RECONSTRUCTING A 3D TRAJECTORY UNDER PERSPECTIVE PROJECTION

Park H.S.*, Shiratori T.†, Matthews I.†, Sheikh Y.*

*Carnegie Mellon University, †Disney Research, Pittsburgh

{hyunsoop,yaser}@cs.cmu.edu, {shiratori,iainm}@disneyresearch.com

Carnegie Mellon

Disney Research Pittsburgh

Abstract

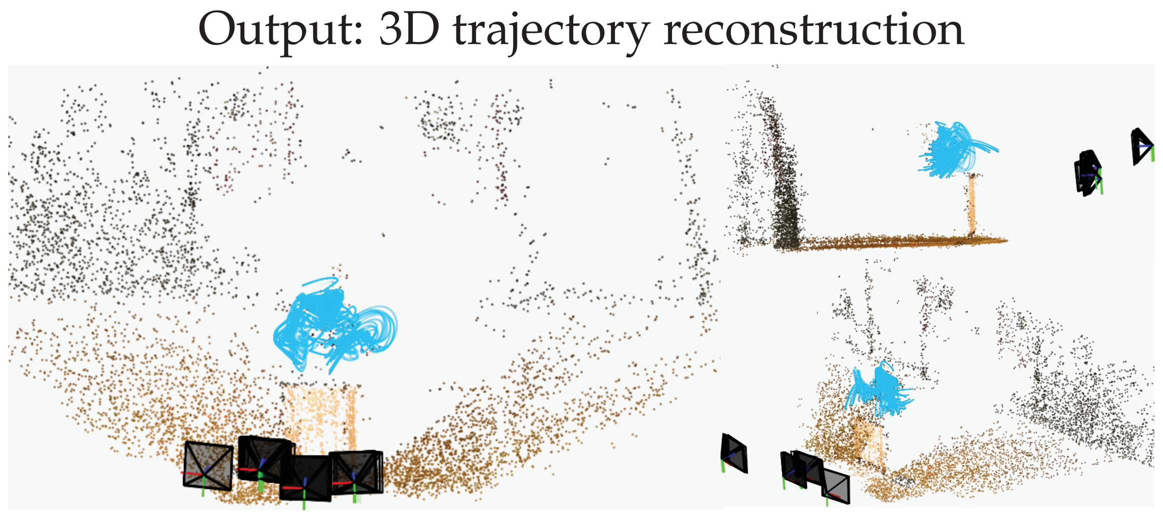
We present an algorithm for reconstructing the 3D trajectory of a moving point from its correspondence in 2D images, given the 3D pose and time of capture of the cameras that produced each image. We solve for the trajectory parameters using linear least squares followed by nonlinear optimization and study a geometric analysis of the problem. This enables us to reconstruct 3D motion from videos or images and to characterize the cases when accurate reconstruction is possible.

Goal

To reconstruct a 3D trajectory given camera poses and correspondences.

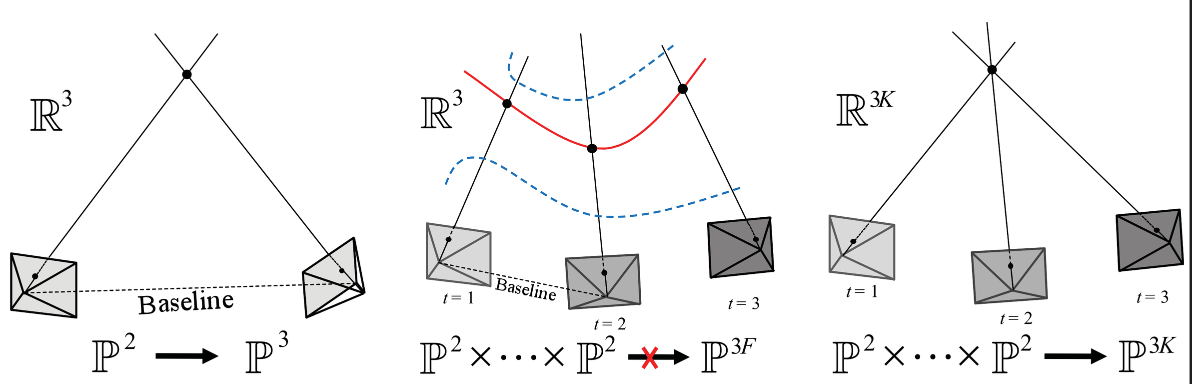
Input: correspondences across an image collection

Output: 3D trajectory reconstruction



Intuition

Unlike static point reconstruction, triangulation cannot be applied to a moving point in $\mathbb{R}^3$ because multiple views of the point may not exist at each instant in time. Instead, we triangulate a trajectory in $\mathbb{R}^{3K}$ where K is the number of trajectory basis such that $3K \leq 2F$ (F is the number of frames).



References

[1] H.S. Park, T. Shiratori, I. Matthews, Y. Sheikh, 3D Reconstruction of a Moving Point from a Series of 2D Projections, in *ECCV*, 2010

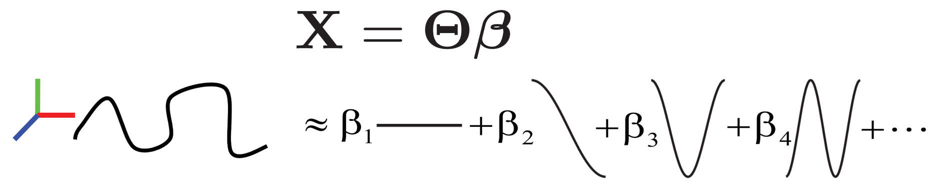
[2] Code and Video: `http://www.andrew.cmu.edu/user/hyunsoop/eccv2010/eccv_project_page.html`

[3] H.S. Park, T. Shiratori, I. Matthews, Y. Sheikh, Reconstructing a 3D Trajectory under Perspective Projection, submitted to PAMI

Method

We parameterize a point trajectory, $\mathbf{X}$, using a linear combination of a predefined trajectory basis, $\mathbf{\Theta}$, where each trajectory basis is a smooth trajectory.

$$\mathbf{X} = \mathbf{\Theta}\boldsymbol{\beta} \quad (1)$$



For each perspective projection, a linear equation can be obtained using direct linear transform given a 2D projection, $\mathbf{x}$, and a camera projection matrix, $\mathbf{P}$.

$$\mathbf{x} \propto \mathbf{P}X, \text{ or } [\mathbf{x}]_\times \mathbf{P}X = \mathbf{0}$$

For all projections,

$$\mathbf{QX} = \mathbf{q}, \quad (2)$$

where $\mathbf{Q}$ and $\mathbf{q}$ are a matrix and a vector encoded 2D projections and camera matrices, respectively and $\mathbf{X}$ is a 3D point trajectory.
By combining Eq. (1) and (2), we can solve for an unknown trajectory parameter, $\boldsymbol{\beta}$,

$$\mathbf{Q\Theta}\boldsymbol{\beta} = \mathbf{q}. \quad (3)$$

Eq. (3) is a least squares system if $3K \leq 2F$.
We refine the linearly reconstructed trajectory by minimizing reprojection error, *i.e.*,

$$\min_{\boldsymbol{\beta}} \sum_{i=1}^{F} d(\widehat{\mathbf{x}}_i, \mathbf{x}_i)^2,$$

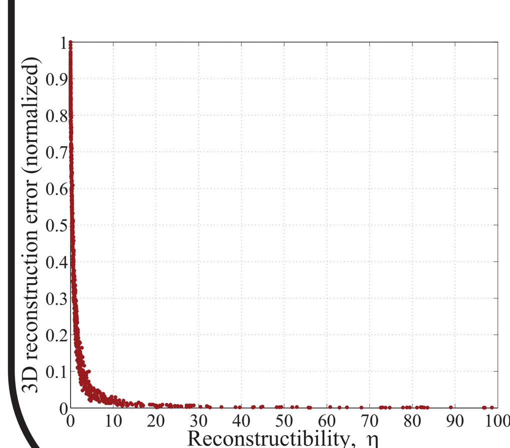
where $\widehat{\mathbf{x}}_i$ is reprojection of the 3D point at the ith frame and $d(\cdot,\cdot)$ is Euclidean distance between two arguments.

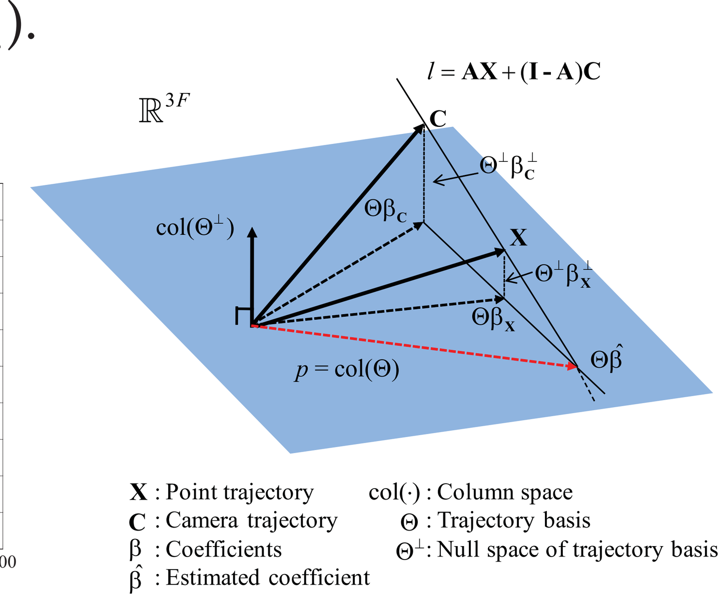
Reconstructibility

We show that 3D trajectory reconstruction is fundamentally limited by the correlation between the 3D trajectory of a point and the 3D trajectory of the camera center. **Reconstructibility** which captures the relation between them characterizes the cases when reconstruction is possible. Accurate point trajectory reconstruction is achievable when the point trajectory can be spanned by the trajectory basis while the camera trajectory cannot.

$$\eta = \frac{\text{Smoothness of a point trajectory}}{\text{Smoothness of a camera trajectory}} = \frac{\|\mathbf{\Theta}^\perp \boldsymbol{\beta}_{\mathbf{C}}^\perp\|}{\|\mathbf{\Theta}^\perp \boldsymbol{\beta}_{\mathbf{X}}^\perp\|}$$

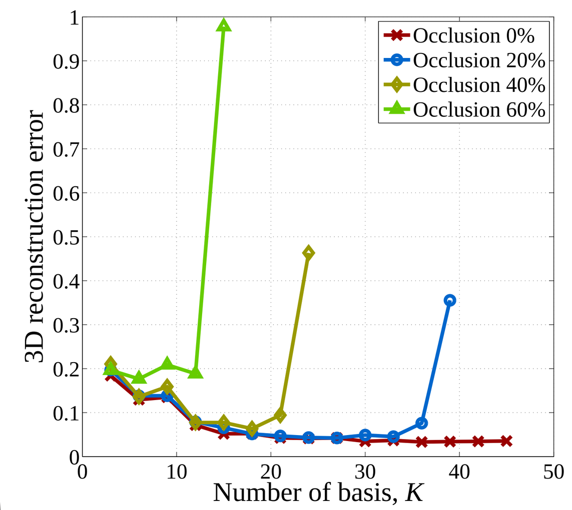
As $\eta \to \infty$, a solution approaches to the ground truth (*i.e.*, $\widehat{\boldsymbol{\beta}} \to \boldsymbol{\beta}_{\mathbf{X}}$).

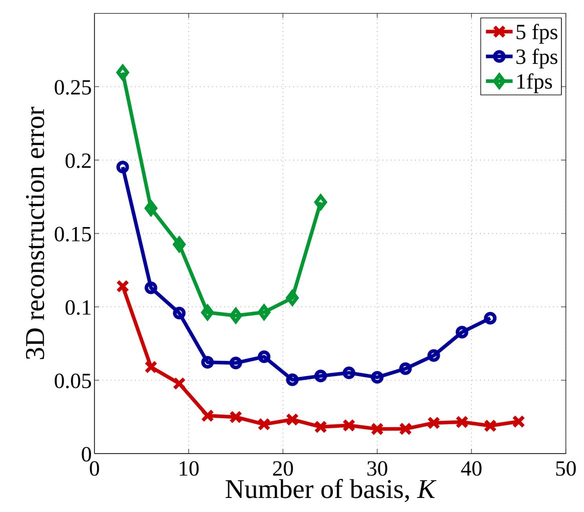




Quantitative Result

We test for effects of missing data and low frame rate with high reconstructibility from walking motion capture data. As long as the linear system of equations is overconstrained, our solution is robust to moderate occlusion and low frame rate.



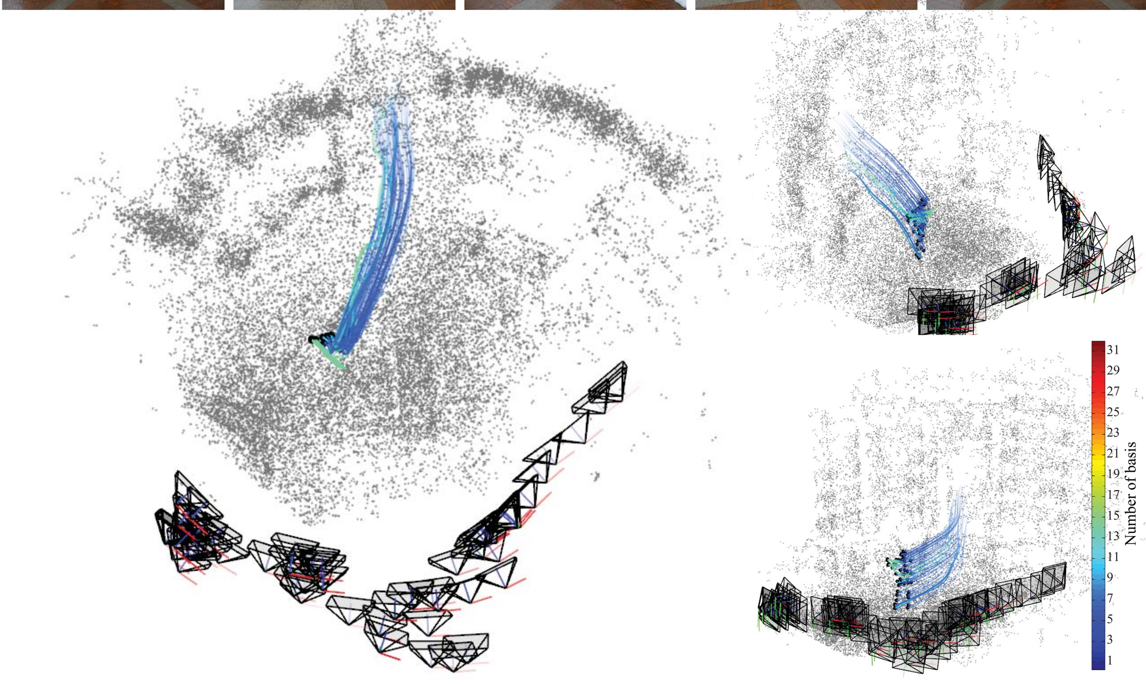


Occlusion — Frame rate

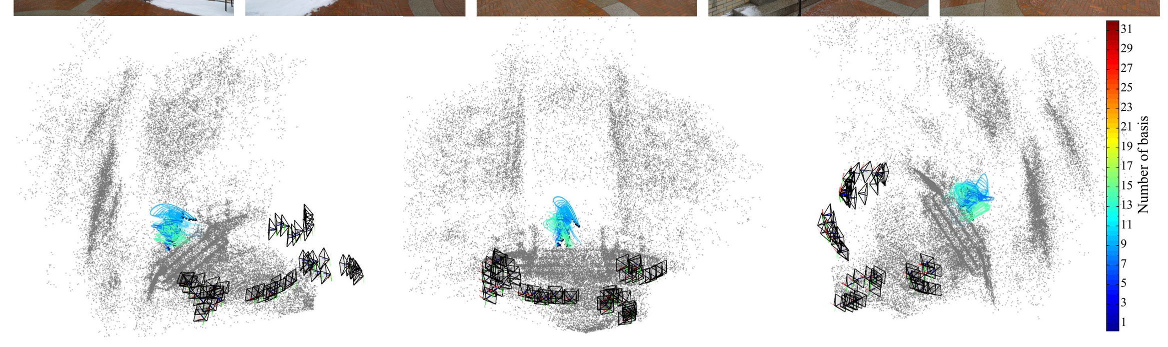
Results on Real Images

We use the DCT trajectory basis and select the number of basis, K, using a cross validation scheme. The results are best seen in video [2].

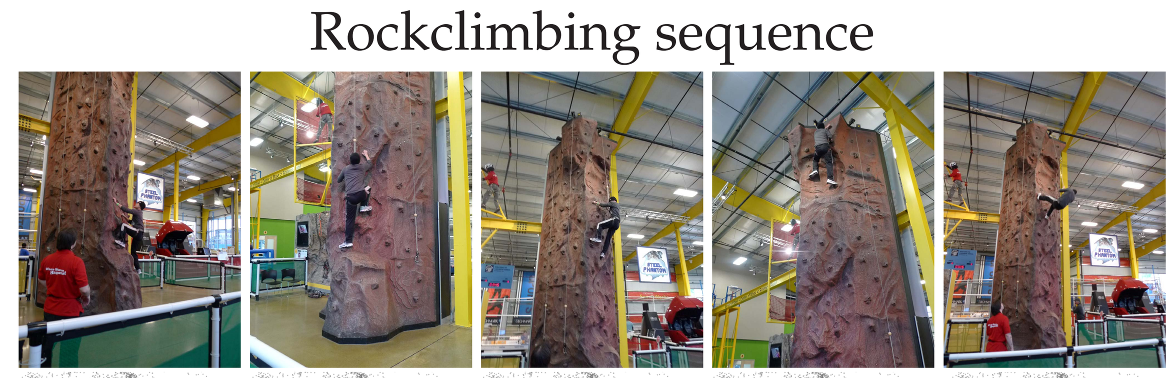
Greeting sequence

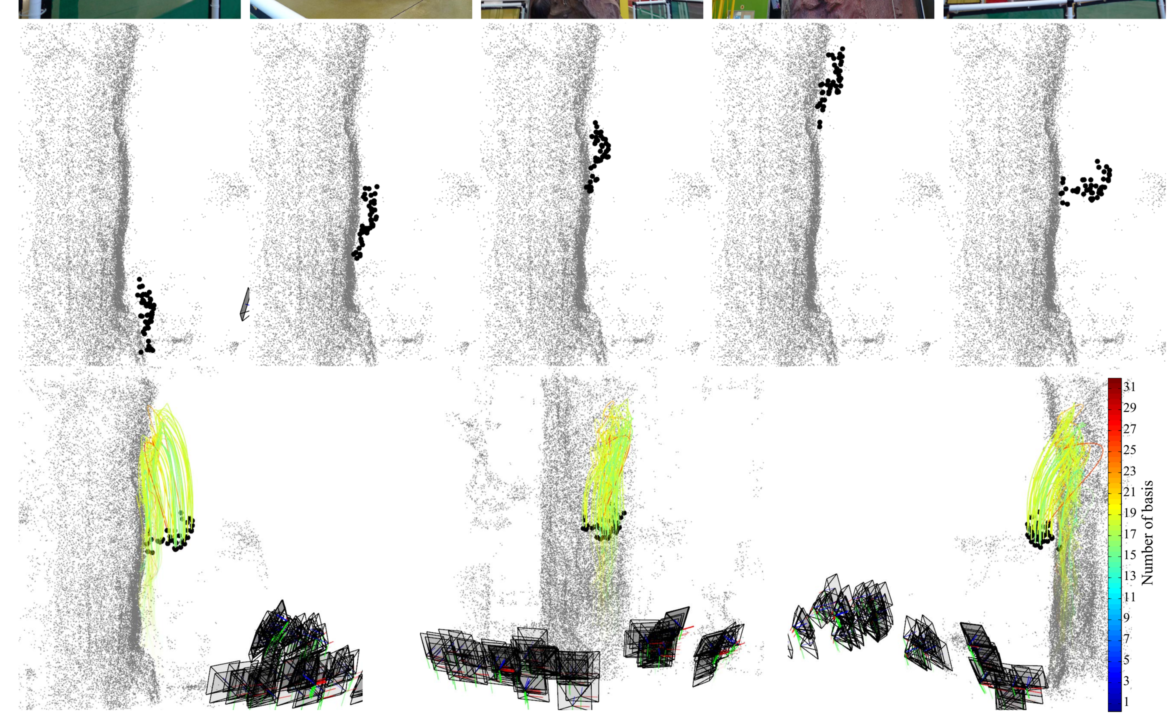


Dance sequence



Rockclimbing sequence





All codes and data are available online [2].