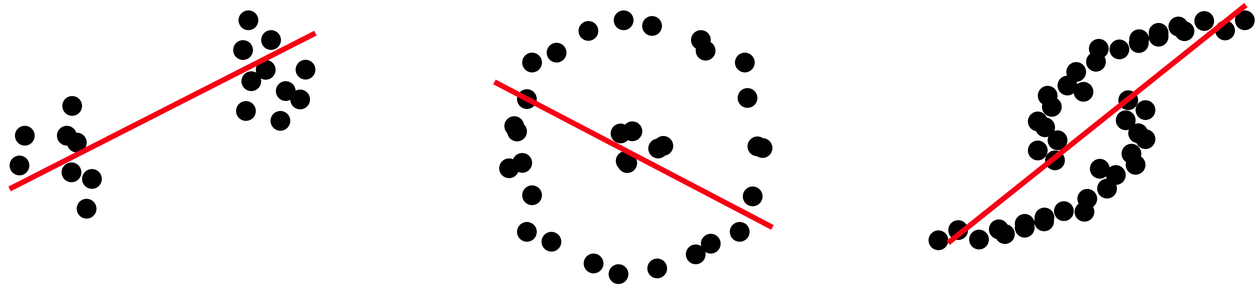
Manifold Learning in Computer Vision Part 2

Hui Wu

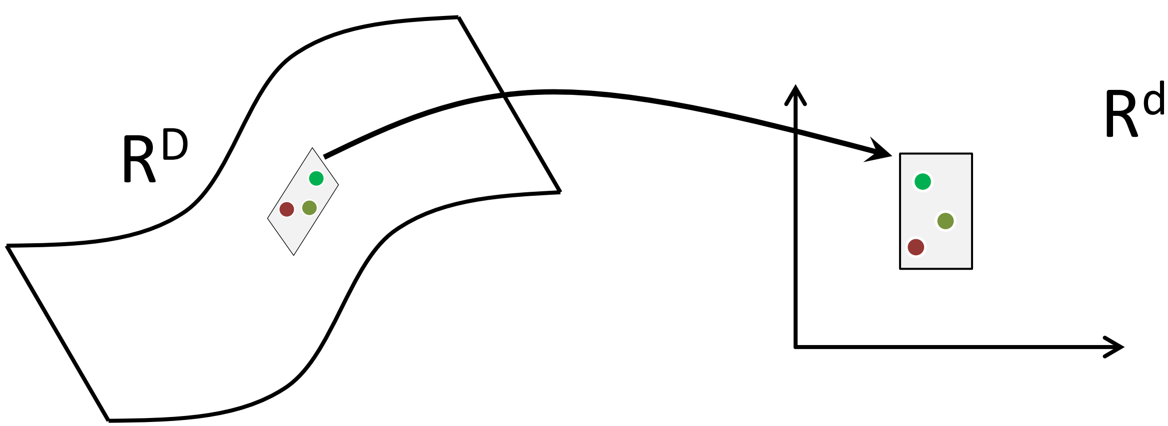
IBM Research

Review: Manifold Learning for Nonlinear Dimensionality Reduction

- PCA can not approximate nonlinear data variation



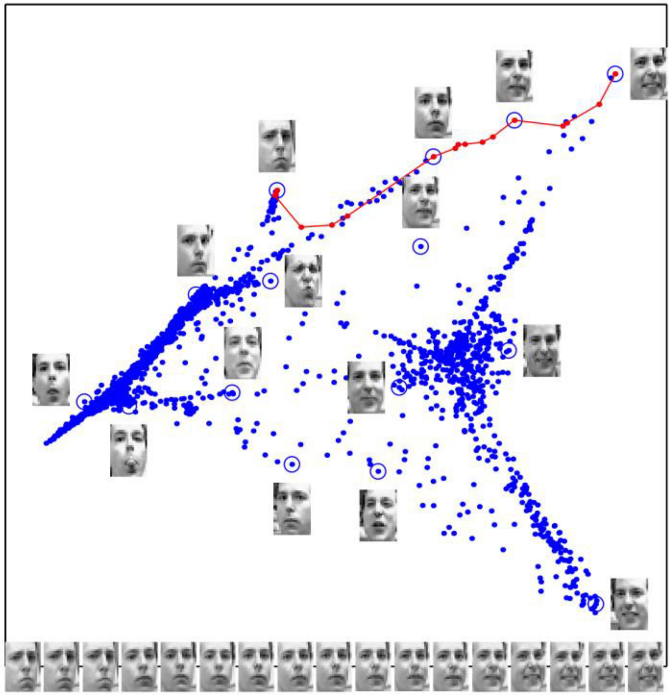
- Manifold is commonly used to model nonlinear structures in data



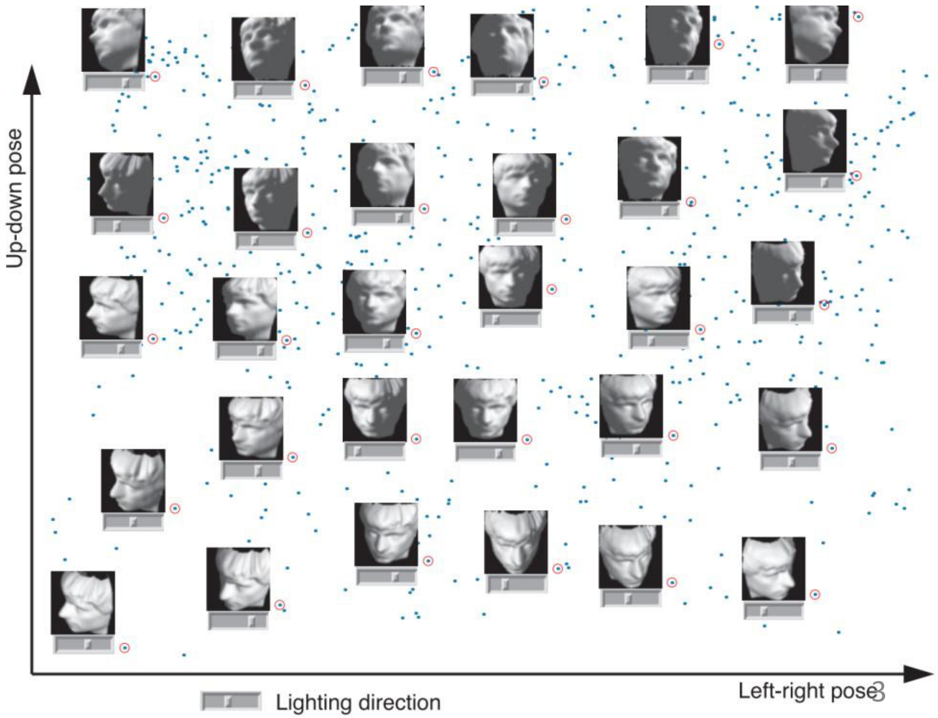
Review: Manifold Learning for Nonlinear Dimensionality Reduction

- We covered two most well known manifold learning methods in the last lecture

LLE

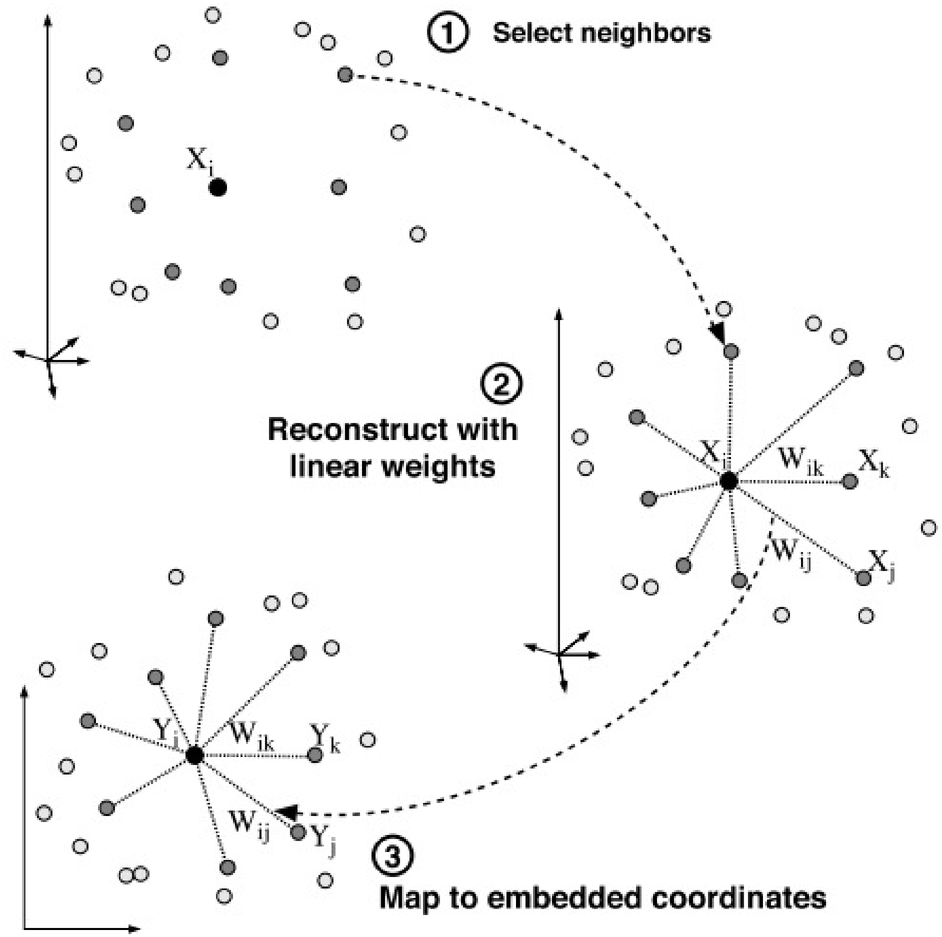


Isomap



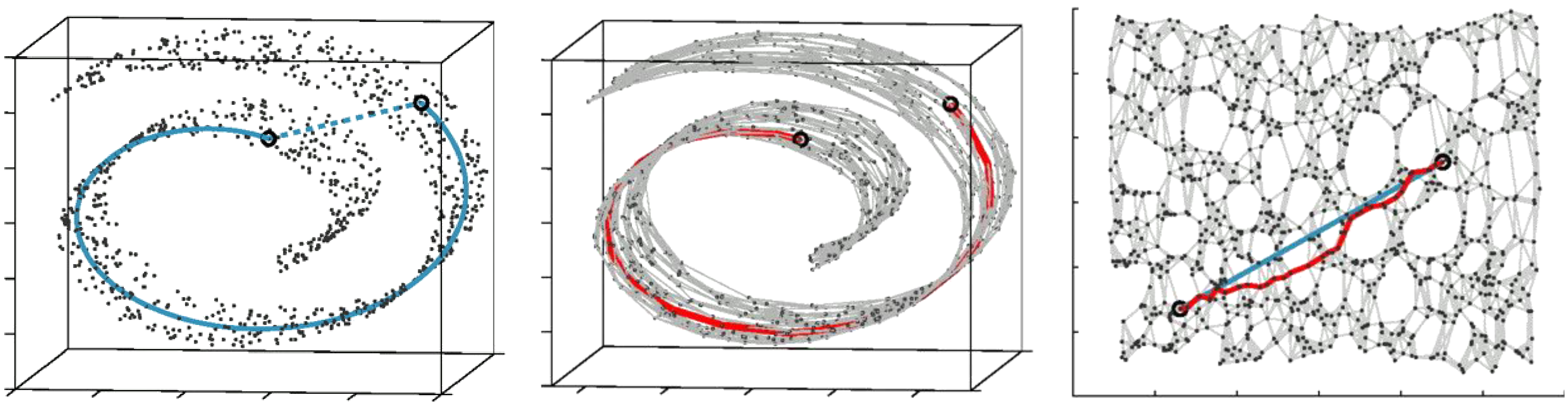
Review: Locally Linear Embedding (LLE)

- LLE utilizes the locally linear property of manifolds, and assumes:
    1. A reference point can be represented as the linear combination of its neighbors
    2. The weights of the linear combination are preserved in the low-dimensional space



Review: Isomap

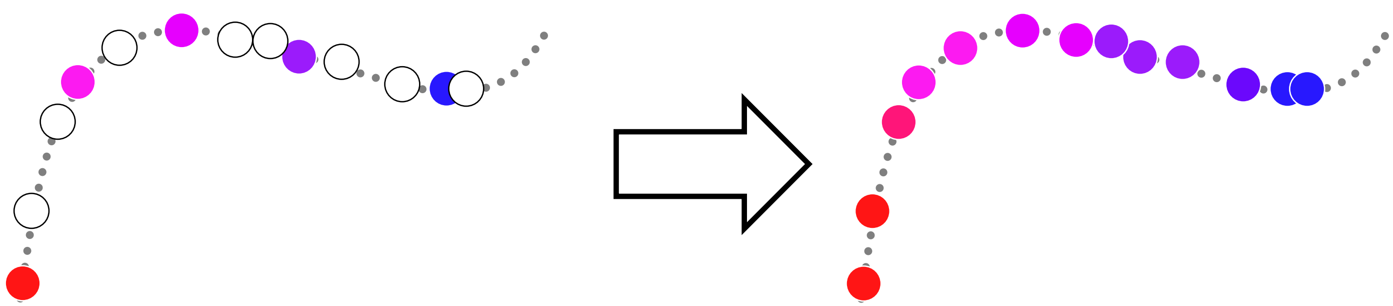
- Isomap also utilizes the locally linear property
  - Local geodesic distances are close to Euclidean distances
  - Global geodesic distances are estimated using the shortest chain of local geodesic distances
- MDS is applied to obtain the low-dimensional coordinates given the estimated geodesic distances



Applications of Manifold Learning in Computer Vision

Review: Semi-supervised Regression on Manifolds

- Input: labeled and unlabeled images
- Output: labels on all images

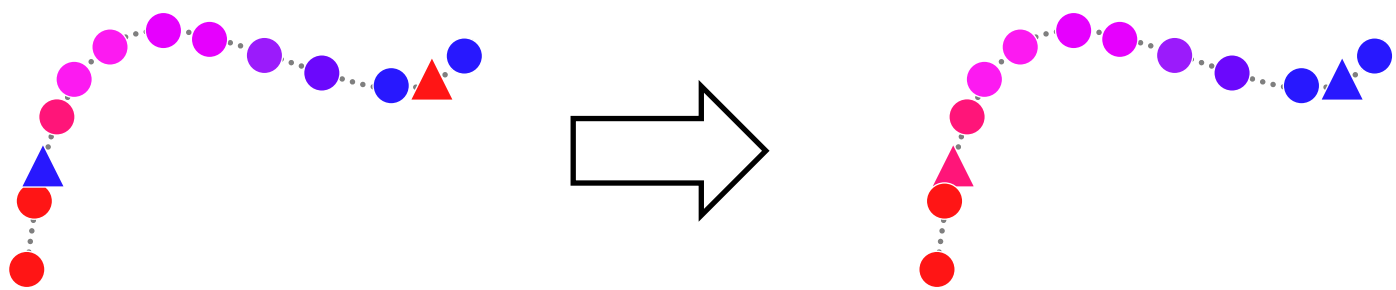


- The objective function minimizes two terms:
  - Manifold regularizer: labels should change smoothly on the manifold
  - Empirical loss: penalizes for changing the values of input labels

$$\underset{\hat{\mathbf{y}}}{argmin} \quad \underbrace{\hat{\mathbf{y}}^T \mathbf{B} \hat{\mathbf{y}}}_{\text{Hessian regularizer}} + \lambda \underbrace{\sum_{i=1}^{l} (\mathbf{y}_i - \hat{\mathbf{y}}_i)^2}_{\text{Empirical loss}}$$

Review: Robust Manifold Regression for Image Label Denoising

- Input: images and noisy labels
- Output: cleaned labels



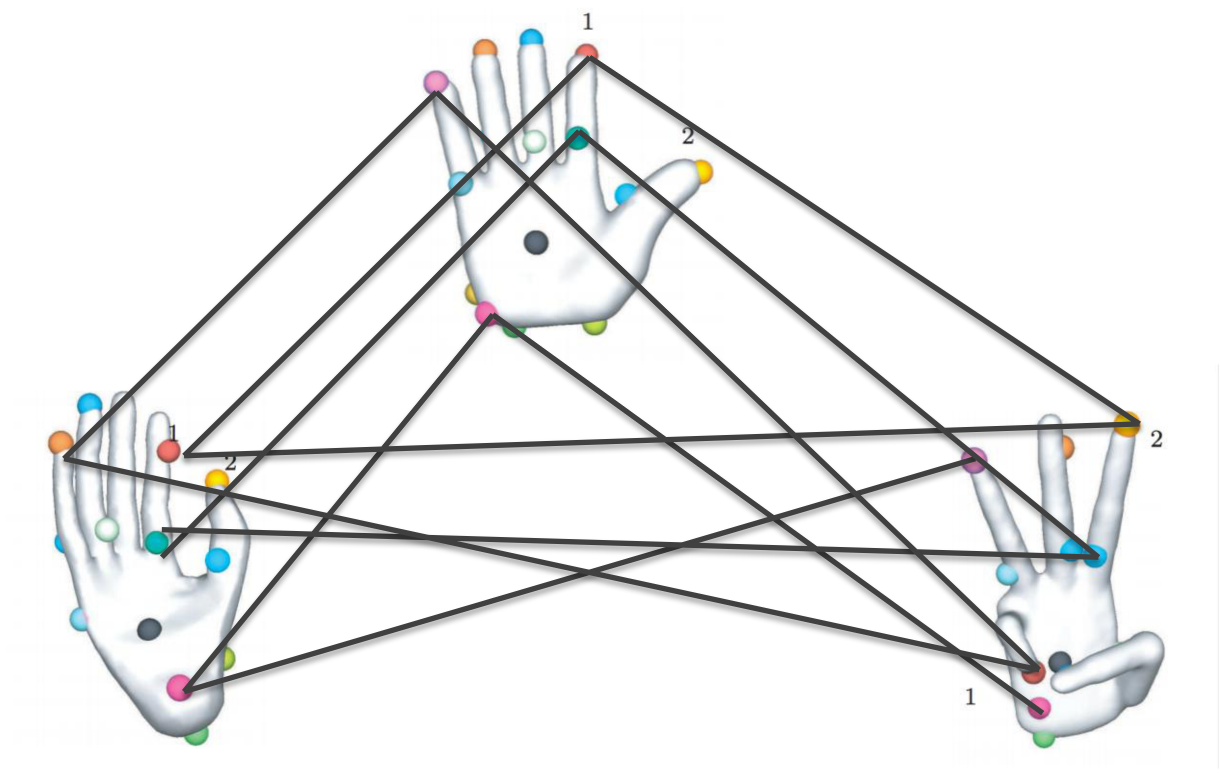
- A regularized empirical risk minimization framework
  - Manifold regularizer: labels should change smoothly on the manifold
  - Empirical loss: L1 norm is robust to high variance in noise

$$\underset{\hat{\mathbf{y}}}{\operatorname{argmin}} \quad \underbrace{\hat{\mathbf{y}}^\intercal \mathbf{B} \hat{\mathbf{y}}}_{\text{Hessian regularizer}} + \lambda \underbrace{\|\hat{\mathbf{y}} - \mathbf{y}\|_1}_{\text{L1 loss term}}$$

Multiple Shape Matching based on Manifold Learning

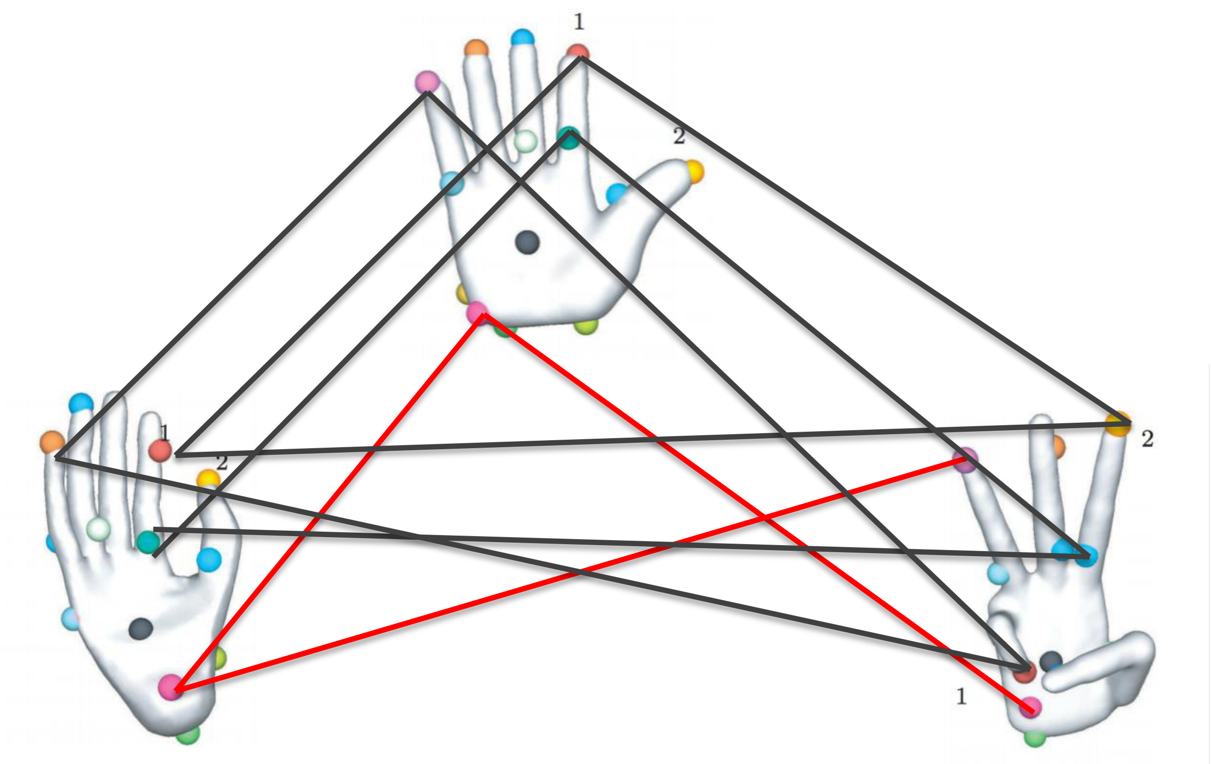
Existing Multiple 3D Shape Matching Methods

- Pairwise matching is usually the first step



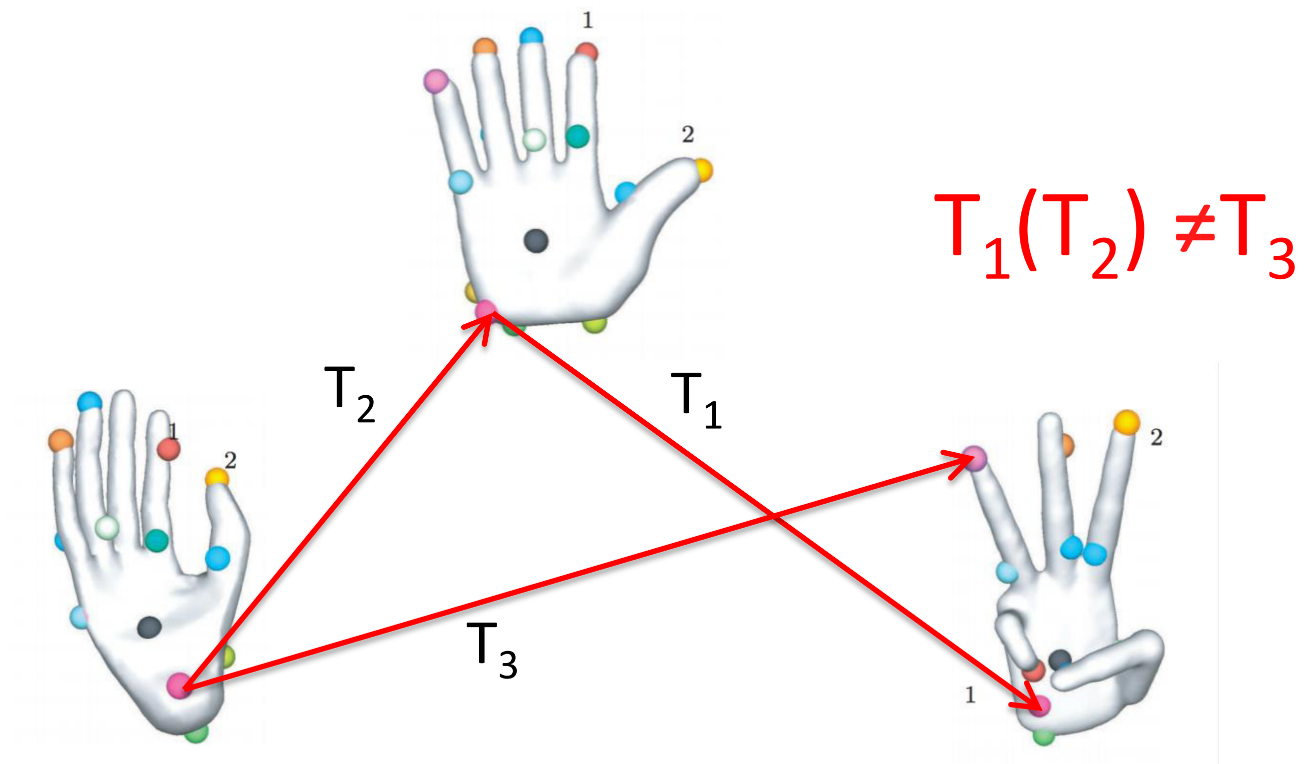
Existing Multiple 3D Shape Matching Methods

- Pairwise matching is usually the first step
  - There are erroneous pairwise correspondences



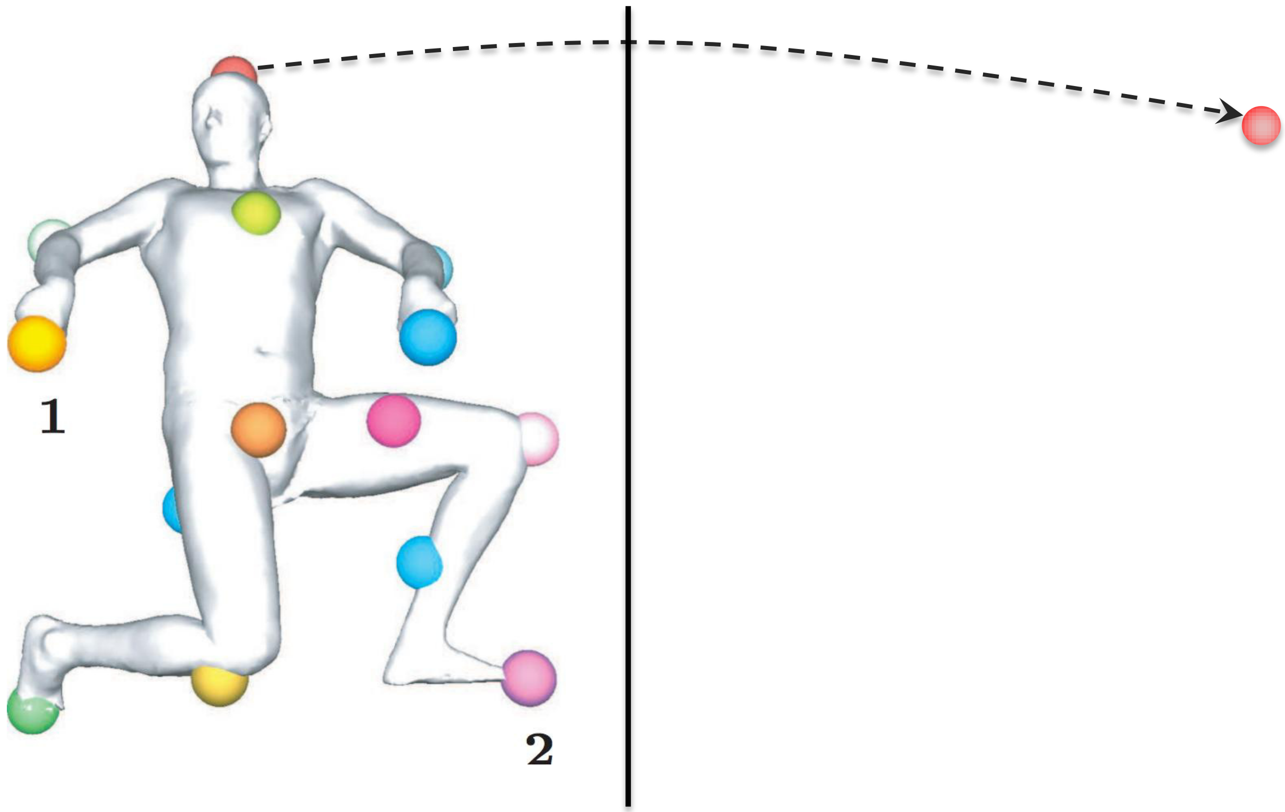
Existing Multiple 3D Shape Matching Methods

- Cycle consistency is used to refine incorrect matches



Manifold Assumption of 3D Shape Feature Points

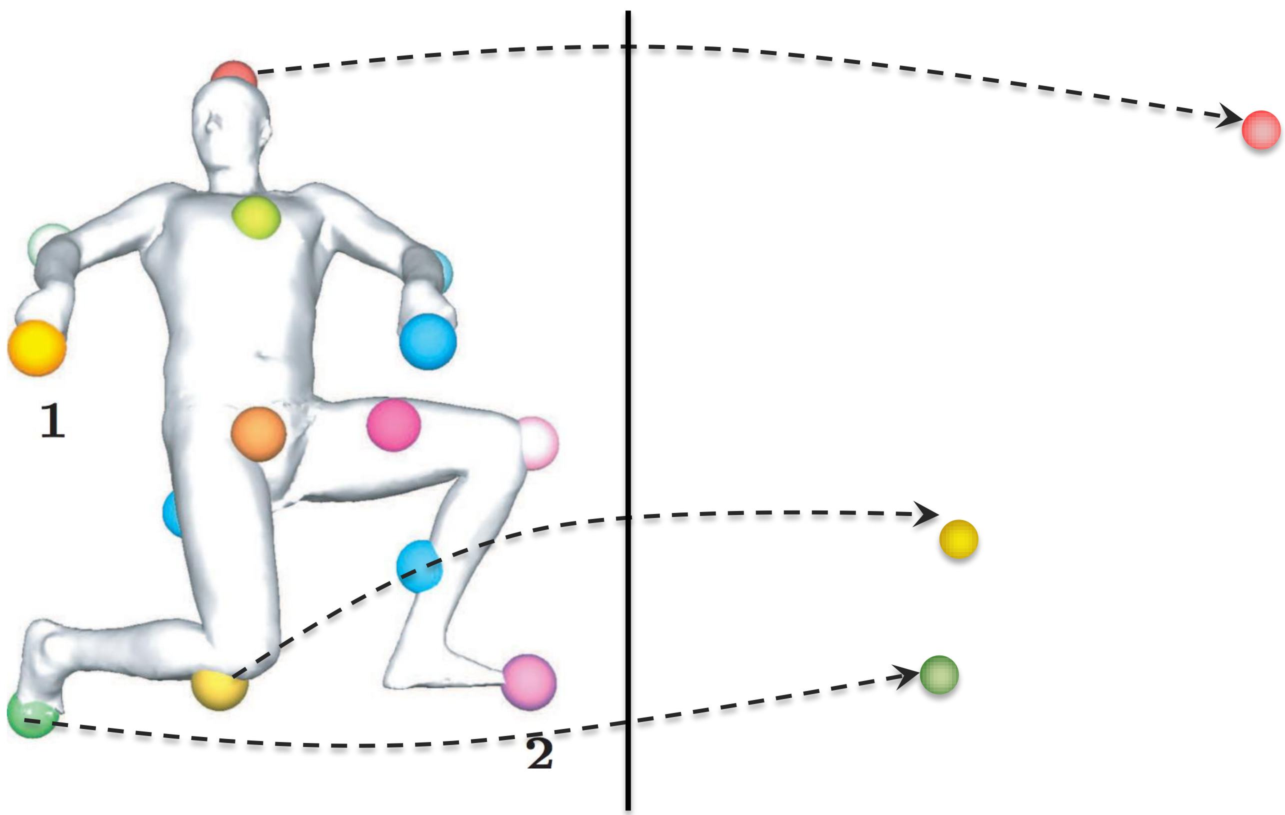
- Assume that the feature points extracted on each shape form a low-dimensional manifold



Feature space
R^D

Manifold Assumption of 3D Shape Feature Points

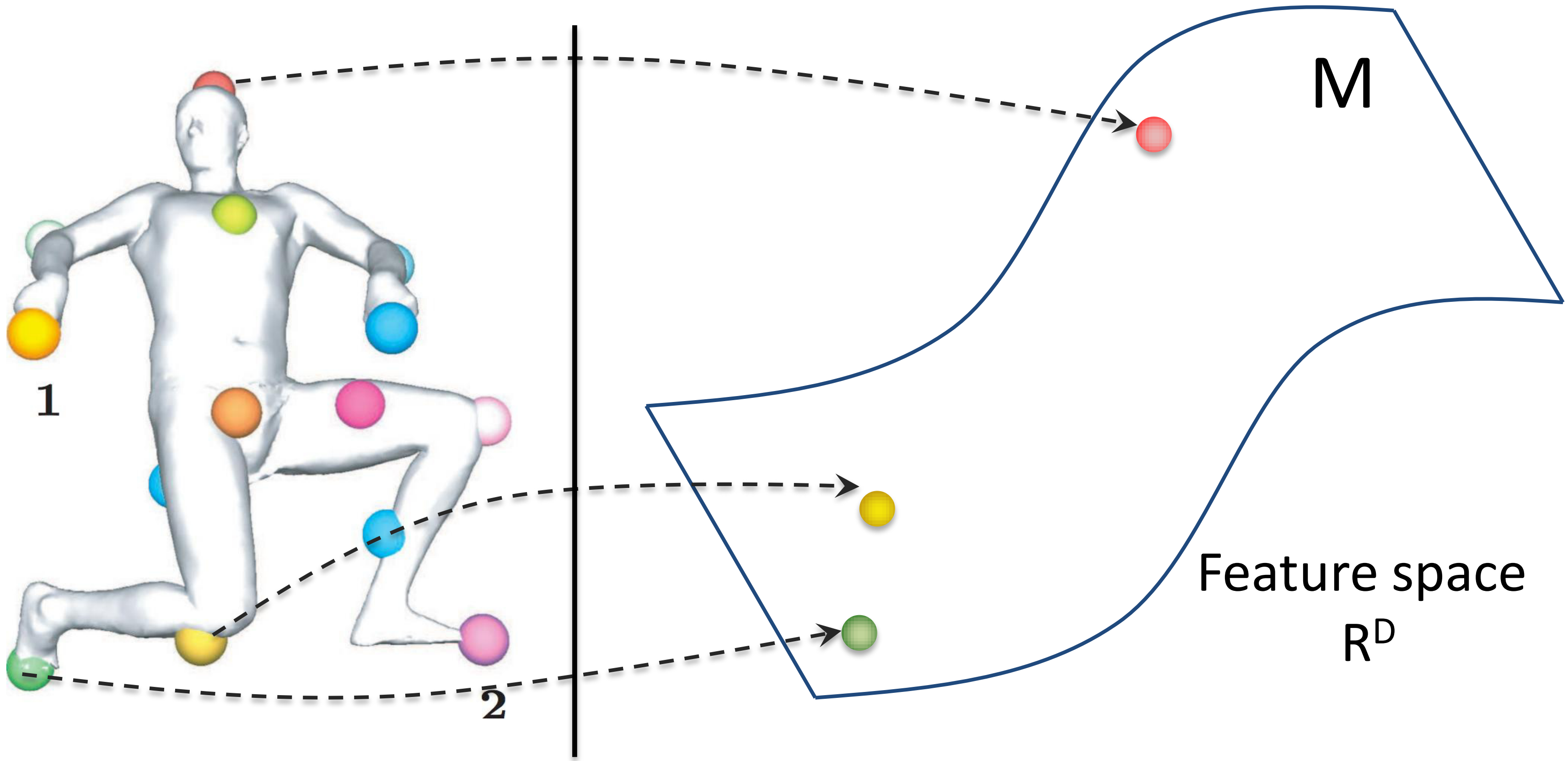
- Assume that the feature points extracted on each shape form a low-dimensional manifold



Feature space
R^D

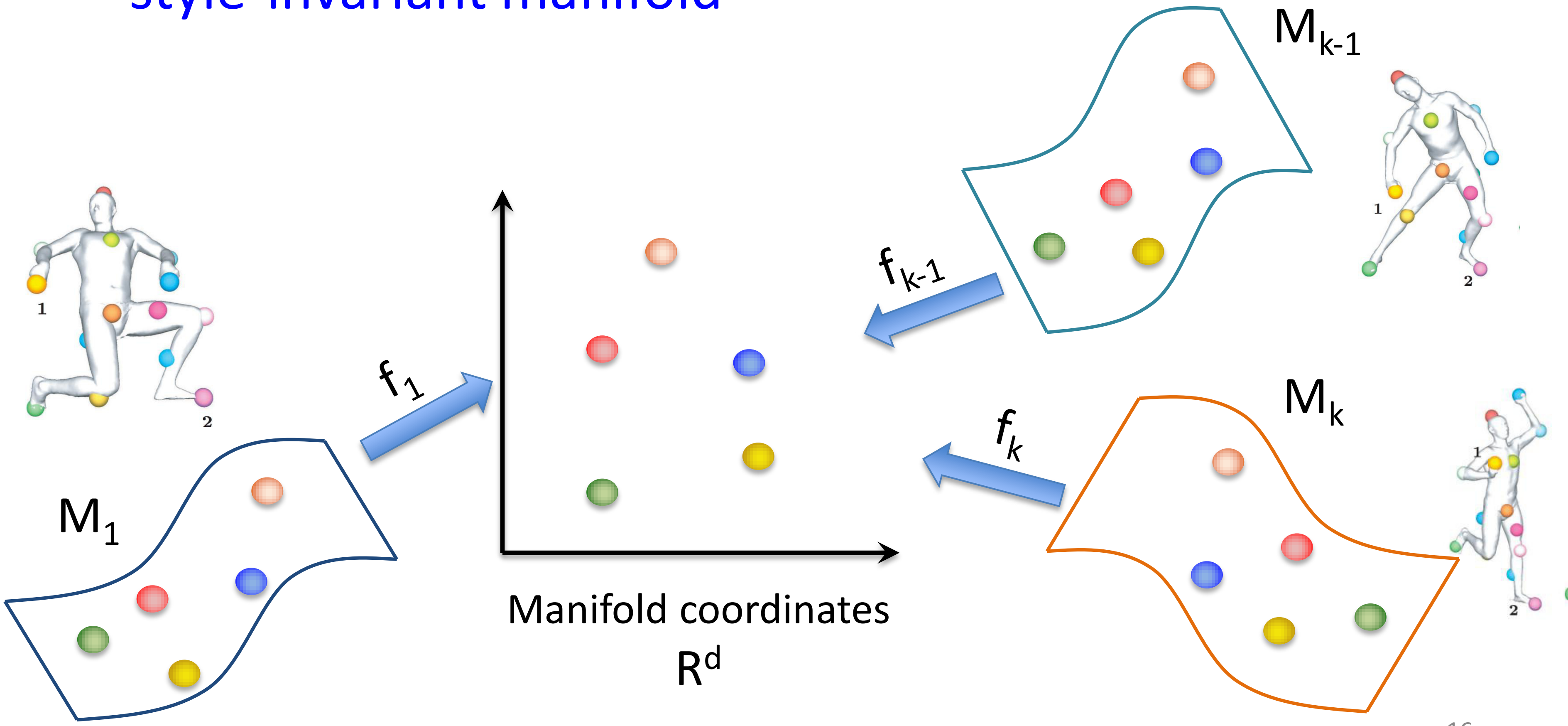
Manifold Assumption of 3D Shape Feature Points

- Assume that the feature points extracted on each shape form a low-dimensional manifold



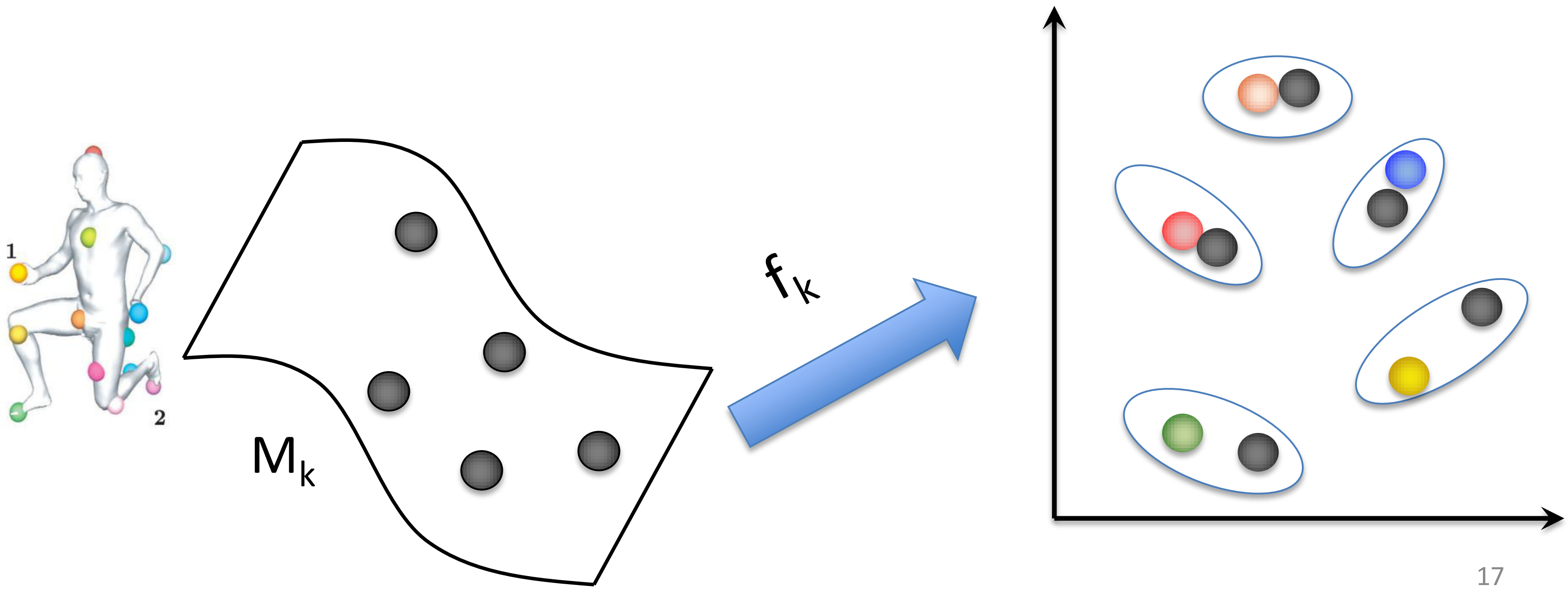
Latent Mean Manifold

- Each shape manifold is an instance of an underlying, style-invariant manifold



Shape Matching using Mean Manifold

- Project each shape manifold to the low-dimensional manifold coordinate system
- Match points based on the distances in the low-D space



Differences from Existing Manifold Learning Methods

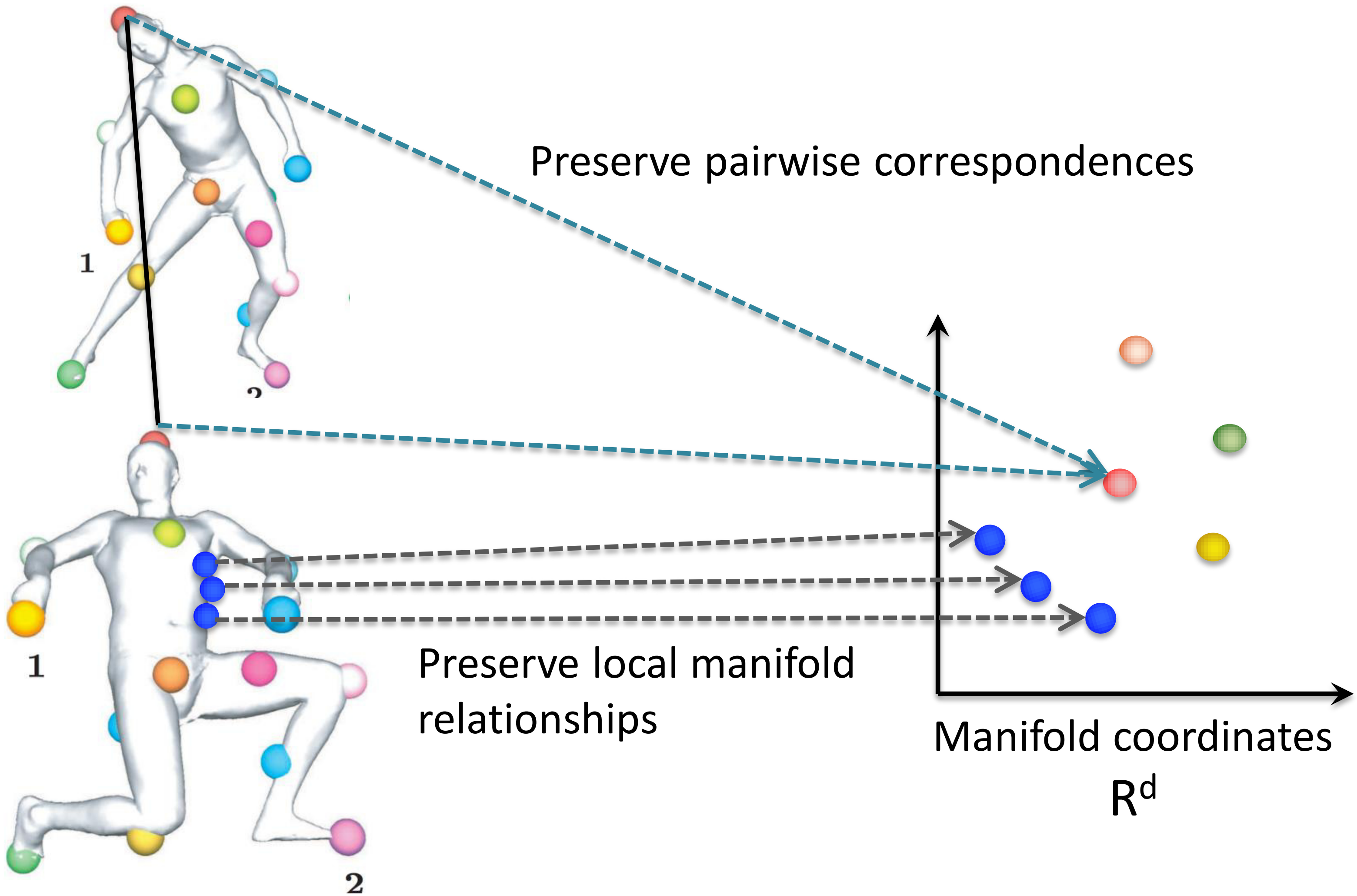
- Classic manifold learning techniques (e.g. LLE, Isomap)
  - **Input**: one set of unorganized high-D points
  - **Output**: low-D coordinates of the points on the latent manifold
  - Parametric mapping is generally unavailable

- Learning the mean manifold for shape matching
  - **Input**: multiple sets of high-D points with their pairwise (potentially noisy) correspondences
  - **Output**: low-D coordinates of the points on the mean manifold
  - Parametric mapping is necessary for mapping new feature points or matching unseen shapes to existing shapes

Learning the Mean Manifold

- The goal is to learn $f_k(\bullet)$, which maps feature points from each shape to a unified manifold representation
- The manifold regularizer
  - The learned mapping should preserve the local relationships on the original manifolds $\{M_k\}$
  - For example, preserving locally linear relationships (LLE), or preserving local proximity (Laplacian Eigenmaps)
- Pairwise correspondence constraint
  - Most input pairwise correspondences are correct
  - Originally matched points should also be close on the mean manifold

Illustration of Learning the Mean Manifold



Summary

- We discussed a preliminary idea of matching multiple shapes by learning a mean manifold
- Deep metric learning can provide parametric mappings from each shape to the mean manifold
  - Encourages matched points to be close in the mapped space
- How to incorporate the manifold regularizer in the cost function?
  - Most manifold regularizers are evaluated locally
  - Forward a set of neighborhood points through the network and evaluate the overall cost term