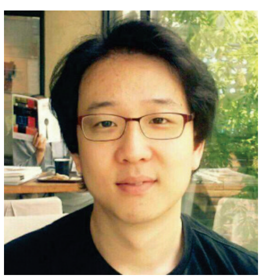
# Procrustean Normal Distribution for Non-Rigid Structure from Motion


Minsik Lee, Jungchan Cho, Chong-Ho Choi, Songhwai Oh



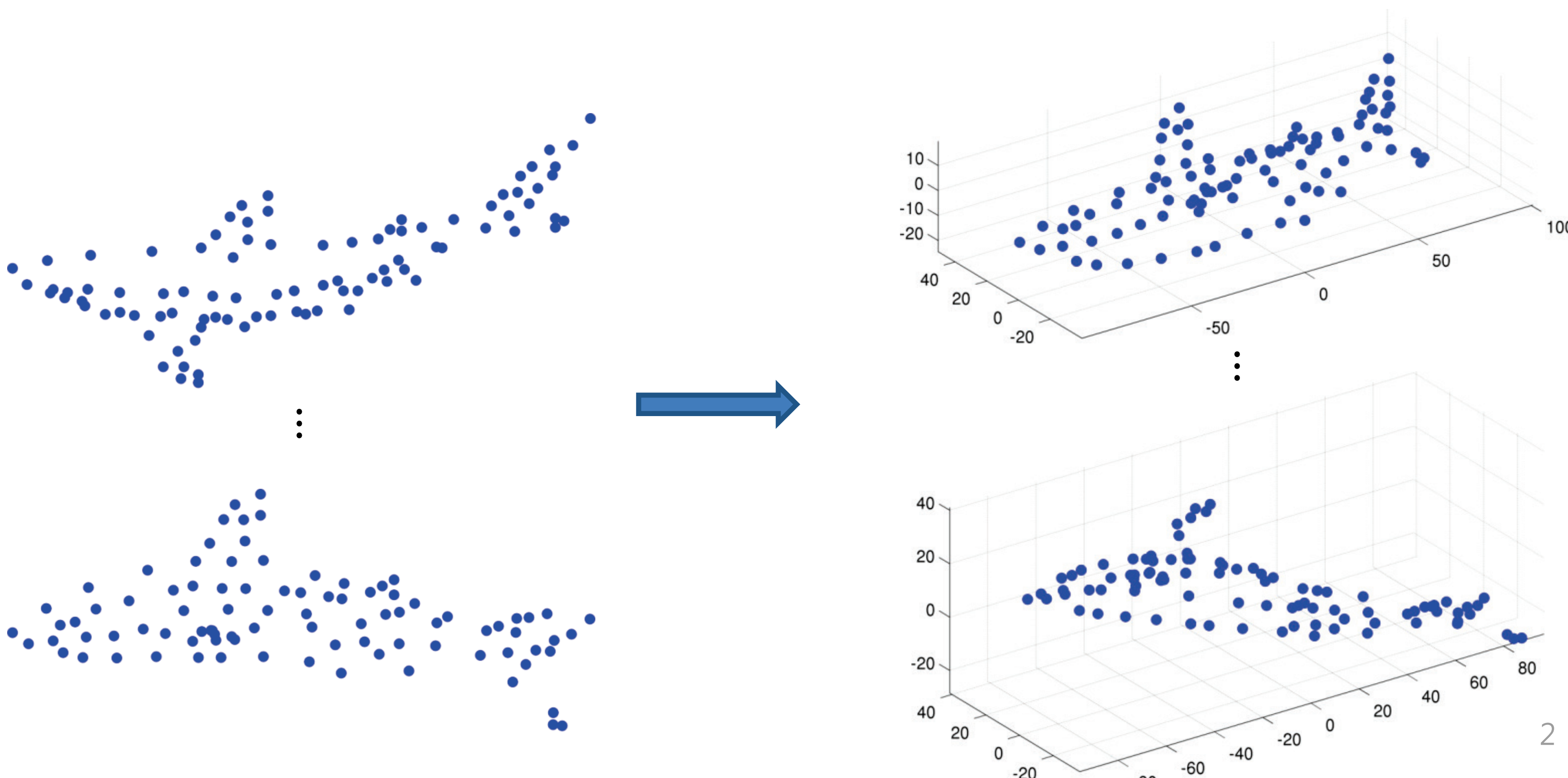

Seoul National University, Korea


CVPR 2013 Portland, Oregon June 23-28

# Non-Rigid Structure from Motion

- A fundamental problem in computer vision
  - App.: biometrics, human-computer interaction …

Problems that deals with deforming objects



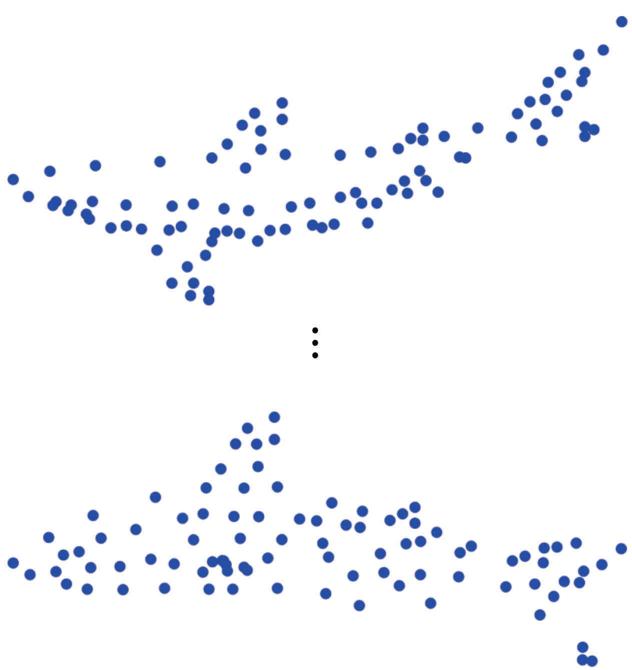
# Non-Rigid Structure from Motion

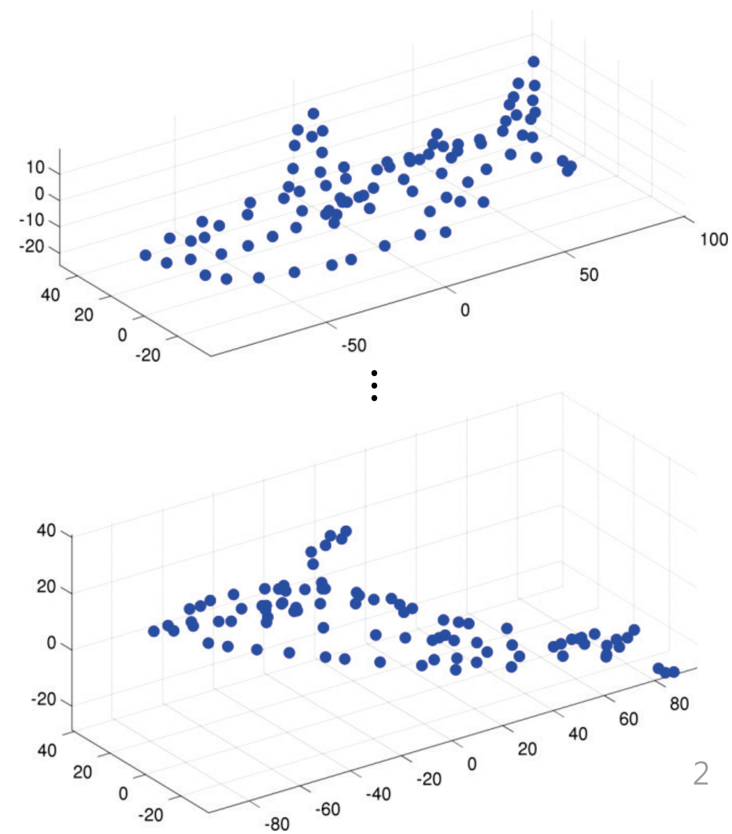
Bregler et al.
Torresani et al.
Bartoli et al.
Xiao et al.
Paladini et al.
Akhter et al.
Wang et al.
Dai et al.

Fayad et al.
Varrol et al.
Ferreira et al.
Del Blue et al.
Park et al.
Gotardo and Martinez

...and many other researchers

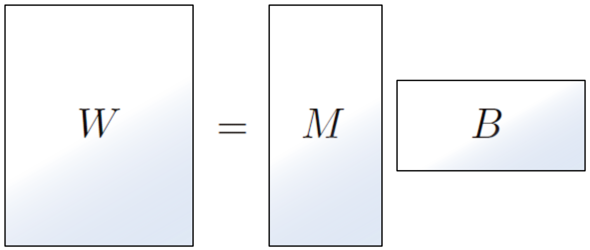
Tons of proposals,
and yet research is still going on.





# Existing methods

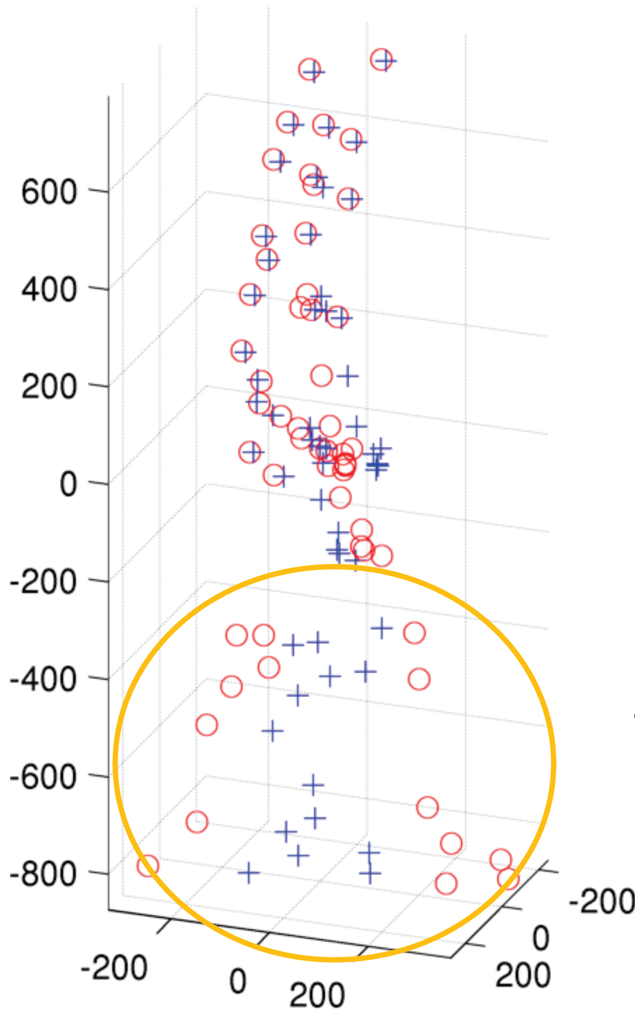
- Based on factorization technique
  [Xiao et al., 2004], [Gotardo and Martinez, 2011], [Akhter et al., 2011], [Dai et al., 2012], etc.
  Extended from [Tomasi and Kanade, 1992]




- Unique solution under the constraints;
  - Rotation (orthogonality + rank-3) + Deformation (low rank).
    [Akhter et al., 2009]

- It can be simply calculated.
  - Optimal solution using convex rank-minimization technique
    [Dai et al., 2012]

# But, still not very satisfactory: Why?




○: ground truth
+: reconstructed by [Dai et al., 2012]

The gap between ○ and + is not small.

Walking sequence [Torresani et al., 2008]

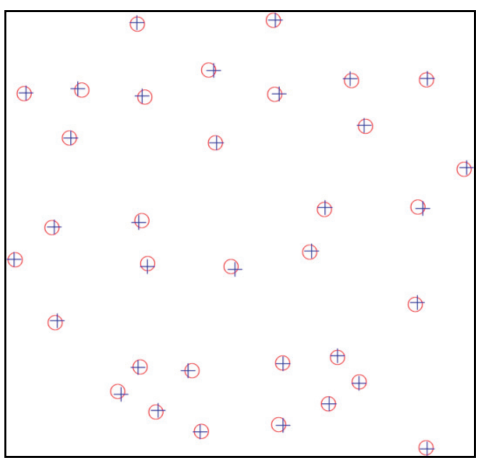
# Main problem: Rotation calculation

- What existing methods do:
  - Enforcing {orthogonality, rank-3} constraints
    = minimizing algebraic error

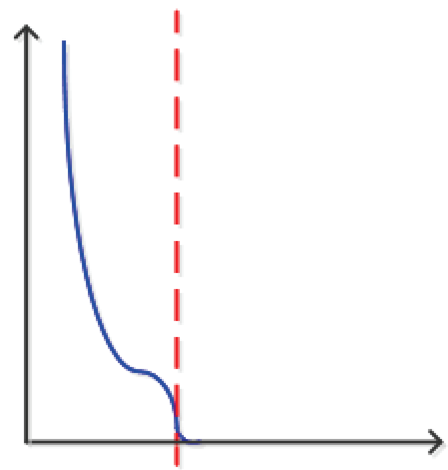
$$M = \tilde{M} \; G$$

$$B = G^{-1} \; \tilde{B}$$

- May work well if;



Very small noise

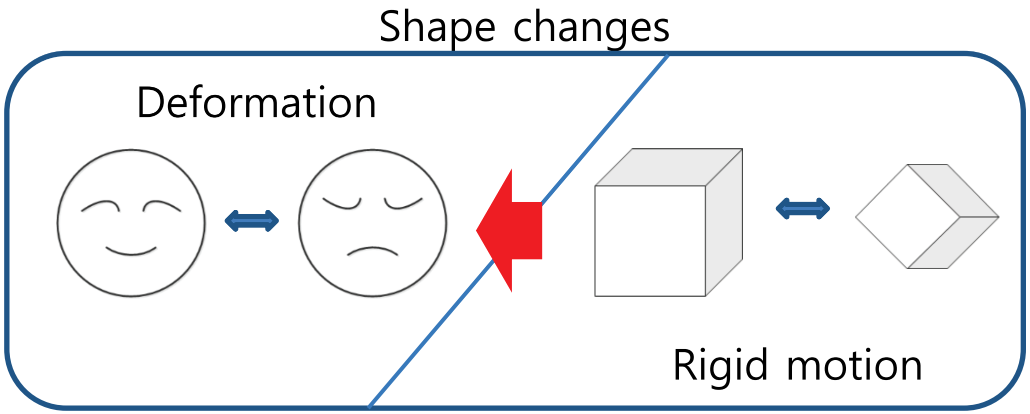


Correct number of bases

# Motivation

- Rigid motions (including rotations) need to be accurately computed in NRSfM.

  ← Non-rigid part is also affected by rigid motions.




- Our contribution
  - More robust constraints for rigid motions → a new prior dist.
  - Inspired by Generalized Procrustes Analysis (GPA).

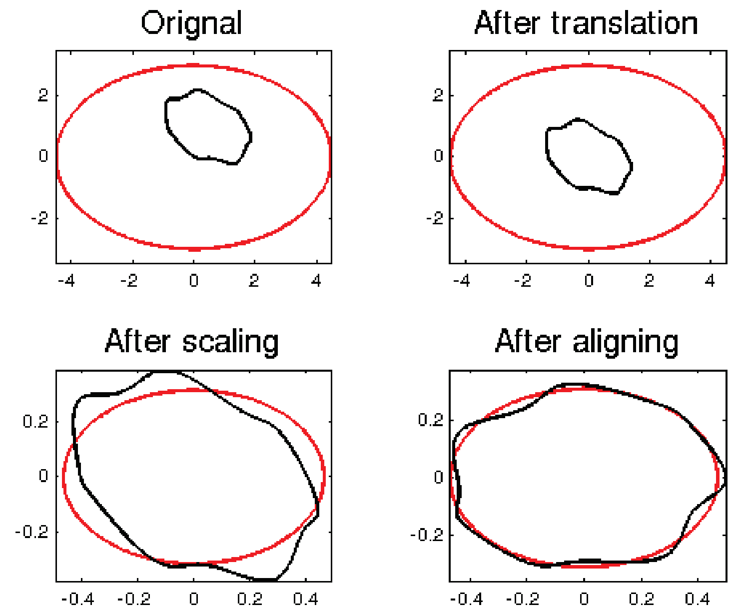
# Generalized Procrustes analysis

- "Standard" for rigid alignment

$$\min_{s_i, \mathbf{R}_i, \overline{\mathbf{X}}} \quad \sum \left\| s_i \mathbf{R}_i \mathbf{X}_i - \overline{\mathbf{X}} \right\|^2$$

$$\text{s.t.} \qquad \mathbf{R}_i \mathbf{R}_i^T = \mathbf{I}, \quad \| s_i \mathbf{X}_i \| = 1.$$

  - $s_i$(scale), $R_i$(rotation), $\bar{X}$(mean)




[Townsend, 2011]

- Properties
  - Gaussian noise assumption
  - Rotations determined by the mean shape

[Gower, 1975]

# GPA in NRSfM?

- Pros:
  - May find rotations more robustly.

- Cons:

  C1. Solution lies on a nonlinear manifold.

  (Procrustes shape space)

  C2. Itself is another optimization problem.

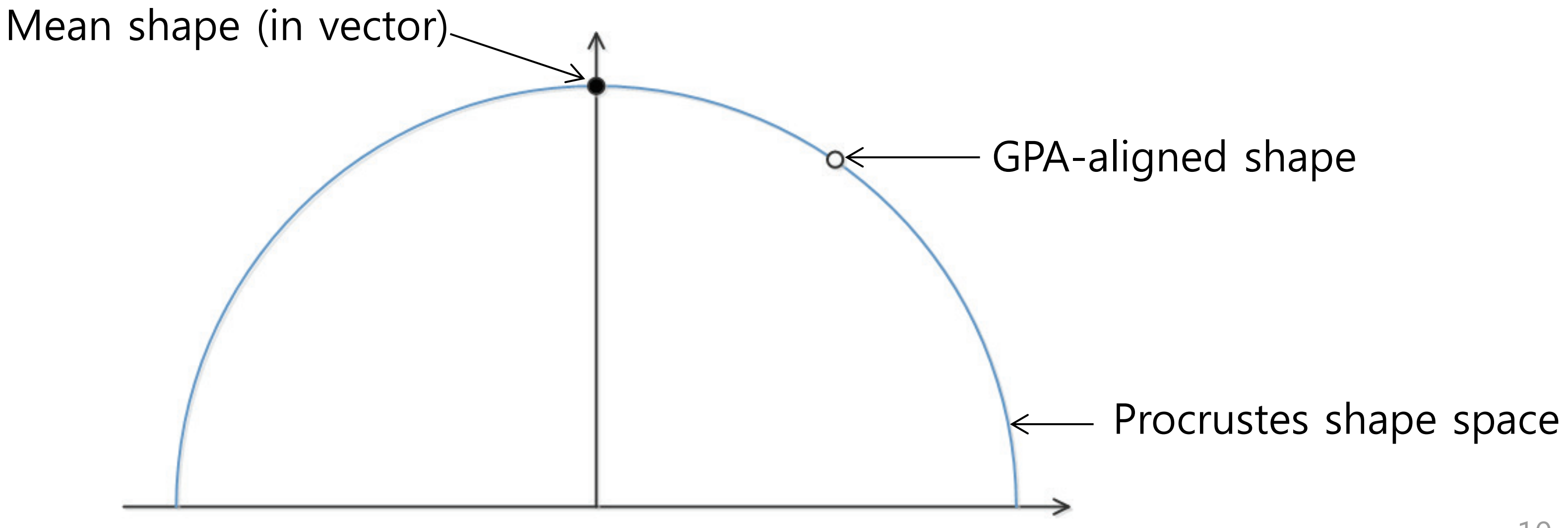
- How to overcome C1 and C2?

# C1 ← Linearize the solution space

- By changing the scale constraint

$$\|s_i \mathbf{X}_i\| = 1$$

"Each shape should be **unit-norm**."



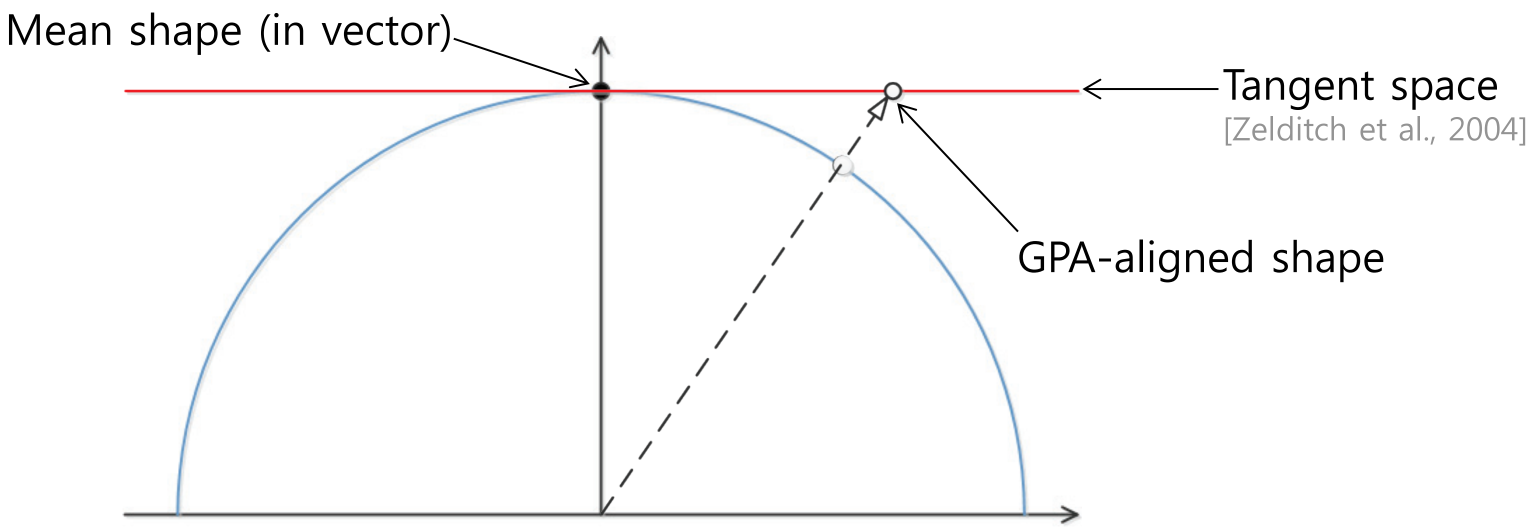
# C1 ← Linearize the solution space

- By changing the scale constraint

$$\|s_i \mathbf{X}_i\| = 1 \quad \Longrightarrow \quad \mathrm{vec}\left(s_i \mathbf{R}_i \mathbf{X}_i - \overline{\mathbf{X}}\right)^T \mathrm{vec}\left(\overline{\mathbf{X}}\right) = 0.$$

"Each shape should be **unit-norm**."

"Shape differences should be **perpendicular to the mean shape**."



# C2 ← Constraints from the optimality of GPA

- Optimality conditions in terms of aligned shapes $Y_i$ $(= s_i R_i X_i)$

$$\|\overline{\mathbf{Y}}\|^2 = 1, \quad \mathrm{tr}\left(\mathbf{Y}_i \overline{\mathbf{Y}}^T\right) = 1, \quad \mathbf{Y}_i \mathbf{1} = \mathbf{0}, \quad \mathbf{Y}_i \overline{\mathbf{Y}}^T - \overline{\mathbf{Y}} \mathbf{Y}_i^T = \mathbf{0}.$$

mean normalization

scale

translation

rotation

⇓

$$\mathbf{Q}_N^T \,\mathrm{vec}\left(\mathbf{Y}_i - \overline{\mathbf{Y}}\right) = \mathbf{0}$$

# C2 ← Constraints from the optimality of GPA

- Optimality conditions in terms of aligned shapes $Y_i$

  $(= s_i R_i X_i)$

$$\|\overline{\mathbf{Y}}\|^2 = 1, \quad \mathrm{tr}\left(\mathbf{Y}_i \overline{\mathbf{Y}}^T\right) = 1, \quad \mathbf{Y}_i \mathbf{1} = \mathbf{0}, \quad \mathbf{Y}_i \overline{\mathbf{Y}}^T - \overline{\mathbf{Y}} \mathbf{Y}_i^T = \mathbf{0}.$$

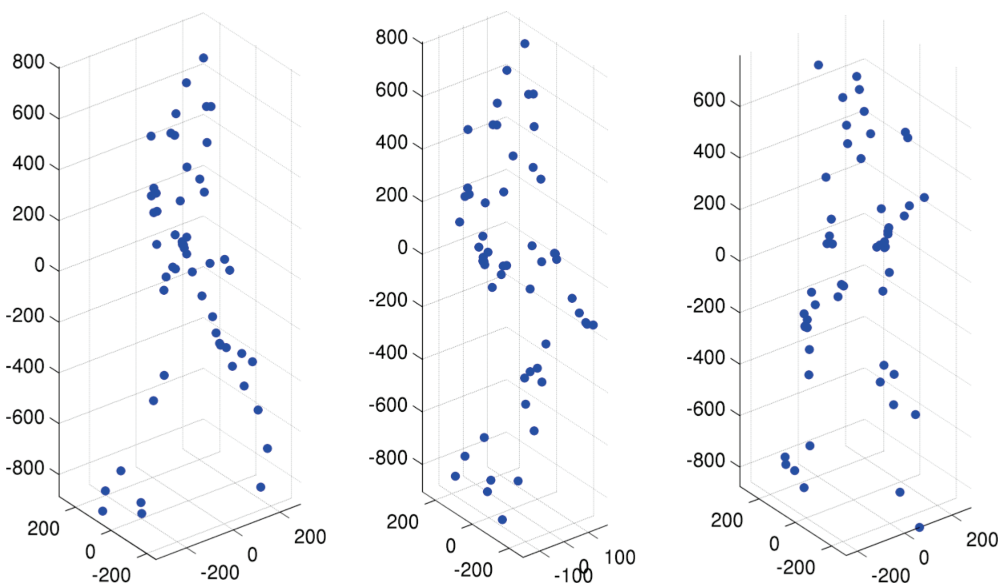
mean normalization | scale | translation | rotation

$$\mathbf{Q}_N^T \, \mathrm{vec}\left(\mathbf{Y}_i - \overline{\mathbf{Y}}\right) = \mathbf{0}$$ (modified GPA constraint)

- Function of $\overline{Y}$ (mean shape)
- 7-dimensional
  = DOF of similarity transform!
- Null space of deformation

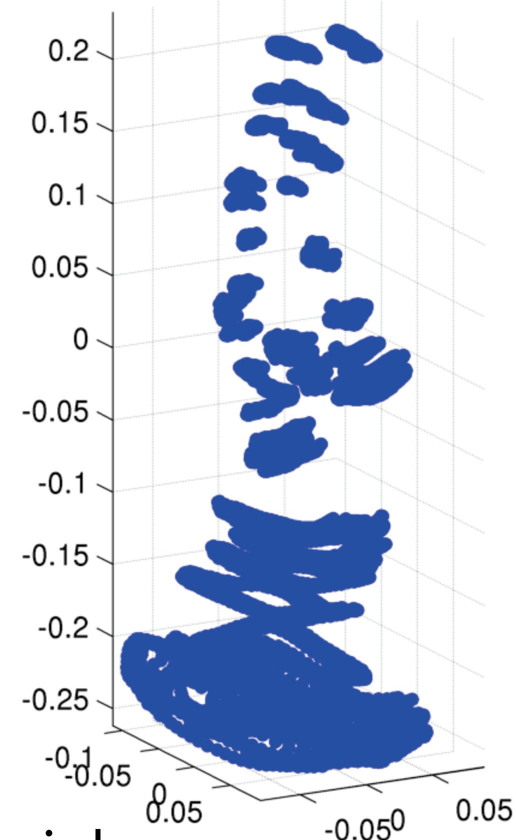
# GPA-aligned shapes

- Concentrated in a small region of the subspace satisfying $\mathbf{Q}_N^T \, \mathrm{vec}(\mathbf{Y}_i - \overline{\mathbf{Y}}) = \mathbf{0}$




Walking sequence [Torresani et al., 2008]

(modified) GPA

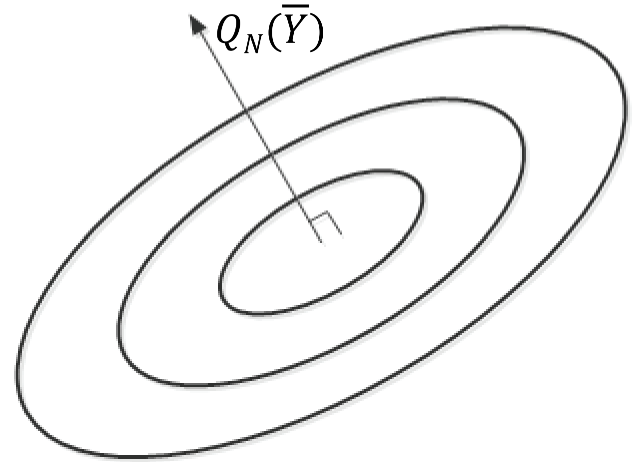



- Can be modeled as a dist. with some special properties. → Procrustean Normal Distribution

# Procrustean Normal Distribution

- A normal distribution with modified GPA constraint

$$\mathbf{Y} \sim \mathcal{N}_P\left(\overline{\mathbf{Y}}, \boldsymbol{\Sigma}\right)$$




Distribution of $Y$

PND variable

Mean shape (unit-norm)

$$\text{vec}(\mathbf{Y}) = \mathbf{Q}\mathbf{u} + \text{vec}\left(\overline{\mathbf{Y}}\right),$$

$$\mathbf{u} \sim \mathcal{N}(\mathbf{0}, \boldsymbol{\Sigma}_{\mathbf{u}}), \quad \mathbf{Q}^T\mathbf{Q}_N = \mathbf{0}.$$

Gaussian r.v.

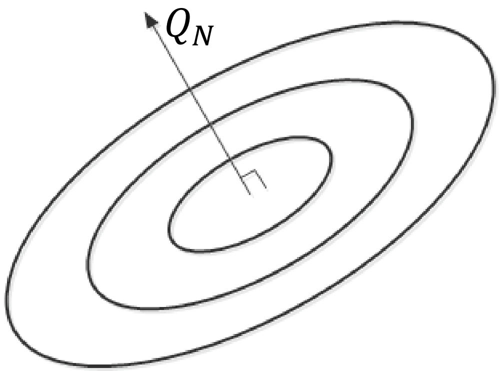
functions of $\overline{Y}$

$$\mathbf{Q}_N^T \text{vec}\left(\mathbf{Y}_i - \overline{\mathbf{Y}}\right) = \mathbf{0}$$

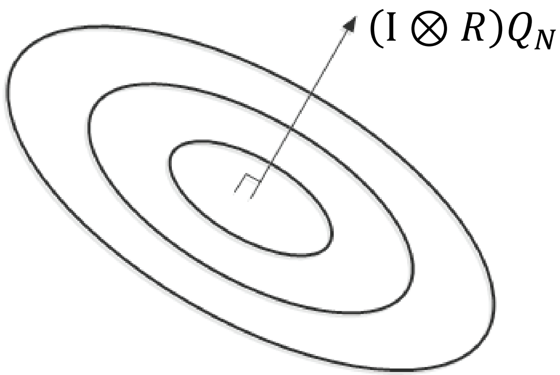
# PND: Properties

- A general model of deformation
- Satisfies the modified GPA constraint
  - Determined by the mean shape
  - Stronger than orthonormality constraint
- No rank restriction on deformation
- Rotation of a PND is also a PND

**Proposition 2.** *Let* $\mathbf{Y} \sim \mathcal{N}_P(\overline{\mathbf{Y}}, \mathbf{\Sigma})$ *be a PND random matrix, and* $\mathbf{Y}' = \mathbf{R}\mathbf{Y}$ *for an orthogonal matrix* $\mathbf{R}$. *Then,* $\mathbf{Y}'$ *is also a PND random matrix as* $\mathbf{Y}' \sim \mathcal{N}_P(\mathbf{R}\overline{\mathbf{Y}}, (\mathbf{I} \otimes \mathbf{R})\,\mathbf{\Sigma}\,(\mathbf{I} \otimes \mathbf{R}^T))$.



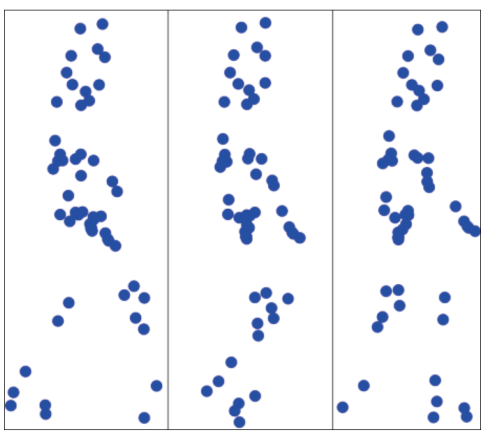

Distribtuion of $Y$




Distribution of $R\,Y$

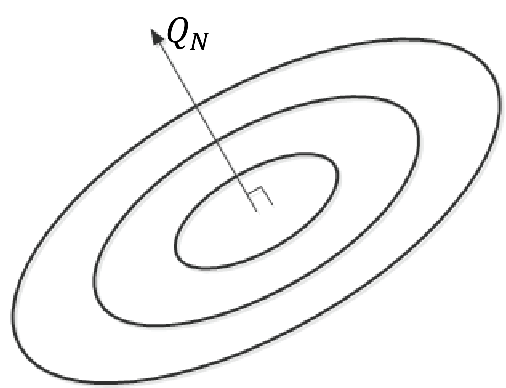
# NRSfM: Fitting PND to input data

| | | |
|---|---|---|
| 2D input | = | observation |
| z coordinates | = | missing data |
| PND | = | underlying distribution |



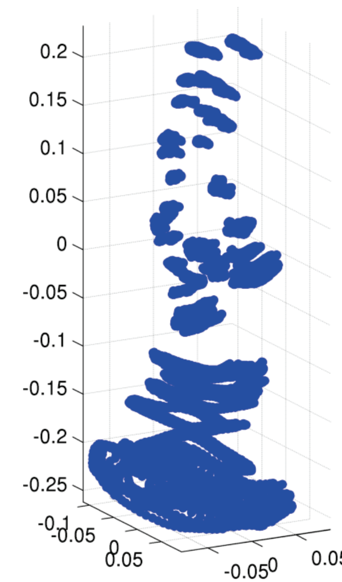
2D input
Walking sequence
[Torresani et al., 2008]

fitting




PND

=




Dist. Of 3D shapes

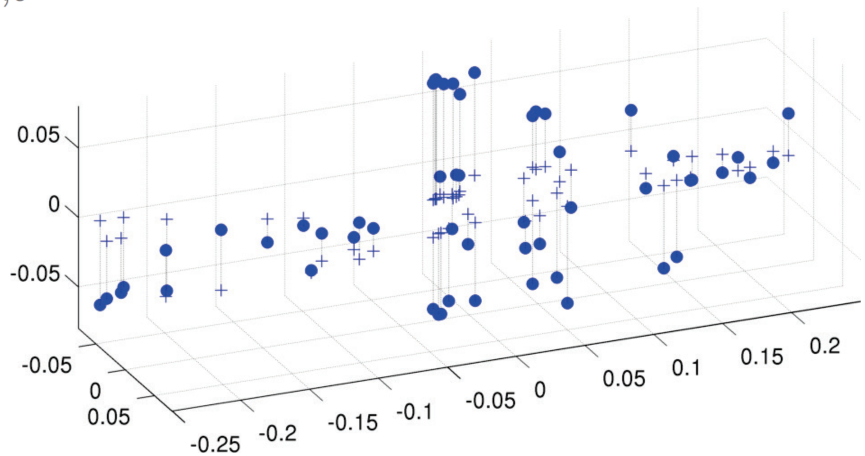
# EM formulation

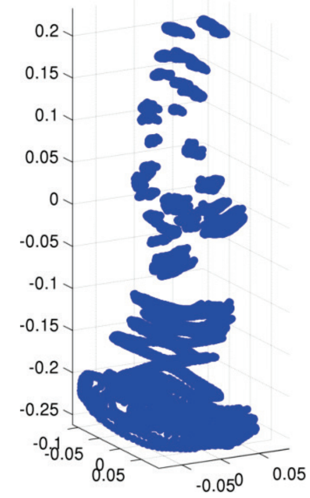
| | | |
|---|---|---|
| 2D input data | – | $D_i$ |
| Hidden variable (3D shape) | – | $X_i$ |
| Parameters | – | $s_i$, $R_i$, $\bar{X}$, $\Sigma$ (covariance), $\sigma$ (observation noise) |

- 3D shape follows PND ($s_i \mathbf{R}_i \mathbf{X}_i \sim \mathcal{N}_P(\overline{\mathbf{X}}, \mathbf{\Sigma})$)

[Data term] [Prior term]

$$\min_{s_i, \mathbf{R}_i, \overline{\mathbf{X}}, \mathbf{\Sigma}, \sigma} \sum \log(p(\mathbf{D}_i | \mathbf{X}_i, \sigma)) + \log\left(p\left(\mathbf{X}_i \middle| s_i, \mathbf{R}_i, \overline{\mathbf{X}}, \mathbf{\Sigma}\right)\right)$$



# EM procedure

- E-step: Find the distribution of $X_i$.

$$p_E(\mathbf{X}_i) = p_E(\mathrm{vec}(\mathbf{X}_i)) \sim \mathcal{N}(\mathbf{m}_i, \mathbf{C}_i)\,,$$
$$\mathbf{C}_i = \mathbf{H}_i^+, \quad \mathbf{m}_i = \frac{1}{\sigma^2}\mathbf{C}_i\,\mathrm{vec}(\mathbf{D}_i)\,.$$

- M-step: Maximize the expectation of log-likelihood.

$$\overline{\mathbf{X}} = \sum s_i \mathbf{R}_i \mathbf{M}_i / \left\| \sum s_i \mathbf{R}_i \mathbf{M}_i \right\|$$
$$\mathbf{M}_i \overline{\mathbf{X}}^T = \mathbf{U}_i \mathbf{\Lambda}_i \mathbf{V}_i^T, \quad \mathbf{R}_i = \mathbf{V}_i \mathbf{U}_i^T,$$
$$s_i = 1/\,\mathrm{tr}\left(\mathbf{R}_i \mathbf{M}_i \overline{\mathbf{X}}^T\right) = 1/\,\mathrm{tr}(\mathbf{\Lambda}_i)\,,$$
$$\mathbf{\Sigma}_R = \frac{1}{n_s} \sum \mathbf{h}_i \mathbf{h}_i^T + s_i^2 \mathbf{Q}^T (\mathbf{I} \otimes \mathbf{R}_i)\, \mathbf{C}_i \left(\mathbf{I} \otimes \mathbf{R}_i^T\right) \mathbf{Q}.$$
$$\sigma^2 = \frac{\alpha}{\sum n_i^W} \sum \|\mathrm{vec}(\mathbf{D}_i) - \mathbf{F}_i \mathbf{m}_i\|^2 + \mathrm{tr}(\mathbf{F}_i \mathbf{C}_i)\,.$$

# EM-PND chooses a PND with maximum log-likelihood

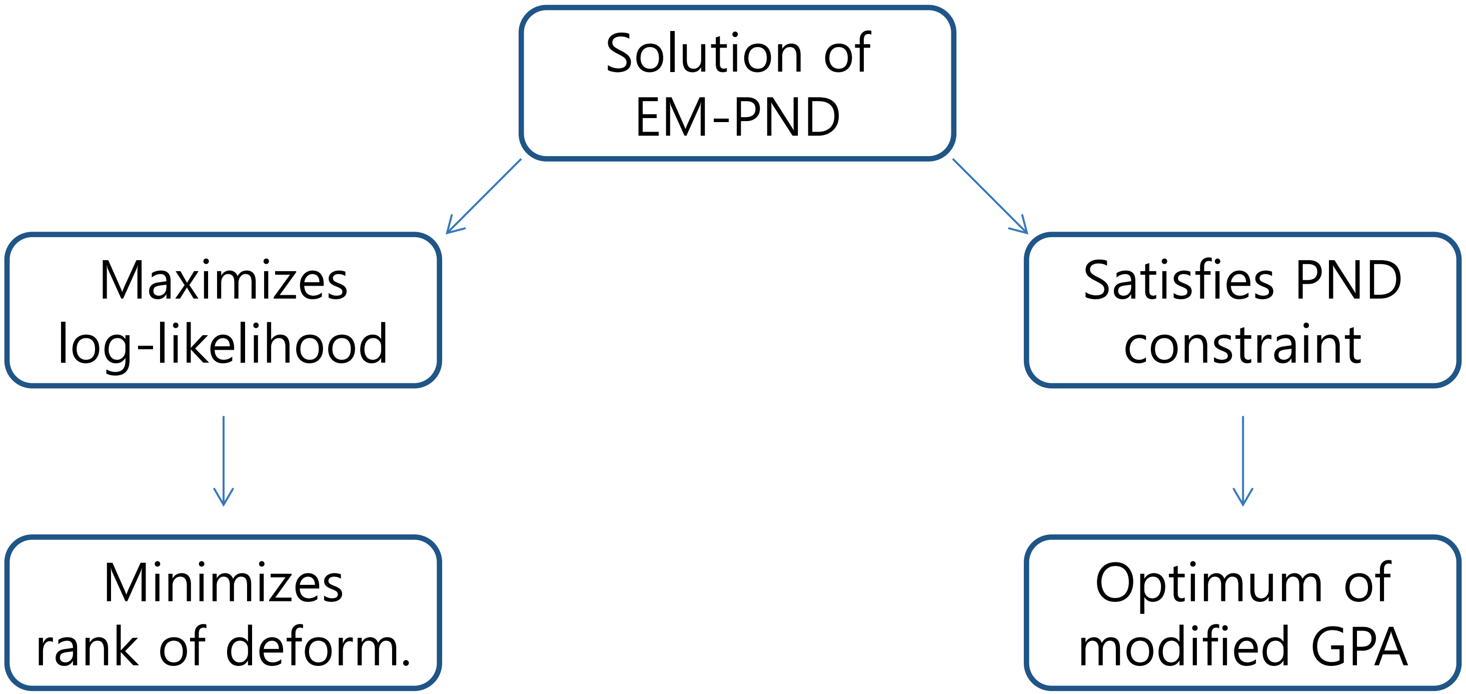
- Log-likelihood Gaussian (PND) prior = $\log|\Sigma|$

$$\sum \log|\Sigma| + \mathrm{tr}\left(x_i^T \Sigma^{-1} x_i\right) \xrightarrow{\text{Optimal } \Sigma = \frac{1}{N}\sum x_i x_i^T} N\left(\log|\Sigma| + 1\right)$$

- $\log|\Sigma|$ is a smooth surrogate of $\mathrm{rank}(\Sigma)$.
  - [Rao et al., 2008]

∴ EM-PND minimizes the rank of shape covariance.
(=rank of shape bases)

# What EM-PND does




- EM-PND finds GPA-aligned shapes,
  that minimize the rank of shape bases.

# Some notes on the optimization

- Cost function is highly complex.
  - Numerous parameters : $s_i$, $R_i$, $\bar{X}$, $\Sigma$, $\sigma$
  - $Q$ is a function of $\overline{X}$.
  - $s_i$ depends on $\overline{X}$.
  - $R_i X_i \overline{X}^T$ should be symmetric.
  - Null space issues.

→ A few tricks in optimization

- In fact, quite fast
  - Converges within 0.5~2 minutes.
    (MATLAB, w/o MEX function / 3GHz dual-core PC)

# Practical considerations

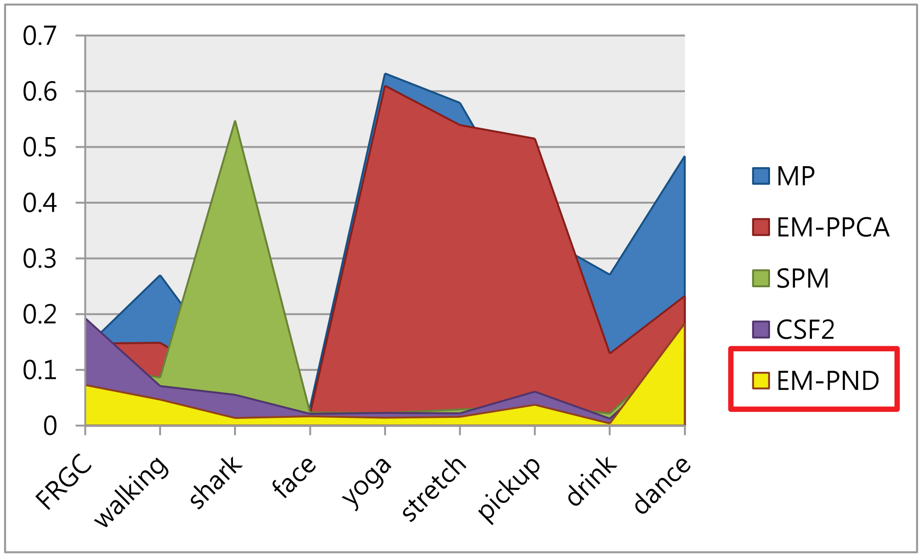
- Highly nonlinear → Initial values are crucial
  - Especially rotations
- An existing initialization scheme is good enough.
  - [Gotardo and Martinez, 2011]
  - Test every rank, pick the most "orthogonal" rotations.
  - No tuning parameters
- Pre-iteration: By fixing some parameters
  - Reduced processing time
  - Improved performance for some data sets

# Results

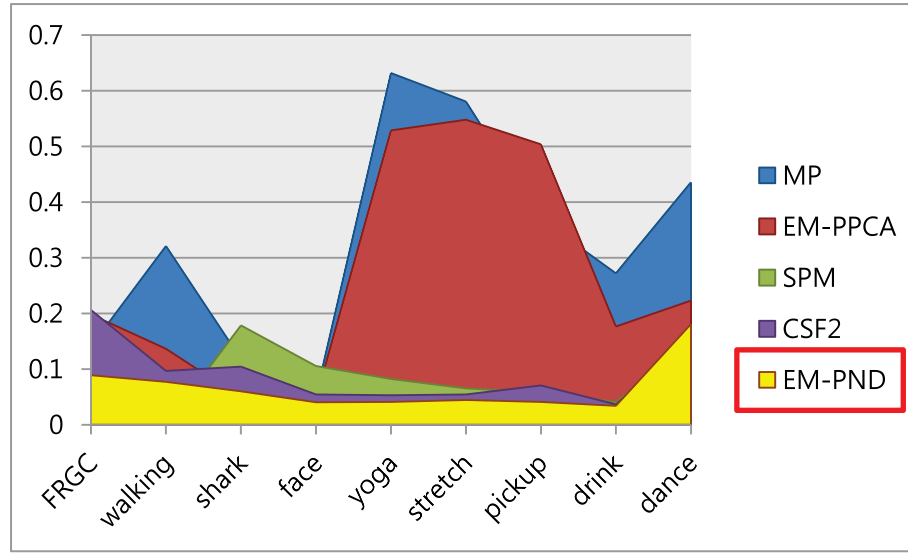
Data sets: [Torresani et al., 2008], [Akhter et al., 2011] / Other methods: [Torresani et al., 2008], [Paladini et al., 2009], [Gotardo and Martinez, 2011], [Dai et al., 2012]

# Normalized reconstruction error

**w/o noise**




**w/ noise**

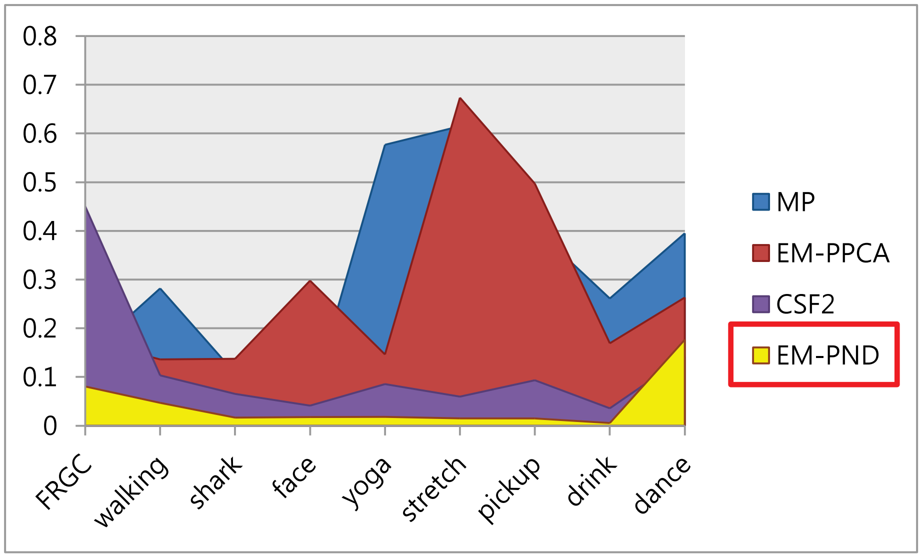



EM-PND outperforms the other methods.

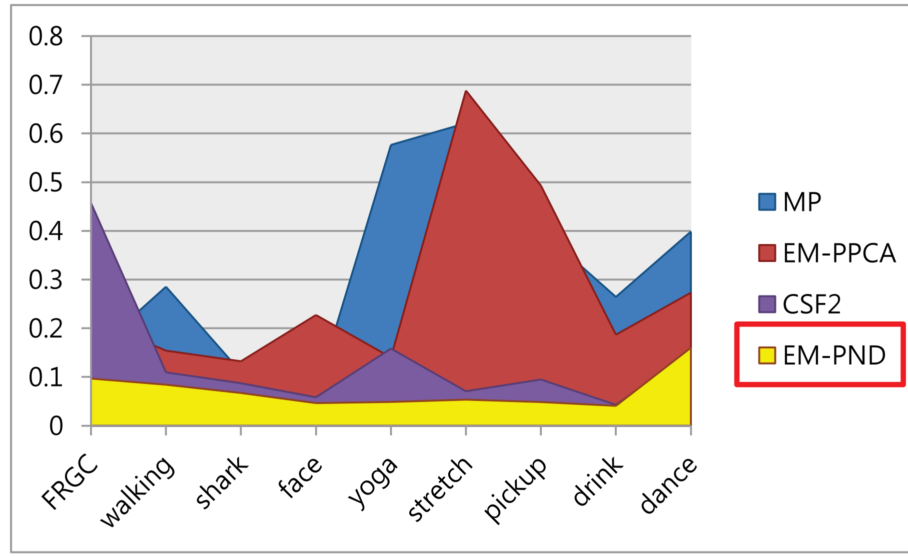
(Up to 75% smaller error than the 2nd best)

# With 30% missing points

**w/o noise**




**w/ noise**

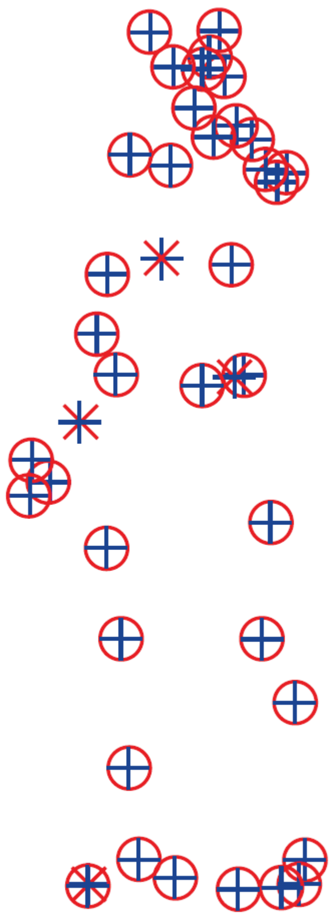



EM-PND is robust to missing points.

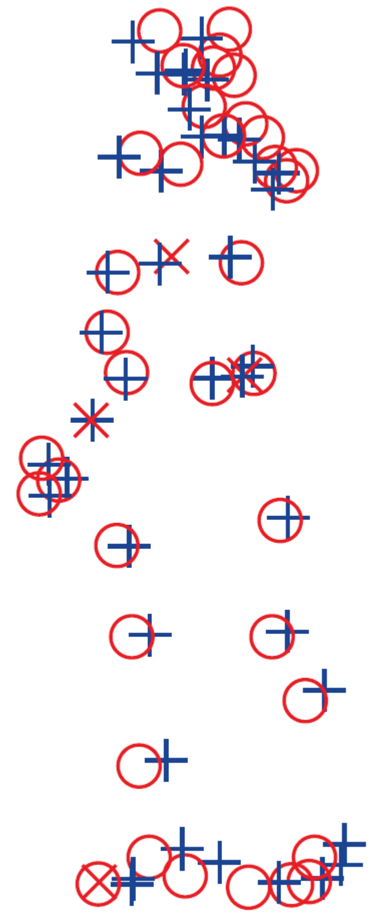
(Up to 85% smaller error than the 2nd best)

# Closer look - Drink

○: ground truth
×: missing points
+: reconstructed

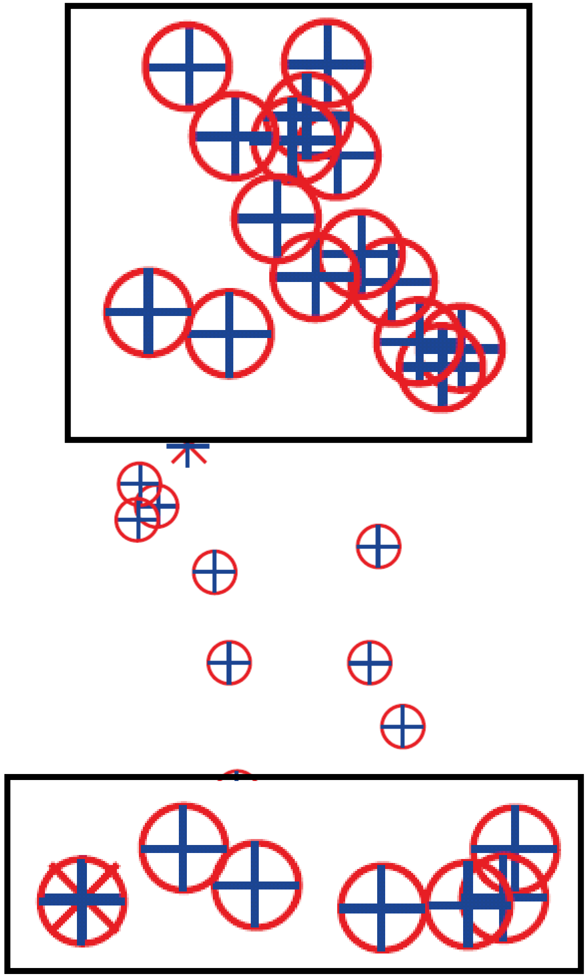


EM-PND

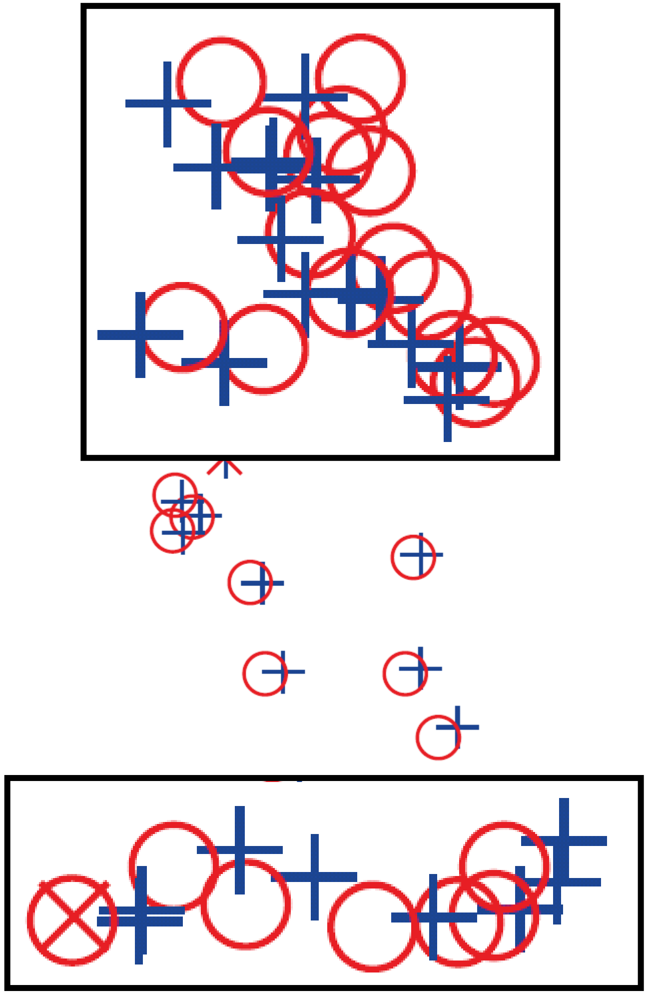


CSF2 (2nd best)

# Closer look - Drink

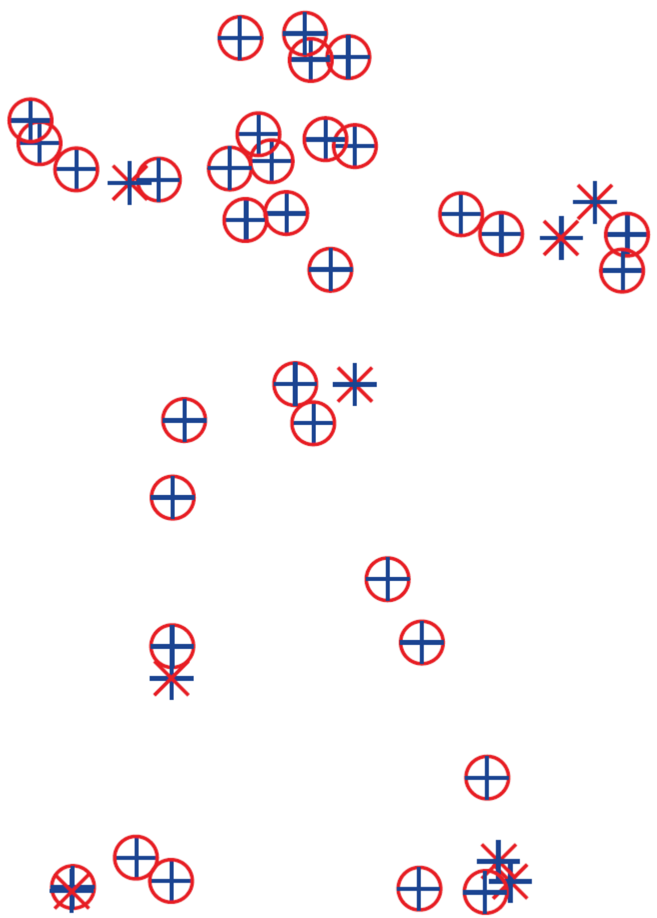
EM-PND



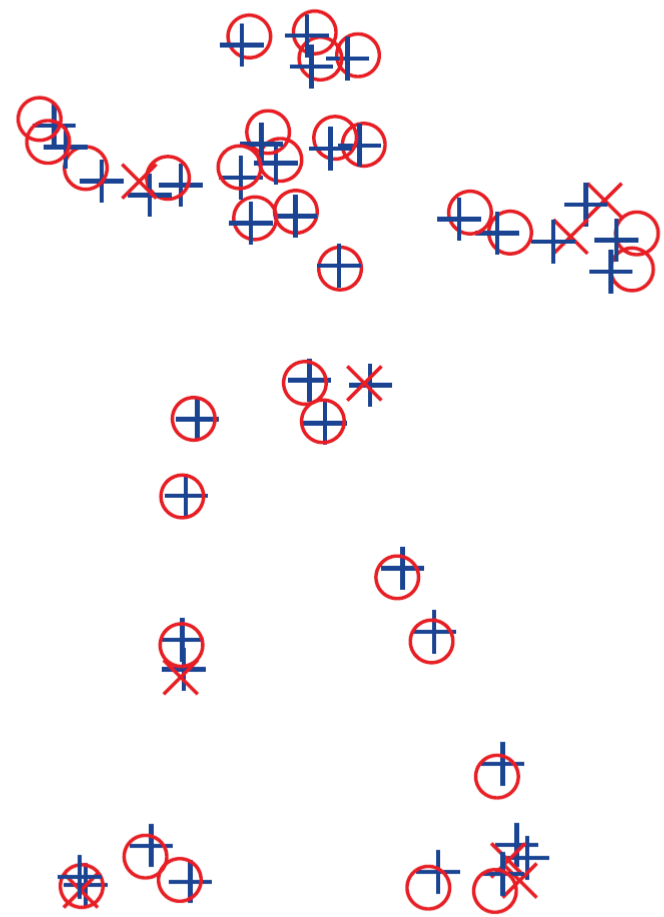
CSF2 (2nd best)

# Closer look - Yoga

○: ground truth
×: missing points
+: reconstructed



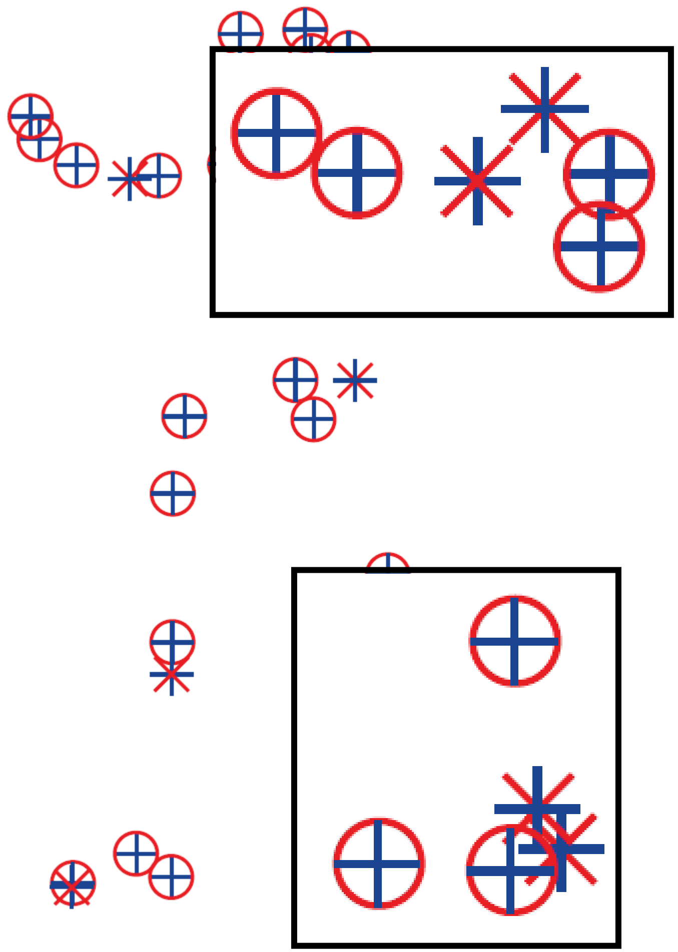
EM-PND



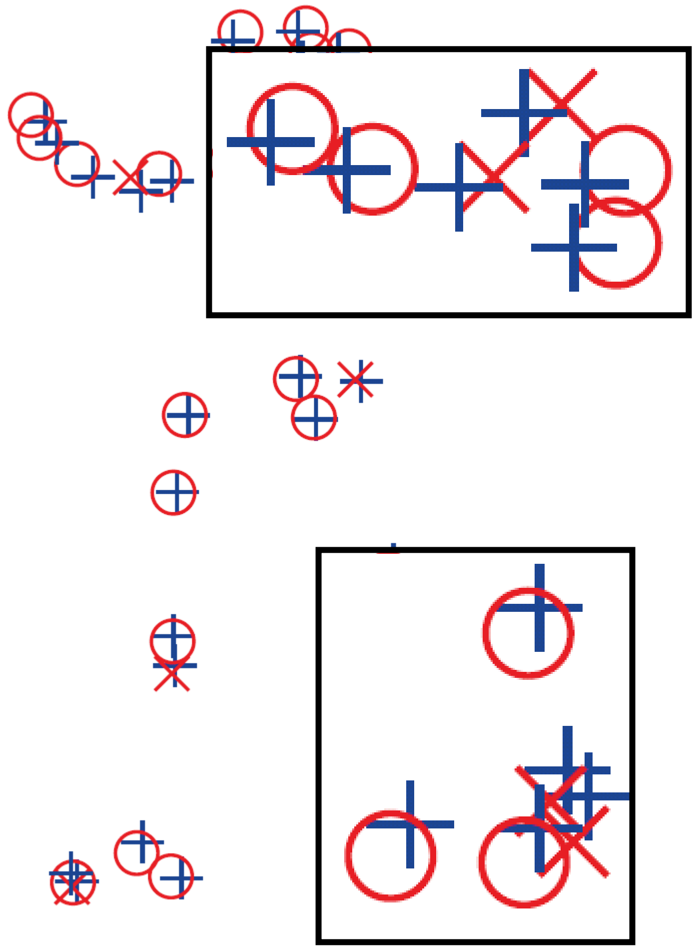
CSF2 (2nd best)

# Closer look - Yoga

○: ground truth
×: missing points
+: reconstructed



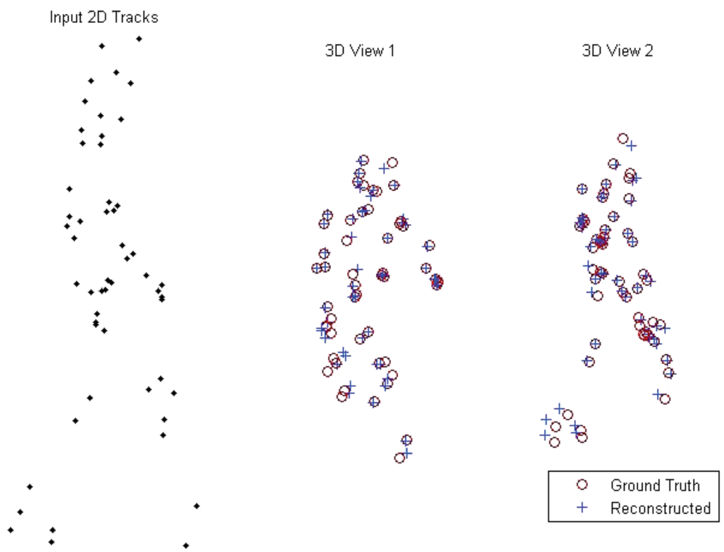
EM-PND



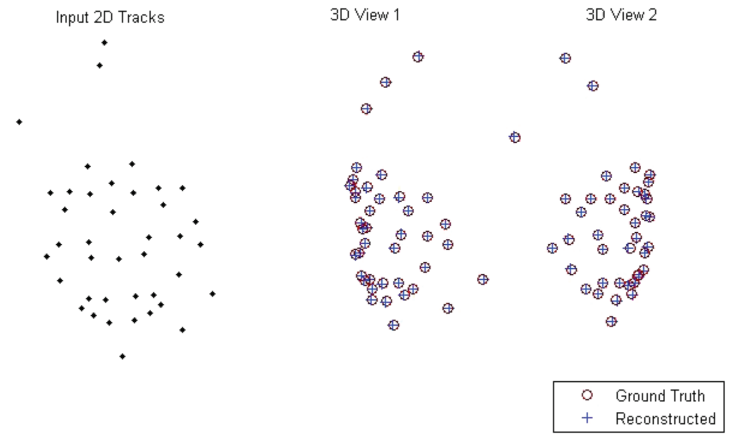
CSF2 (2nd best)

# Videos

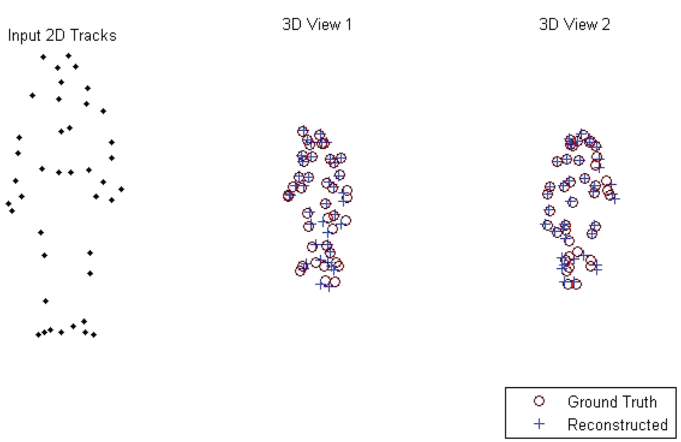
Walking



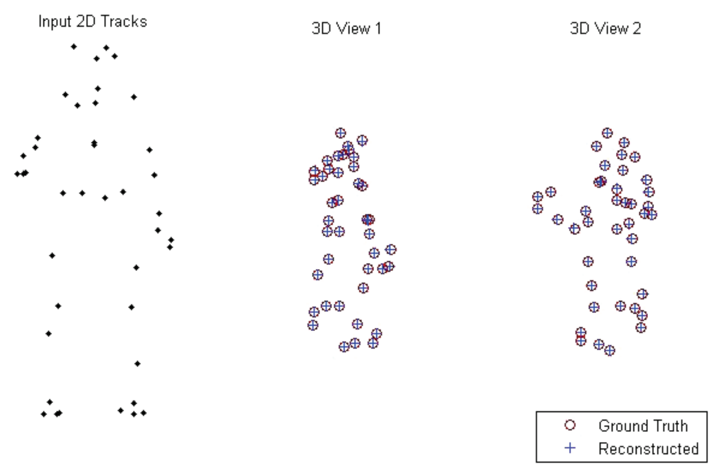

Face




Stretch




Drink



# Conclusion

- “Uniqueness only” is not good enough.
  - A model should work well for real situations.
- PND: GPA + Normal distribution
  - A general deformation model
- EM-PND: NRSfM based on PND
  - Outperforms the state-of-the-art.
  - Complex procedure, yet fast convergence
- Future work
  - Outlier case, articulated objects, temporal dependence, etc.