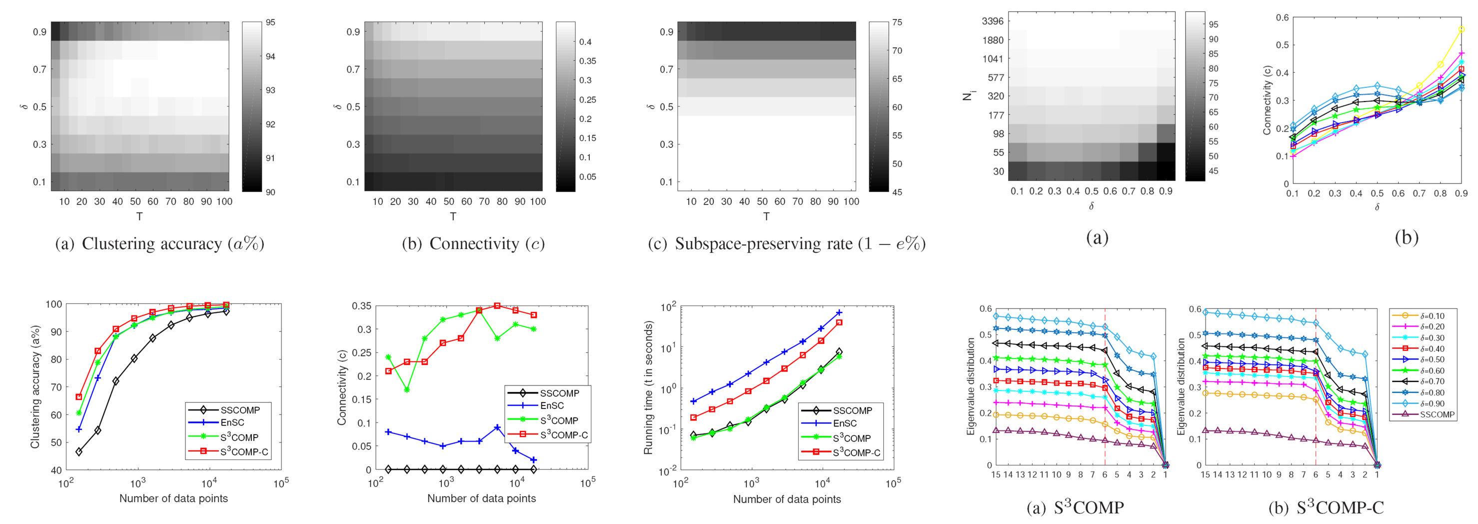
by alternating the following two steps:

Each subproblem can be solved in parallel

- **Step 1**: **Fixed c_j**, update $\{b_j^{(t)}\}_{i=1}^T$ by solving T subproblems via **Damped OMP**
- **Step 2**: **Fixed** $\{b_j^{(t)}\}_{i=1}^T$, update c_j by taking an average over $\{b_j^{(t)}\}_{i=1}^T$



Experiments



➤ Clustering accuracy compared with scalable subspace clustering methods

Dataset	# data	ESC	SR-SSC	SSCOMP	EnSC	S³COMP-C
Extended Yale B	2,432	**87.58%**	62.11%	77.59%	61.20%	87.41%
COIL100	7,200	56.90%	58.85%	49.88%	63.94%	**78.89%**
GTSRB	12,390	90.16%	78.42%	82.52%	86.05%	**95.54%**
MNIST	70,000	90.87%	87.22%	81.59%	93.67%	**96.32%**

- **Reference**

[1] Ying Chen, Chun-Guang Li, and Chong You, "Stochastic Sparse Subspace Clustering", IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), 2020, pp.4155-4164.
[arXiv version available]
[code available]

➤ For more recent work and released codes, please visit my homepage: www.pris.net.cn/teacher/lichunguang