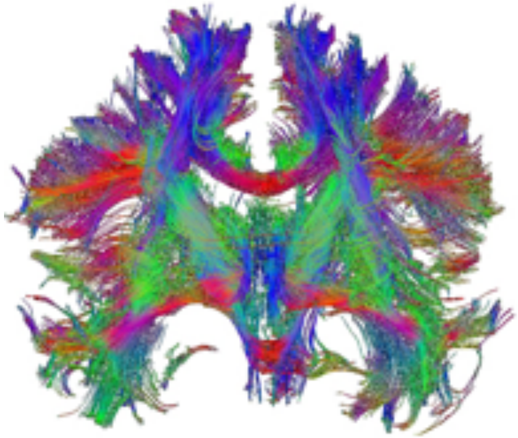
JHU vision lab

Low-Rank and Sparse Modeling for Visual Analytics

René Vidal (JHU), Ehsan Elhamifar (NEU),
Zhouchen Lin (PKU), Jiashi Feng (NUS),
Sheng Li (NEU), Yun Fu (NEU)

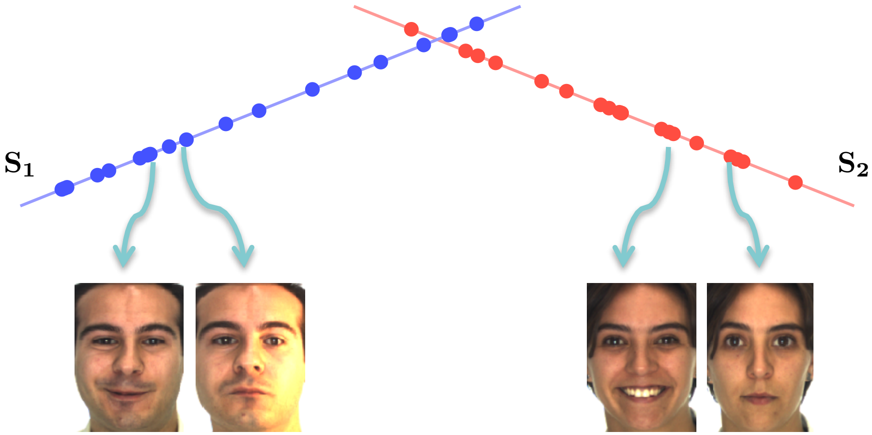
High-Dimensional Data

- In many areas, we deal with high-dimensional data
  - Computer vision
  - Medical imaging
  - Medical robotics
  - Signal processing
  - Bioinformatics

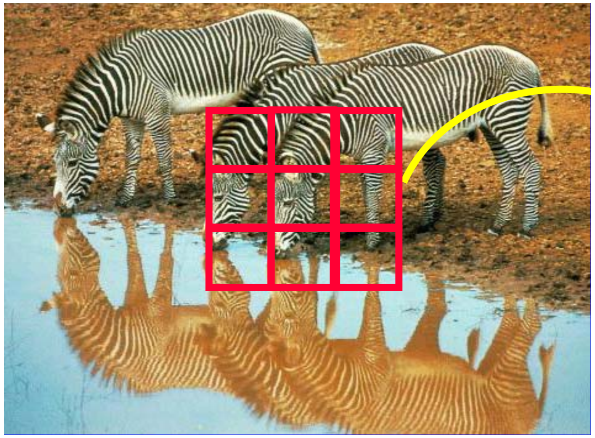


Low-Dimensional Manifolds

- Face clustering and classification

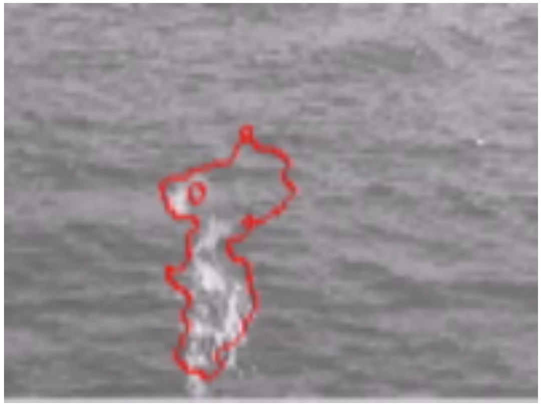


- Lossy image representation



- Motion segmentation

- DT segmentation

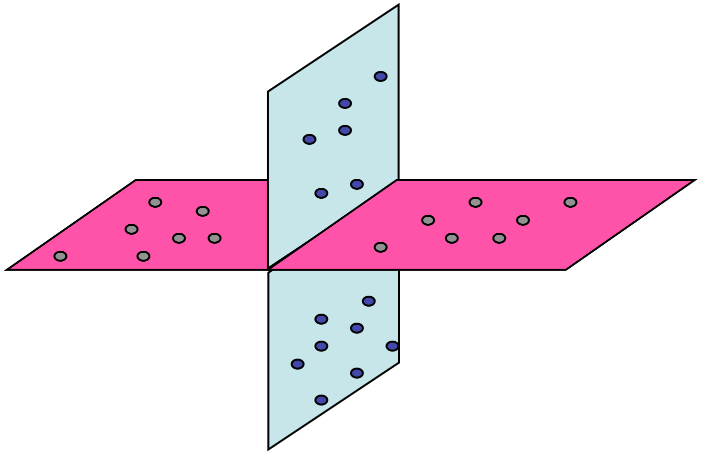


- Video segmentation

Subspace Clustering Problem

- Given a set of points lying in multiple subspaces, identify
  - The number of subspaces and their dimensions
  - A basis for each subspace
  - The segmentation of the data points

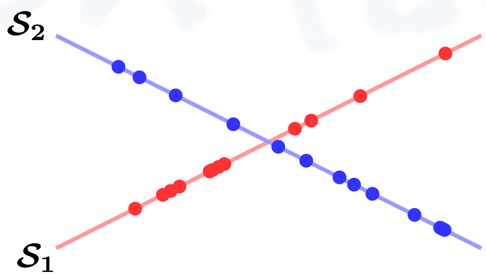
- Challenges
  - Model selection
  - Nonconvex
  - Combinatorial

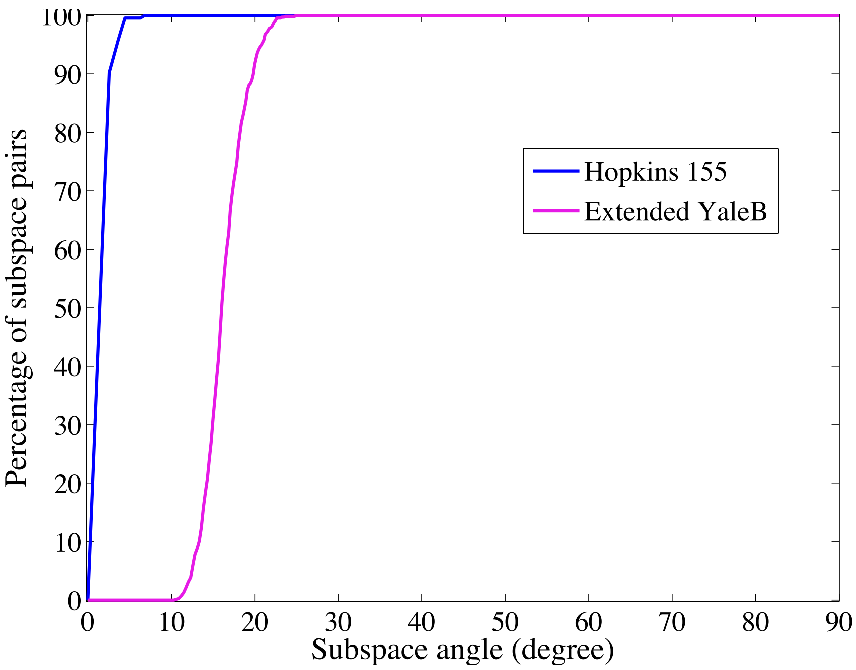


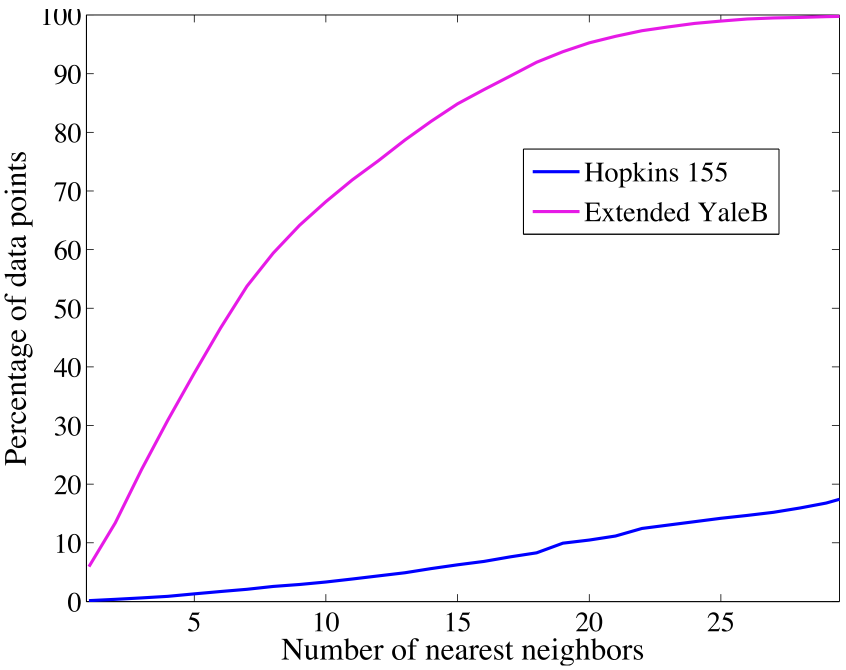
- More challenges
  - Noise
  - Outliers
  - Missing entries

Subspace Clustering Problem: Challenges

- Even more challenges
  - Angles between subspaces are small
  - Nearby points are in different subspaces

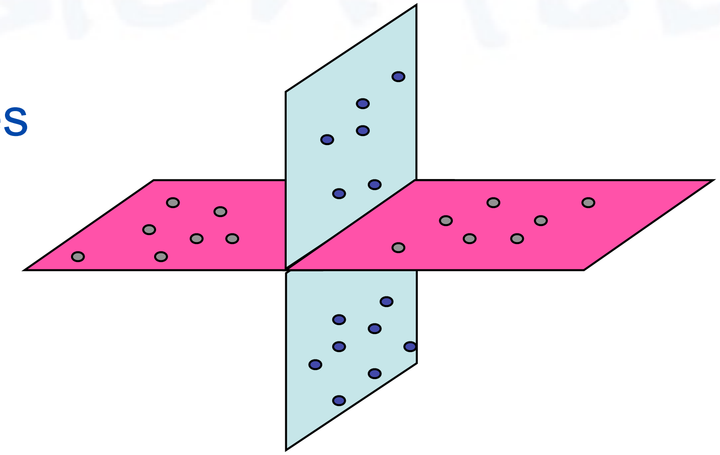






Prior Work: Iterative-Probabilistic Methods

- Approach
  - Given segmentation, estimate subspaces
  - Given subspaces, segment the data
  - Iterate till convergence

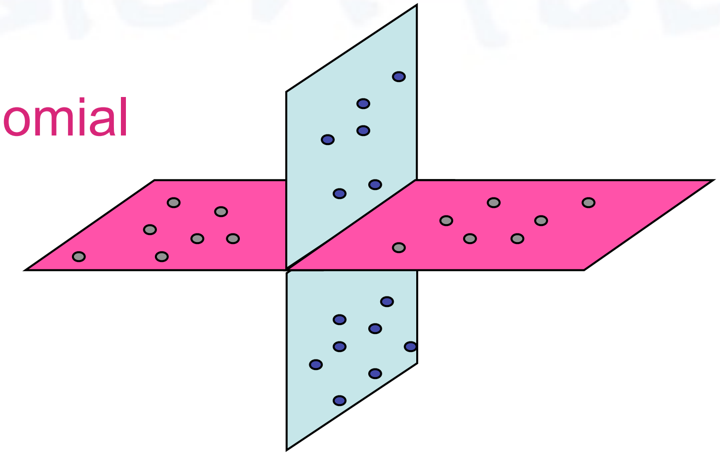


- Representative methods
  - K-subspaces (Bradley-Mangasarian ’00, Kambhatla-Leen ’94, Tseng’00, Agarwal-Mustafa ’04, Zhang et al. ’09, Aldroubi et al. ’09)
  - Mixtures of PPCA (Tipping-Bishop ’99, Grubber-Weiss ’04, Kanatani ’04, Archambeau et al. ’08, Chen ’11)

Advantages	**Disadvantages / Open Problems**
Simple, intuitive	Known number of subspaces and dimensions
Missing data	Sensitive to initialization and outliers

Prior Work: Algebraic-Geometric Methods

- Approach
  - Number of subspaces = degree of polynomial
  - Subspaces = factors of polynomial

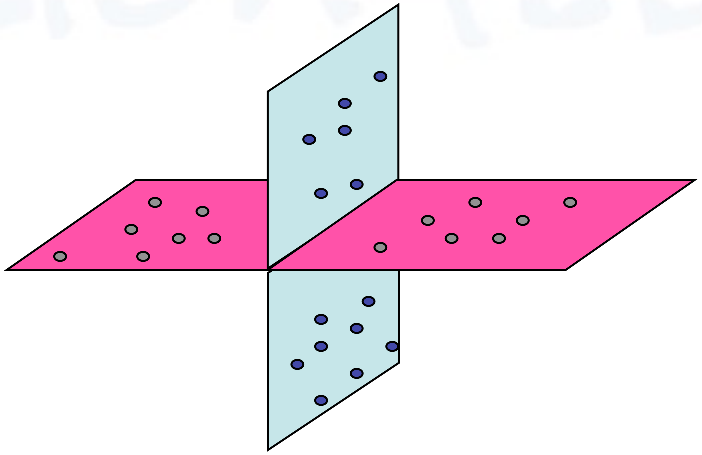


- Representative methods
  - Factorization (Boult-Brown'91, Costeira-Kanade'98, Gear'98, Kanatani et al.'01, Wu et al.'01, Sekmen'13)
  - GPCA (Shizawa-Maze '91, Vidal et al. '03 '04 '05, Huang et al. '05, Yang et al. '05, Derksen '07, Ma et al. '08, Ozay et al. '10, Tsakiris-Vidal '14 '15)

Advantages	**Disadvantages / Open Problems**
Closed form	Complexity
Arbitrary dimensions	Sensitive to noise, outliers, missing entries

Prior Work: Spectral-Clustering Methods

- Approach
  - Data points = graph nodes
  - Pairwise similarity = edge weights
  - Segmentation = graph cut

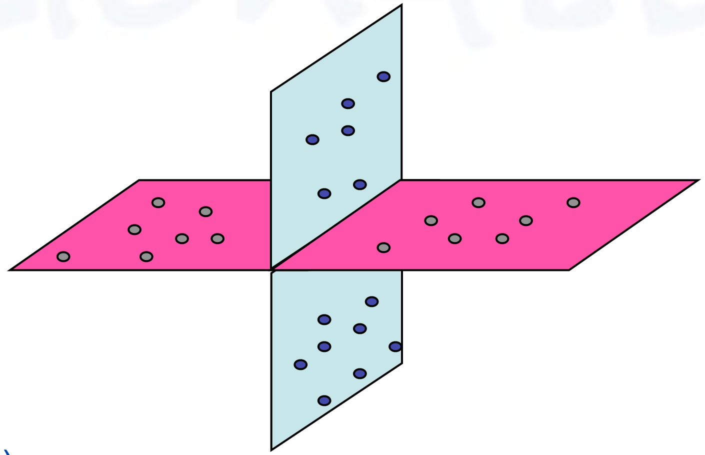


- Representative methods
  - Local (Zelnik-Manor ’03, Yan-Pollefeys ’06, Fan-Wu ’06, Goh-Vidal ’07, Sekmen’12)
  - Global (Govindu ’05, Agarwal et al. ’05, Chen-Lerman ’08, Lauer-Schnorr ’09, Zhang et al. ’10)

Advantages	**Disadvantages / Open Problems**
Efficient	Known number of subspaces and dimensions
Robust	Global methods are complex

Prior Work: Sparse and Low-Rank Methods

- Approach
  - Data are self-expressive
  - Global affinity by convex optimization



- Representative methods
  - Sparse Subspace Clustering (SSC)
    (Elhamifar-Vidal ’09 ’10 ‘13, Candes-Soltanolkotabi ’12 ‘13)
  - Low-Rank Subspace Clustering (LRR and LRSC)
    (Liu et al. ’10 ‘13, Favaro-Vidal ’11 ’13)
  - Least Square Regression (LSR) (Lu ‘12)
  - Sparse + Low-Rank (Wang ‘13), Sparse + Frobenius (Dyer ’13, You ‘16)

Advantages	**Disadvantages / Open Problems**
Efficient, Convex	Low-dimensional subspaces
Robust to noise/corruptions	Missing entries

Tutorial Objective

- Overview state-of-the-art sparse and low-rank subspace clustering methods
  - Representative methods and their theoretical properties
  - Optimization algorithms and their scalability
  - Applications in computer vision

- Unified framework

$$\min_{C,E} f(C) + \lambda g(E) \text{ s.t. } X = XC + E$$

  - Sparse Subspace Clustering: f = l1 g = l1 or Frob
  - Low Rank Representation: f = nuclear g = l21
  - Low Rank Subspace Clustering: f = nuclear g = l1 or Frob
  - Elastic Net Subspace Clustering: f = l1 + Frob g = l1 or Frob

Center for IMAGING SCIENCE

Tutorial Outline

- 08:30-08:45 Introduction to Subspace Clustering
- 08:45-10:00 Sparse Subspace Clustering
- 10:00-10:30 Coffee Break
- 10:30-12:00 Low Rank Subspace Clustering
- 12:00-02:00 Lunch Break
- 02:00-03:30 Algorithms & More Models
- 03:00-03:30 Coffee Break
- 03:30-05:00 Applications