Non-Rigid Structure from Motion for Building 3D Face Model

Diploma Thesis

Chengchao Qu

Advisors:

Dipl.-Inform. Hua Gao
Dr.-Ing. Hazım Kemal Ekenel

Facial Image Processing and Analysis Group (FIPA)
Institute for Anthropomatics (IFA)
Department of Informatics
Karlsruhe Institute of Technology

13.01.2011

Outline

Karlsruhe Institute of Technology

Outline

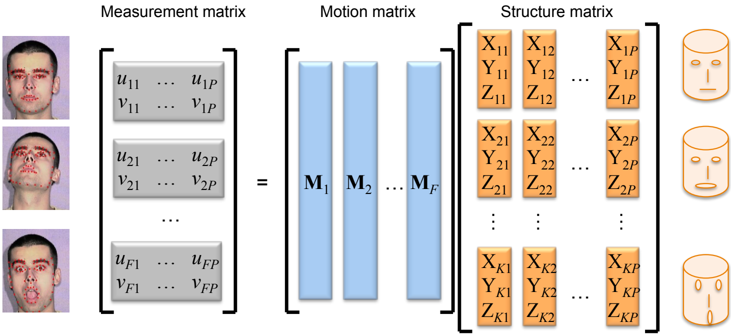
Introduction

Methodology
- Probabilistic PCA
- Probabilistic Relational PCA
- PPCA Optimization on Manifolds

Experiments

Conclusion

Motivation



INPUT

- 2D image features tracked over a time span

OUTPUT

- 3D configurations as point location of the deforming shape at each time span

Challenges

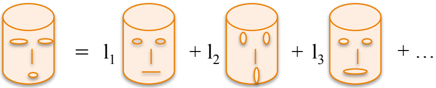
- Reconstruction from a single camera
- No pre-training subspace models
- Object freely moving and articulating
- Ambiguities & degeneracies in matrix factorization
- Noisy data
- Closed form vs. Statistical estimation

Outline

Outline

Probabilistic PCA

- Shape model



$$\mathbf{s}_i = \bar{\mathbf{s}} + \mathbf{V}\mathbf{z}_i$$

- Gaussian prior latent variable

$$\mathbf{z}_i \sim \mathcal{N}(0; \mathbf{I})$$

- 2D projection with zero-mean Gaussian noise

$$\mathbf{p}_i = \mathbf{R}_i\mathbf{s}_i + \mathbf{T}_i + \mathbf{n}_i$$
$$\mathbf{p}_i \sim \mathcal{N}(\mathbf{R}_i\bar{\mathbf{s}} + \mathbf{T}_i; \mathbf{R}_i\mathbf{V}\mathbf{V}^\top\mathbf{R}_i^\top + \sigma^2\mathbf{I})$$

Probabilistic PCA

- Estimate the PPCA model with EM algorithm

$$\begin{aligned} q(\mathbf{z}_i) &= p(\mathbf{z}_i|\mathbf{p}_i, \mathbf{R}_i, \mathbf{T}_i, \bar{\mathbf{s}}, \mathbf{V}, \sigma^2) \\ &= \mathcal{N}(\mathbf{z}_i|\mu; \mathbf{\Sigma}) \end{aligned}$$

- E-step: compute distribution over the latent variable $\mathbf{z}_i$

$$\begin{aligned} \mu &\equiv \mathbb{E}[\mathbf{z}] \\ \phi &\equiv \mathbb{E}[\mathbf{z}\mathbf{z}^\top] = \mathbf{\Sigma} + \mu\mu^\top \end{aligned}$$

- M-step: update the other motion parameters by minimizing negative log likelihood

$$Q \equiv (\mathbf{R}, \mathbf{T}, \bar{\mathbf{s}}, \mathbf{V}, \sigma^2) = \arg\min_{Q} \mathbb{E}[-\log p(\mathbf{p}|\mathbf{R}, \mathbf{T}, \bar{\mathbf{s}}, \mathbf{V}, \sigma^2)]$$

Probabilistic PCA

BUT. . .

- useful information between frames are discarded by the i.i.d. assumption
- the performance with noise is still not satisfying

Outline

Probabilistic Relational PCA

- Relational matrix Δ in covariance

$$\mathbf{z} \sim \mathcal{N}(0; \mathbf{\Phi}), \mathbf{\Phi} = \Delta^{-1}$$

- E-step

$$\mu \equiv \mathbb{E}[\mathbf{z}]$$
$$\phi' \equiv \mathbb{E}[\mathbf{z}\Delta\mathbf{z}^\top] = \mathbf{\Sigma} + \mu\Delta\mu^\top$$

- Relational matrix Δ
  - Apply Procrustes alignment & PCA to the input 2D data
  - The first two eigenvectors represent pan & tilt
  - Framewise distance of the remaining as shape relation $\mathbf{A}$
  - $\Delta = \gamma\mathbf{I} + (\mathbf{I} + \mathbf{A})(\mathbf{I} + \mathbf{A})^\top$

Outline

Camera Rotation Update

Camera rotation is subject to orthonormality constraint[1], therefore no closed form update.

- Original PPCA performs a single step of Gauss-Newton update on Euclidean space, which is not optimized & has only a low convergence rate.
- Smooth manifold optimization is a natural generalization of smooth optimization on Euclidean space, unconstrained optimization on Manifolds & quadratic convergence rate possible.

[1] A rotation matrix in n-dimensions is a $n \times n$ special orthogonal matrix, that is an orthogonal matrix whose determinant is 1: $\mathbf{R}^\top = \mathbf{R}^{-1}, \det \mathbf{R} = 1$.

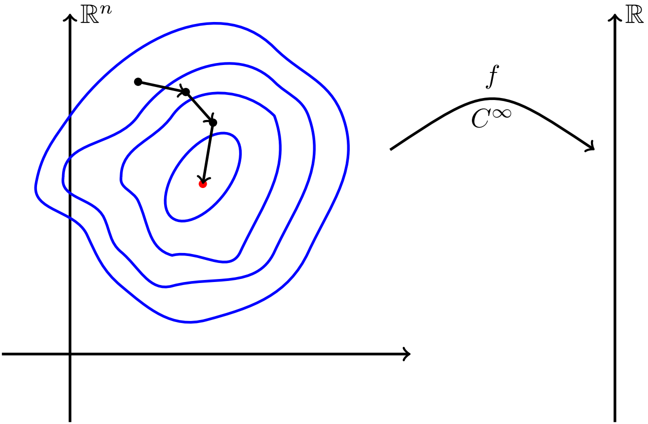
Camera Rotation Update

Camera rotation is subject to orthonormality constraint[1], therefore no closed form update.

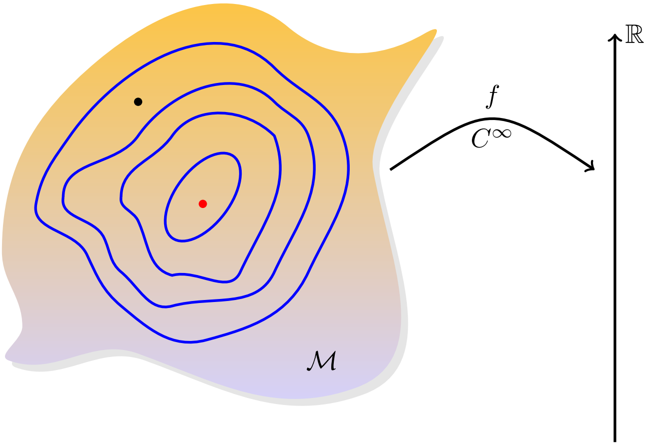
- Original PPCA performs a single step of Gauss-Newton update on Euclidean space, which is not optimized & has only a low convergence rate.
- Smooth manifold optimization is a natural generalization of smooth optimization on Euclidean space, unconstrained optimization on Manifolds & quadratic convergence rate possible.

[1] A rotation matrix in n-dimensions is a $n \times n$ special orthogonal matrix, that is an orthogonal matrix whose determinant is 1: $\mathbf{R}^\top = \mathbf{R}^{-1}, \det \mathbf{R} = 1$.

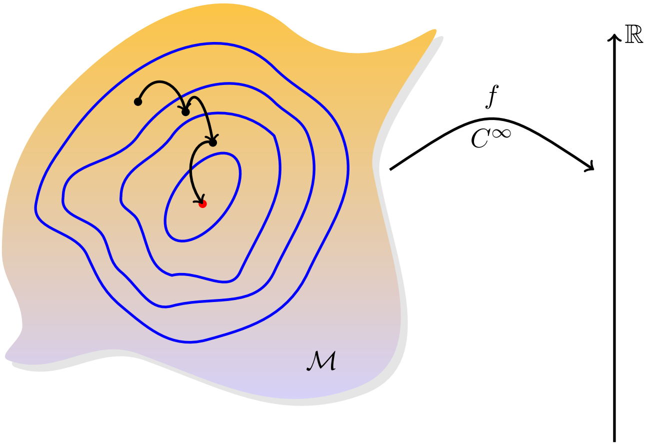
Optimization on Manifolds



Optimization on Manifolds



Optimization on Manifolds



Smooth manifold optimization is a natural generalization of smooth optimization on Euclidean space.

Generalization of the Newton Method

$$x_{k+1} = x_k - [f''(x_k)]^{-1} f'(x_k)$$

- Locally quadratic rate of convergence to a local minimum in general
- Convergence in a single iteration for quadratic functions
- Knowledge up to only second order of the function at the current point required

GENERALIZATION

- Update direction
- Metric & update path
- Gradient & Hessian

Generalization of the Newton Method

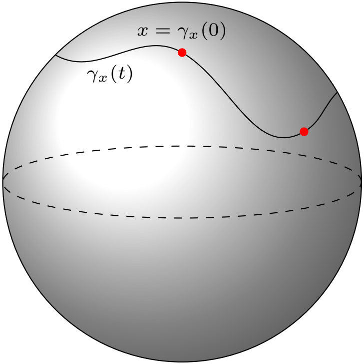
$$x_{k+1} = x_k - [f''(x_k)]^{-1} f'(x_k)$$

- Locally quadratic rate of convergence to a local minimum in general
- Convergence in a single iteration for quadratic functions
- Knowledge up to only second order of the function at the current point required

GENERALIZATION

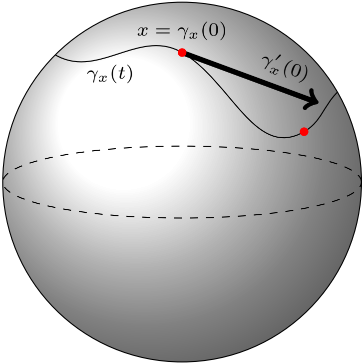
- Update direction
- Metric & update path
- Gradient & Hessian

Tangent Space



If an inner product $g_x : T_x\mathcal{M} \times T_x\mathcal{M} \to \mathbb{R}$ is defined on the tangent space $T_x\mathcal{M}$ then $\mathcal{M}$ is a Riemannian manifold.

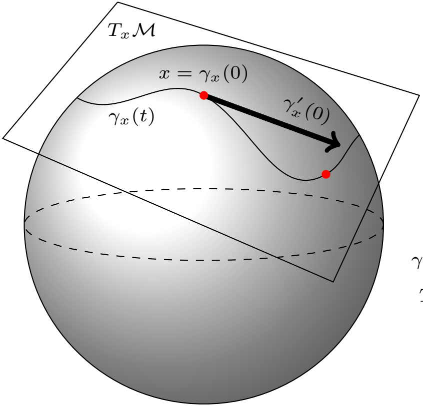
Tangent Space



$$\gamma_x'(0) : f \in C^\infty(x) \mapsto \frac{\mathrm{d}}{\mathrm{d}\,t} f(\gamma(t))|_{t=0} \in \mathbb{R}$$

If an inner product $g_x : T_x\mathcal{M} \times T_x\mathcal{M} \to \mathbb{R}$ is defined on the tangent space $T_x\mathcal{M}$ then $\mathcal{M}$ is a Riemannian manifold.

Tangent Space



$$\gamma_x'(0) : f \in C^\infty(x) \mapsto \frac{\mathrm{d}}{\mathrm{d}\,t} f(\gamma(t))|_{t=0} \in \mathbb{R}$$

$$T_x\mathcal{M} = \{\gamma'(0) : \gamma \text{ curve in } \mathcal{M}, \gamma(0) = x\}$$

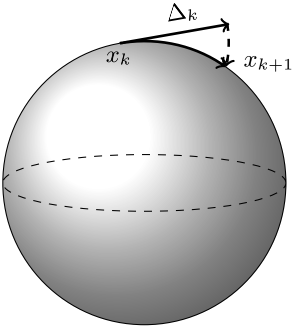
If an inner product $g_x : T_x\mathcal{M} \times T_x\mathcal{M} \to \mathbb{R}$ is defined on the tangent space $T_x\mathcal{M}$ then $\mathcal{M}$ is a Riemannian manifold.

Riemannian manifold

- In case of Riemannian manifold $f'(p)$ & $f''(p)$ in the Newton iteration can be replaced with the gradient & the Hessian.
- Computation of geodesics is possible given the metric structure.
- Hence a Newton iteration $x_{k+1} = x_k - \mathrm{d}\,x$ is done by a unit step along the geodesic in the direction $-\mathrm{d}\,x$.

Riemannian manifold

- In case of Riemannian manifold $f'(p)$ & $f''(p)$ in the Newton iteration can be replaced with the gradient & the Hessian.
- Computation of geodesics is possible given the metric structure.
- Hence a Newton iteration $x_{k+1} = x_k - \mathrm{d}\,x$ is done by a unit step along the geodesic in the direction $-\mathrm{d}\,x$.



Riemannian Newton direction $\Delta_k \in T_x\mathcal{M}$:

$$\mathrm{Hess}\, f(x_k)\Delta_k = -\,\mathrm{grad}\, f(x_k)$$

Properties

- The constraint manifold $SO(3)$ is a instance of Stiefel manifold[2] with the canonical Riemannian metric

$$g(\Delta_1, \Delta_2) = \frac{1}{2}\operatorname{tr}(\Delta_1^\top \Delta_2)$$

- We have explicit formula for the geodesic of $SO(3)$ in the direction $\Delta \in T(SO(3))$[3]

$$\mathbf{R}(t) = \exp(\mathbf{R}, \Delta t) = \mathbf{R}\exp(\hat{\omega} t) = \mathbf{R}\left(\mathbf{I} + \hat{\omega}\sin(t) + \hat{\omega}^2(1-\cos(t))\right)$$

- Gradient & Hessian

$$\mathrm{d}\,F(\Delta) = \left.\frac{\mathrm{d}\,F(\mathbf{R}(t))}{\mathrm{d}\,t}\right|_{t=0}, \operatorname{Hess} F(\Delta,\Delta) = \left.\frac{\mathrm{d}^2\,F(\mathbf{R}(t))}{\mathrm{d}\,t^2}\right|_{t=0}$$

$$\operatorname{Hess} F(\mathbf{X},\mathbf{Y}) = \frac{1}{4}\left(\operatorname{Hess} F(\mathbf{X}+\mathbf{Y},\mathbf{X}+\mathbf{Y}) - \operatorname{Hess} F(\mathbf{X}-\mathbf{Y},\mathbf{X}-\mathbf{Y})\right)$$

[2] The Stiefel manifold $V_k(\mathbb{R}^n)$ is the set of all orthonormal k-frames in $\mathbb{R}^n$ $V_k(\mathbb{R}^n) = \{\mathbf{A} \in \mathbb{R}^{n\times k} : \mathbf{A}^\top\mathbf{A} = \mathbf{I}\}$.

[3] $t \in \mathbb{R}, \hat{\omega} = \mathbf{R}^\top\Delta \in \mathfrak{so}(3)$ (the Lie algebra associated with $SO(3)$)

Properties

- The constraint manifold $SO(3)$ is a instance of Stiefel manifold[2] with the canonical Riemannian metric

$$g(\Delta_1, \Delta_2) = \frac{1}{2}\operatorname{tr}(\Delta_1^\top \Delta_2)$$

- We have explicit formula for the geodesic of $SO(3)$ in the direction $\Delta \in T(SO(3))$[3]

$$\mathbf{R}(t) = \exp(\mathbf{R}, \Delta t) = \mathbf{R}\exp(\hat{\omega} t) = \mathbf{R}\left(\mathbf{I} + \hat{\omega}\sin(t) + \hat{\omega}^2(1-\cos(t))\right)$$

- Gradient & Hessian

$$\mathrm{d}\,F(\Delta) = \left.\frac{\mathrm{d}\,F(\mathbf{R}(t))}{\mathrm{d}\,t}\right|_{t=0}, \operatorname{Hess} F(\Delta, \Delta) = \left.\frac{\mathrm{d}^2\,F(\mathbf{R}(t))}{\mathrm{d}\,t^2}\right|_{t=0}$$

$$\operatorname{Hess} F(\mathbf{X}, \mathbf{Y}) = \frac{1}{4}\left(\operatorname{Hess} F(\mathbf{X}+\mathbf{Y}, \mathbf{X}+\mathbf{Y}) - \operatorname{Hess} F(\mathbf{X}-\mathbf{Y}, \mathbf{X}-\mathbf{Y})\right)$$

[2] The Stiefel manifold $V_k(\mathbb{R}^n)$ is the set of all orthonormal k-frames in $\mathbb{R}^n$ $V_k(\mathbb{R}^n) = \{\mathbf{A} \in \mathbb{R}^{n\times k} : \mathbf{A}^\top\mathbf{A} = \mathbf{I}\}$.

[3] $t \in \mathbb{R}, \hat{\omega} = \mathbf{R}^\top\Delta \in \mathfrak{so}(3)$ (the Lie algebra associated with $SO(3)$)

Properties

- The constraint manifold $SO(3)$ is a instance of Stiefel manifold[2] with the canonical Riemannian metric

$$g(\Delta_1, \Delta_2) = \frac{1}{2}\,\mathrm{tr}(\Delta_1^\top \Delta_2)$$

- We have explicit formula for the geodesic of $SO(3)$ in the direction $\Delta \in T(SO(3))$[3]

$$\mathbf{R}(t) = \exp(\mathbf{R}, \Delta t) = \mathbf{R}\exp(\hat{\omega} t) = \mathbf{R}\left(\mathbf{I} + \hat{\omega}\sin(t) + \hat{\omega}^2(1-\cos(t))\right)$$

- Gradient & Hessian

$$\mathrm{d}\,F(\Delta) = \left.\frac{\mathrm{d}\,F(\mathbf{R}(t))}{\mathrm{d}\,t}\right|_{t=0}, \mathrm{Hess}\,F(\Delta,\Delta) = \left.\frac{\mathrm{d}^2\,F(\mathbf{R}(t))}{\mathrm{d}\,t^2}\right|_{t=0}$$

$$\mathrm{Hess}\,F(\mathbf{X},\mathbf{Y}) = \frac{1}{4}\left(\mathrm{Hess}\,F(\mathbf{X}+\mathbf{Y},\mathbf{X}+\mathbf{Y}) - \mathrm{Hess}\,F(\mathbf{X}-\mathbf{Y},\mathbf{X}-\mathbf{Y})\right)$$

[2]The Stiefel manifold $V_k(\mathbb{R}^n)$ is the set of all orthonormal k-frames in $\mathbb{R}^n$ $V_k(\mathbb{R}^n) = \left\{\mathbf{A} \in \mathbb{R}^{n\times k} : \mathbf{A}^\top\mathbf{A} = \mathbf{I}\right\}$.

[3]$t \in \mathbb{R}, \hat{\omega} = \mathbf{R}^\top\Delta \in \mathfrak{so}(3)$ (the Lie algebra associated with $SO(3)$)

Algorithm

Algorithm 1 Minimize $F(\mathbf{R}_i) = \mathbb{E}[||\mathbf{p}_i - (\mathbf{R}_i(\bar{\mathbf{s}} + \mathbf{V}\mathbf{z}_i) + \mathbf{T}_i)||_F^2]$

At the point $\mathbf{R}_i \in SO(3)$, compute the optimal updating vector $\Delta_i = -\operatorname{Hess}^{-1} G$

1: Choose basis tangent vectors $\mathbf{E}^k = \mathbf{R}_i \hat{\mathbf{e}}_k \in T(SO(3))$ with $\mathbf{e}_k$ for $1 \leq k \leq 3$ the standard basis for $\mathbb{R}^3$.
2: Compute the 3×3 matrix $\mathbf{H}_{kl} = \operatorname{Hess} F(\mathbf{E}^k, \mathbf{E}^l)$, $1 \leq k, l \leq 3$.
3: Compute the 3 dimensional vector $\mathbf{g}_k = \mathrm{d}\, F(\mathbf{E}^k)$, $1 \leq k \leq 3$.
4: Compute the vector $\omega = (\omega_1, \omega_2, \omega_3)^\top \in \mathbb{R}^3$ such that $\omega = -\mathbf{H}^{-1}\mathbf{g}$.
5: Then the optimal updating vector $\Delta_i = -\operatorname{Hess}^{-1} G = \mathbf{R}_i \hat{\omega}$.

Update the rotation $\mathbf{R}_i$

1: Move $\mathbf{R}_i$ in the direction Δ_i along the geodesic to $\exp(\mathbf{R}_i, \Delta_i t)$, where $t = \sqrt{\frac{1}{2}\operatorname{tr}(\Delta_i^\top \Delta_i)}$.

Outline

Introduction

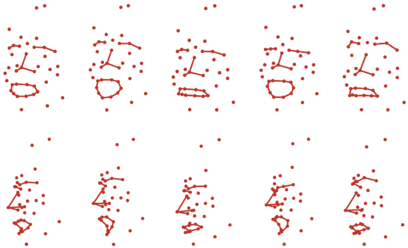
Methodology
- Probabilistic PCA
- Probabilistic Relational PCA
- PPCA Optimization on Manifolds

Experiments

Conclusion

Datasets

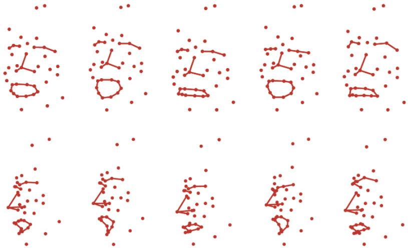
Vicon face data [1]



- 316 frames, 40 points
- Same person
- Captured using markers
- Few noises
- Few pose changes

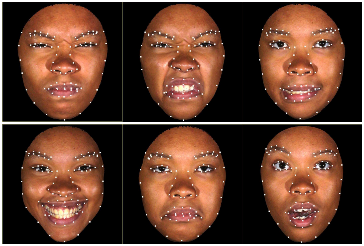
Datasets

Vicon face data [1]



- 316 frames, 40 points
- Same person
- Captured using markers
- Few noises
- Few pose changes

BU-3DFE [4]



- 300 frames, 83 points
- Different persons
- Labeled by hand
- A lot of noises
- Zero-mean Gaussian random pan & tilt added

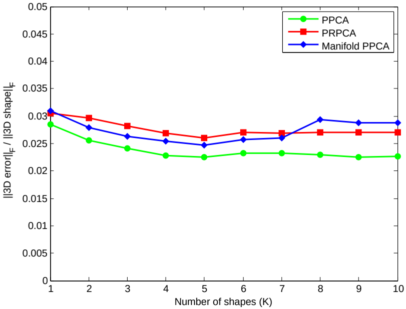
Karlsruhe Institute of Technology

Results

- 50 EM iterations
- 0% noise

Results

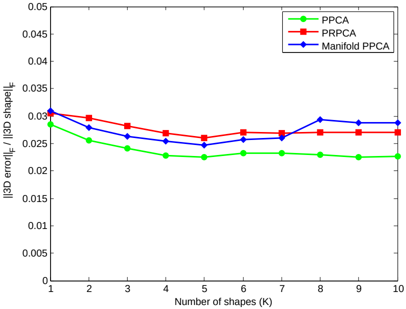
Vicon face data



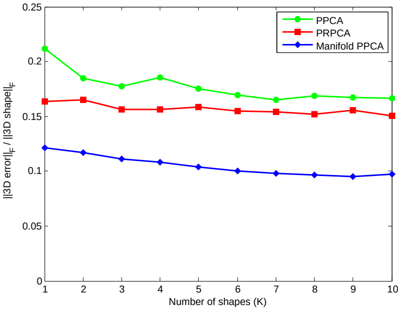
- 50 EM iterations
- 0% noise

Results

Vicon face data



BU-3DFE



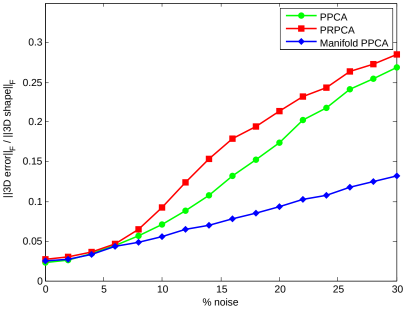
- 50 EM iterations
- 0% noise

Results

- 50 EM iterations
- Up to 30% additive Gaussian noise $\frac{||\text{noise}||_F}{||\text{measurement}||_F}$
- Averaged over 10 runs

Results

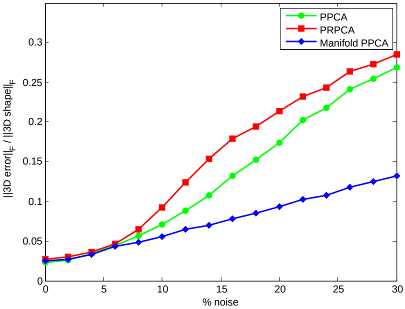
Vicon face data



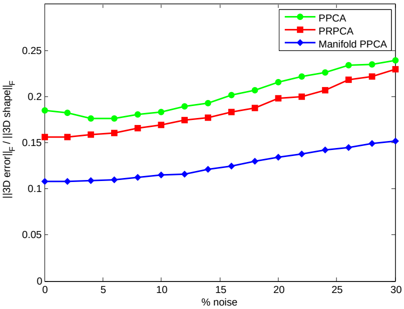
- 50 EM iterations
- Up to 30% additive Gaussian noise $\frac{||\text{noise}||_F}{||\text{measurement}||_F}$
- Averaged over 10 runs

Results

Vicon face data



BU-3DFE



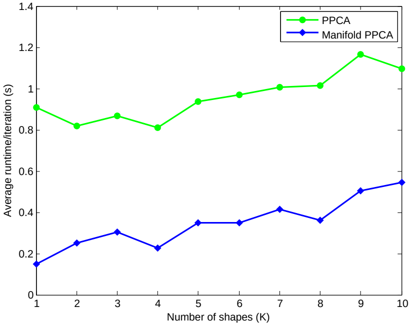
- 50 EM iterations
- Up to 30% additive Gaussian noise $\frac{||\text{noise}||_F}{||\text{measurement}||_F}$
- Averaged over 10 runs

Runtime

- Averaged over 50 EM iterations

Runtime

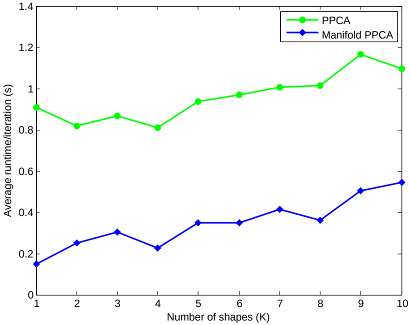
Vicon face data



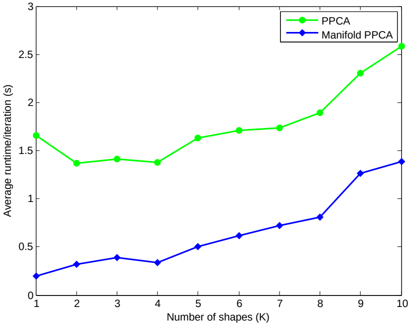
- Averaged over 50 EM iterations

Runtime

Vicon face data



BU-3DFE



- Averaged over 50 EM iterations

Outline

Conclusion

- Statistical NRSFM framework
- PRPCA for shape recovery, slight improvement on BU-3DFE
- Riemannian Newton method for motion recovery, significant improvement on both datasets
  - Orthonormality constraint natively solved by its geometric properties rather than numerical optimizations
  - Better performance under noise
  - Higher computational efficiency

Future Work

- More extensive experiments
- Appearance model, texture...

Conclusion

- Statistical NRSFM framework
- PRPCA for shape recovery, slight improvement on BU-3DFE
- Riemannian Newton method for motion recovery, significant improvement on both datasets
  - Orthonormality constraint natively solved by its geometric properties rather than numerical optimizations
  - Better performance under noise
  - Higher computational efficiency

FUTURE WORK

- More extensive experiments
- Appearance model, texture…

That's All...

Thank you for your attention!
Questions? Suggestions?

Literature

- L. Torresani, A. Hertzmann and C. Bregler, "*Nonrigid structure-from-motion: Estimating shape and motion with hierarchical priors*," IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 30, no. 5, 2008.
- W.-J. Li, D-Y. Yeung and Z. Zhang, "*Probabilistic relational PCA*," Advances in Neural Information Processing Systems 23, MIT Press, 2009.
- Y. Ma, J. Košecká and S. Sastry, "*Optimization criteria and geometric algorithms for motion and structure estimation*," Int'l J. Computer Vision, vol. 44, no. 3, pp. 219-249, 2001.
- L. Yin, X. Wei, Y. Sun, J. Wang, and M. Rosato, "*A 3D facial expression database for facial behavior research*," The 7th International Conference on Automatic Face and Gesture Recognition, pp. 211-216, 2006.