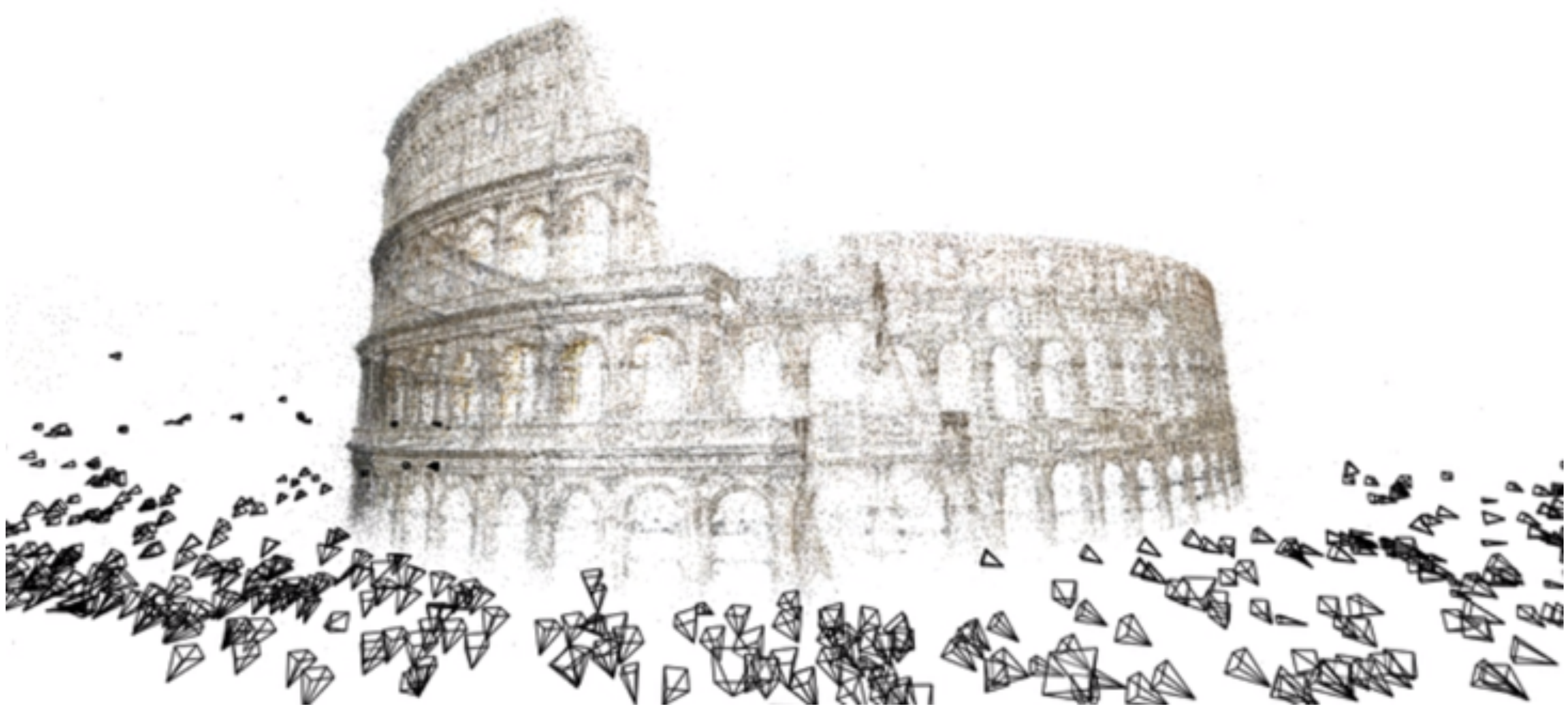
# MAP Visibility Estimation for Large-Scale Dynamic 3D Reconstruction

Hanbyul Joo   Hyun Soo Park   Yaser Sheikh

Carnegie Mellon University

# Large-scale 3D Reconstruction

## Utilizing a Large Number of Images



Dense

Accurate

Covering large area

[Snavely et al., SIGGRAPH 2006]
[Agarwal et al., ICCV 2009]

# Large-scale 3D **Event** Reconstruction

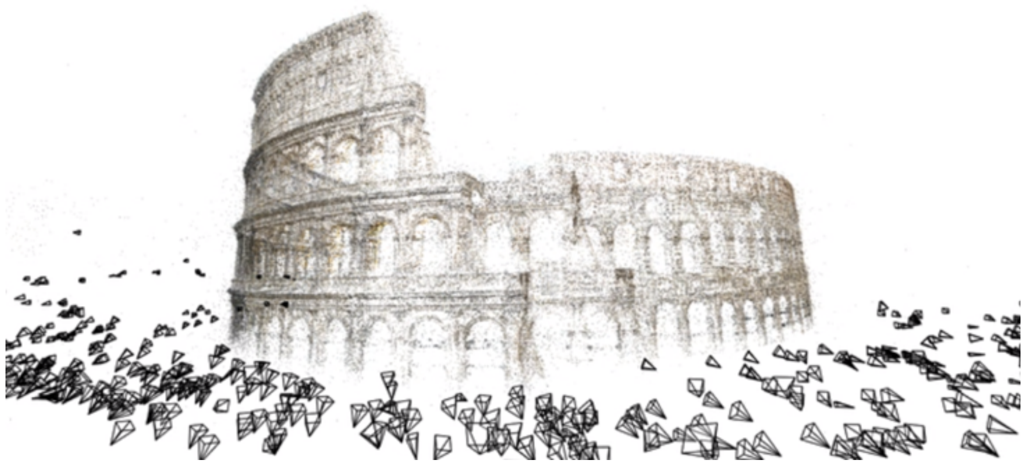
## New Opportunity from a Large Number of Videos

What can we reconstruct in dynamic scenes?

# Large-scale Dynamic **Event** Reconstruction

## What to reconstruct

**Static Scene**



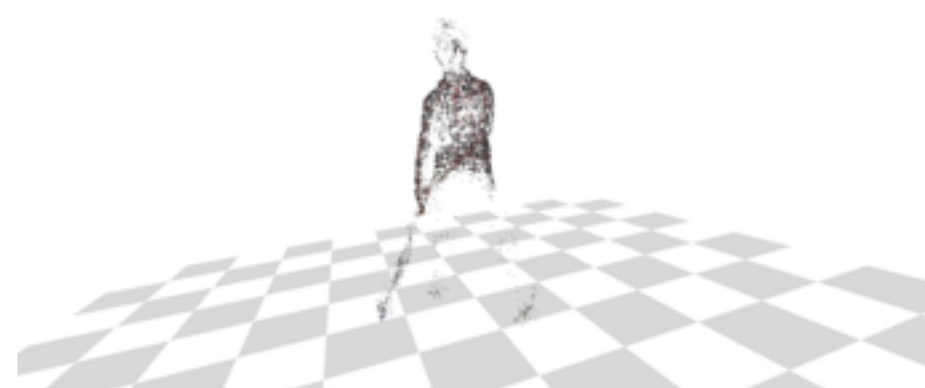
3D Point cloud
(3D shape)

Dense
Accurate
Covering Large area

**Dynamic Scene**

Dense
Accurate
Long-term

VS

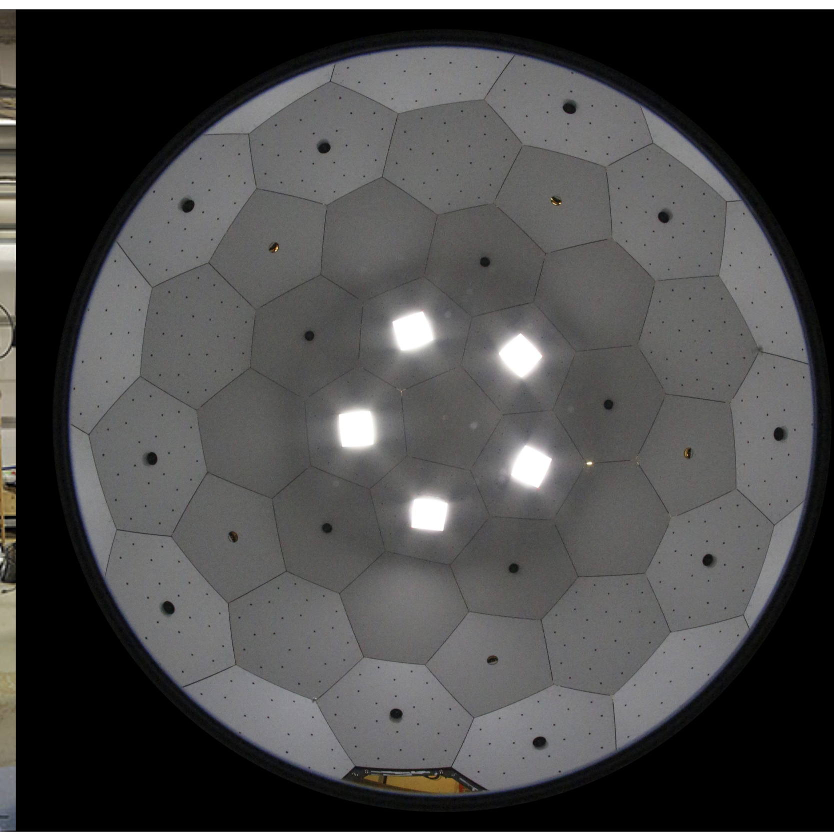


**Trajectory stream**
(3D shape + **3D motion**)

# CMU Panoptic Studio

## A System to Simulate Crowd Capture Videos

Geodesic Dome Exterior



Spherical Image (Interior)

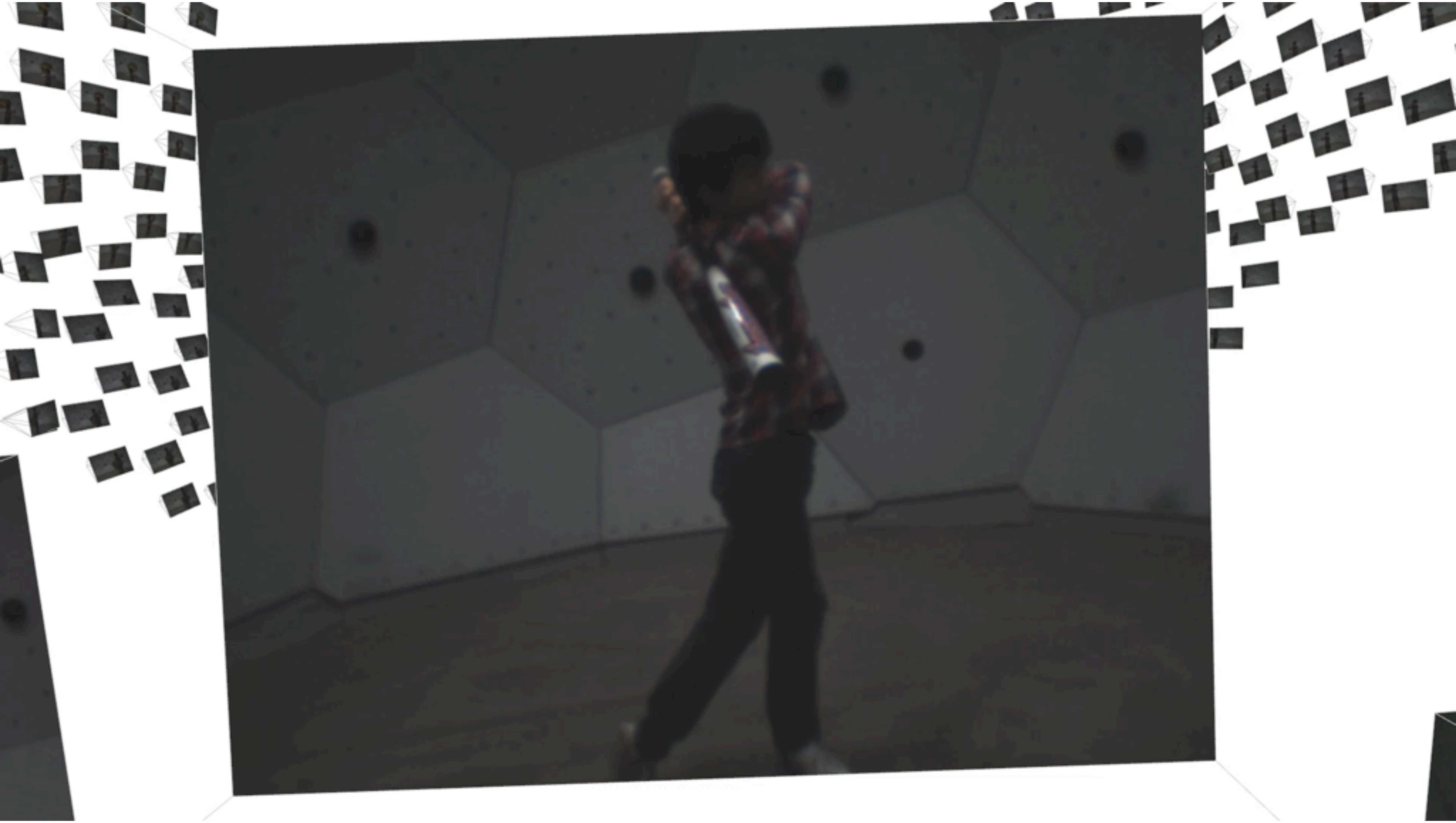
Looking in

# Input for Dynamic Event Reconstruction

## An Example View

# Input for Dynamic Event Reconstruction

480 Unique Viewpoints

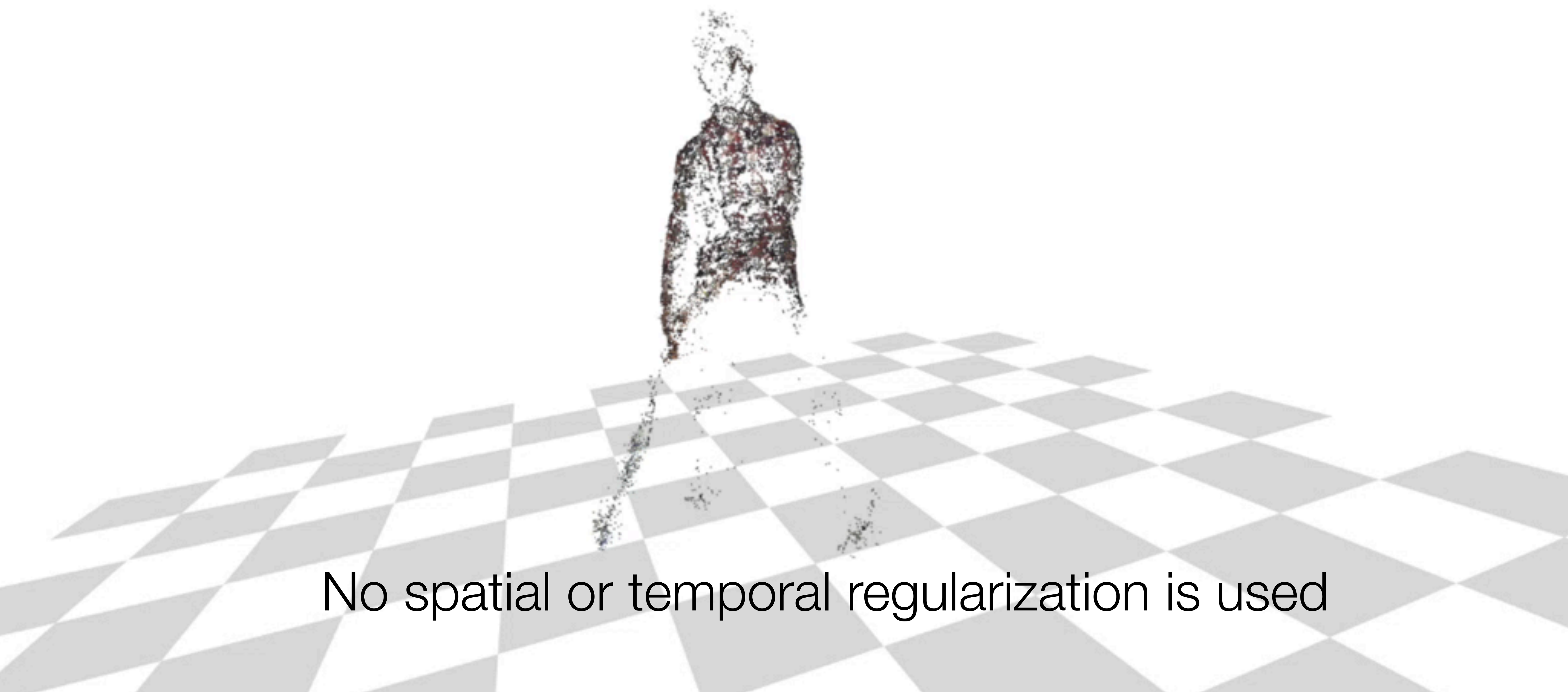


# Input for Dynamic Event Reconstruction

## All 480 Input Videos

# Large-scale Dynamic 3D Reconstruction

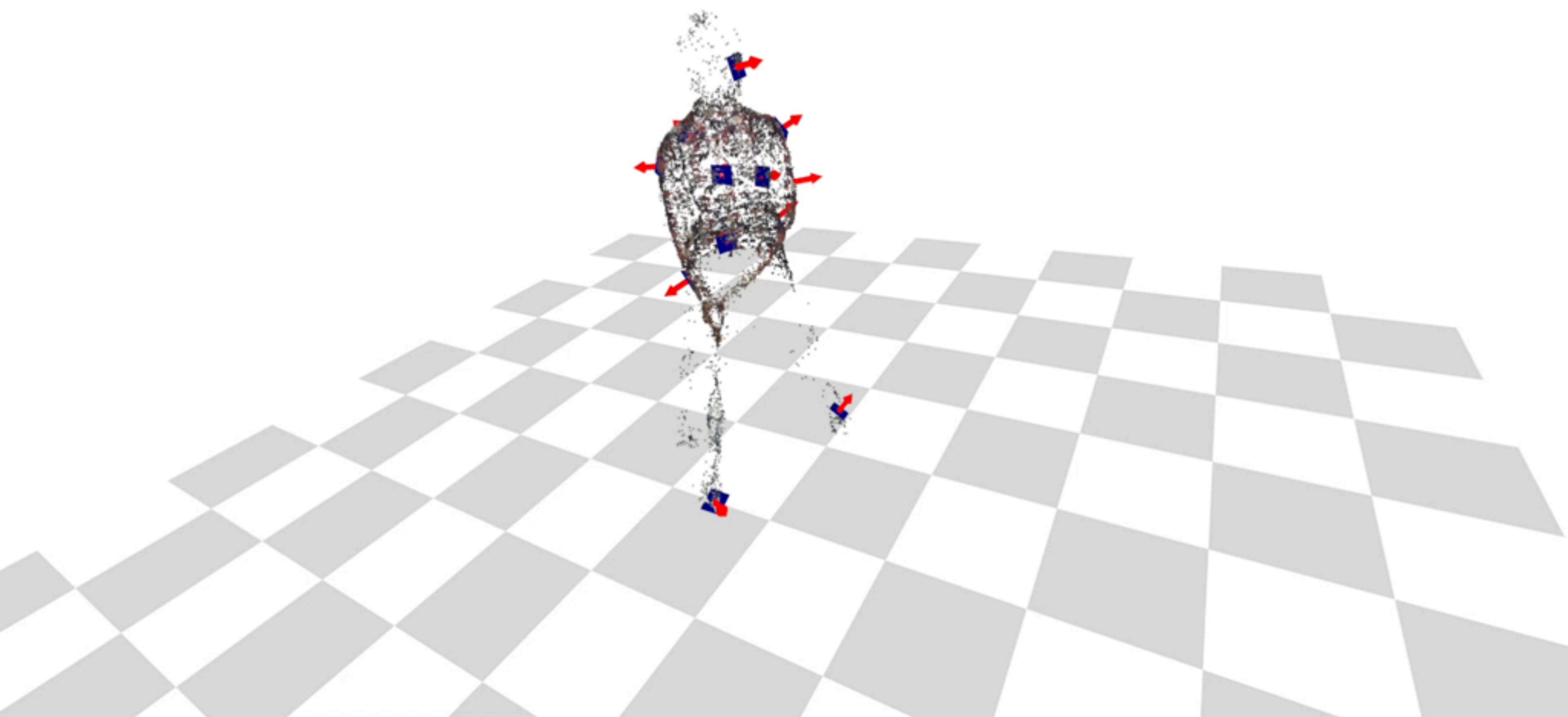
## 100,000 Trajectories over Hundreds of Frames



No spatial or temporal regularization is used

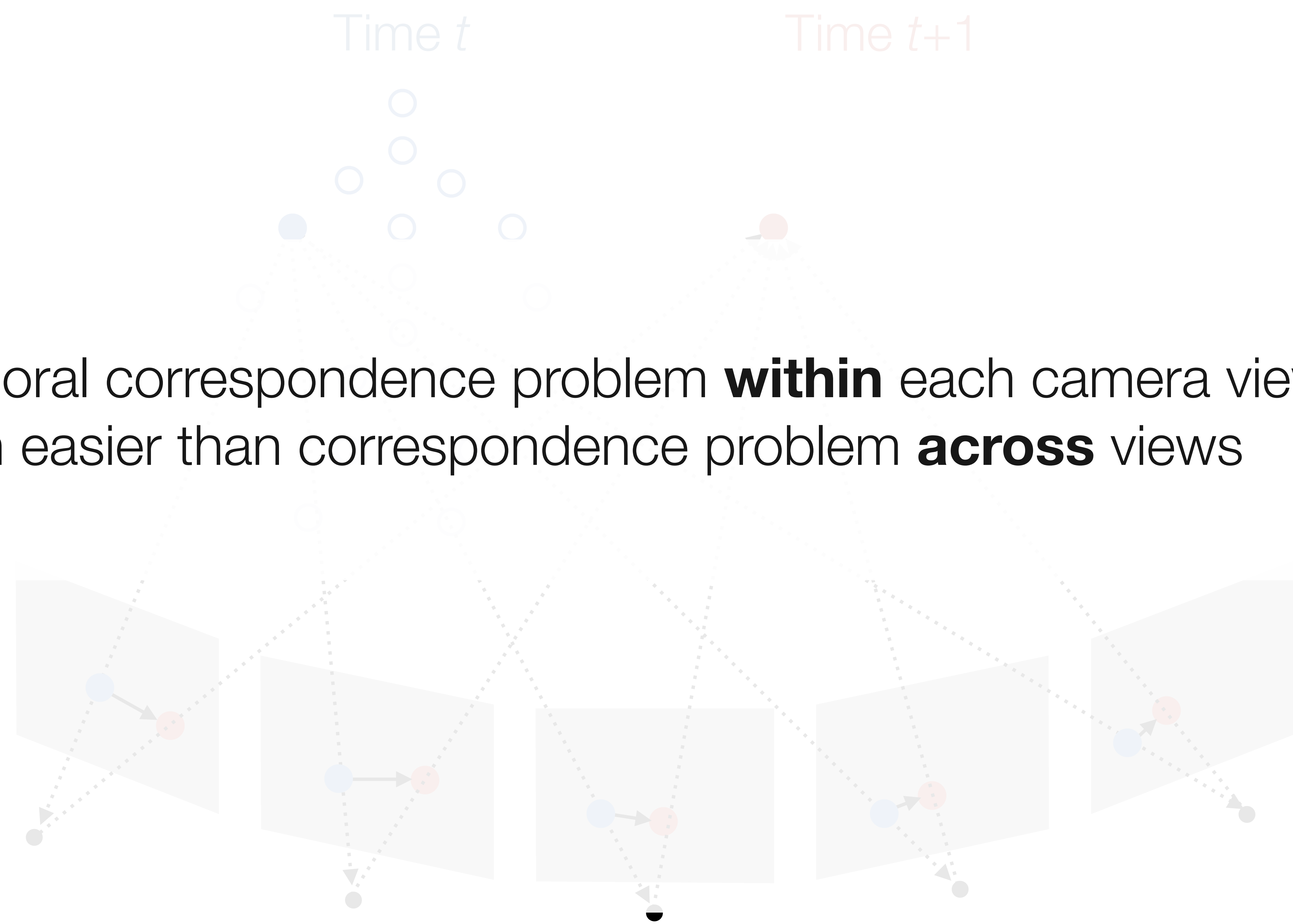
# Large-scale Dynamic 3D Reconstruction

## A Detailed View of Selected Patches



# Reconstructing 3D Trajectory

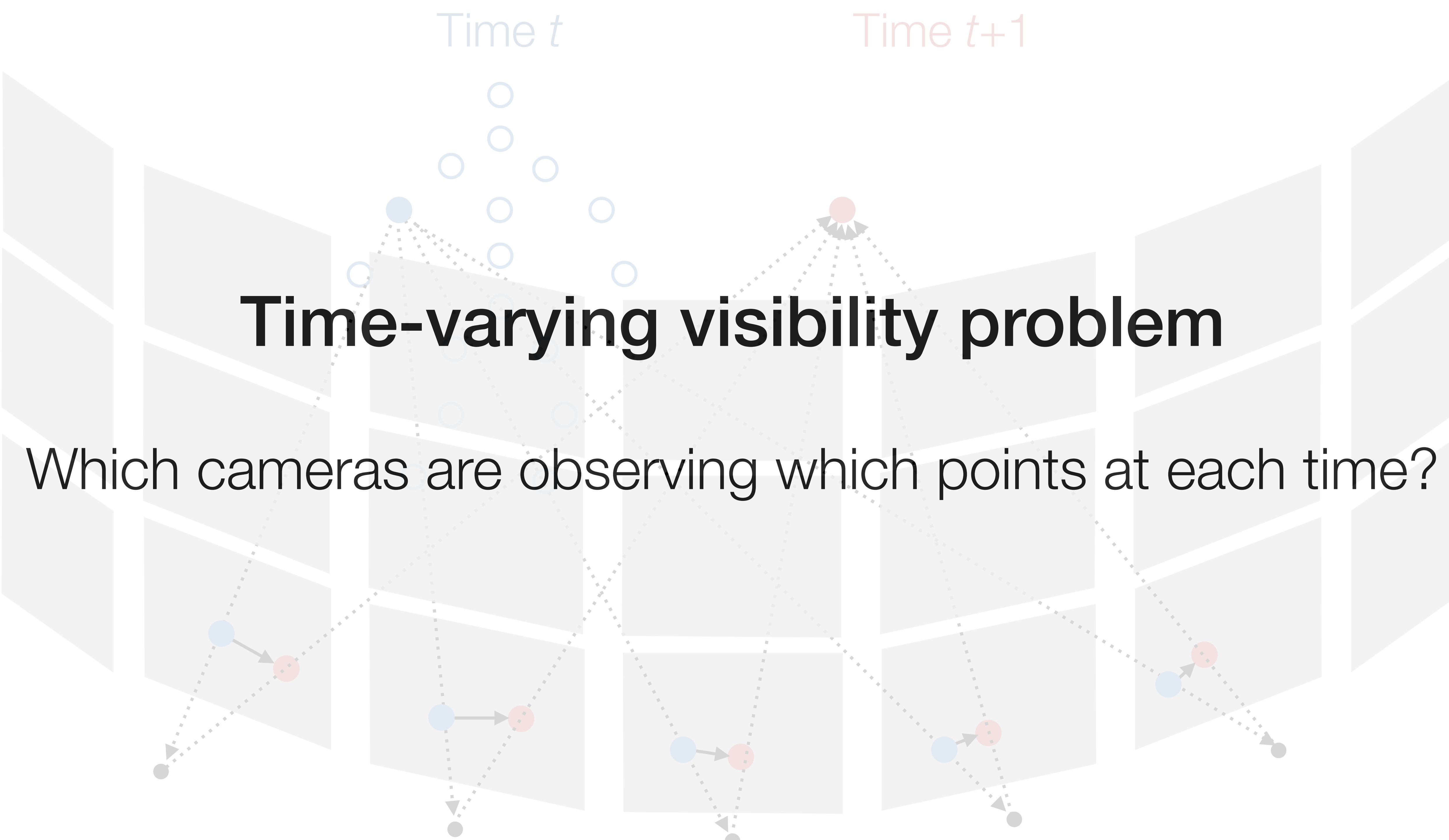
## 2D Flow-based Method




Temporal correspondence problem **within** each camera view is much easier than correspondence problem **across** views

# Reconstructing 3D Trajectory

## Key Issue To Leverage a Large Number of Views




**Time-varying visibility problem**

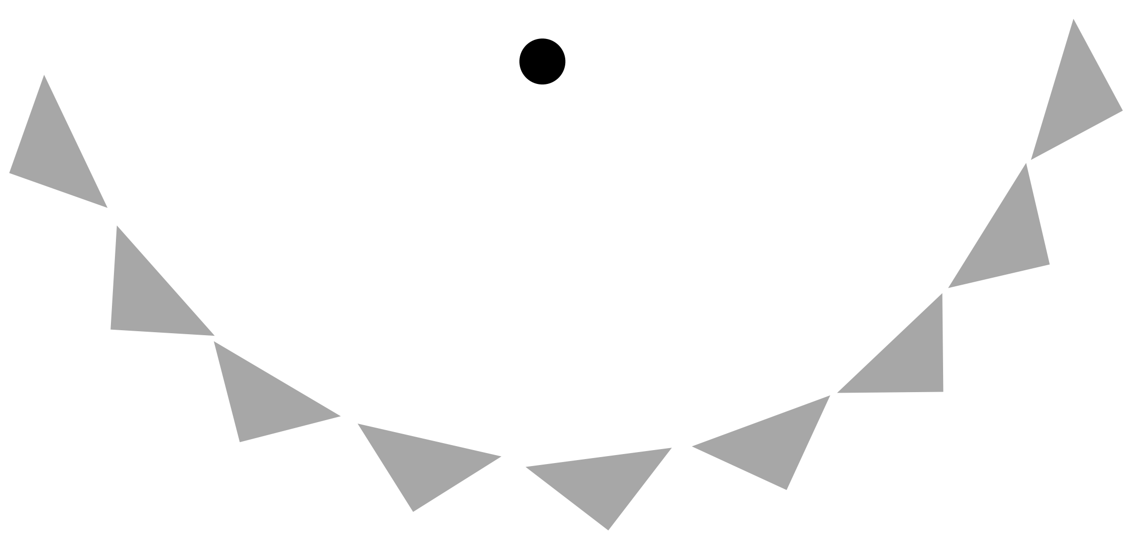
Which cameras are observing which points at each time?

# Time-varying Visibility Reasoning

## Why Is It Important in Dynamic 3D Reconstruction?

**Static Scene**

Point cloud reconstruction

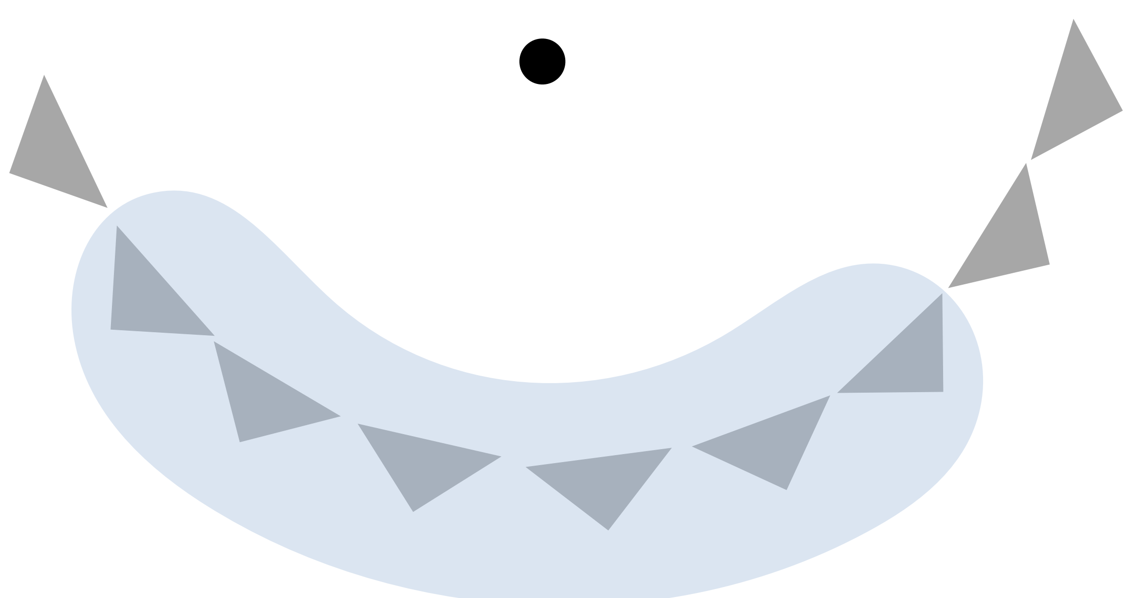


# Time-varying Visibility Reasoning

## Why Is It Important in Dynamic 3D Reconstruction?

**Static Scene**

Point cloud reconstruction

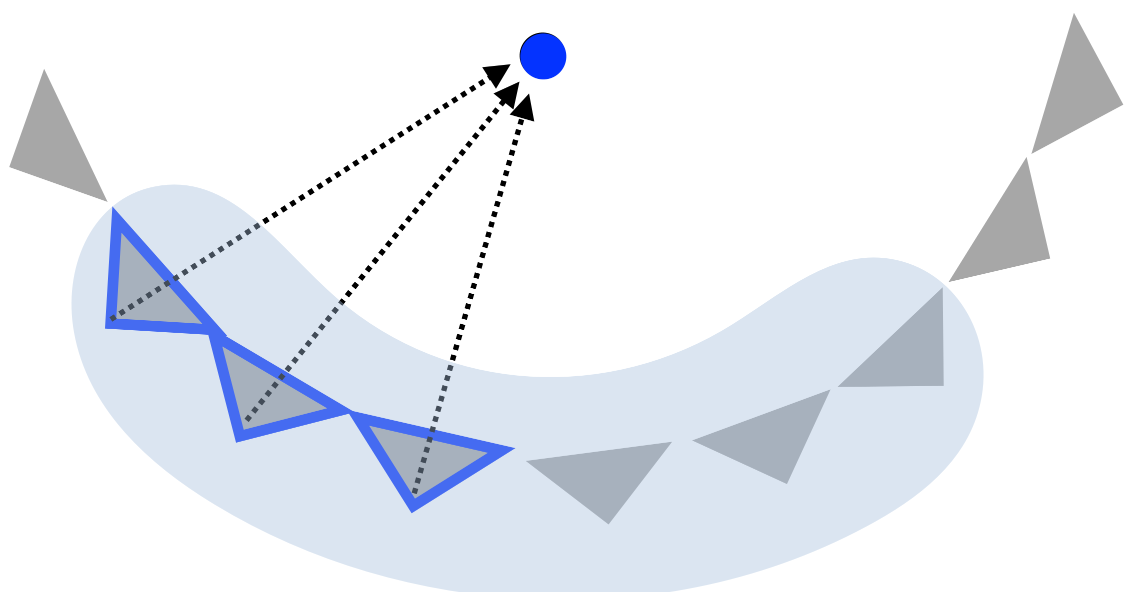


# Time-varying Visibility Reasoning

## Why Is It Important in Dynamic 3D Reconstruction?

**Static Scene**

Point cloud reconstruction



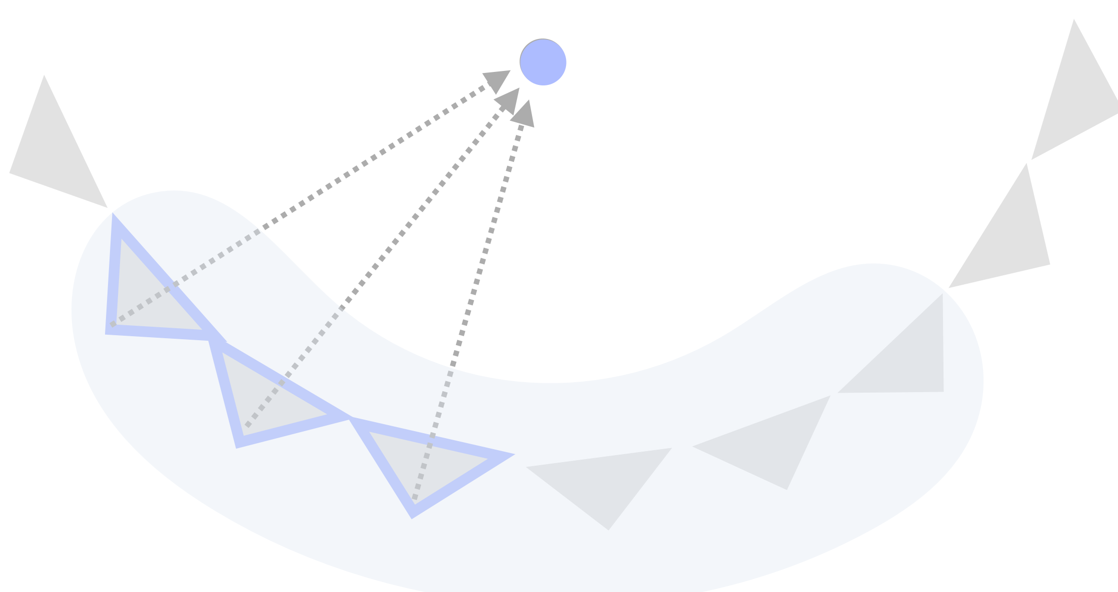
Error in visibility reasoning

# Time-varying Visibility Reasoning

## Why Is It Important in Dynamic 3D Reconstruction?

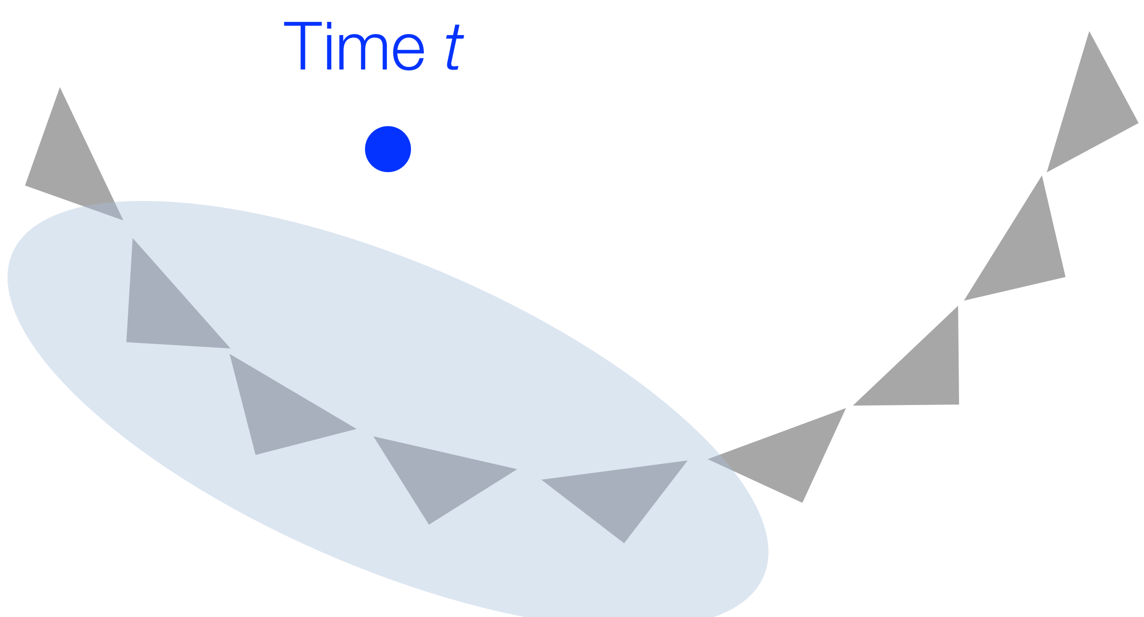
**Static Scene**

Point cloud reconstruction



**Dynamic Scene**

Trajectory stream reconstruction

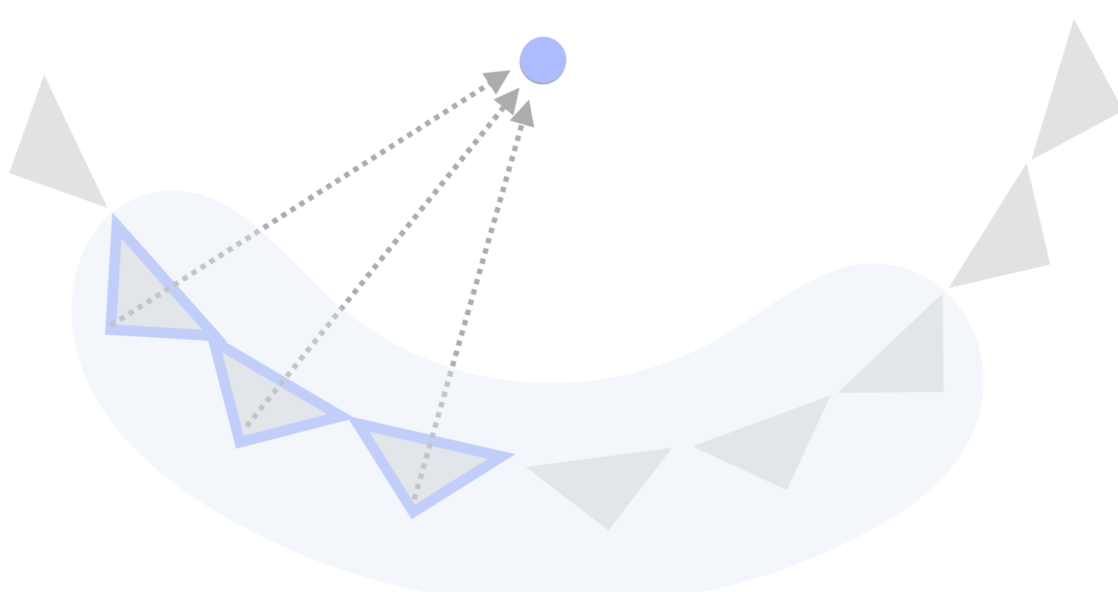


# Time-varying Visibility Reasoning

## Why Is It Important in Dynamic 3D Reconstruction?

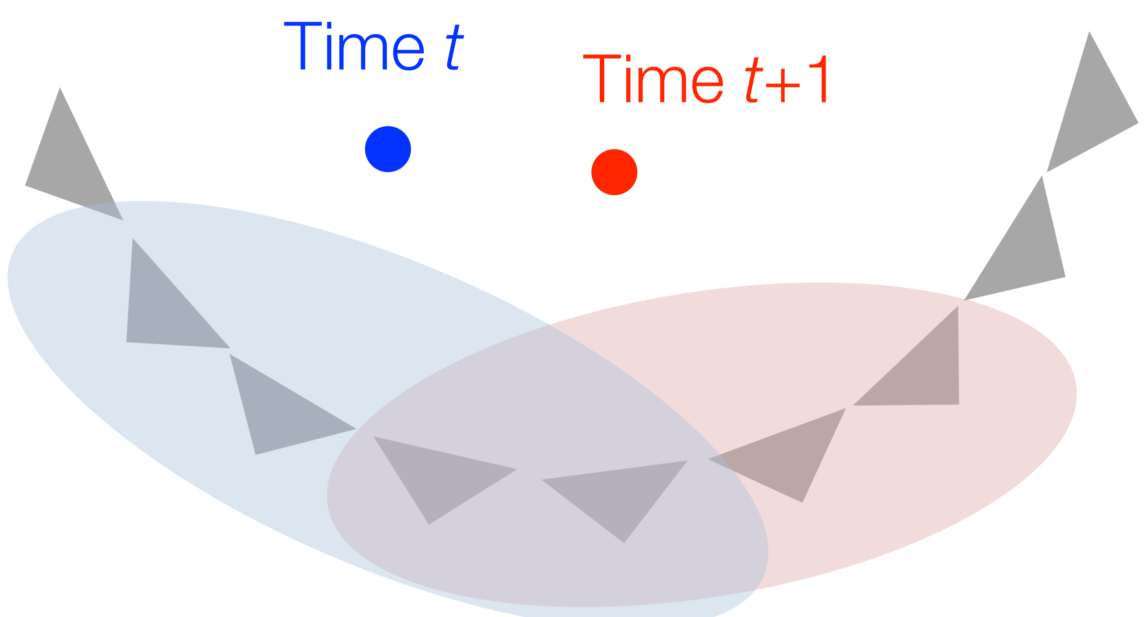
**Static Scene**

Point cloud reconstruction



**Dynamic Scene**

Trajectory stream reconstruction

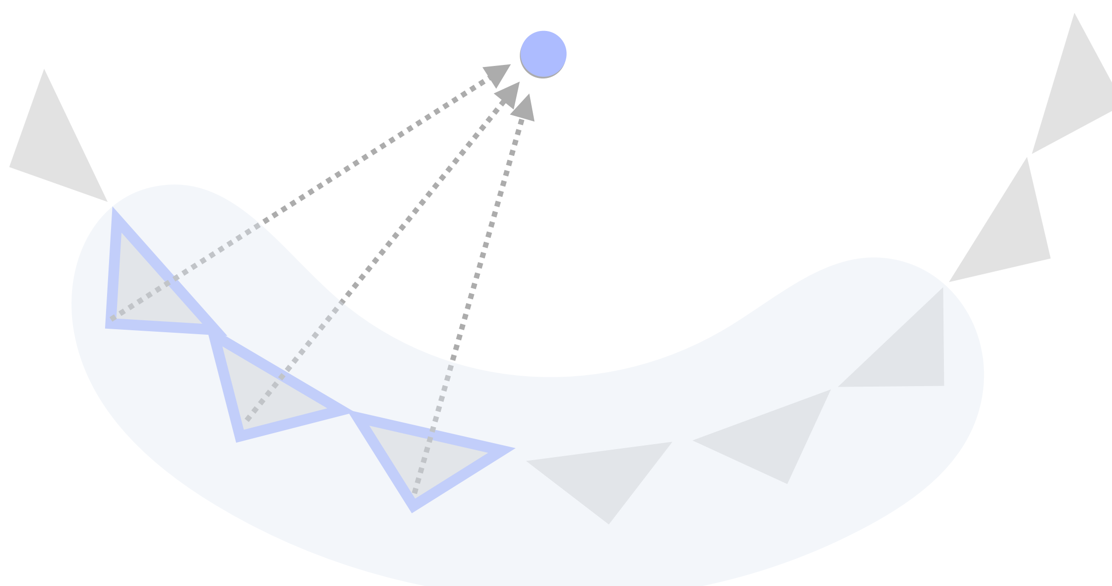


# Time-varying Visibility Reasoning

## Why Is It Important in Dynamic 3D Reconstruction?

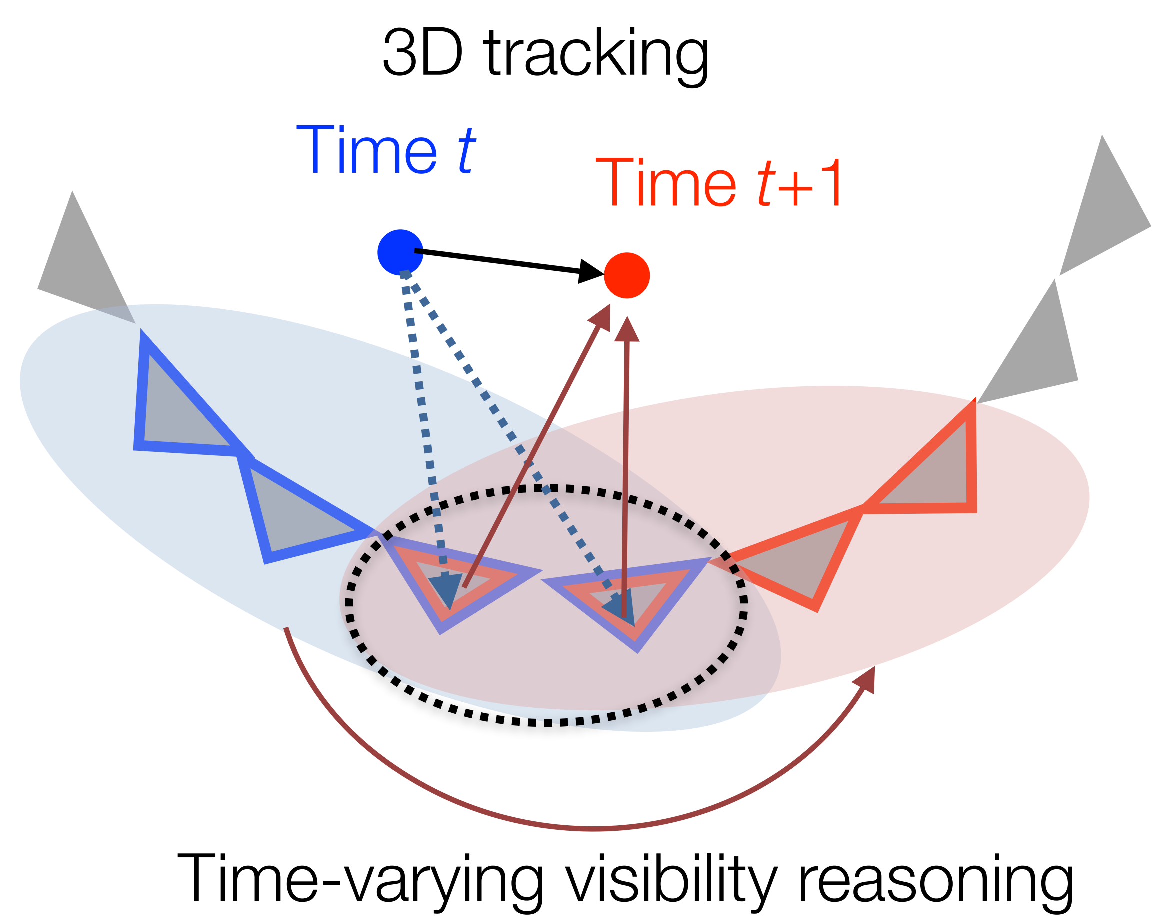
**Static Scene**

Point cloud reconstruction



**Dynamic Scene**

Trajectory stream reconstruction

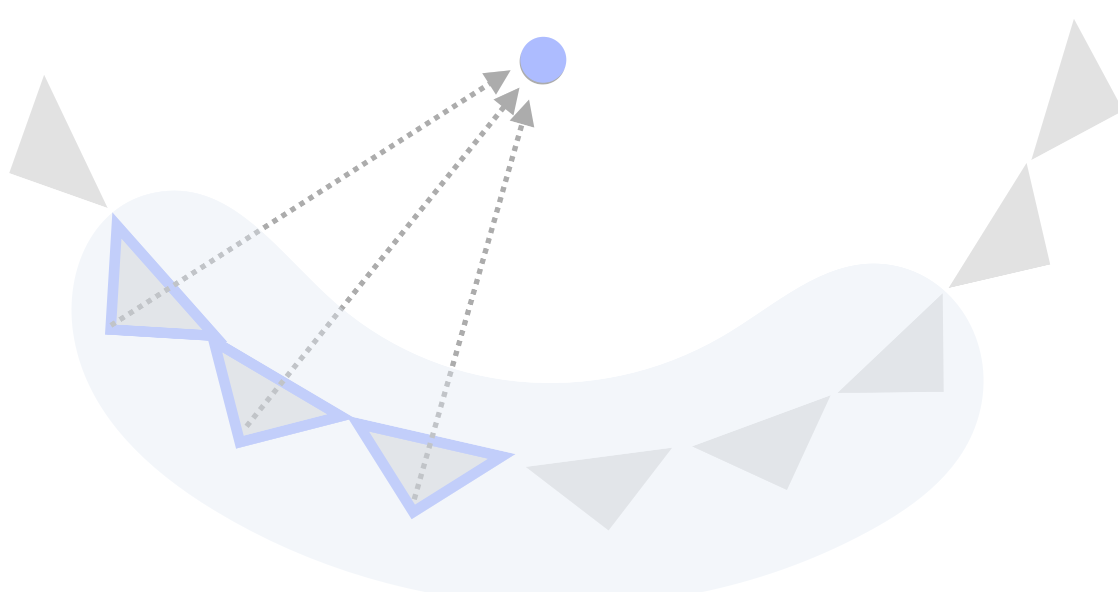


# Time-varying Visibility Reasoning

## Why Is It Important in Dynamic 3D Reconstruction?

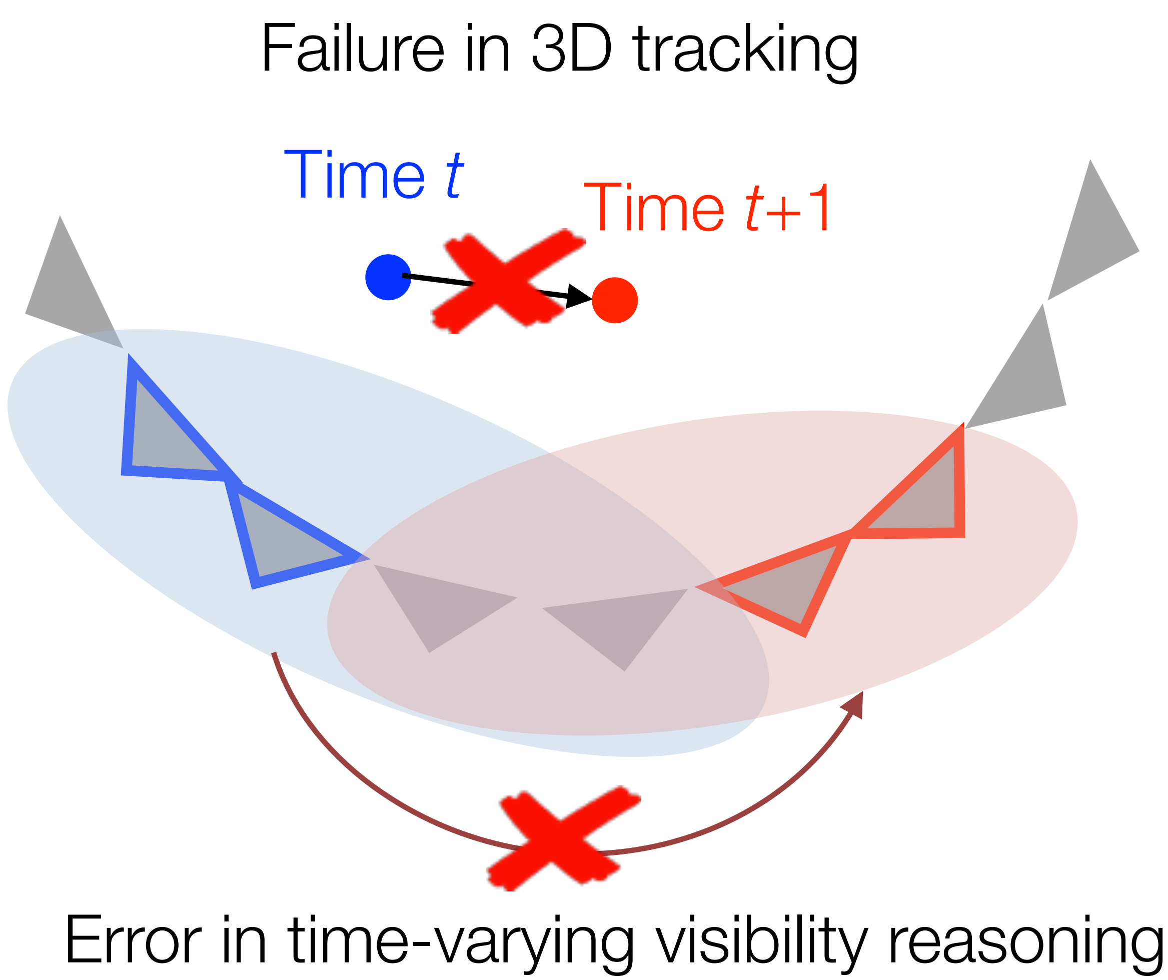
**Static Scene**

Point cloud reconstruction



**Dynamic Scene**

Trajectory stream reconstruction

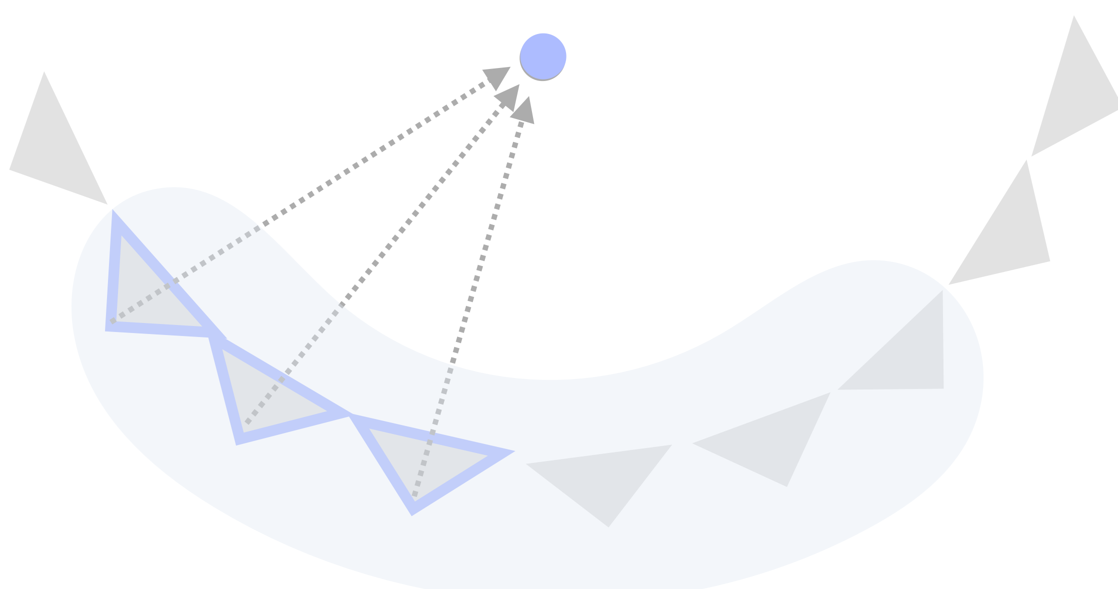


# Time-varying Visibility Reasoning

## Why Is It Important in Dynamic 3D Reconstruction?

**Static Scene**

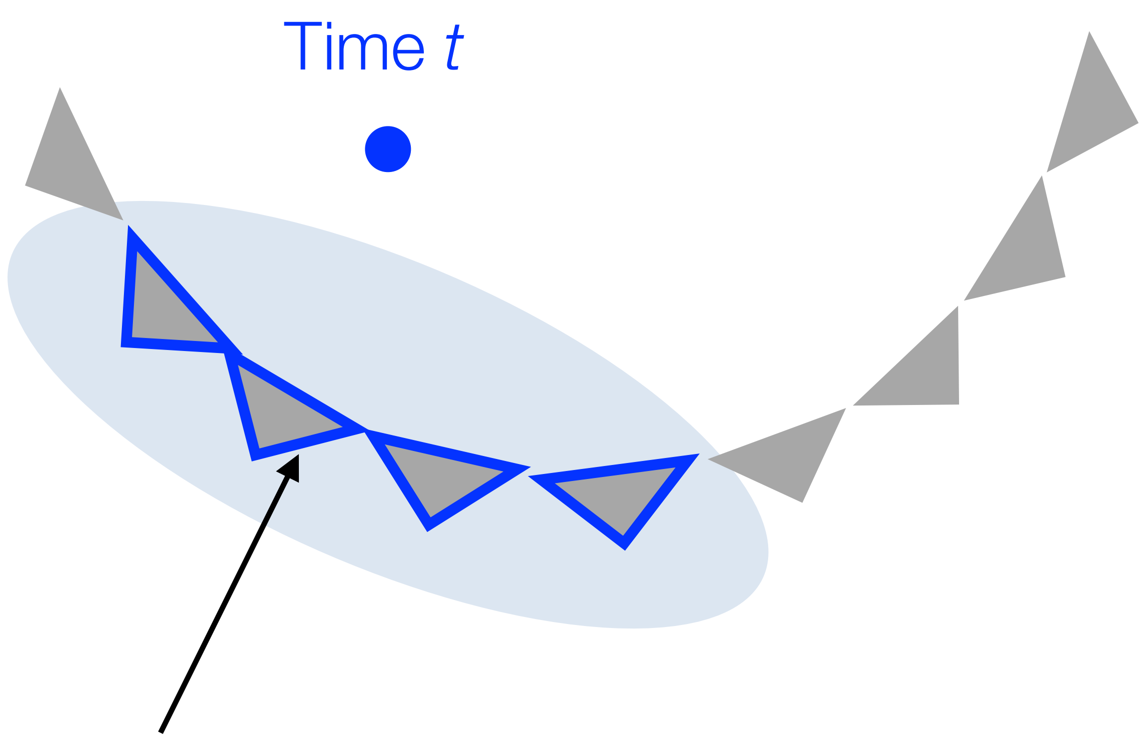
Point cloud reconstruction



**Dynamic Scene**

Trajectory stream reconstruction

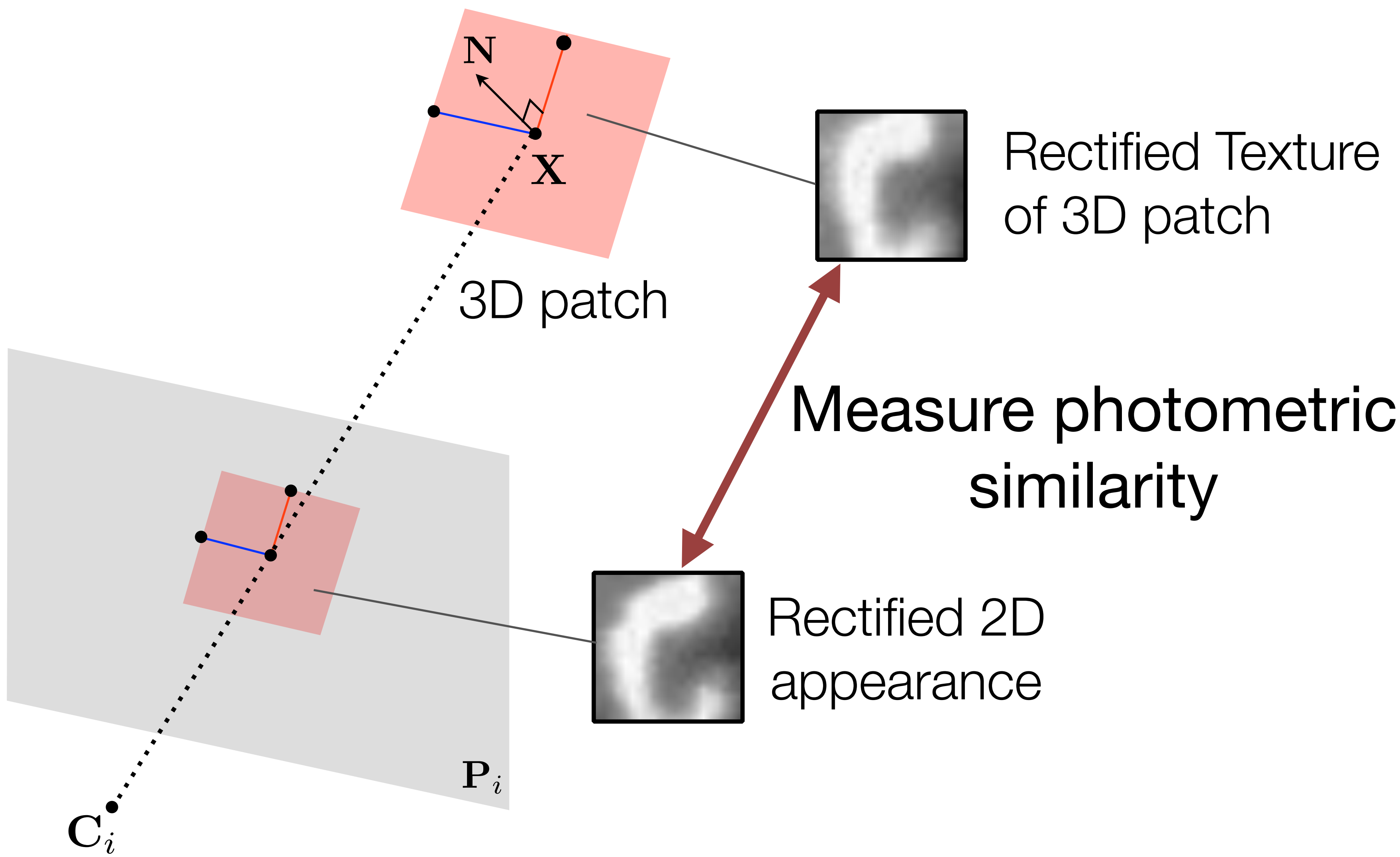
Failure in 3D tracking




As large and accurate visibility set as possible

# Photometric Consistency

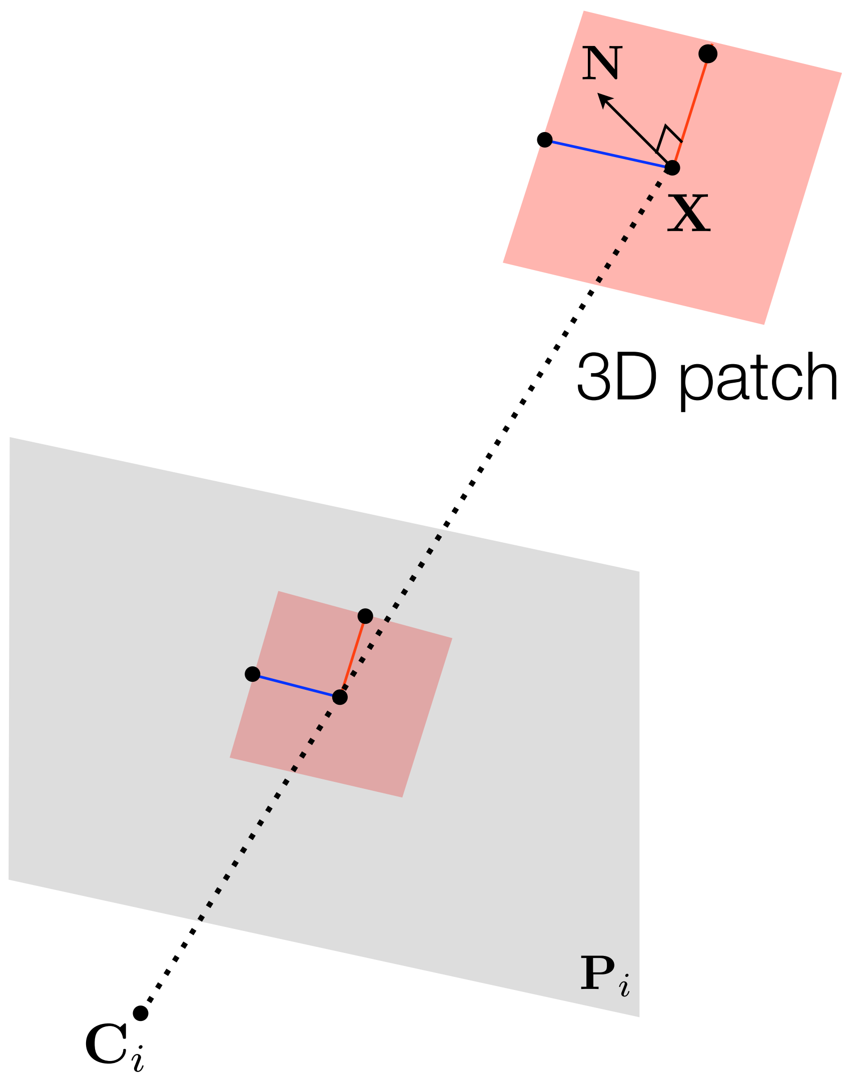
## A Common Cue for Static Scene Reconstruction

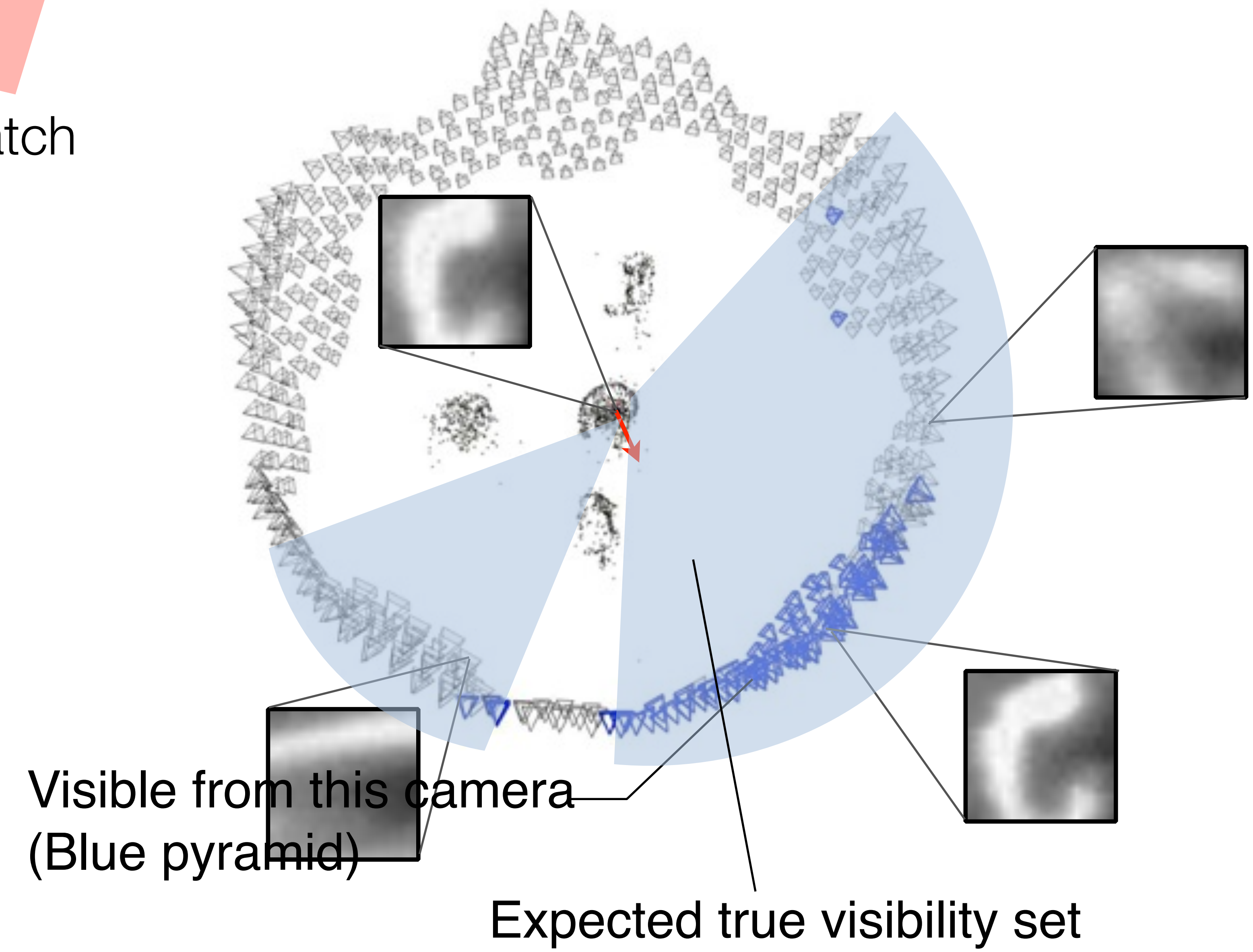



Accurate 3D patch shape and its texture are required

# Photometric Consistency

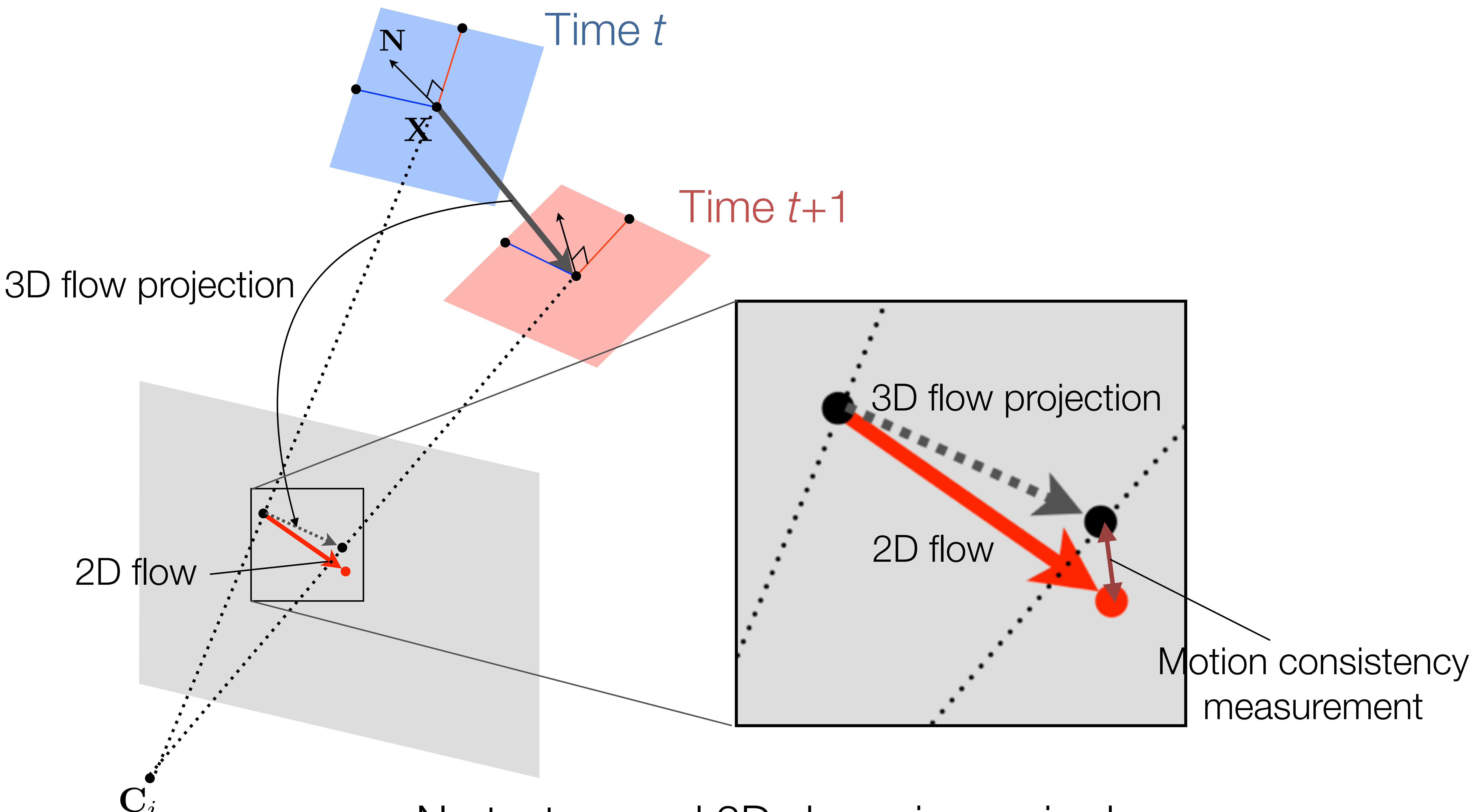
## A Common Cue for Static Scene Reconstruction



# Motion Consistency

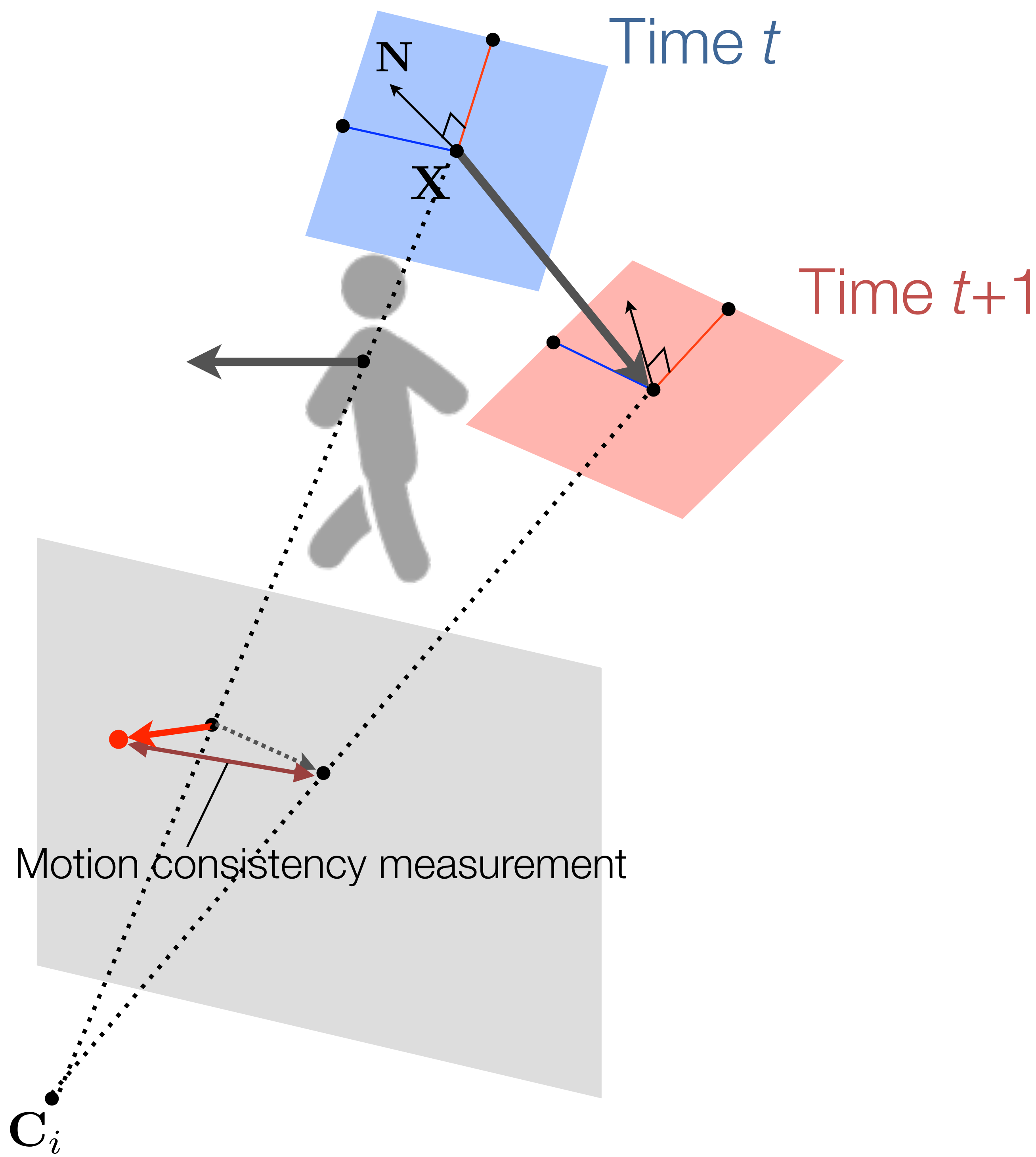
## A Novel Cue in Dynamic Scene




No texture and 3D shape is required

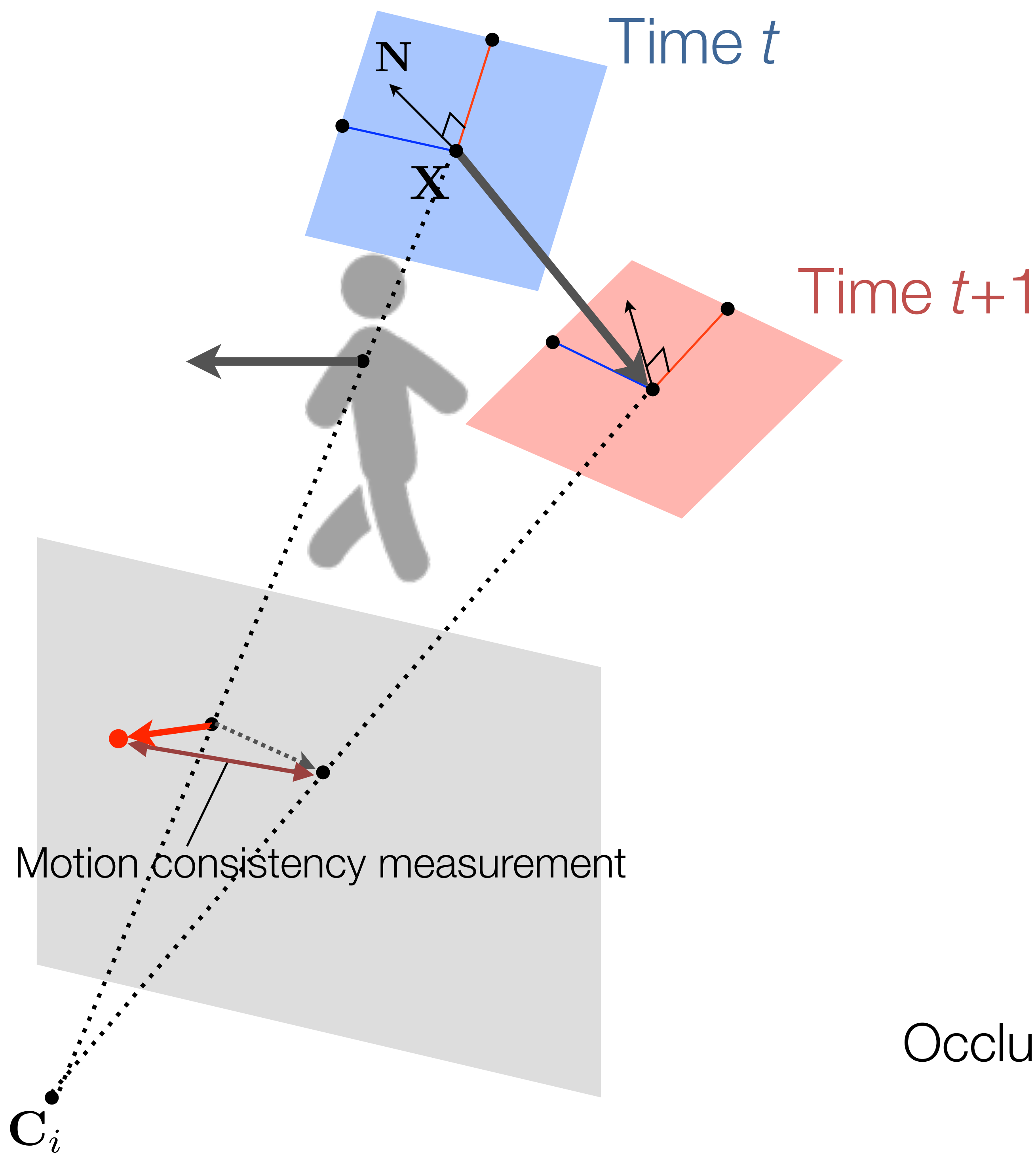
# Motion Consistency

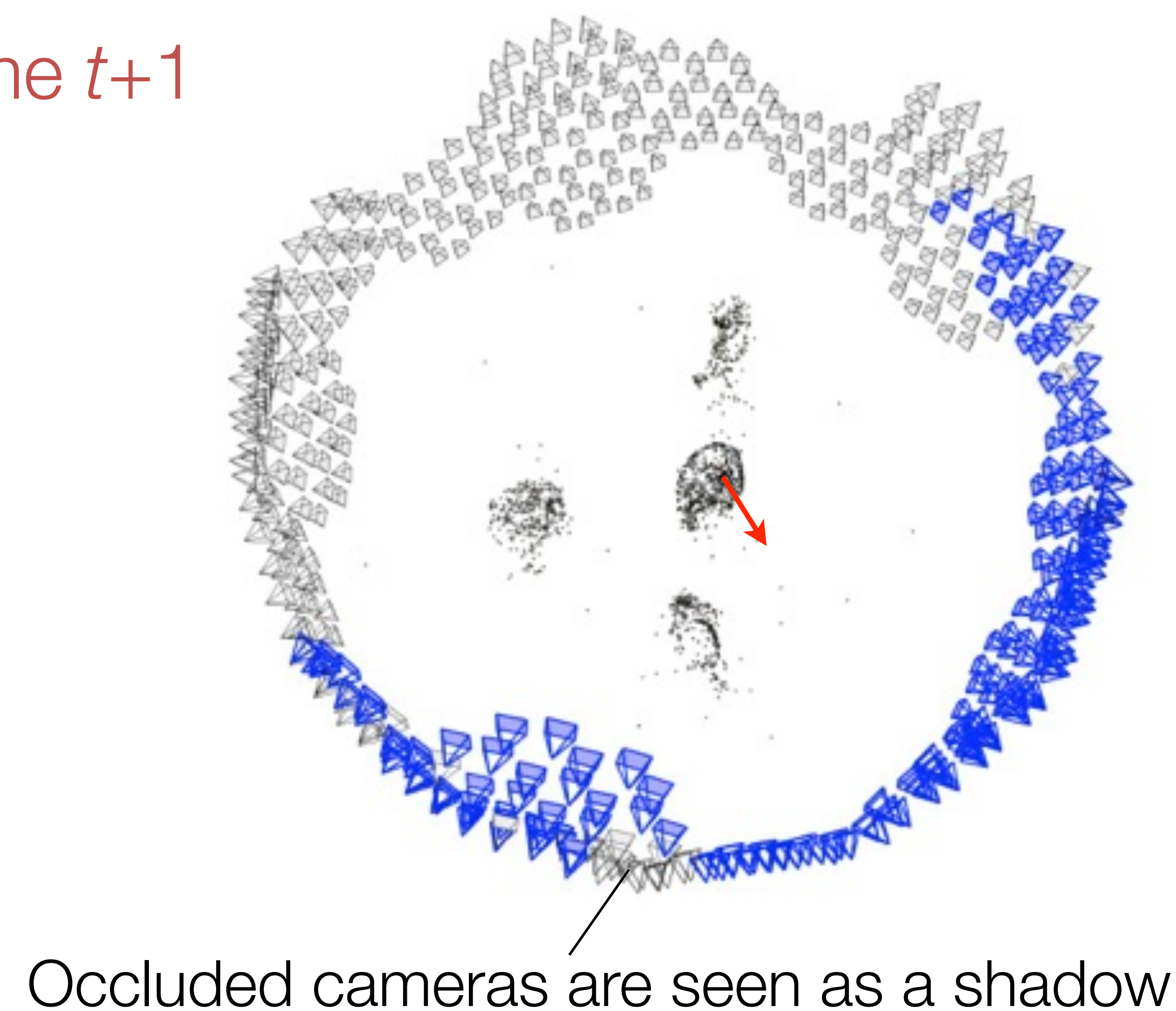
## A Novel Cue in Dynamic Scene



# Motion Consistency

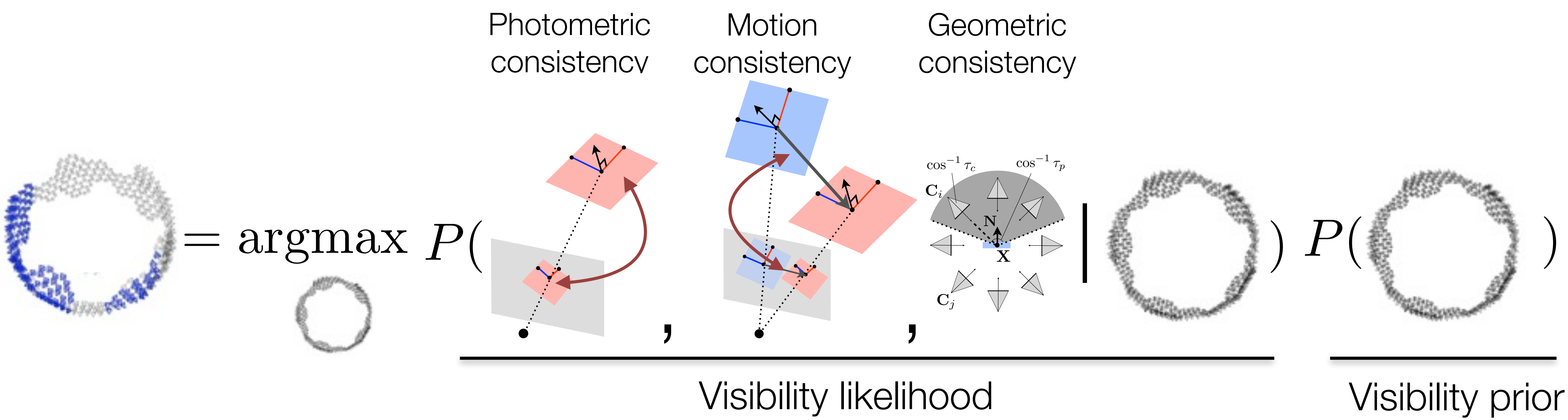
## A Novel Cue in Dynamic Scene



# MAP Visibility Estimate

## Visibility Likelihood and Visibility Prior



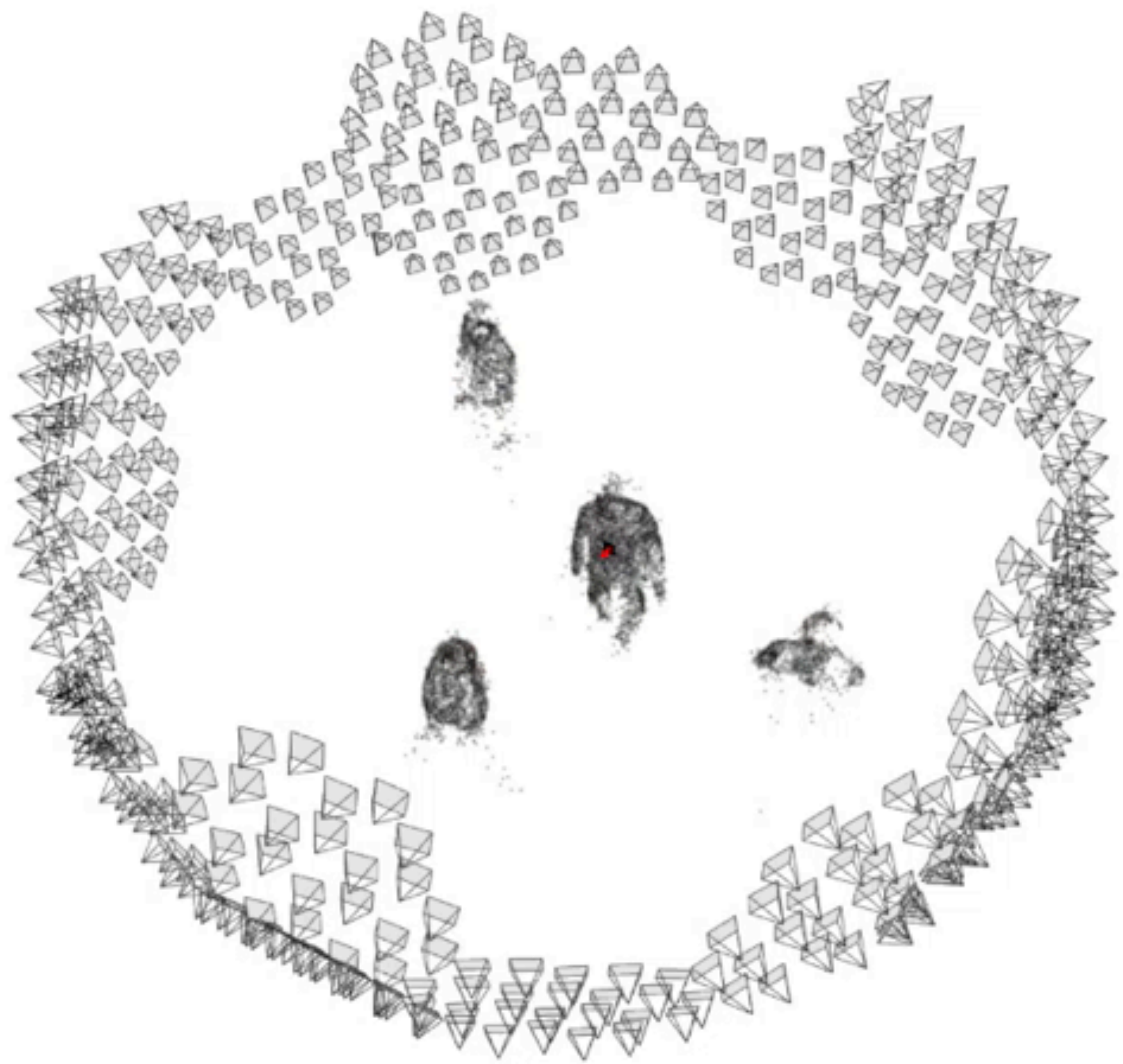
# Result

# Trajectory Stream Reconstruction Result

## The Circular Motion Sequence

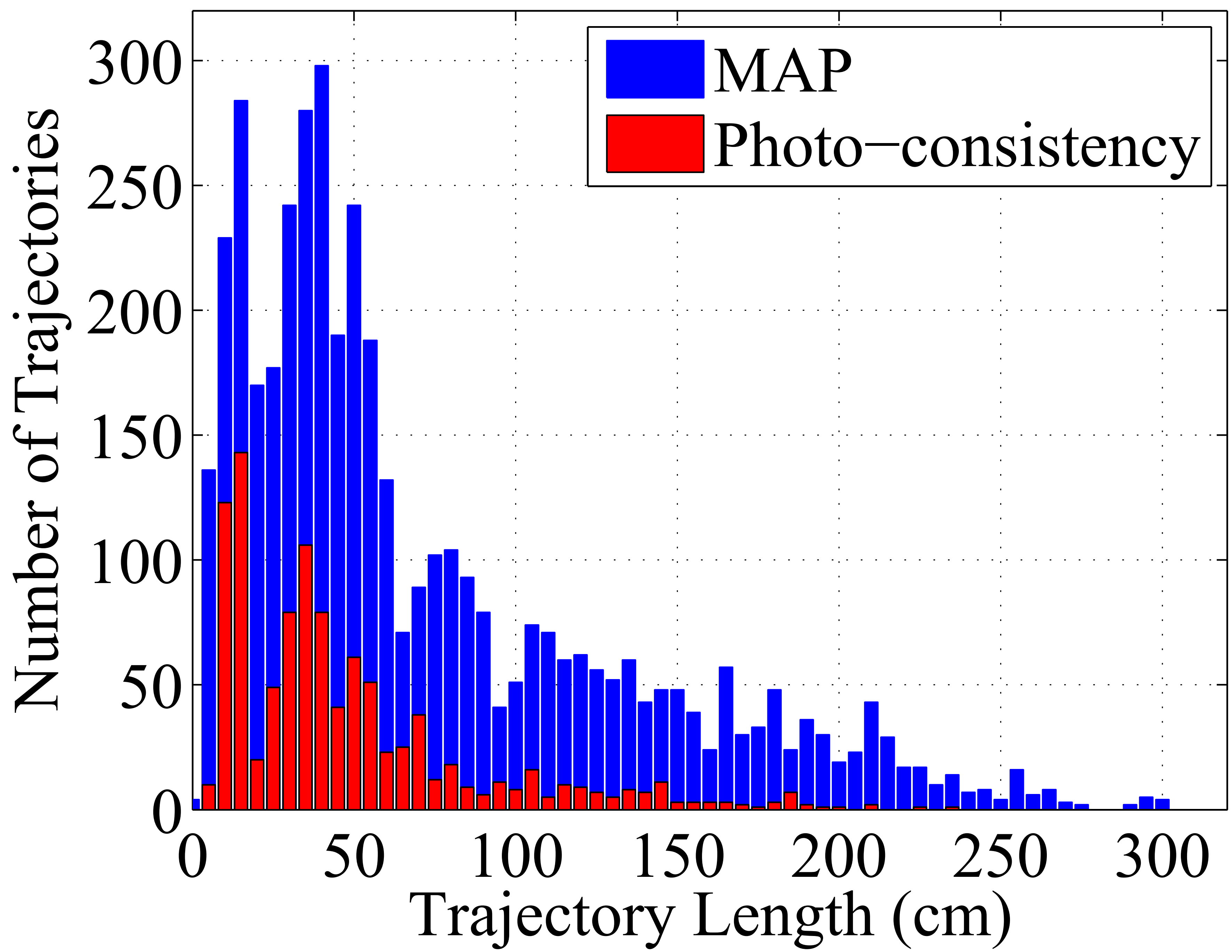
# Time-varying Visibility Reasoning

## Our Result



# Dynamic 3D Reconstruction Result

## Quantitative Comparison

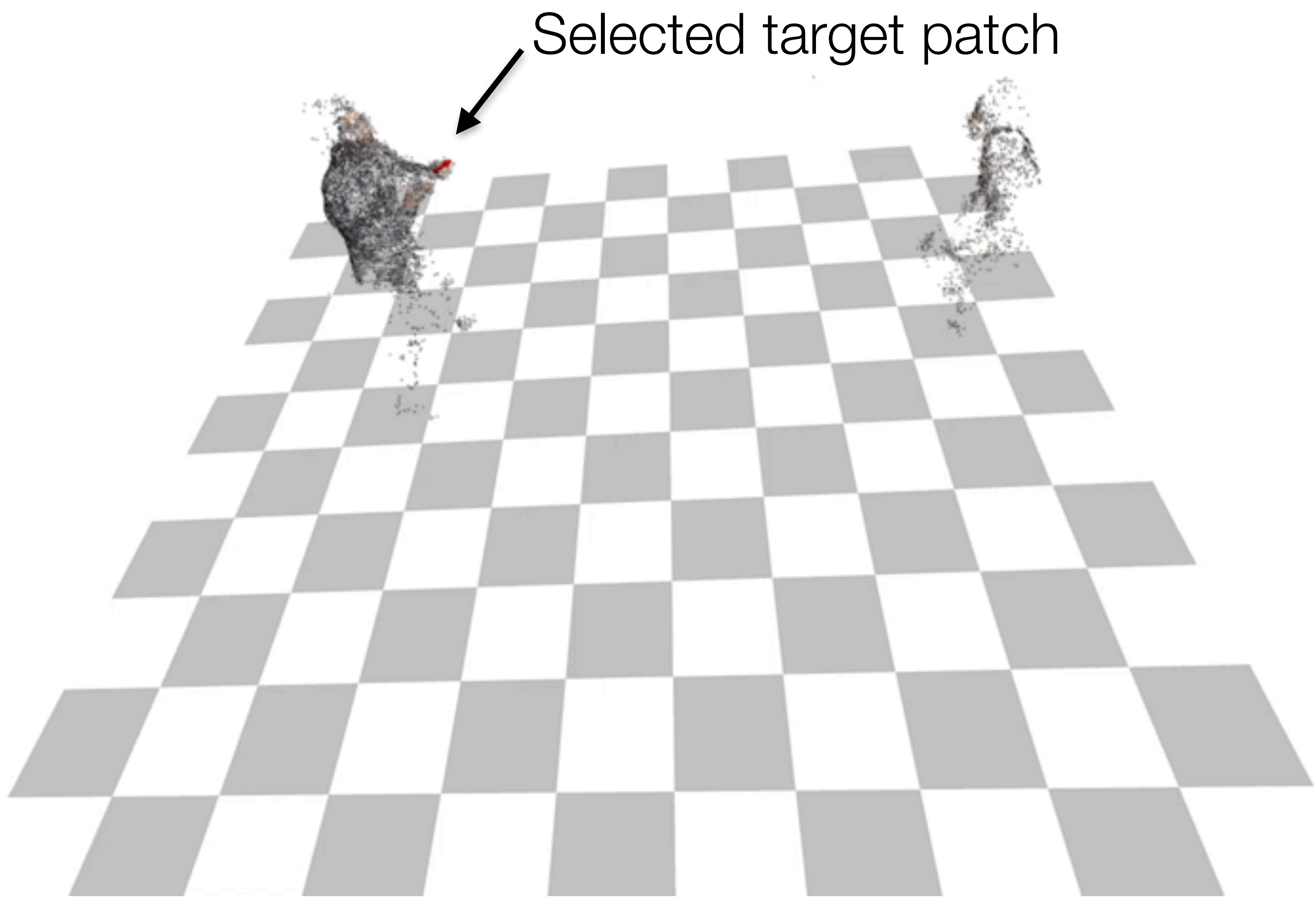


# Trajectory Stream Reconstruction Result

## The Volleyball Sequence

# Trajectory Stream Reconstruction Result

## The Volleyball Sequence: a Detail View



# Trajectory Stream Reconstruction Result

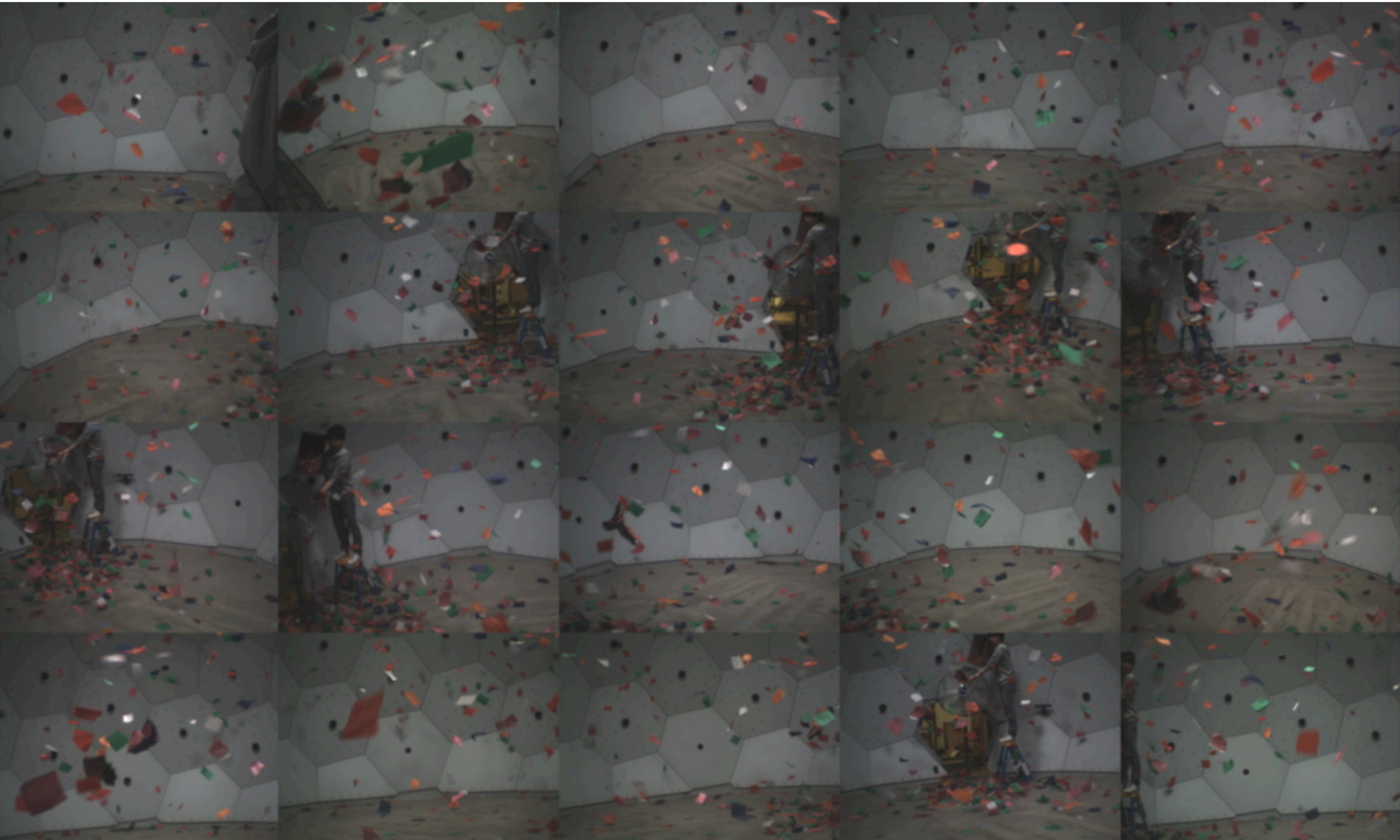
## The Falling Boxes Sequence

# Trajectory Stream Reconstruction Result

## The Confetti Sequence

# Trajectory Stream Reconstruction Result

## The Fluid Motion Sequence



# Future Work

Moving cameras

Motion analysis

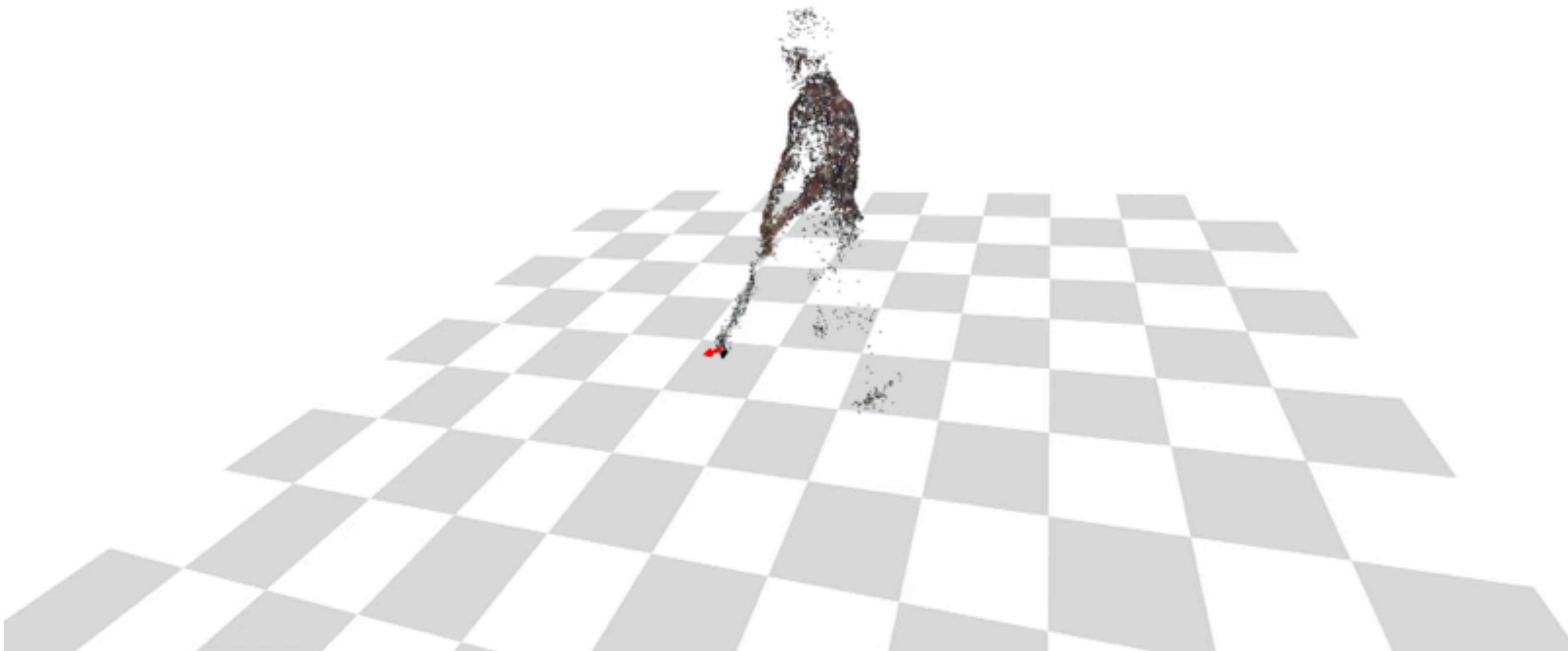
Social interactions

# Thank you

Please visit our poster (O-2A-5)

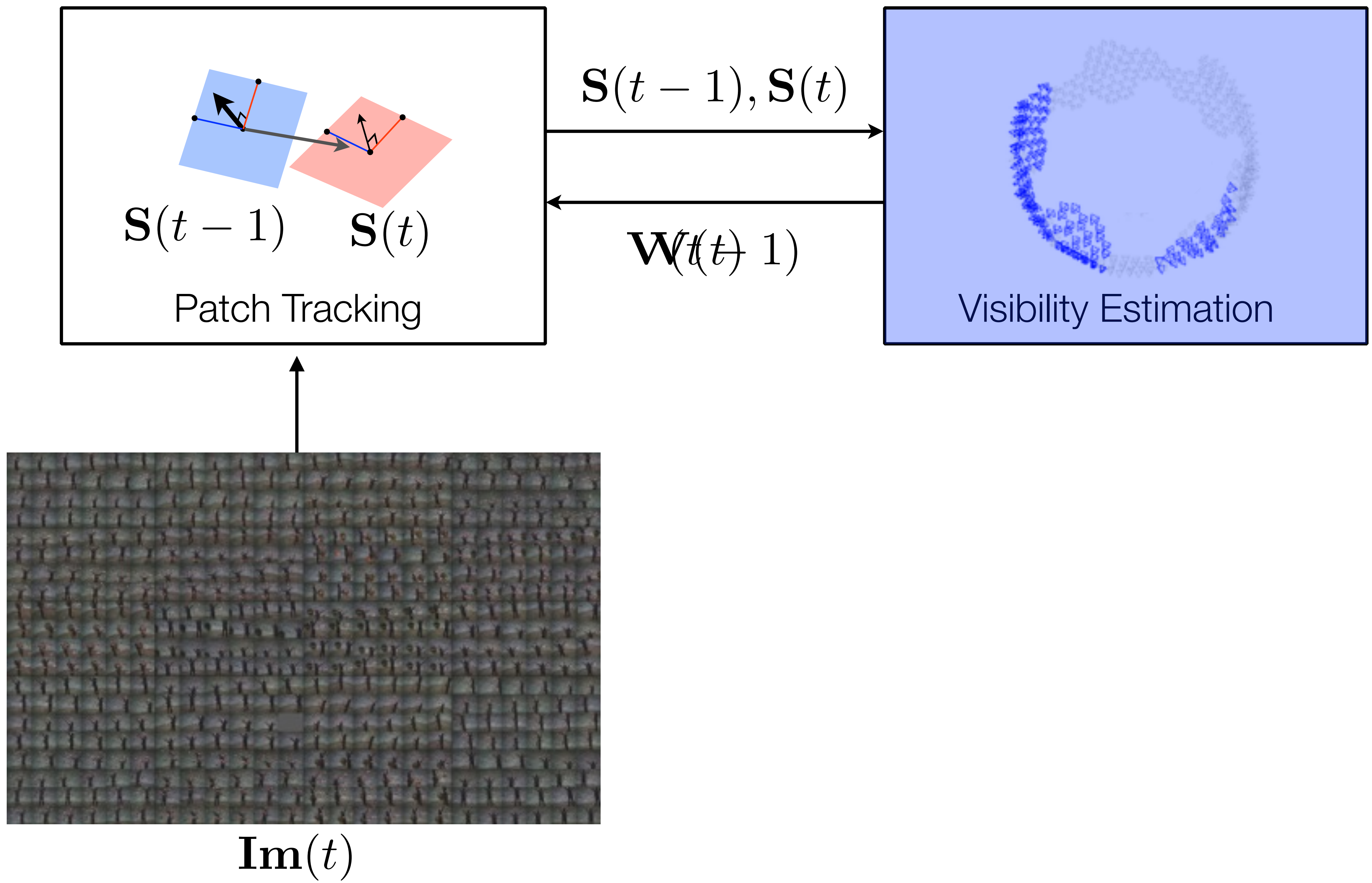
Dataset will be available at our project website:

http://www.cs.cmu.edu/~hanbyulj/14/visibility.html

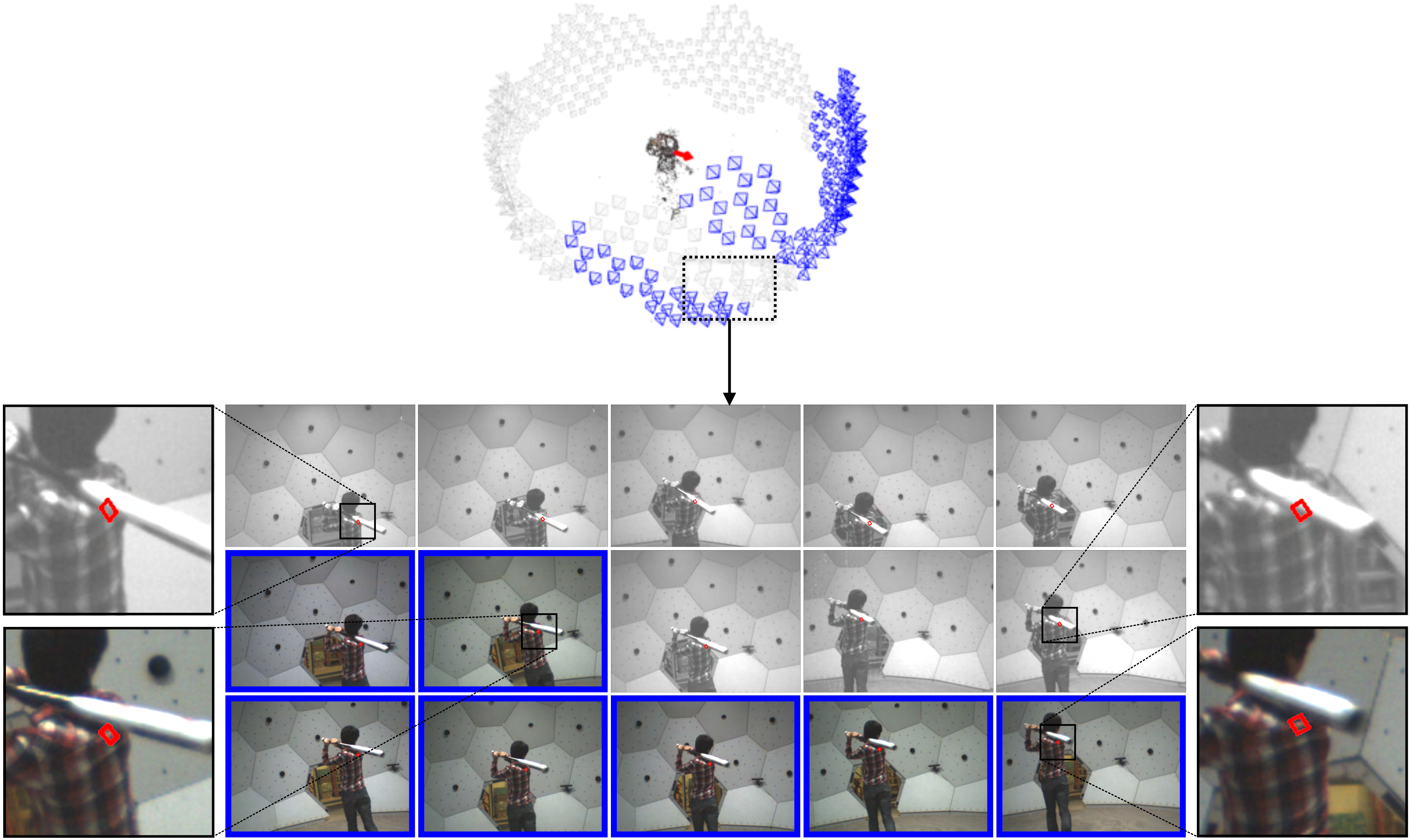


# Backup Slides

# Algorithm Overview
## Patch Tracking and Visibility Estimation

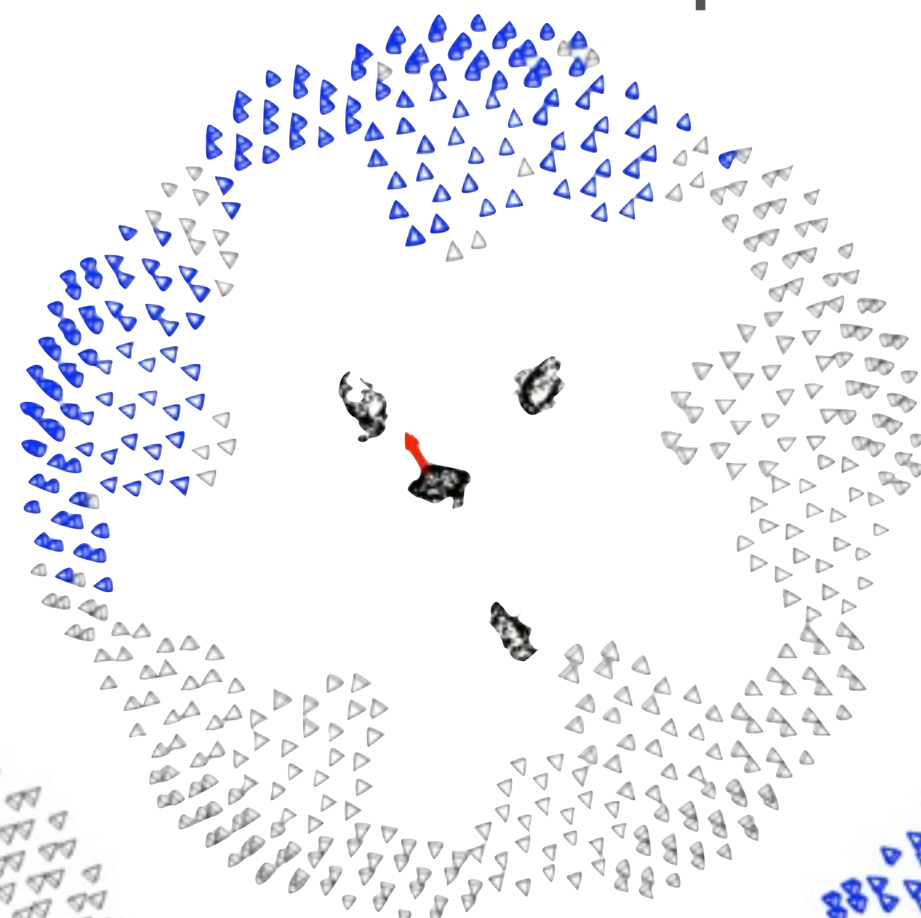


# Detailed Views of Visibility Reasoning Result

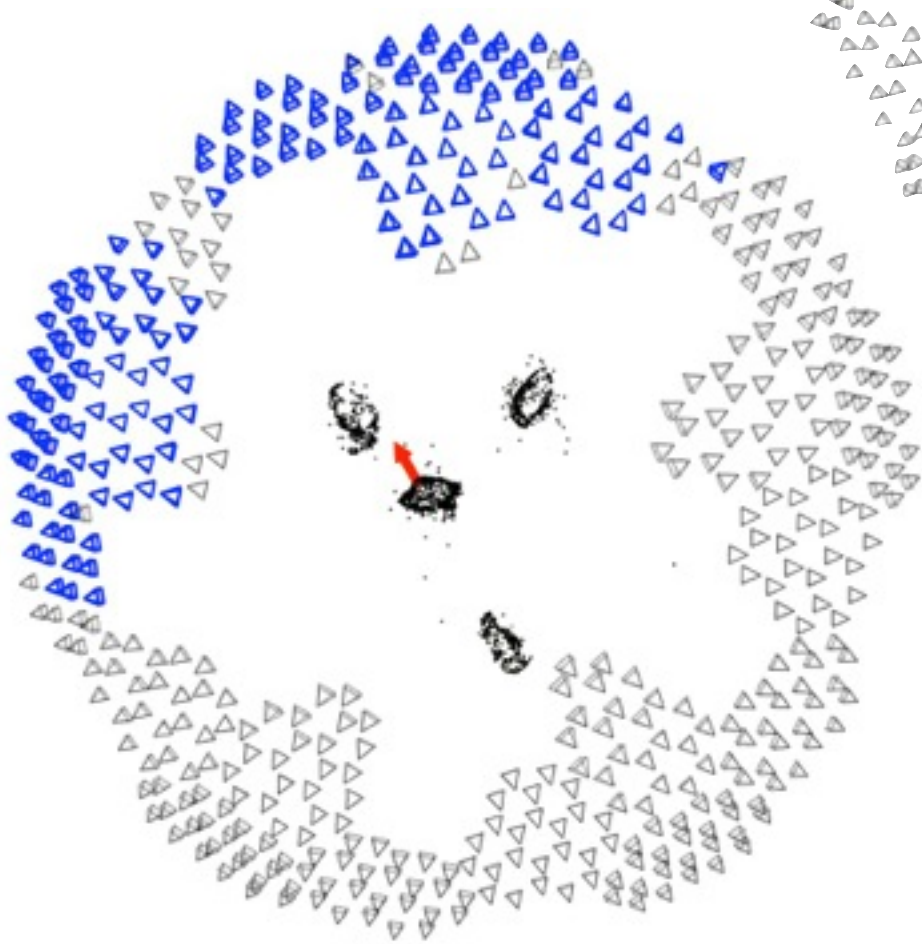


# Visibility Reasoning Result

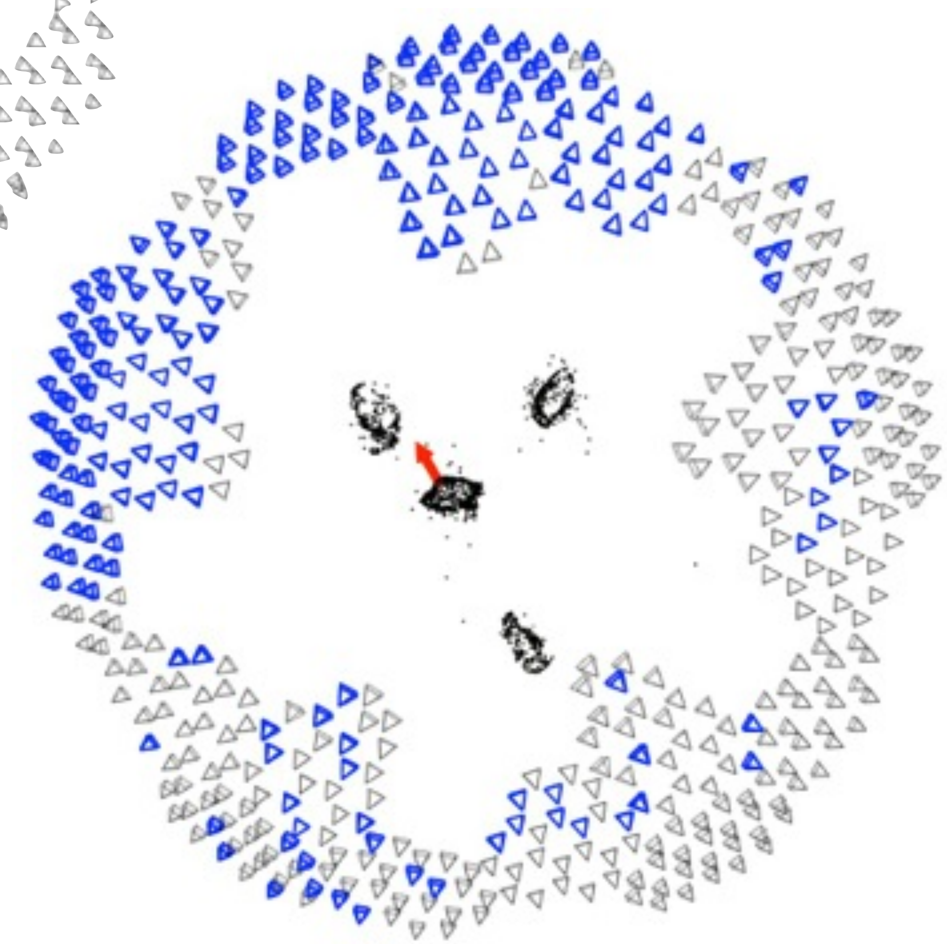
## Qualitative Comparison



Ground truth



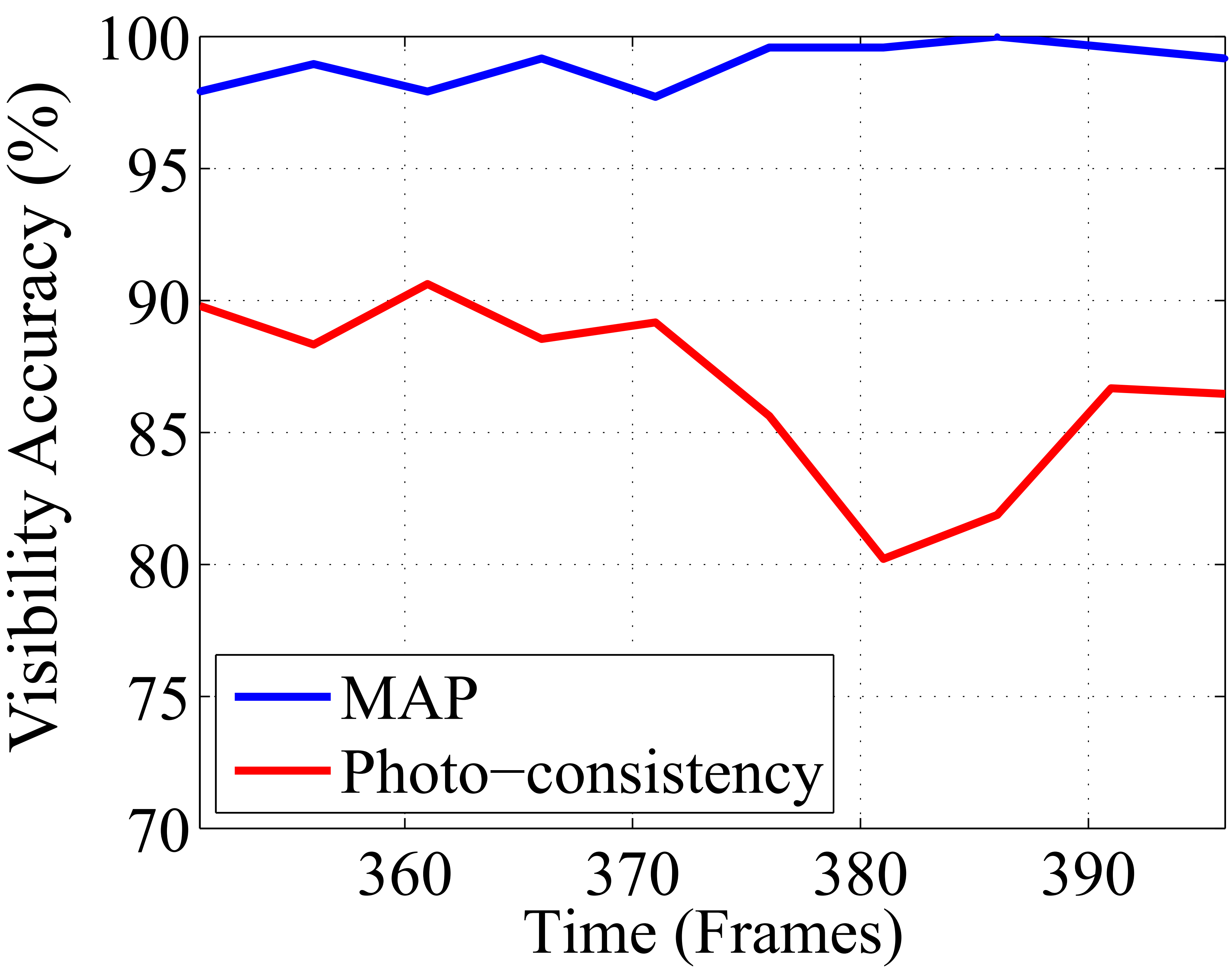
Our method



Photometric consistency

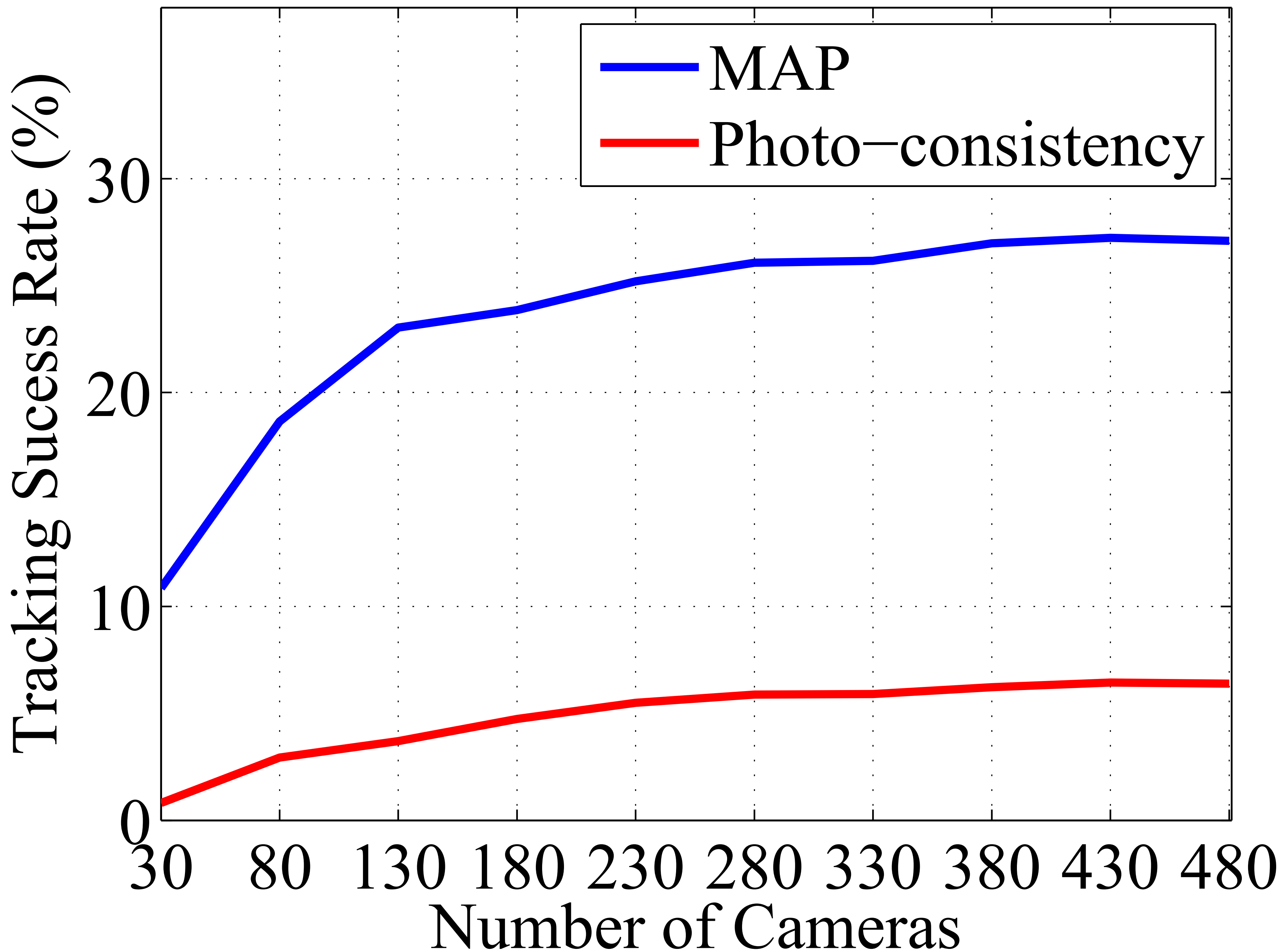
# Visibility Reasoning Result

## Quantitative Comparison



# Dynamic 3D Reconstruction Result

## Quantitative Comparison



# Summary of the Datasets

| Sequence | Frames | Duration | # of points | Av. traj. length |
|---|---|---|---|---|
| Circ. Movement | 250 | 10.0 sec | 10433 | 404.9 cm |
| Volleyball | 210 | 8.4 sec | 8422 | 326.4 cm |
| Bat Swing | 200 | 8.0 sec | 3849 | 224.1 cm |
| Falling Boxes | 160 | 6.4 sec | 17934 | 164.7 cm |
| Confetti | 200 | 8.0 sec | 10345 | 103.0 cm |
| Fluid Motion | 200 | 8.0 sec | 3153 | 123.1 cm |

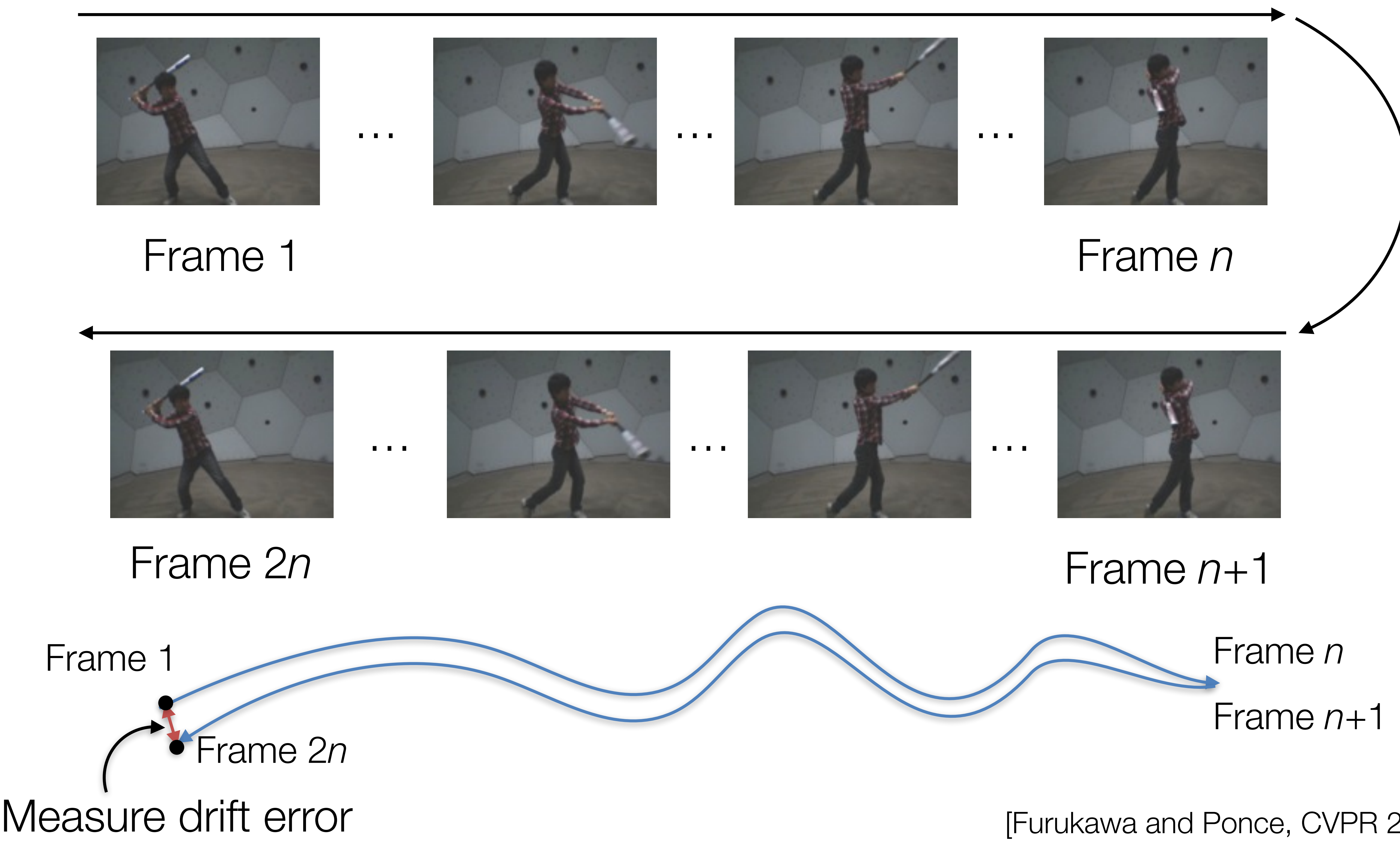
Trajectory reconstruction
from one time instance

Frame rate: 25frame / sec
Data size: 220GB / min

# Computation Time

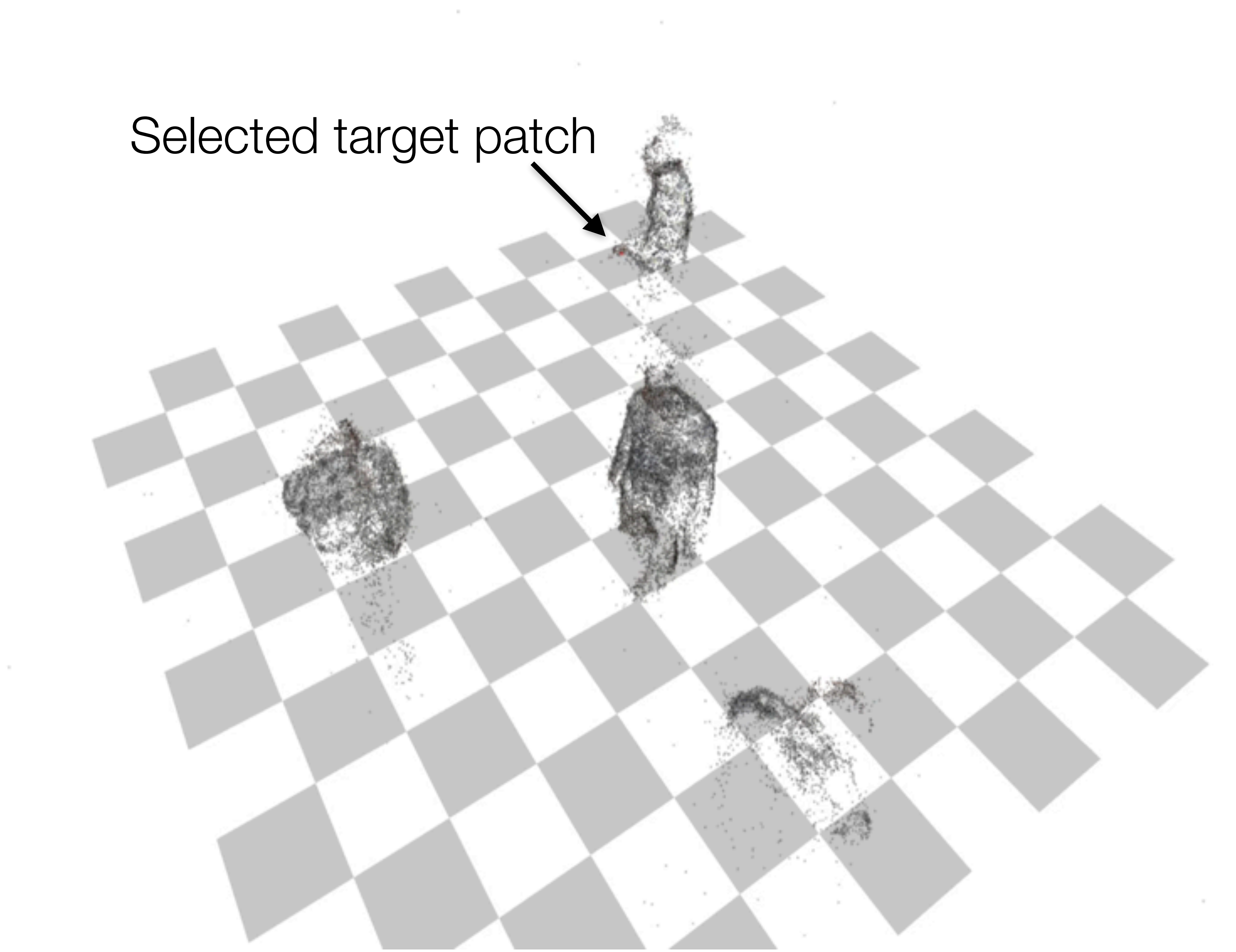
- 10,000 points over 8 sec
- Using 100 cores
- 12 hours
- 10~15 starting frames for each sequence — 1 week

# Quantitative Evaluation Method

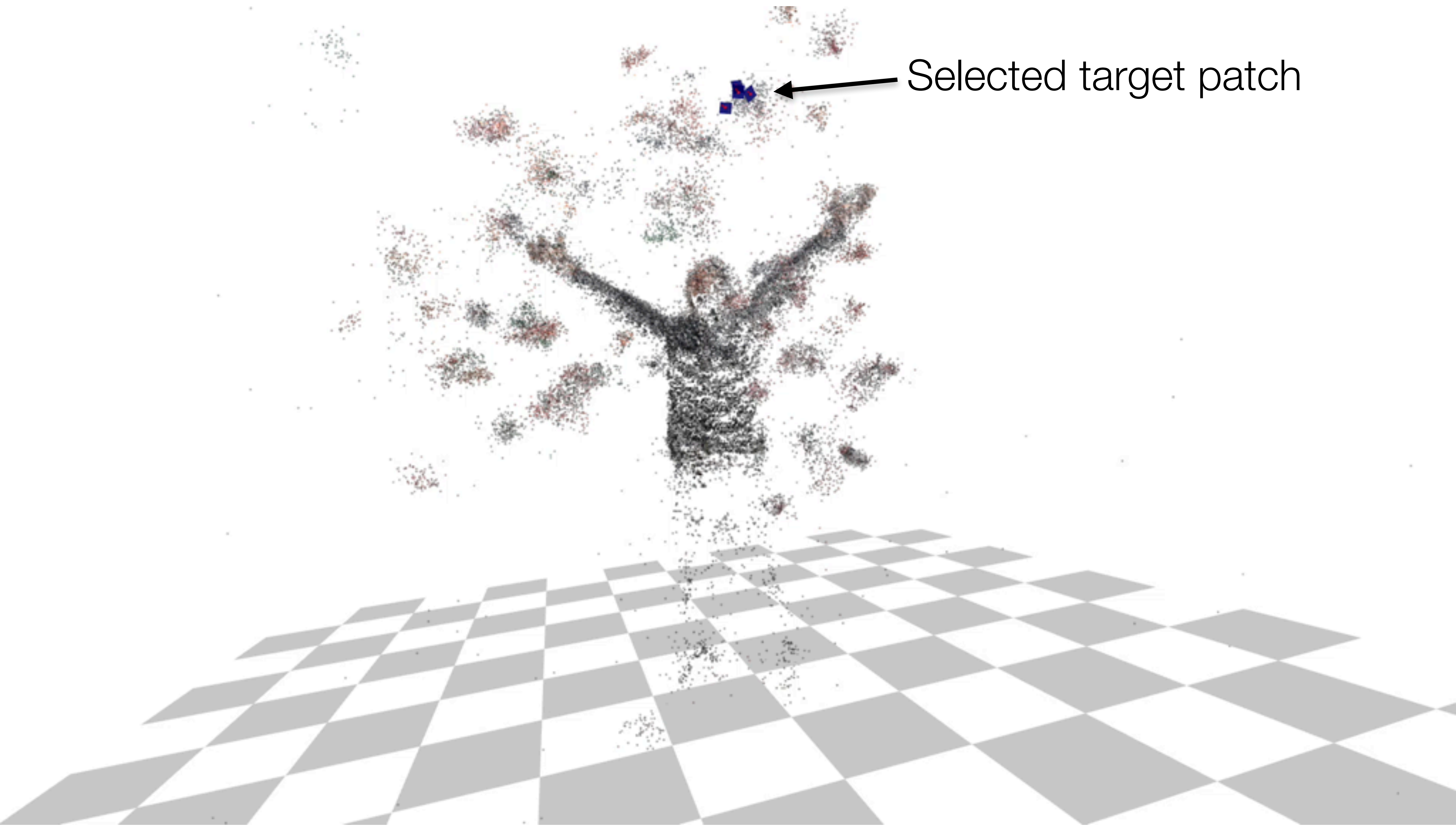



[Furukawa and Ponce, CVPR 2008]

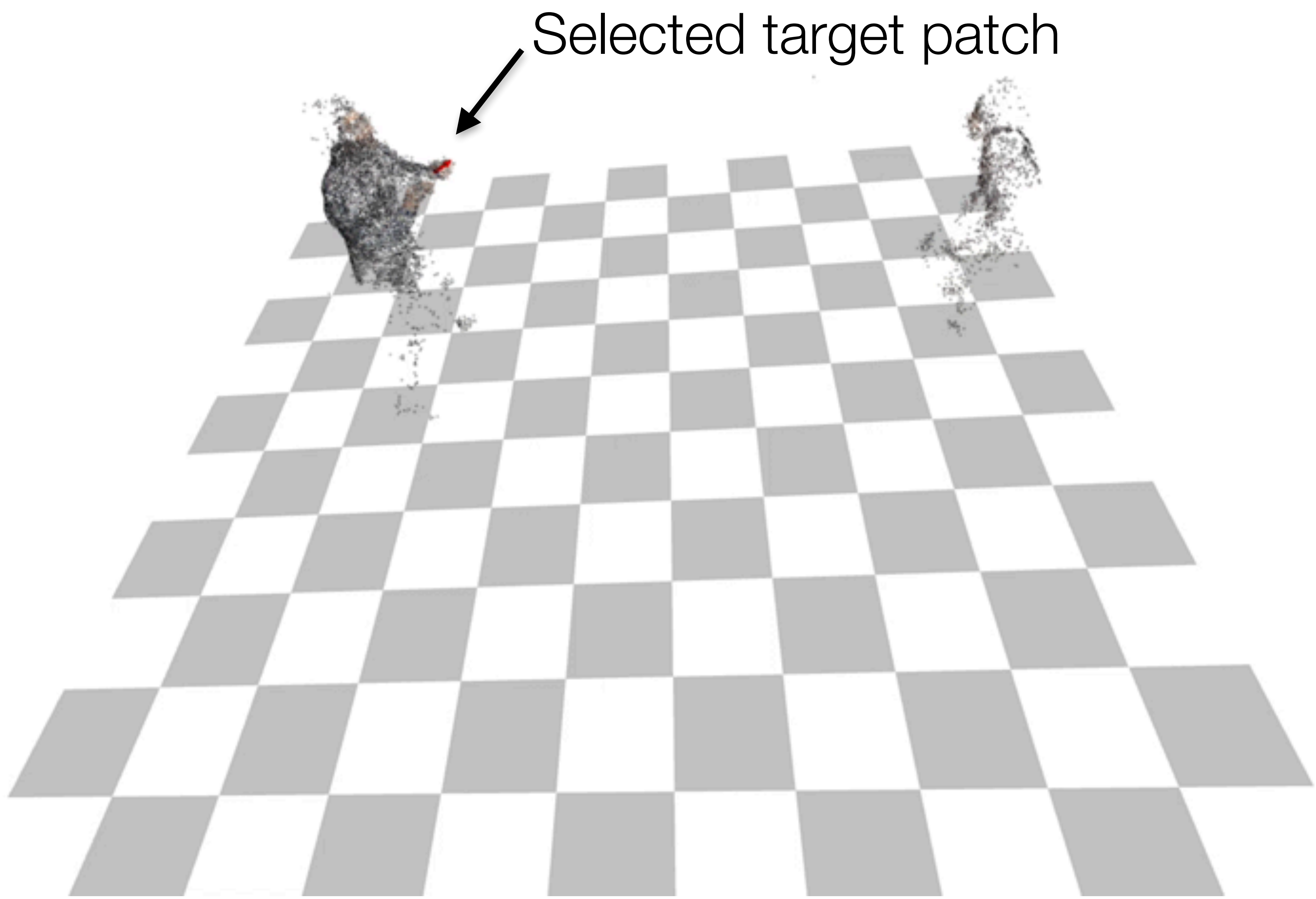
# A Detailed View of A Selected Patch



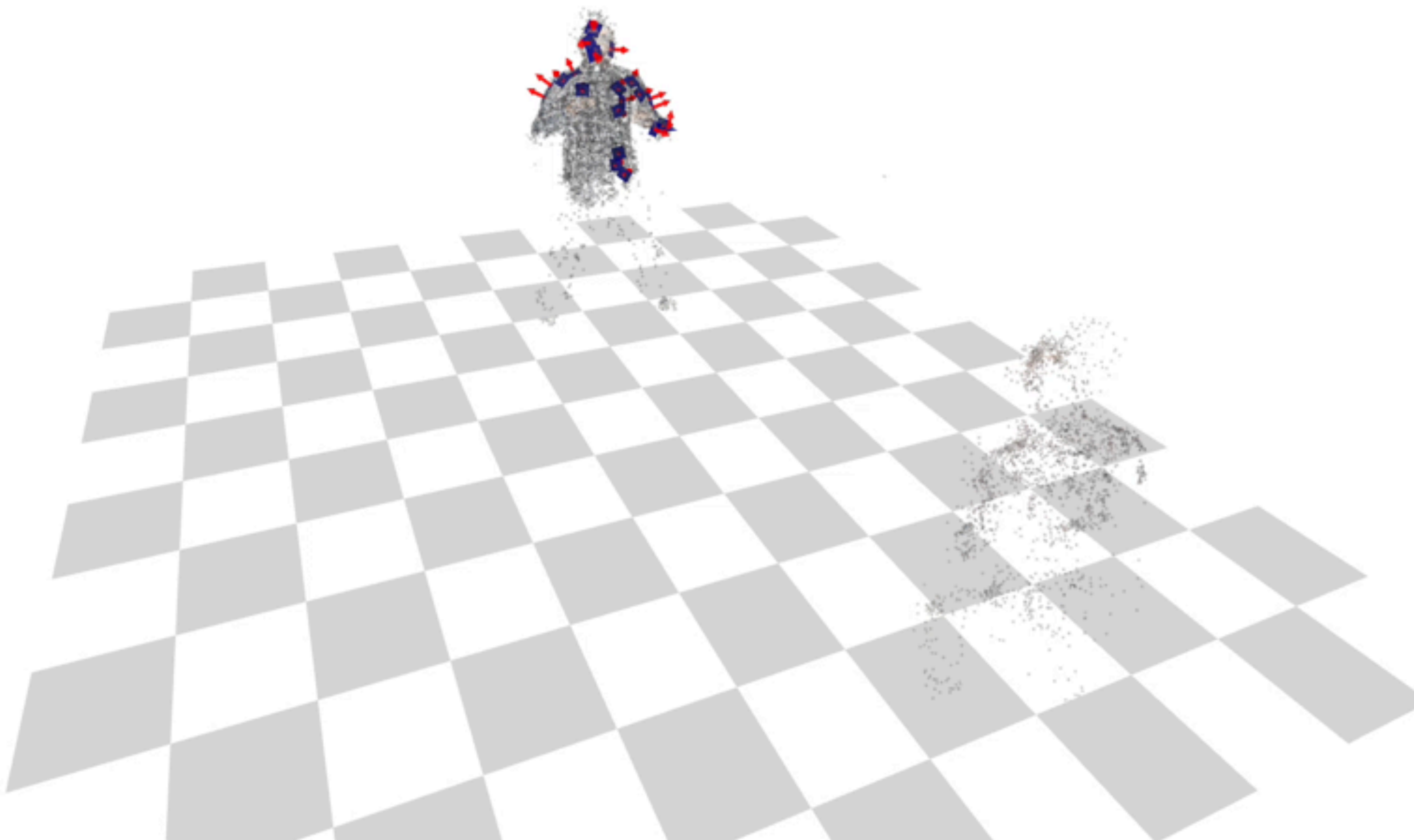
# A Detailed View of Selected Patches



# A Detailed View of A Selected Patch



# A Detailed View of Selected Patches



# References


F. Huguet and F. Devernay. A variational method for scene flow estimation from stereo sequences. In ICCV, 2007.

T. Basha, Y. Moses, and N. Kiryati. Multi-view Scene Flow Estimation: A View Centered Variational Approach. IJCV, 2012.

Y. Furukawa and J. Ponce. Dense 3D motion capture from synchronized video streams. In CVPR, 2008

N. Snavely, S. M. Seitz, and R. Szeliski. Photo tourism: Exploring photo collections in 3D. TOG, 2006.

S. Agarwal, N. Snavely, I. Simon, S. M. Seitz, and R. Szeliski  Building Rome in a Day, ICCV, 2009

S. Vedula, S. Baker, P. Rander, R. Collins, and T. Kanade. Three- dimensional scene flow. TPAMI, 2005.

R. Carceroni and K. Kutalakos. Multi-view scene capture by surfel sampling: From video streams to non-rigid 3D motion, shape and reflectance. IJCV, 2002.

F. Devernay,D.Mateus,and M.Guilbert. Multi-CameraSceneFlow by Tracking 3-D Points and Surfels. In CVPR, 2006

L. Guan, J. S. Franco, E. Boyer, M. Pollefeys. Probabilistic 3D occupancy flow with latent silhouette cues. In CVPR 2010.

T. Tung and T. Matsuyama. Dynamic surface matching by geodesic mapping for 3D animation transfer. In CVPR, 2010

K. Varanasi, A. Zaharescu, E. Boyer, R. Horaud,Temporal surface tracking using mesh evolution. In ECCV, 2008.

E. de Aguiar, C. Theobalt, C. Stoll, and H.-P. Seidel. Marker-less deformable mesh tracking for human shape and motion capture. In CVPR, 2007