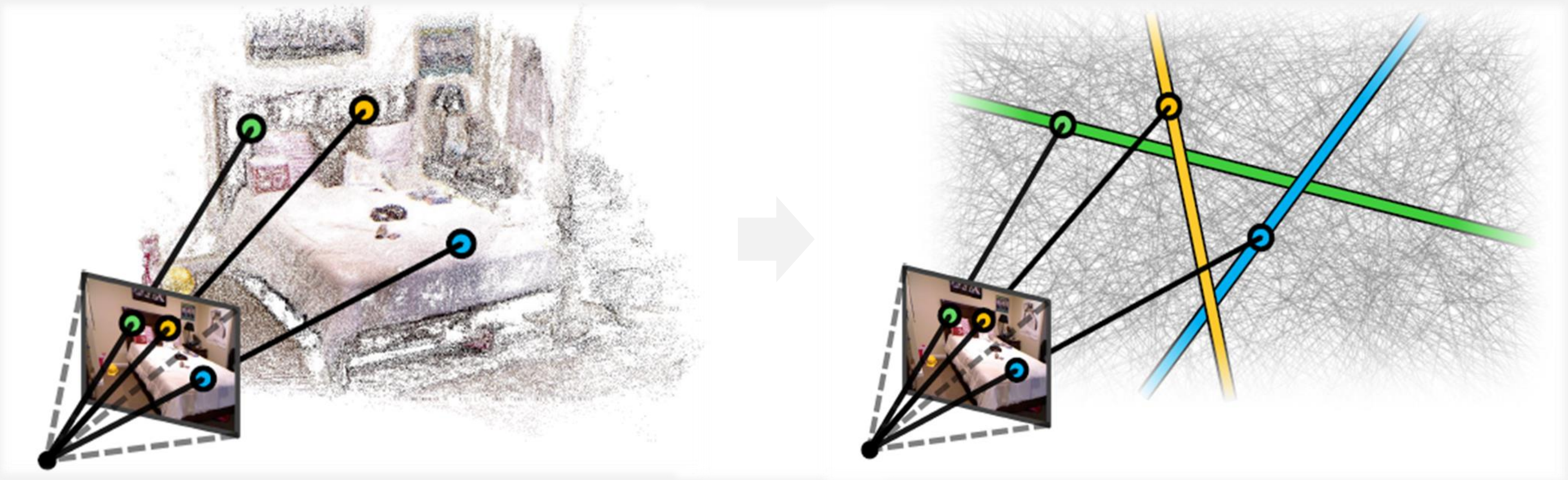


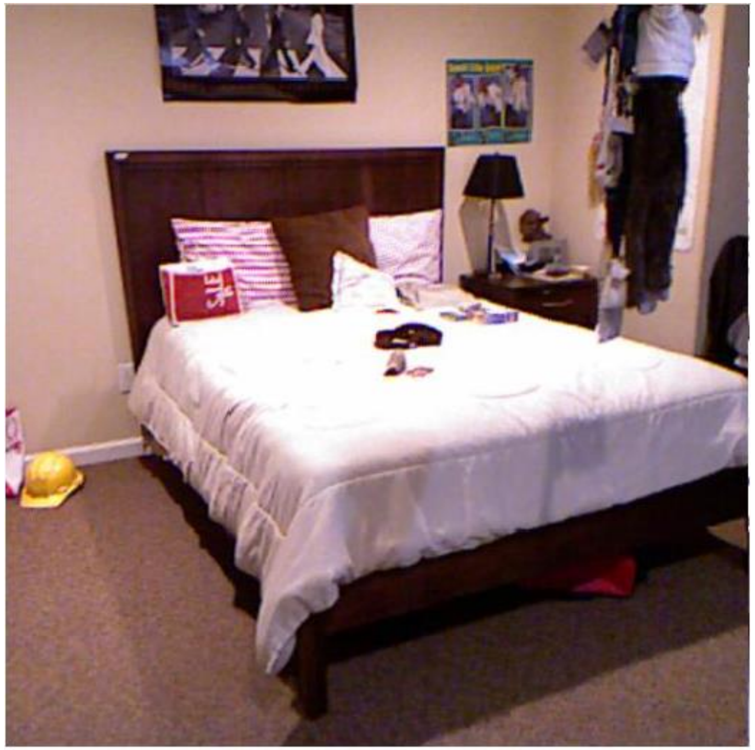
Privacy-Preserving Image-based Localization

Sudipta N. Sinha
Microsoft Research, Redmond, USA
(joint work with Francesco Pittaluga, Sanjeev Koppal, Sing Bing Kang, Pablo Speciale, Johannes Schonberger and Marc Pollefeys)

Osaka University seminar talk, October 25, 2019

Image-based Localization

Query Image



3D Point Cloud Maps

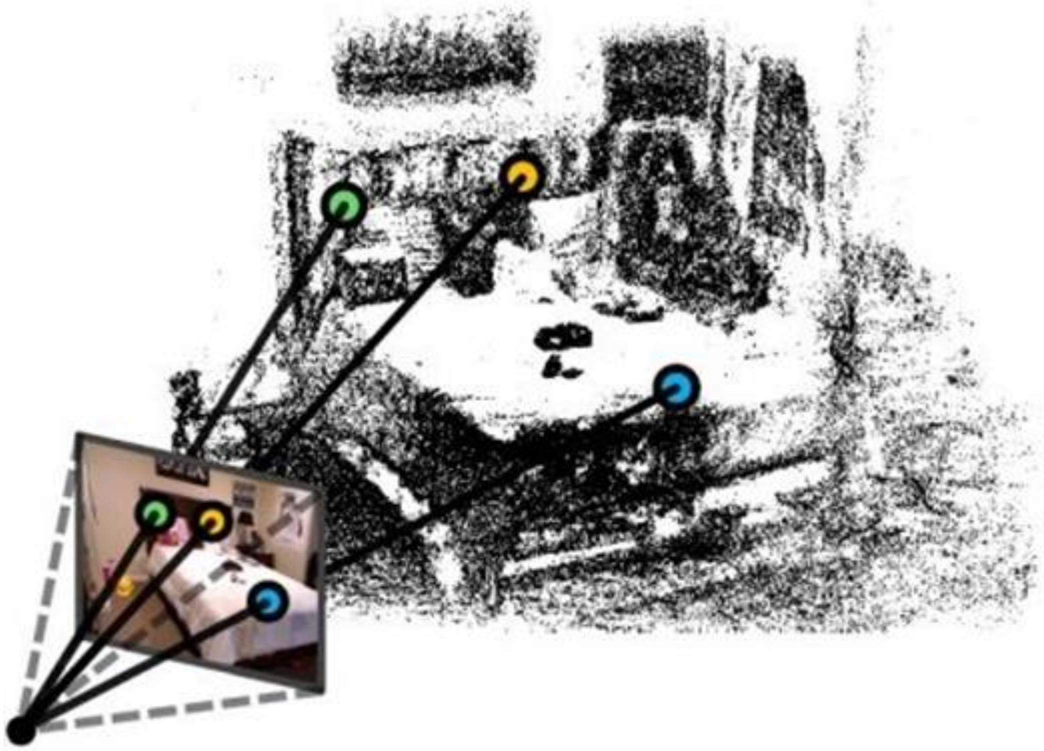


Image-based Localization

- Arth+ 2009
- Irschara+ 2009
- Sattler+ 2011
- Li+ 2012
- Lim+ 2012
- Ventura+ 2014
- Zeisl+ 2015
- Sattler+ 2015
- Lynen+ 2015
- Kendall+ 2015
- Weyand+ 2016
- …

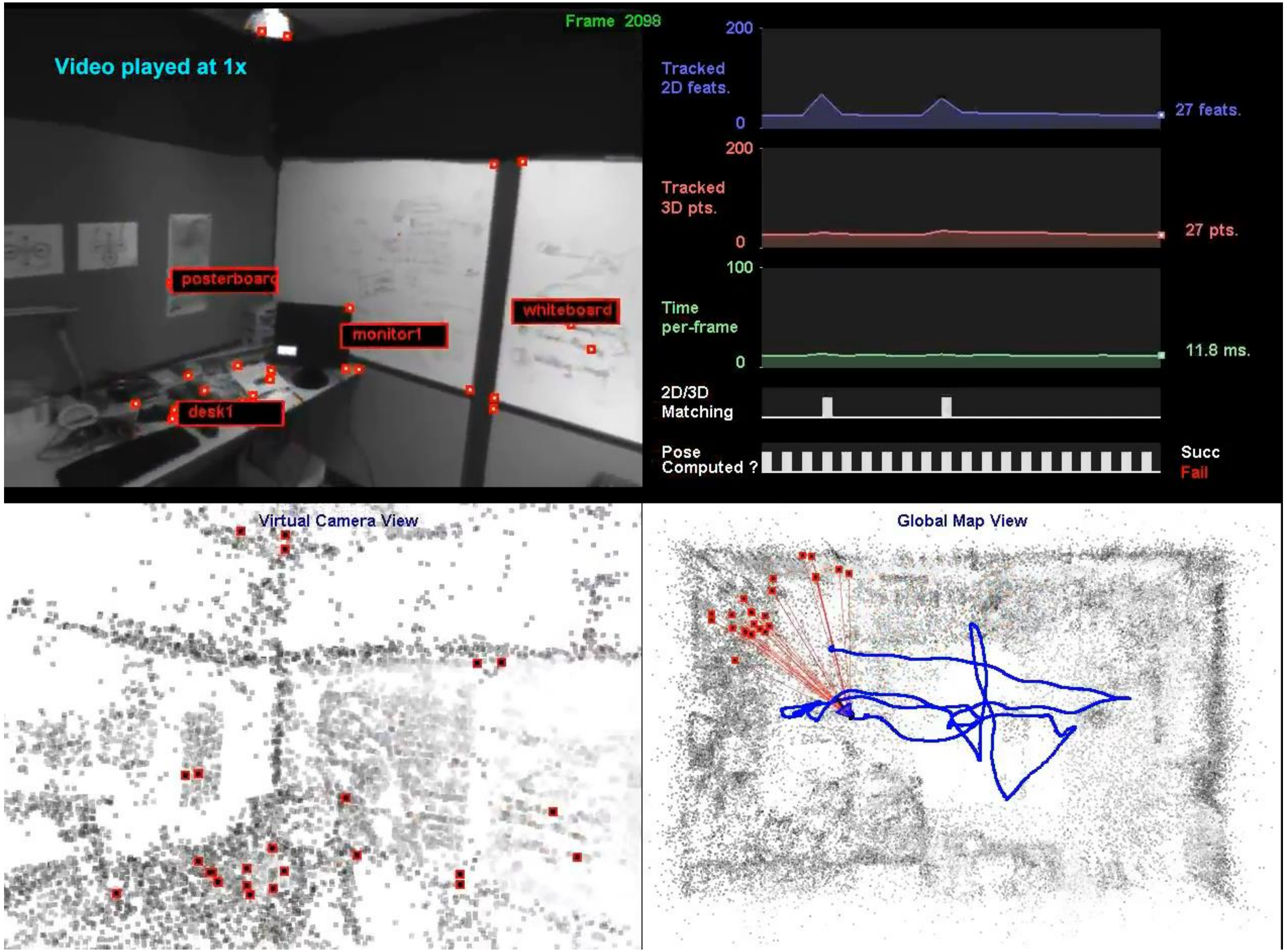


Image-based Localization

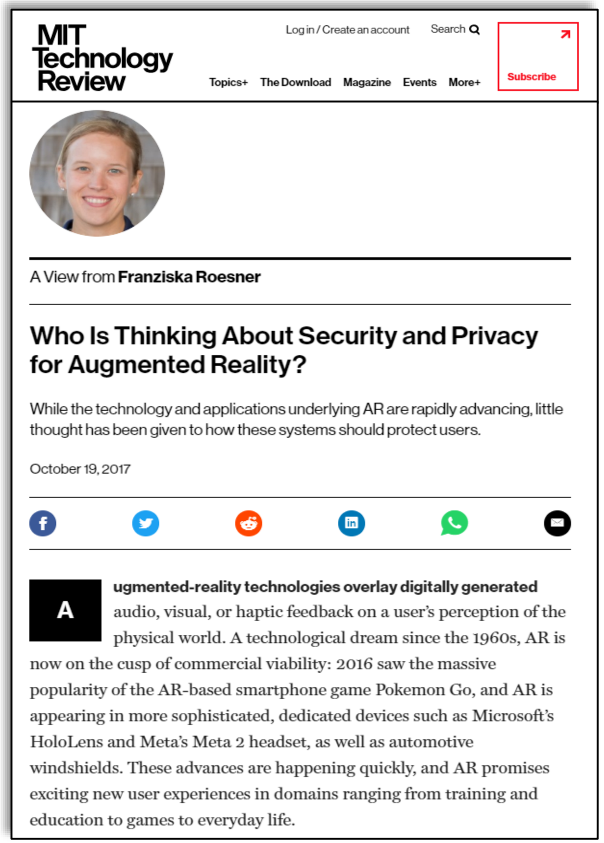
Microsoft HoloLens

Drone Navigation (UAVs)

Google AR Core / Apple ARKit

A crucial task in Augmented Reality (AR) & Robotics applications.

New Privacy Concerns for AR


MIT Technology Review

Log in / Create an account Search

Subscribe

Topics+ The Download Magazine Events More+

A View from **Franziska Roesner**

Who Is Thinking About Security and Privacy for Augmented Reality?

While the technology and applications underlying AR are rapidly advancing, little thought has been given to how these systems should protect users.

October 19, 2017

Augmented-reality technologies overlay digitally generated audio, visual, or haptic feedback on a user's perception of the physical world. A technological dream since the 1960s, AR is now on the cusp of commercial viability: 2016 saw the massive popularity of the AR-based smartphone game Pokemon Go, and AR is appearing in more sophisticated, dedicated devices such as Microsoft's HoloLens and Meta's Meta 2 headset, as well as automotive windshields. These advances are happening quickly, and AR promises exciting new user experiences in domains ranging from training and education to games to everyday life.


Privacy Manifesto for AR Cloud Solutions

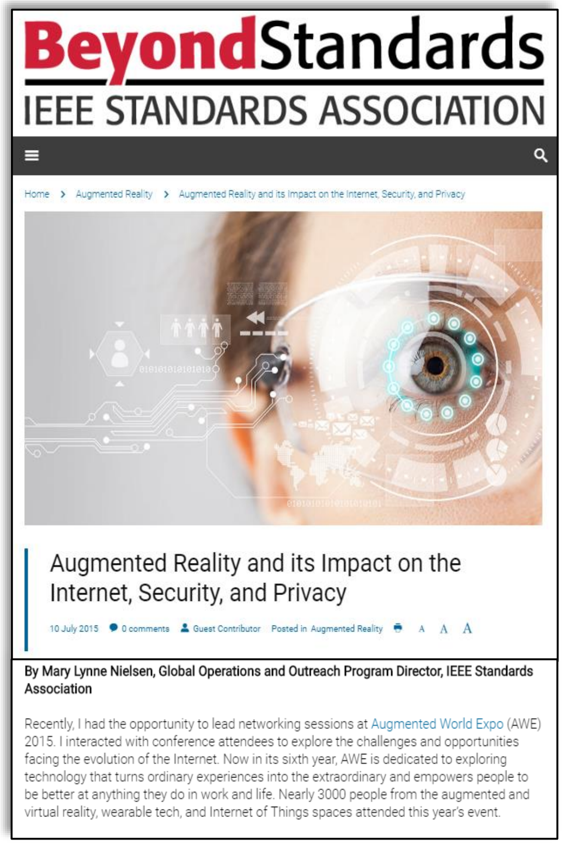
DRAFT v 0.1.3, October 18th at AWE EU Münich

Jan-Erik Vinje Follow

Oct 17, 2018 · 5 min read

The next era of computing is upon us. New technology to capture and combine physical spaces with digital content has the potential to profoundly improve the way we see and interact with the world and each other.

As the world's largest companies and organizations race to create the required "AR Cloud" infrastructure to build and fuel these systems, we face unprecedented new challenges and risks to privacy and individual's rights


BeyondStandards

IEEE STANDARDS ASSOCIATION

Home › Augmented Reality › Augmented Reality and its Impact on the Internet, Security, and Privacy

Augmented Reality and its Impact on the Internet, Security, and Privacy

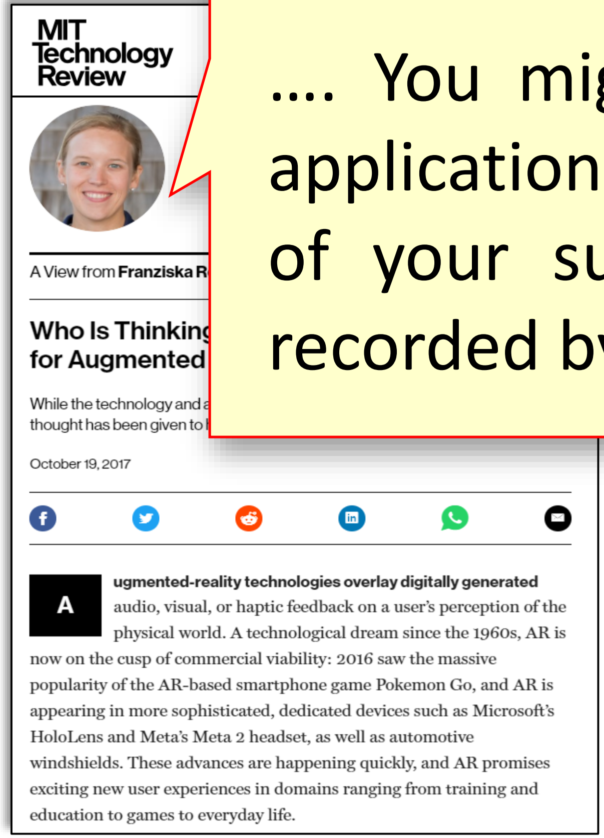
10 July 2015 · 0 comments · Guest Contributor · Posted in Augmented Reality

By Mary Lynne Nielsen, Global Operations and Outreach Program Director, IEEE Standards Association

Recently, I had the opportunity to lead networking sessions at Augmented World Expo (AWE) 2015. I interacted with conference attendees to explore the challenges and opportunities facing the evolution of the Internet. Now in its sixth year, AWE is dedicated to exploring technology that turns ordinary experiences into the extraordinary and empowers people to be better at anything they do in work and life. Nearly 3000 people from the augmented and virtual reality, wearable tech, and Internet of Things spaces attended this year's event.

https://beyondstandards.ieee.org/augmented-reality/augmented-reality-and-its-impact-on-the-internet-security-and-privacy/
https://www.technologyreview.com/s/609143/who-is-thinking-about-security-and-privacy-for-augmented-reality/
https://medium.com/openarcloud/privacy-manifesto-for-ar-cloud-solutions-9507543f50b6

New Privacy Concerns for AR


MIT Technology Review

A View from Franziska R

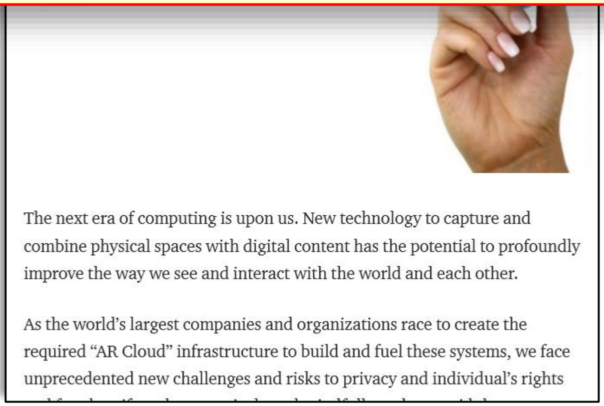
Who Is Thinking
for Augmented

While the technology and a
thought has been given to

October 19, 2017

Augmented-reality technologies overlay digitally generated audio, visual, or haptic feedback on a user's perception of the physical world. A technological dream since the 1960s, AR is now on the cusp of commercial viability: 2016 saw the massive popularity of the AR-based smartphone game Pokemon Go, and AR is appearing in more sophisticated, dedicated devices such as Microsoft's HoloLens and Meta's Meta 2 headset, as well as automotive windshields. These advances are happening quickly, and AR promises exciting new user experiences in domains ranging from training and education to games to everyday life.

.... You might find it a bit creepy that the device and its applications have access to a constant video and audio feed of your surroundings, not to mention that you're being recorded by other people's devices. – *F. Roesner 2017*


The next era of computing is upon us. New technology to capture and combine physical spaces with digital content has the potential to profoundly improve the way we see and interact with the world and each other.

As the world's largest companies and organizations race to create the required "AR Cloud" infrastructure to build and fuel these systems, we face unprecedented new challenges and risks to privacy and individual's rights


Augmented Reality and its Impact on the Internet, Security, and Privacy

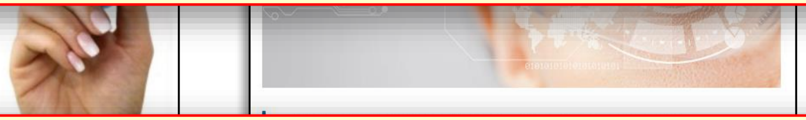
10 July 2015 · 0 comments · Guest Contributor · Posted in Augmented Reality

By Mary Lynne Nielsen, Global Operations and Outreach Program Director, IEEE Standards Association

Recently, I had the opportunity to lead networking sessions at Augmented World Expo (AWE) 2015. I interacted with conference attendees to explore the challenges and opportunities facing the evolution of the Internet. Now in its sixth year, AWE is dedicated to exploring technology that turns ordinary experiences into the extraordinary and empowers people to be better at anything they do in work and life. Nearly 3000 people from the augmented and virtual reality, wearable tech, and Internet of Things spaces attended this year's event.

https://beyondstandards.ieee.org/augmented-reality/augmented-reality-and-its-impact-on-the-internet-security-and-privacy/
https://www.technologyreview.com/s/609143/who-is-thinking-about-security-and-privacy-for-augmented-reality/
https://medium.com/openarcloud/privacy-manifesto-for-ar-cloud-solutions-9507543f50b6

New Privacy Concerns for AR



…. You might find it a bit creepy that the device and its applications have access to a constant video and audio feed of your surroundings, not to mention that you're being recorded by other people's devices. – *F. Roesner 2017*

…….. the risk that such *(private)* data could be collected, analyzed, transmitted and stored in databases or distributed and sold to third parties without the explicit consent of users or worse, unsuspecting citizens that happen to be within sensor range of mixed reality enabled devices. – *Jan-Erik Vinje 2018*

Outline

- Revealing Scenes by Inverting Structure from Motion Reconstructions

  Francesco Pittaluga, Sanjeev Koppal, Sing Bing Kang and Sudipta N. Sinha

  CVPR 2019

- Privacy-Preserving Image-based Localization

  Pablo Speciale, Johannes L. Schönberger, Sing Bing Kang, Sudipta N. Sinha and Marc Pollefeys

  CVPR 2019

- Privacy-Preserving Image Queries for Camera Localization

  Pablo Speciale, Johannes L. Schönberger, Sudipta N. Sinha and Marc Pollefeys

  ICCV 2019

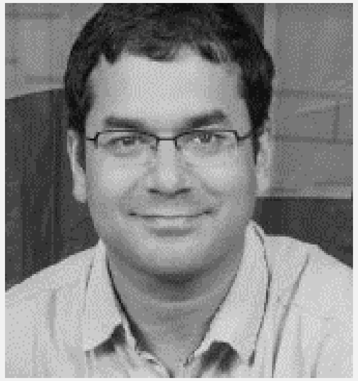
Revealing Scenes by Inverting Structure from Motion Reconstructions

CVPR 2019


Francesco Pittaluga [1]

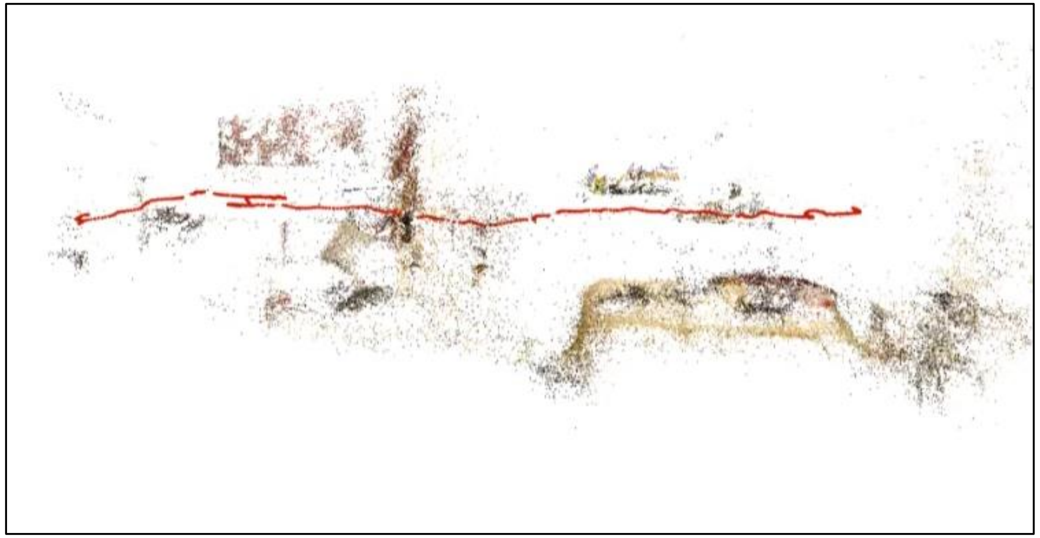

Sanjeev Koppal [1]


Sing Bing Kang[2]

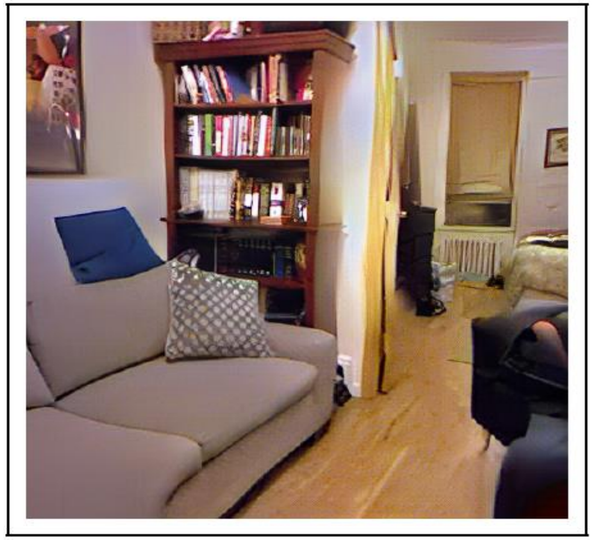

Sudipta N. Sinha[2]

[1] University of Florida [2] Microsoft Research

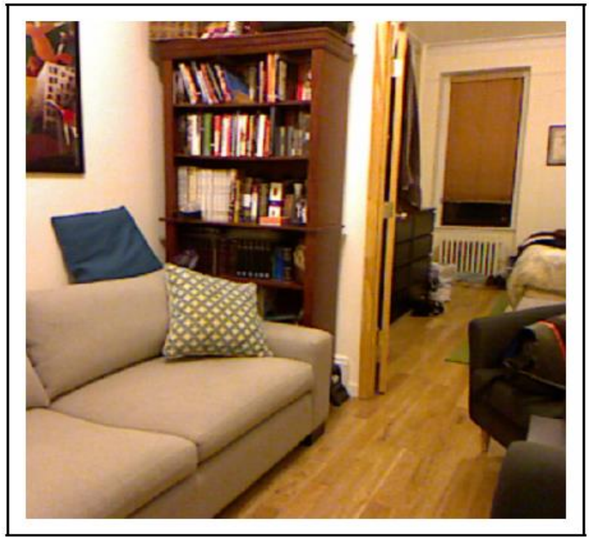
New Privacy Attack on 3D Maps



3D point cloud map



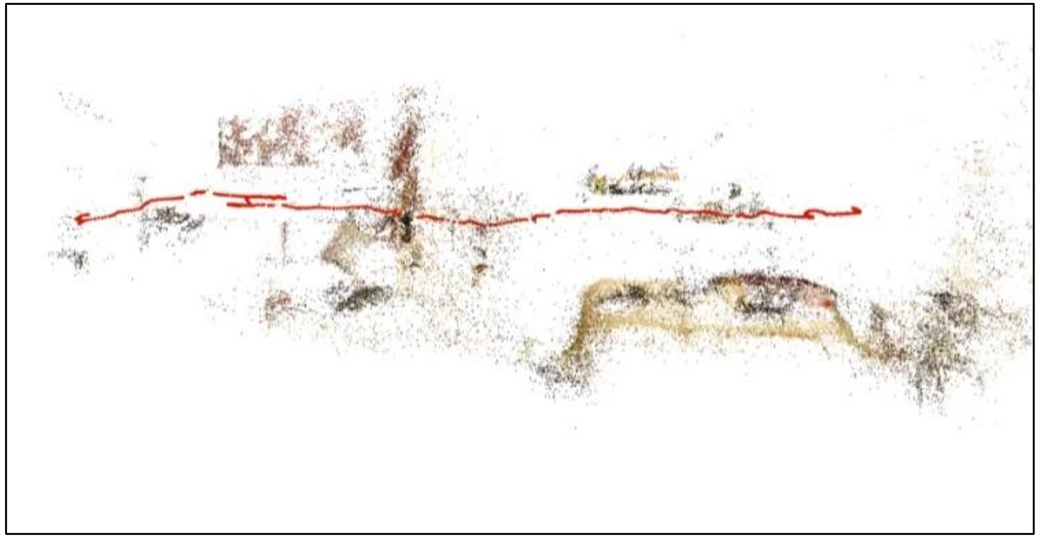
Our Result



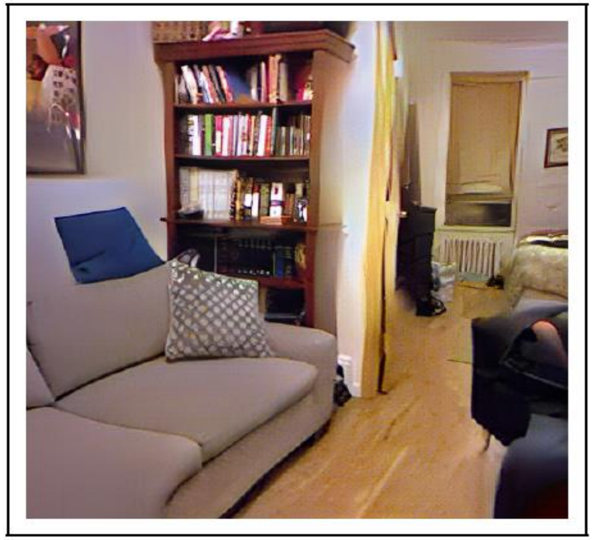
Original

3D Point Cloud Map Inversion

New Privacy Attack on 3D Maps



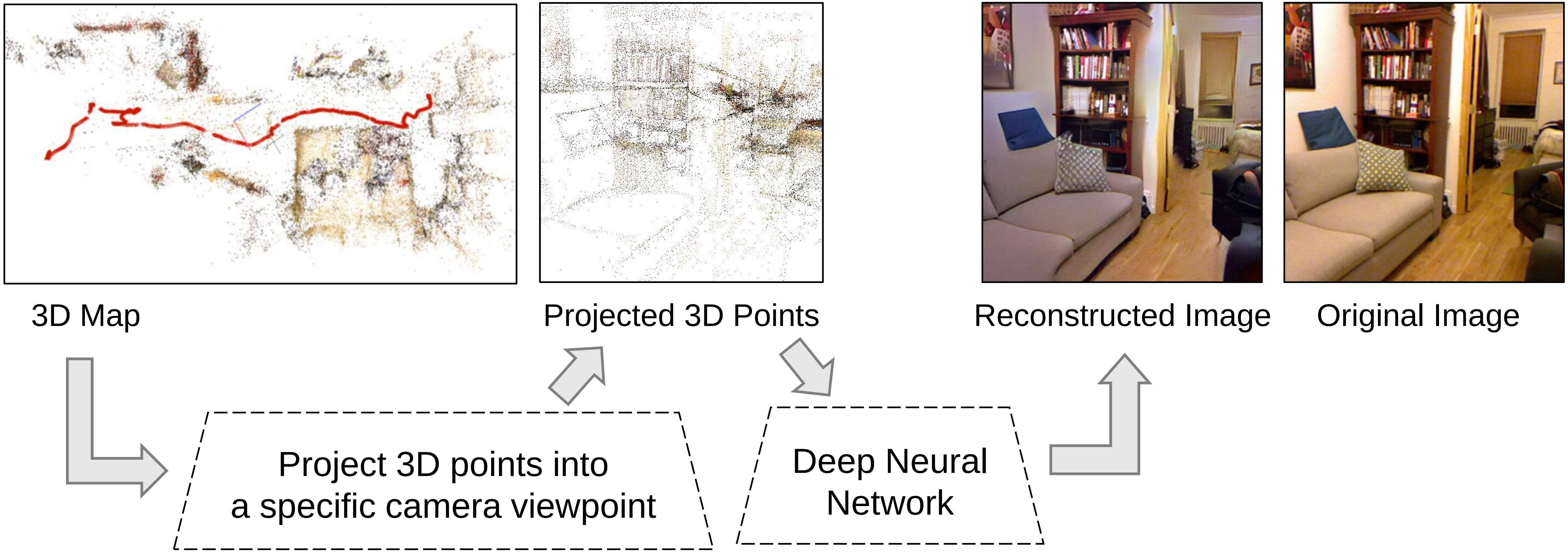
3D point cloud map



Our Result

3D Point Cloud Map Inversion

Problem Definition: 3D Map Inversion



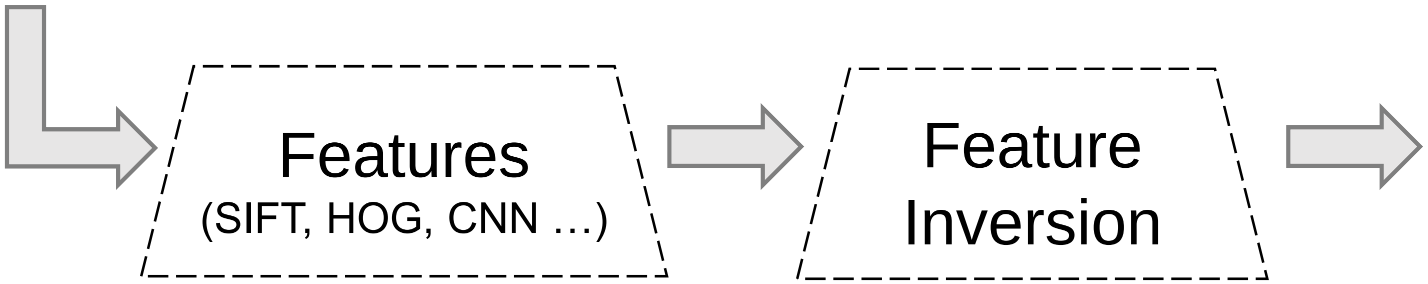
Specifically, the attacker's goal is to reconstruct a color image of a scene from 2D projections of sparse 3D points and descriptors. We assume that the attacker will do that using a deep neural network.

Previous Work: Single Image Feature Inversion

Weinzaepfel et al. 2011
Vondrick et al. 2013
Kato & Harada, 2014
….
Dosovitskiy & Brox, CVPR 2016
Dosovitskiy & Brox, NIPS 2016

Original image

Visualization of SIFT Keypoints

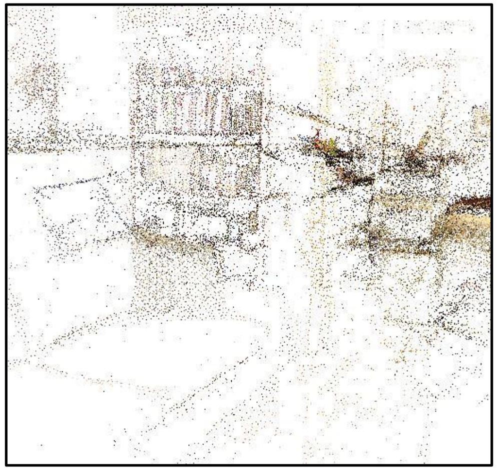


Reconstructed Image

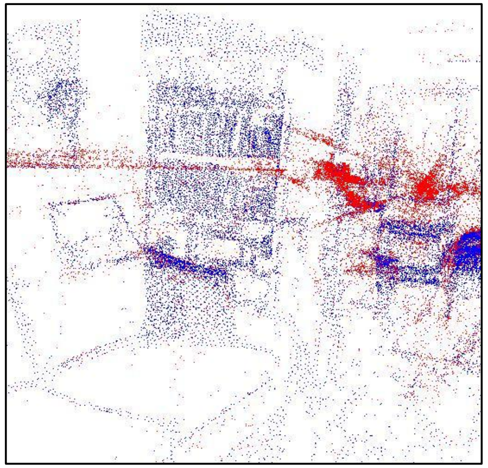
3D Map Inversion: Challenges

- Visibility of 3D points unknown
- All feature attributes are not stored; hence unavailable for inversion
  - SIFT keypoint orientation
  - SIFT keypoint scale
  - SIFT descriptor image source
- 3D point cloud distributions are often quite sparse and irregular

3D Map



Projected 3D Points

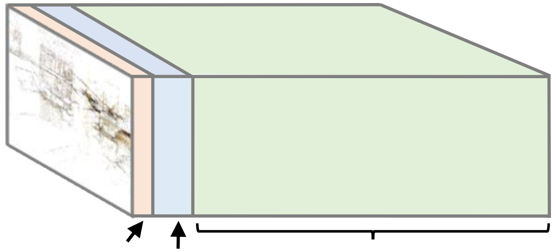


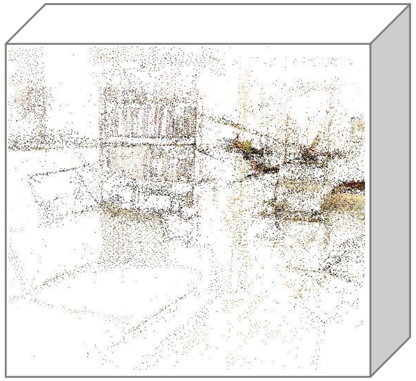
Visibility Map

U-Net Architecture

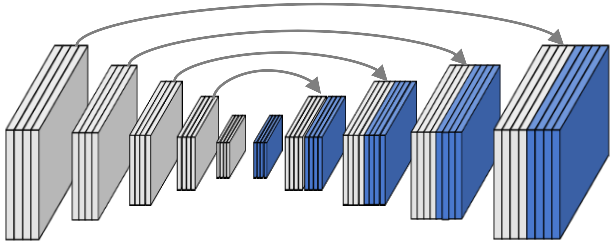
- Model: U-Net with skip connections
- Loss: Reconstruction Loss (L1)
- Dataset: SFM pre-processing on
  - MegaDepth (Li and Snavely, 2018)
  - NYU v2 (Silberman et al. 2012)
- Initialization: Random weights

nD
Input Tensor





nD Input



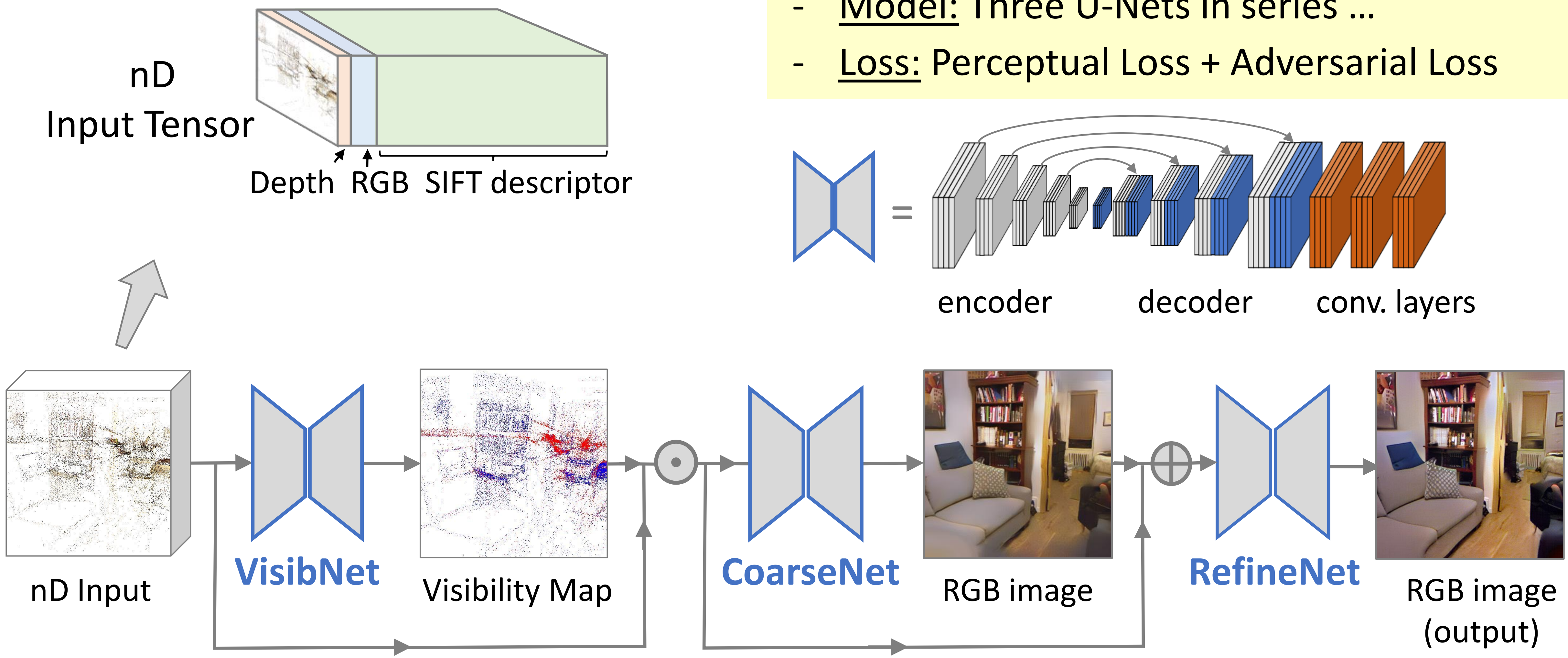
CoarseNet

(architecture similar to Dosovitskiy and Brox, CVPR 2016)

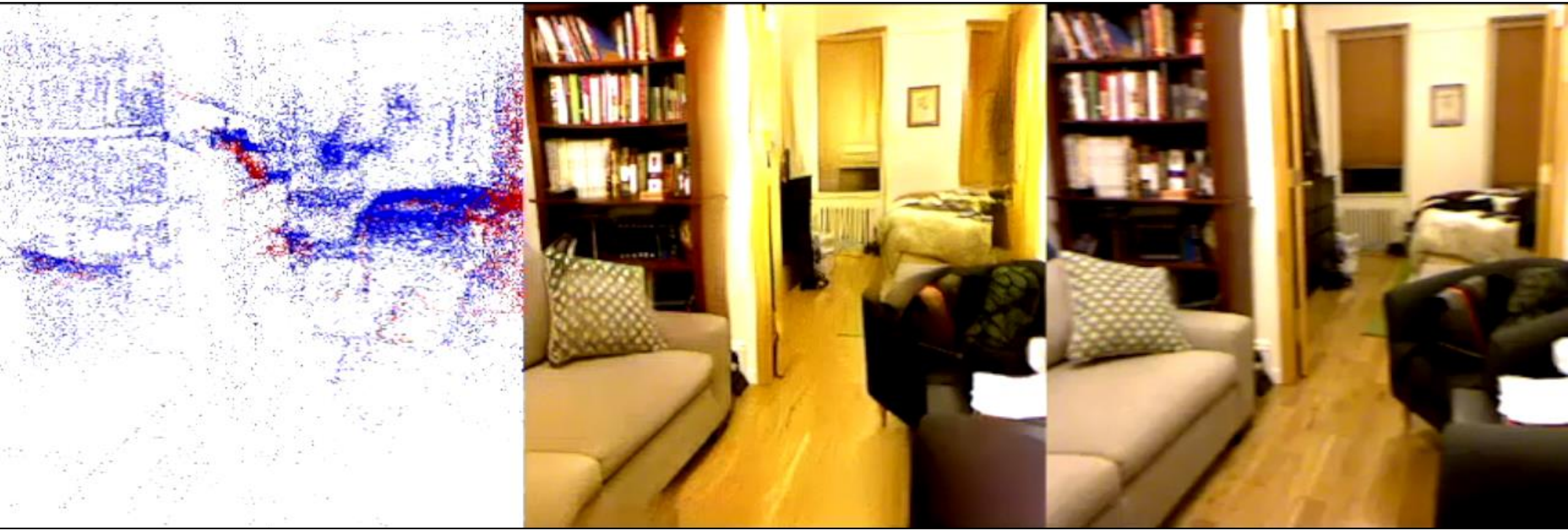
RGB image (output)

Final Network Architecture

- <u>Model:</u> Three U-Nets in series …
- <u>Loss:</u> Perceptual Loss + Adversarial Loss

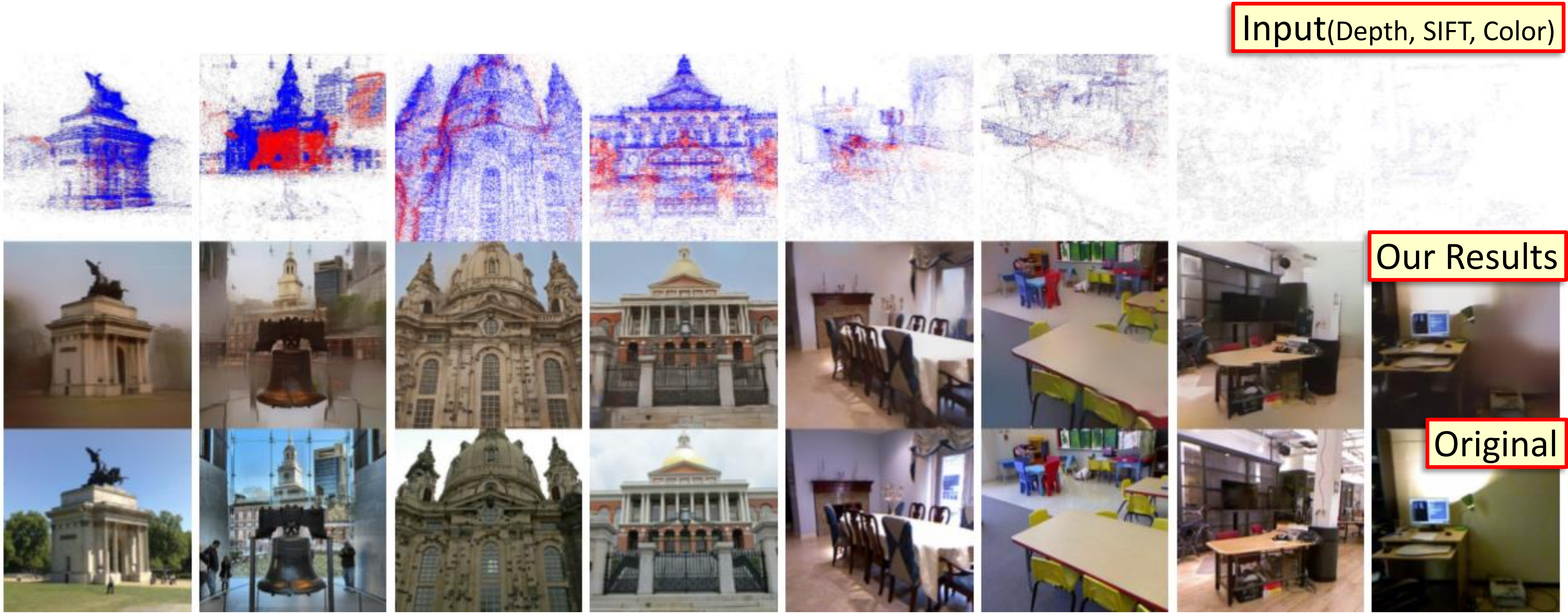


Reconstruction of Source Video used in Mapping

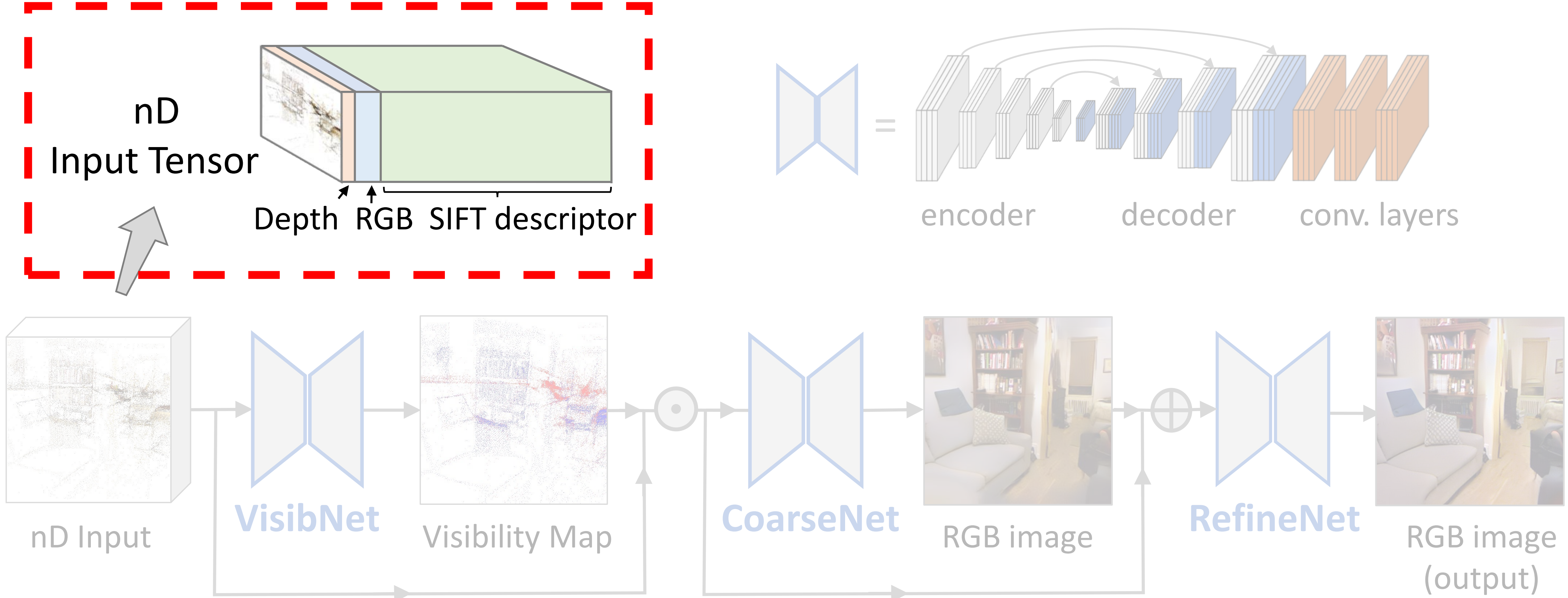


Novel Views Rendered from a Virtual Camera Path

Results



Effect of Input Attributes

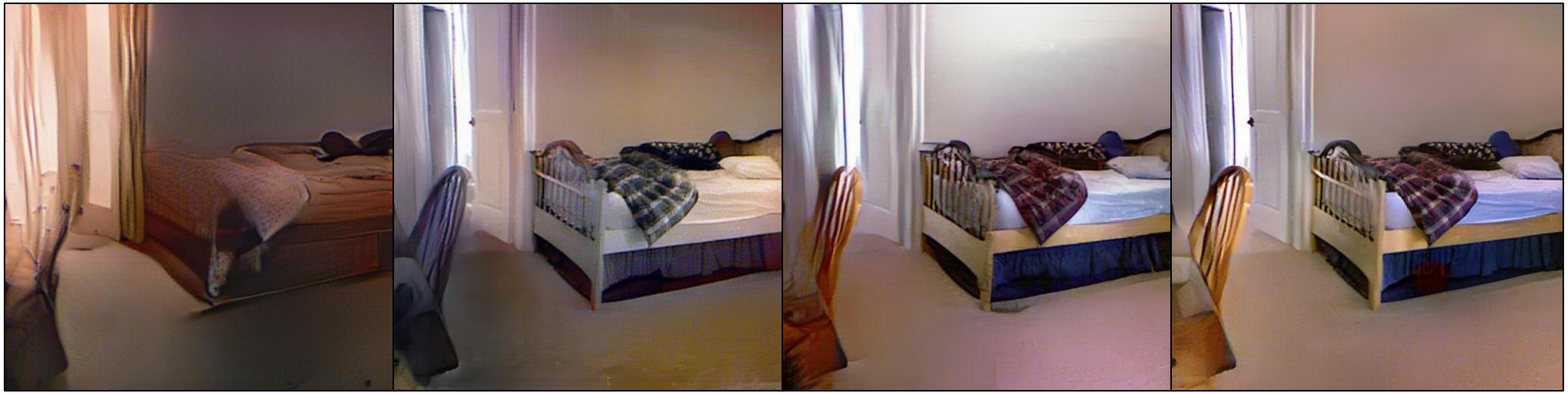


Effect of Input Attributes

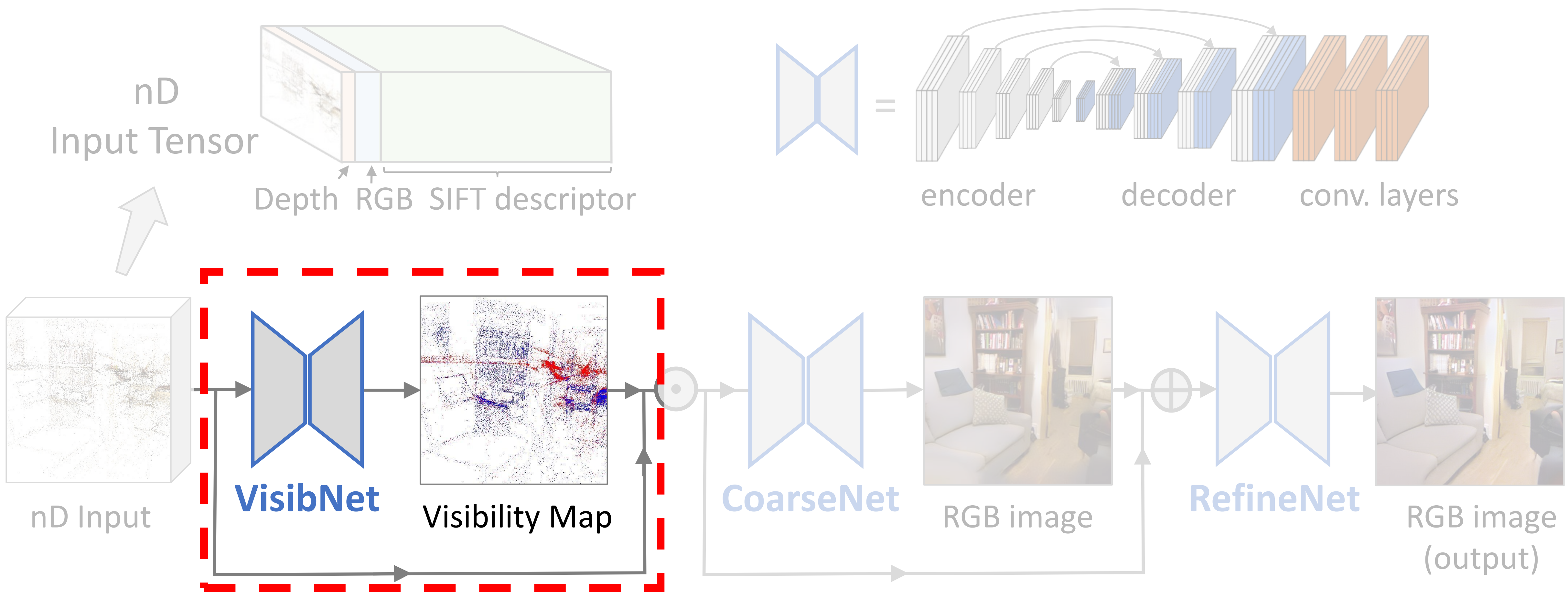
Network Input:	Only Depth	Depth + SIFT	Depth + RGB	Depth + SIFT +RGB

Example 1
(Bathroom Scene)

Example 2
(Bedroom Scene)



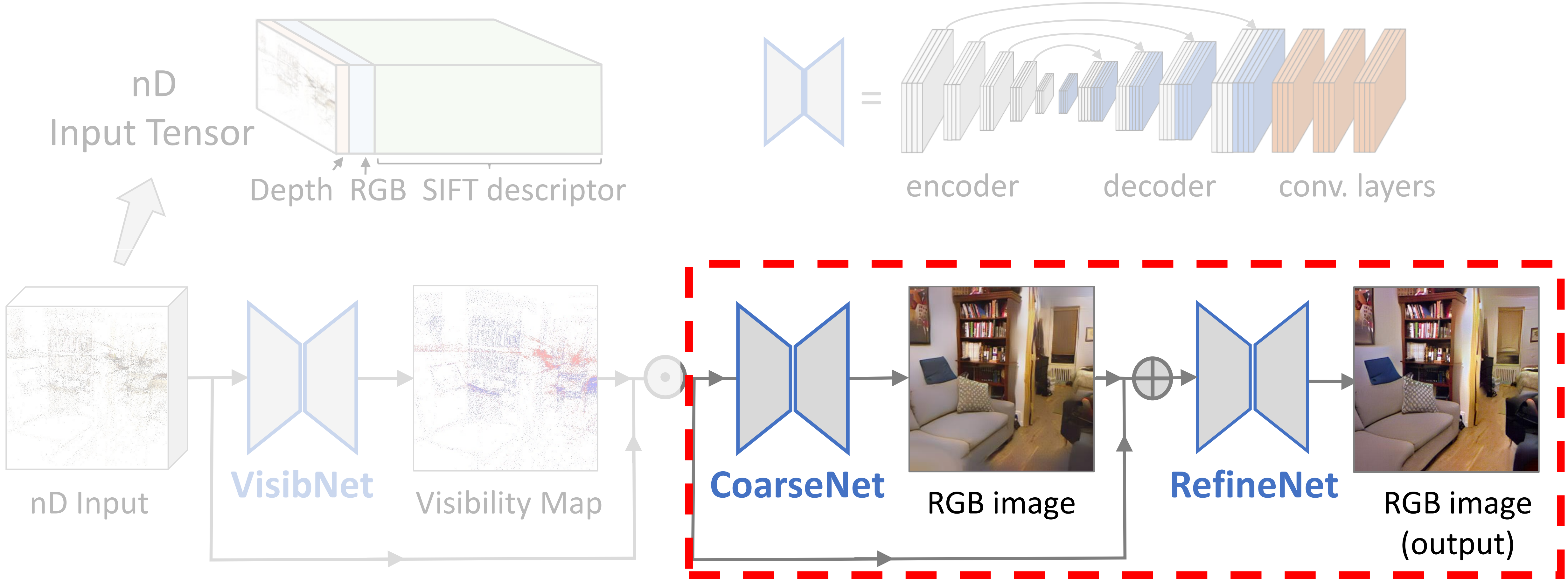
Importance of VisibNet



Importance of VisibNet

output of VisibNet (red: predicted as occluded) | without VisibNet | with VisibNet

Importance of RefineNet



Importance of RefineNet

Input: Depth + SIFT

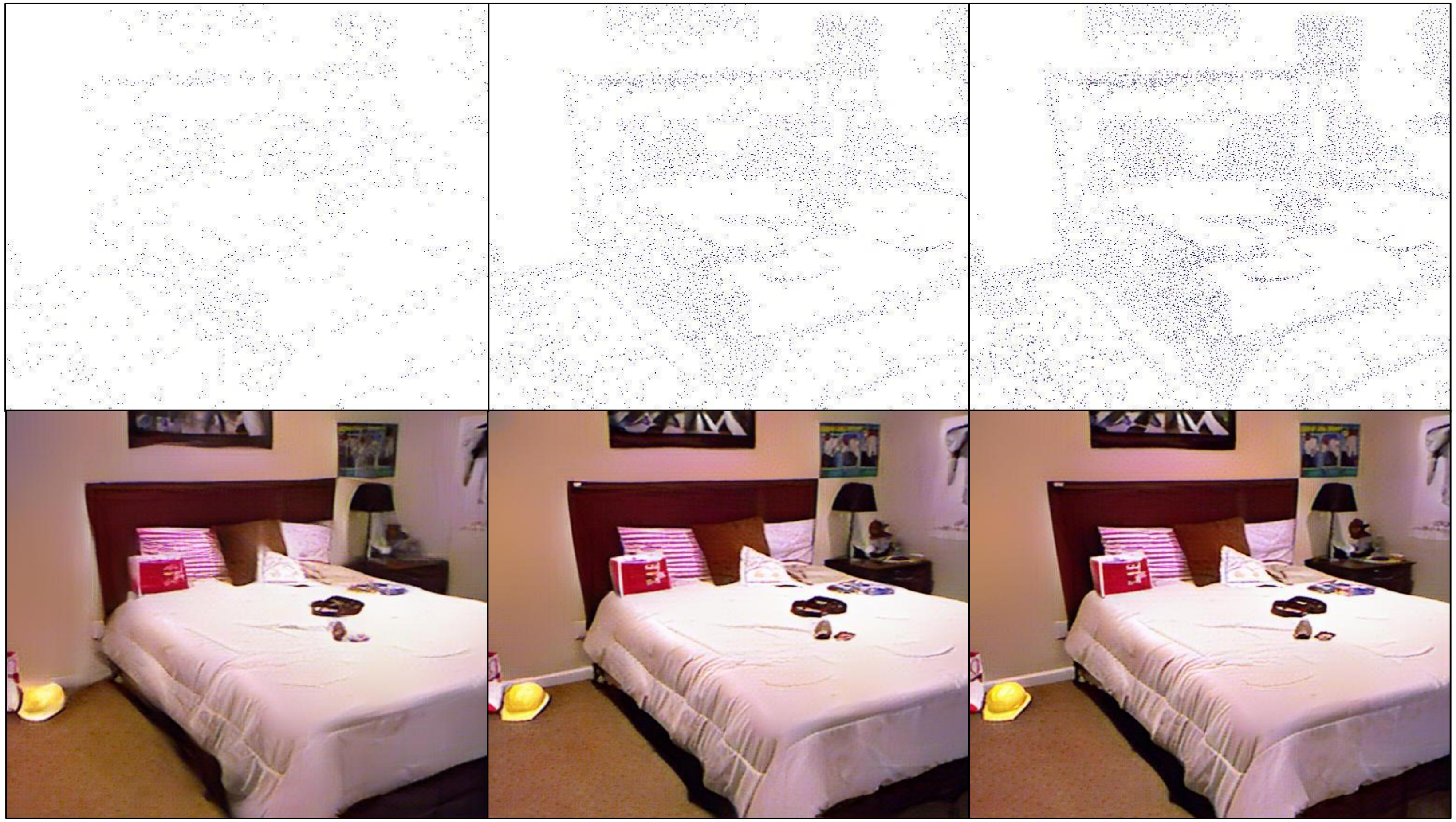
Output of CoarseNet

Output of RefineNet

Effect of Input Sparsity

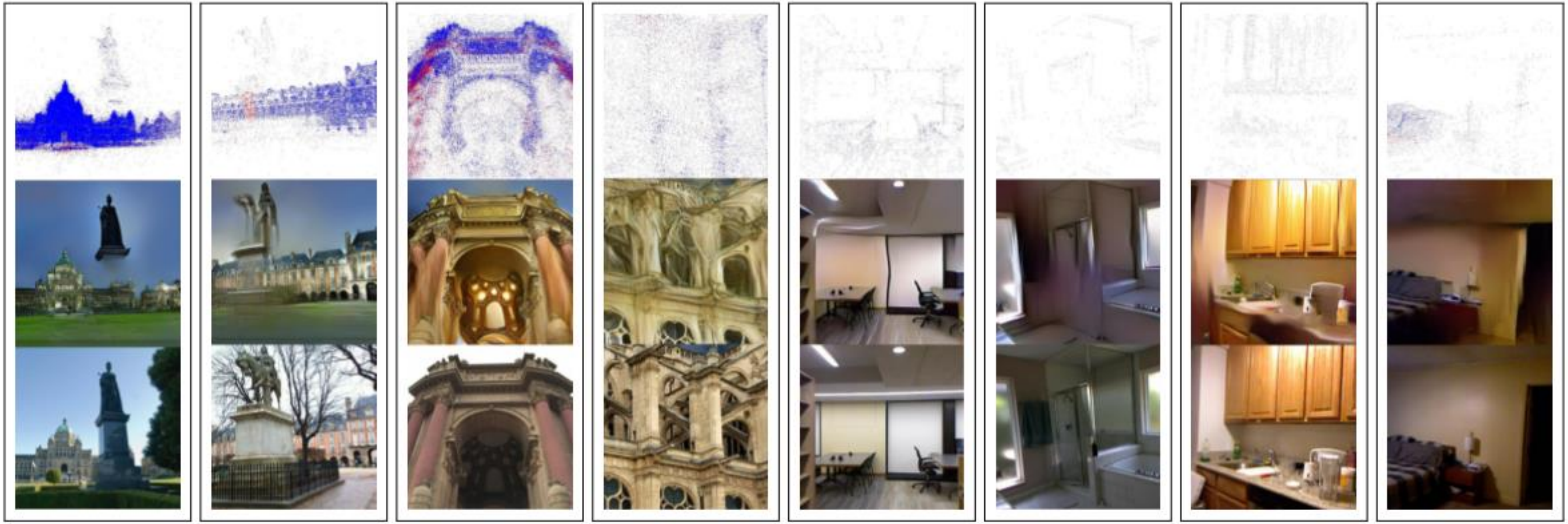
Visualization of Input Sparsity

Reconstruction Results



Input Sparsity (% of SFM points)	**20%**	**60%**	**100%**

Failures Cases and Artifacts



- Incorrect visibility estimation (foreground objects disappear)
- Straight lines becomes wavy
- Highly occluded scenes are difficult
- Phantom structures and erroneous 3D points in point cloud

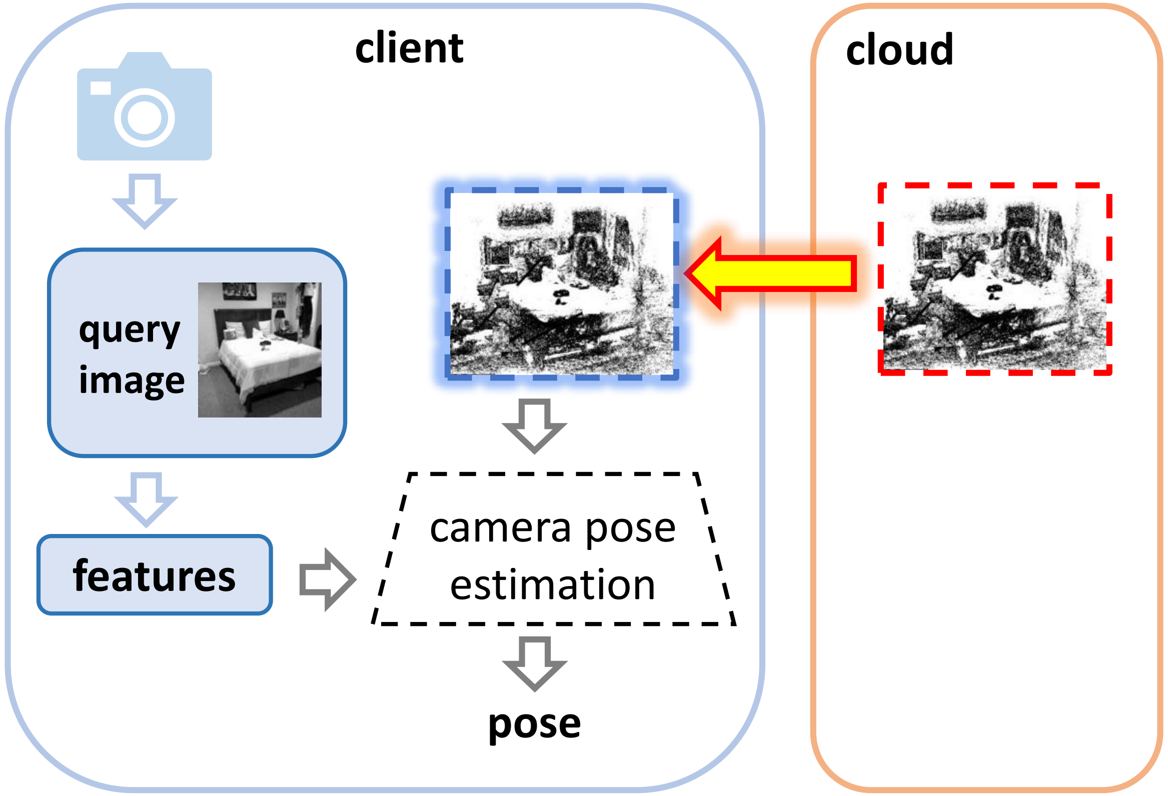
Conclusions

- We show that detailed images can be recovered from SFM point clouds, such as those used for camera localization.
- The attack seems seem quite effective even when very little information is available.
- Empirical analysis and ablation studies.

Our work highlights **potential privacy implications as spatial mapping and localization for AR, Robotics** etc. becomes widely adopted in homes, workplaces, other sensitive environments.

Privacy Implications for Camera Localization

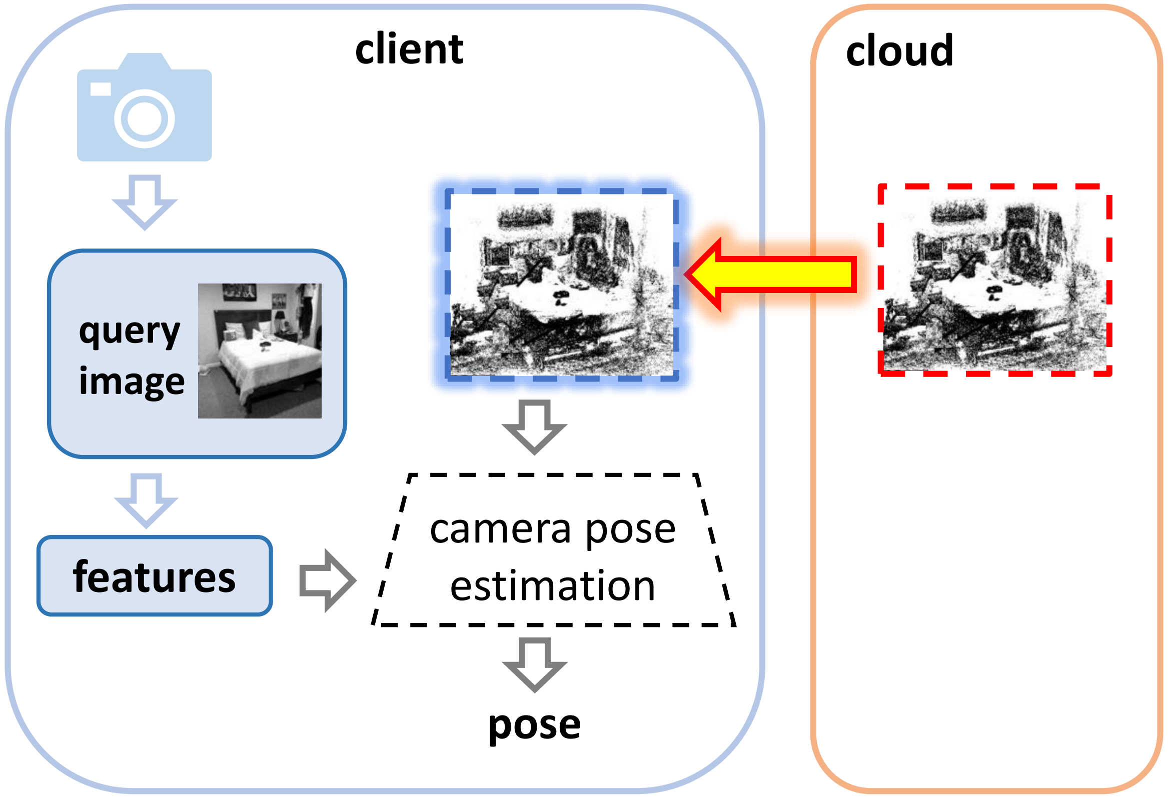
Processing on Client



- Server shares map with clients
- Privacy of map data is a concern

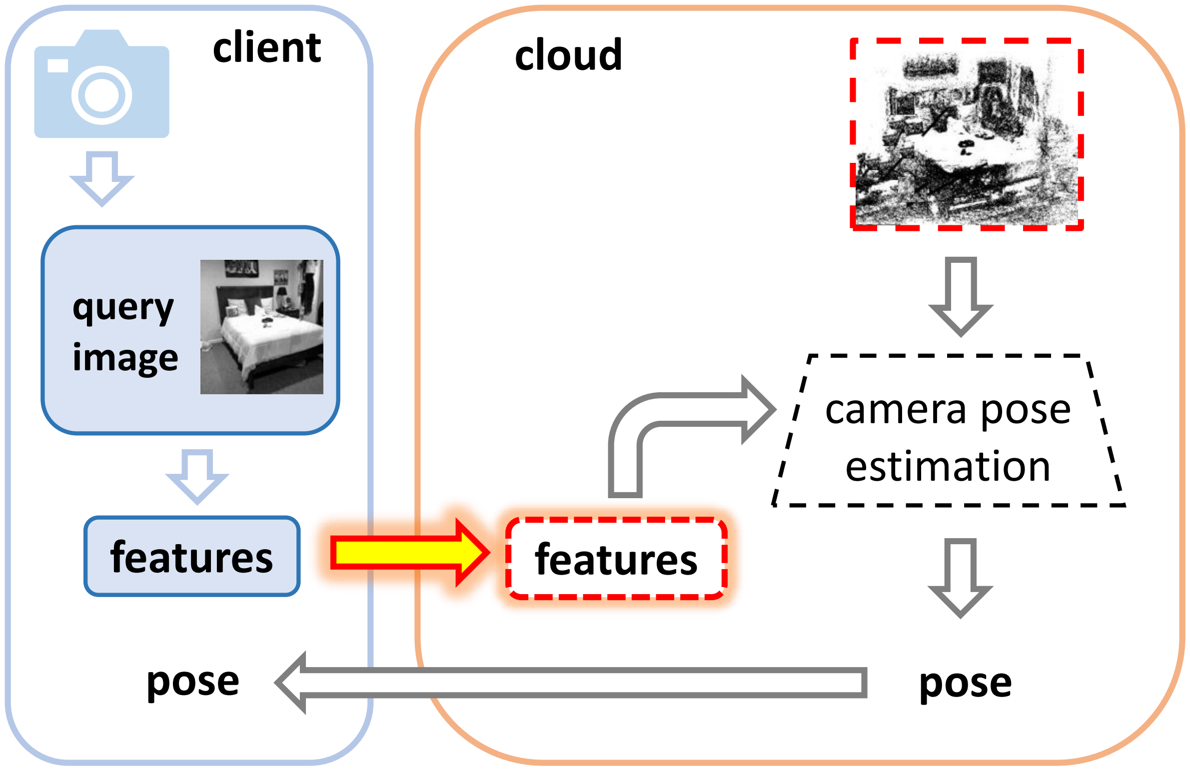
Privacy Implications for Camera Localization

Processing on Client



- Server shares map with clients
- Privacy of map data is a concern

Processing on Server

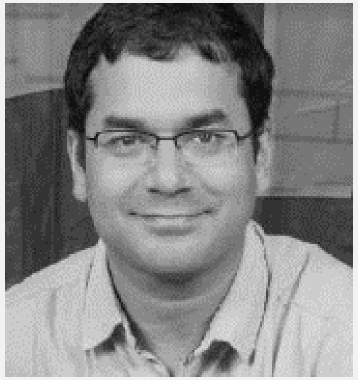


- Client shares image features with server
- Privacy of query image is a concern

Privacy-Preserving Image-based Localization

CVPR 2019




Pablo Speciale[1]

Johannes L. Schönberger[1]

Sing Bing Kang[2]

Sudipta N. Sinha[2]

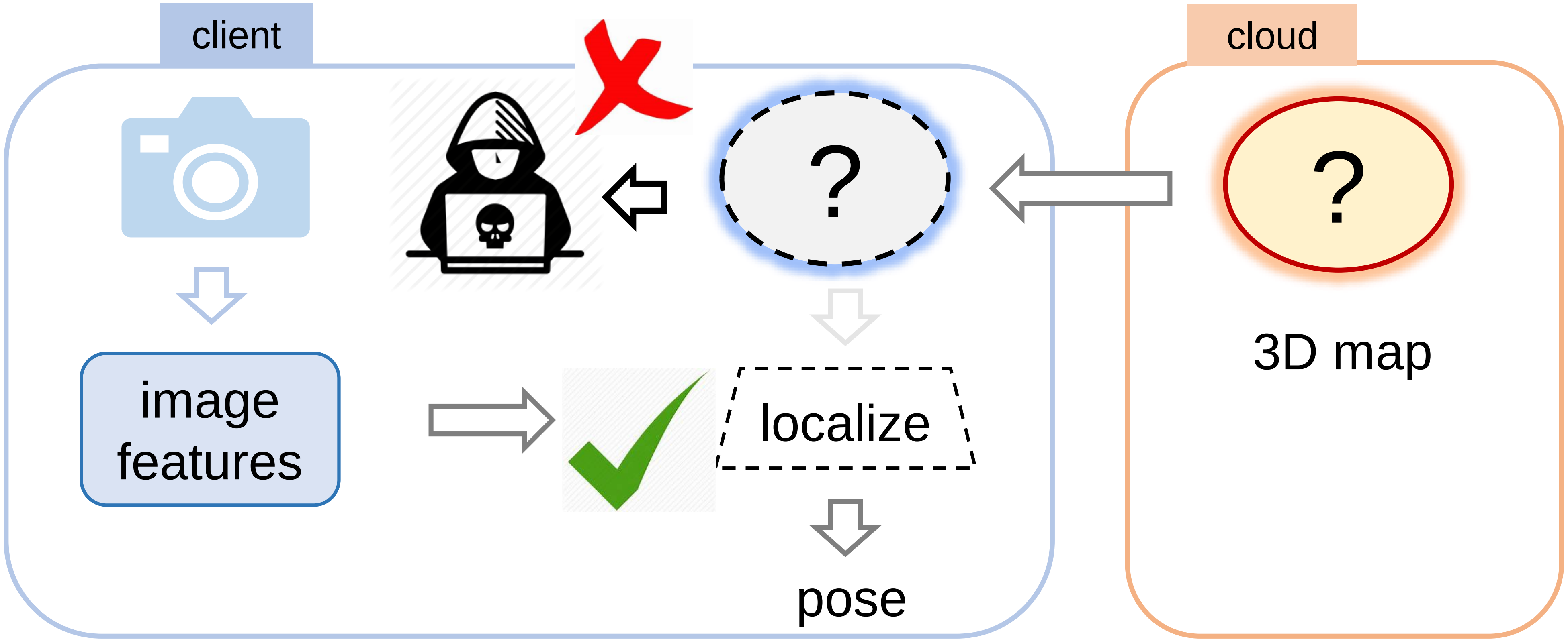
Marc Pollefeys[1,3]

[1] Microsoft Mixed Reality & AI Group, Zurich

[2] Microsoft Research Redmond

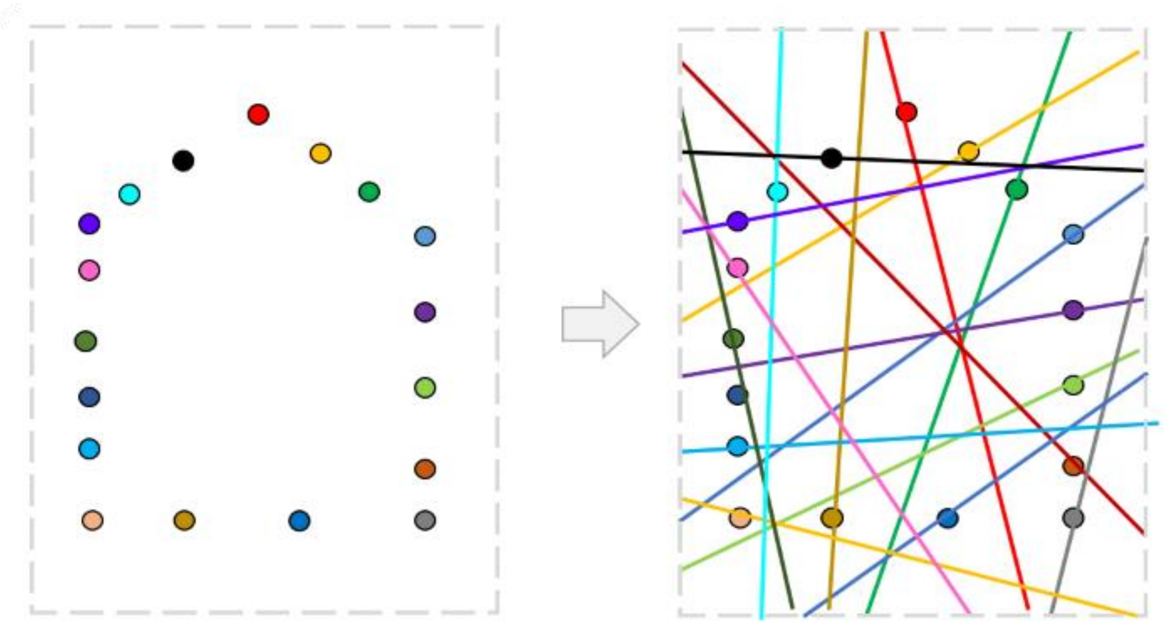
[3] ETH Zurich

Goal: Keep the Map Confidential



1. Conceal the 3D map; prevent inversion attacks.
2. Yet, somehow allow camera pose estimation!

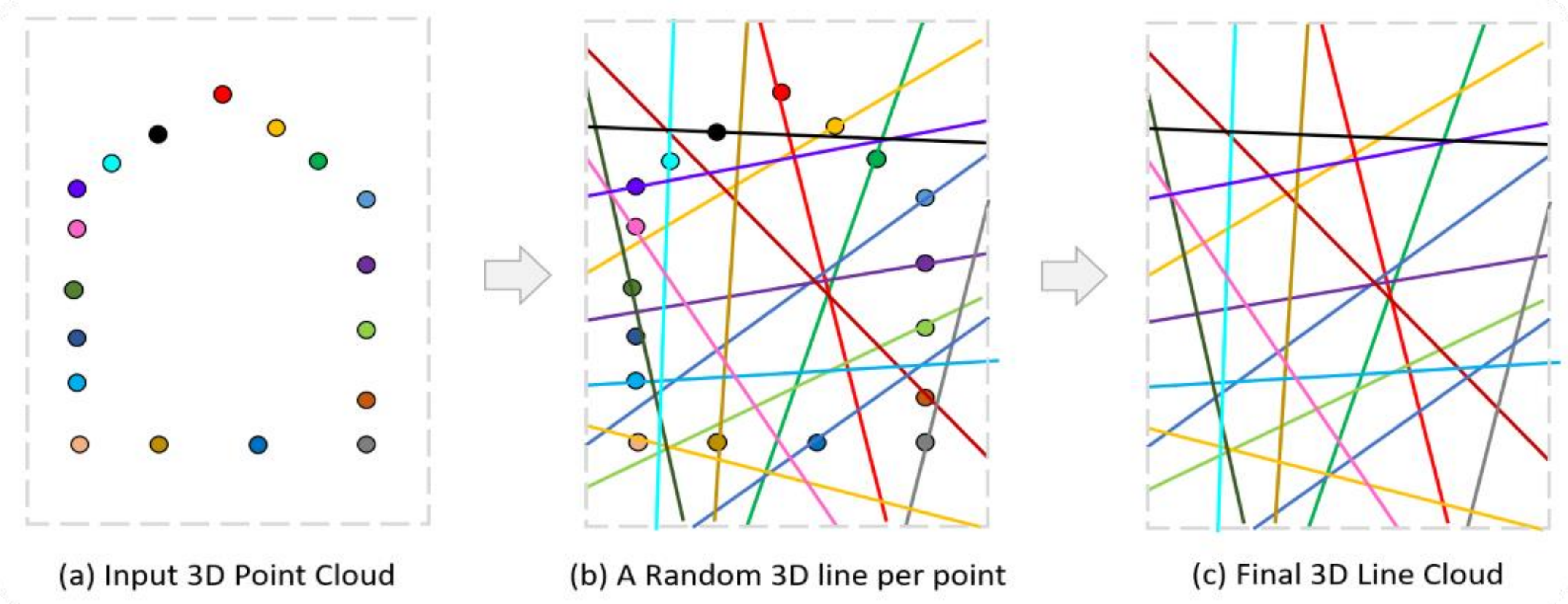
New Map Representation



(a) Input 3D Point Cloud (b) A Random 3D line per point

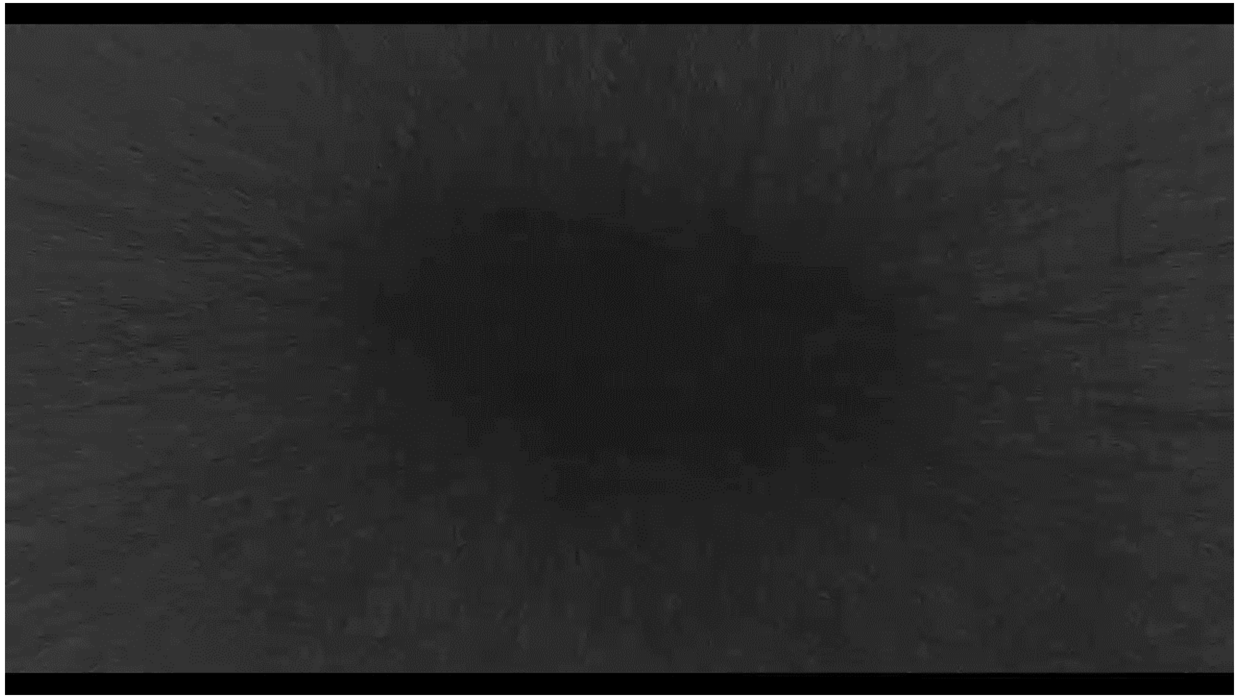
- For each 3D point, pick a randomly oriented 3D line passing through the point.

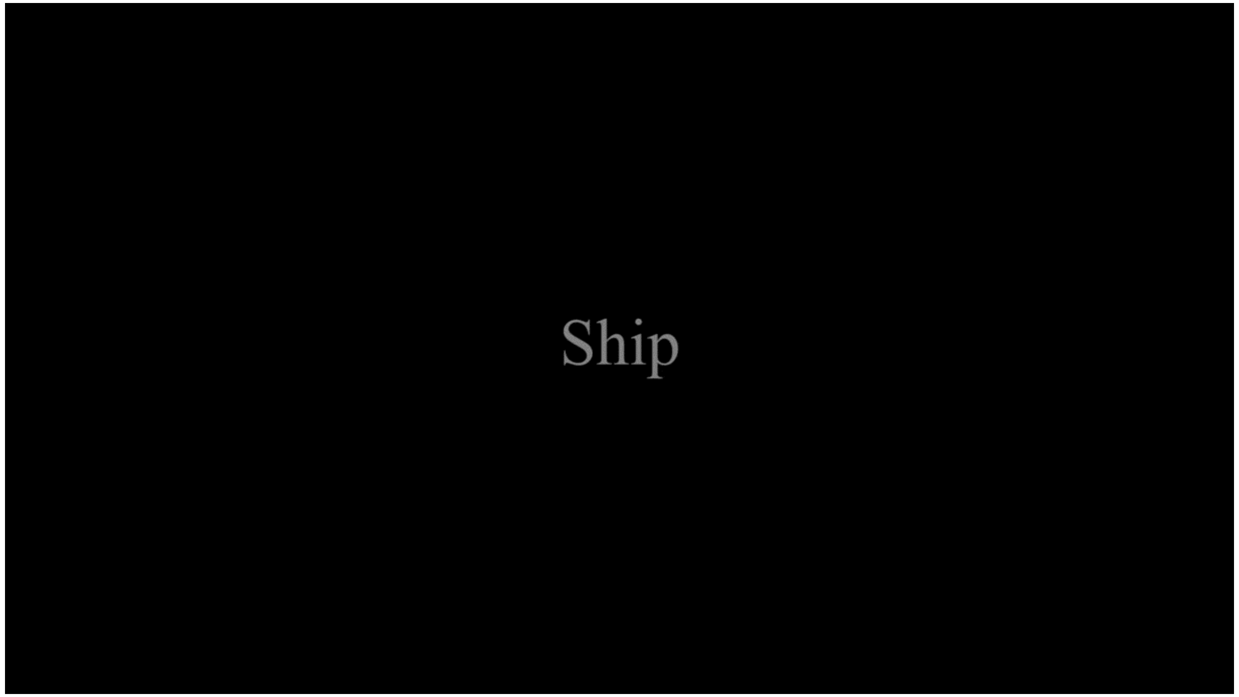
New Map Representation



- For each 3D point, pick a randomly oriented 3D line passing through the point.
- Then **discard** the 3D point.

Proposed Idea

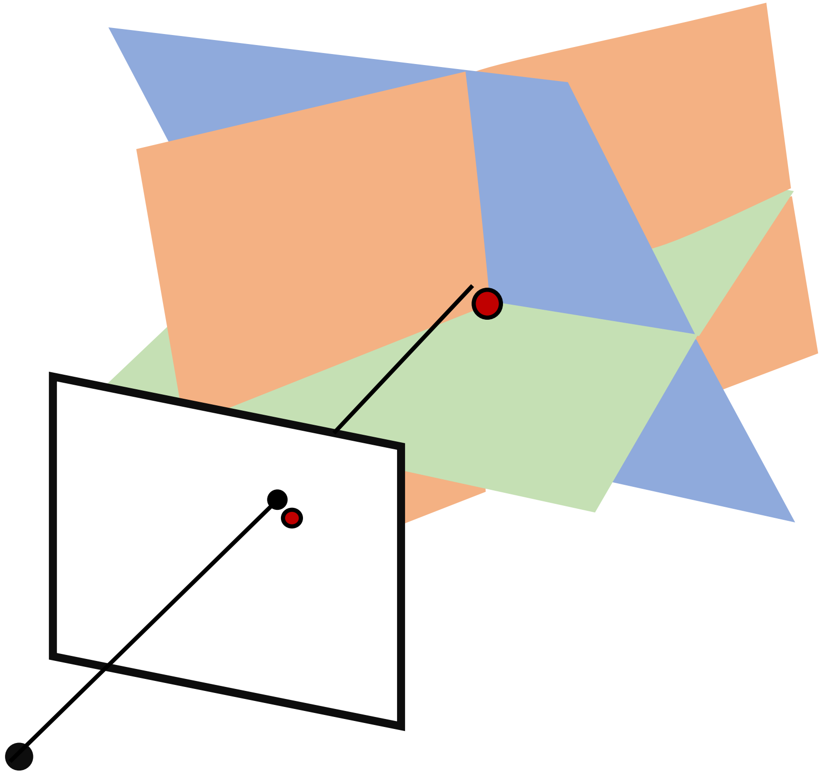




- For each 3D point, pick a randomly oriented 3D line passing through the point.
- Then **discard** the 3D point.

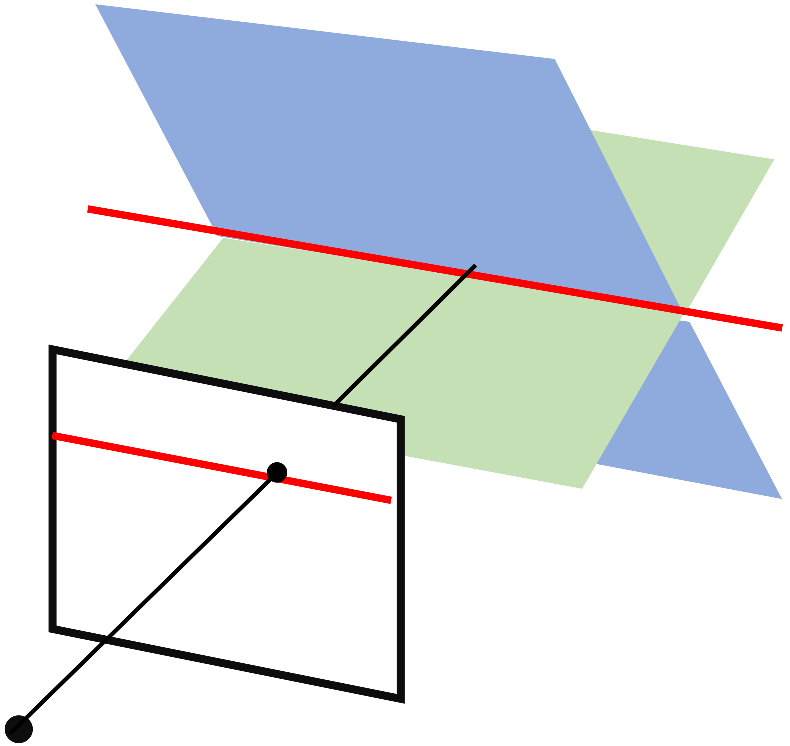
Key Insight for Camera Pose Estimation

3D point as intersection of *three planes*



Two reprojection constraints per correspondence

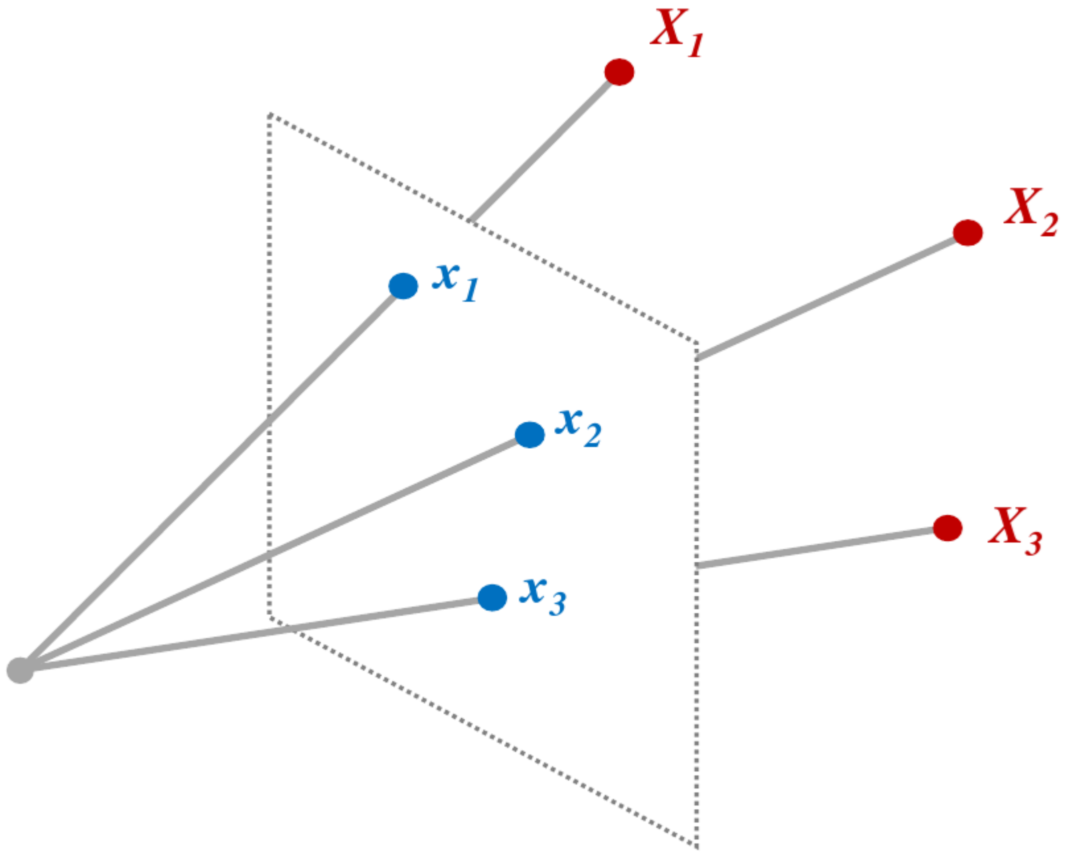
3D point as intersection of ***two planes***



One re-projection constraint per correspondence

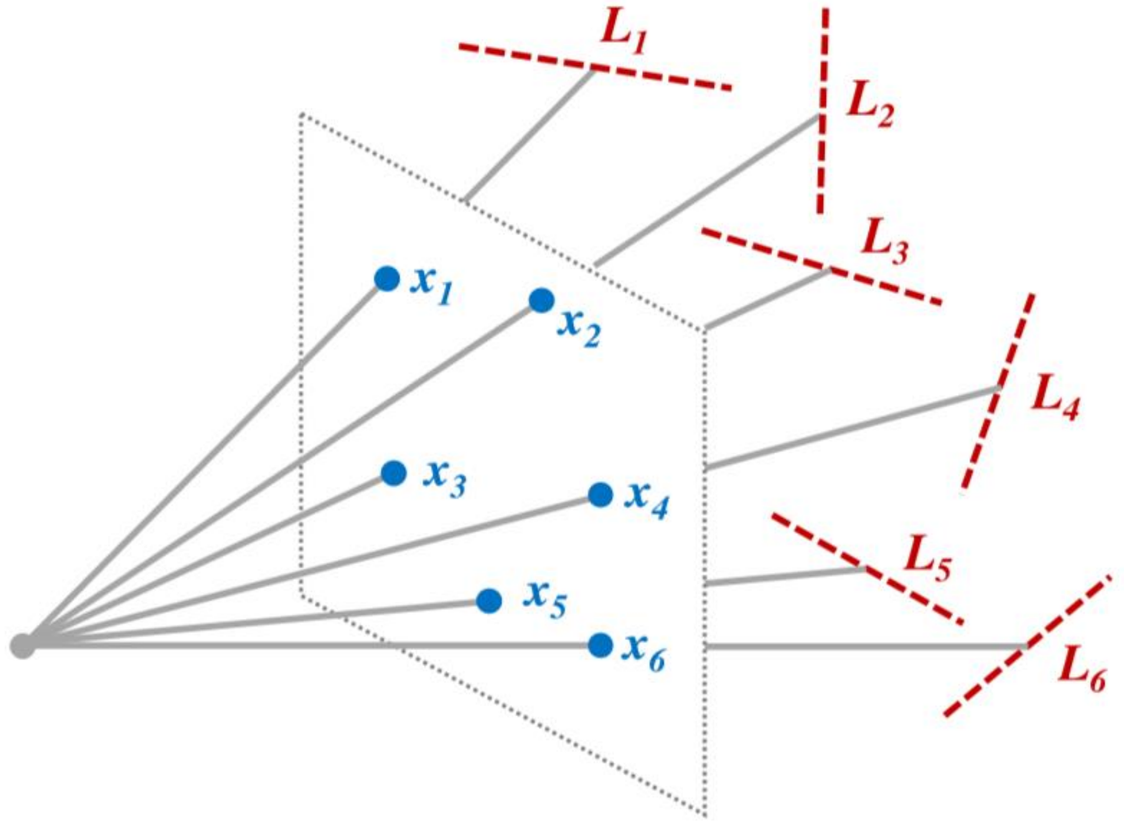
Camera Pose Estimation

Traditional Method ($p3P$)



Three "image point"–3D point correspondences

Proposed Method ($p6L$)



Six "image point"–3D line correspondences

Camera Pose Estimation

1. Our minimal problem can be cast as **generalized relative pose problem** [1].
   [1] Stewenius et al. 2005

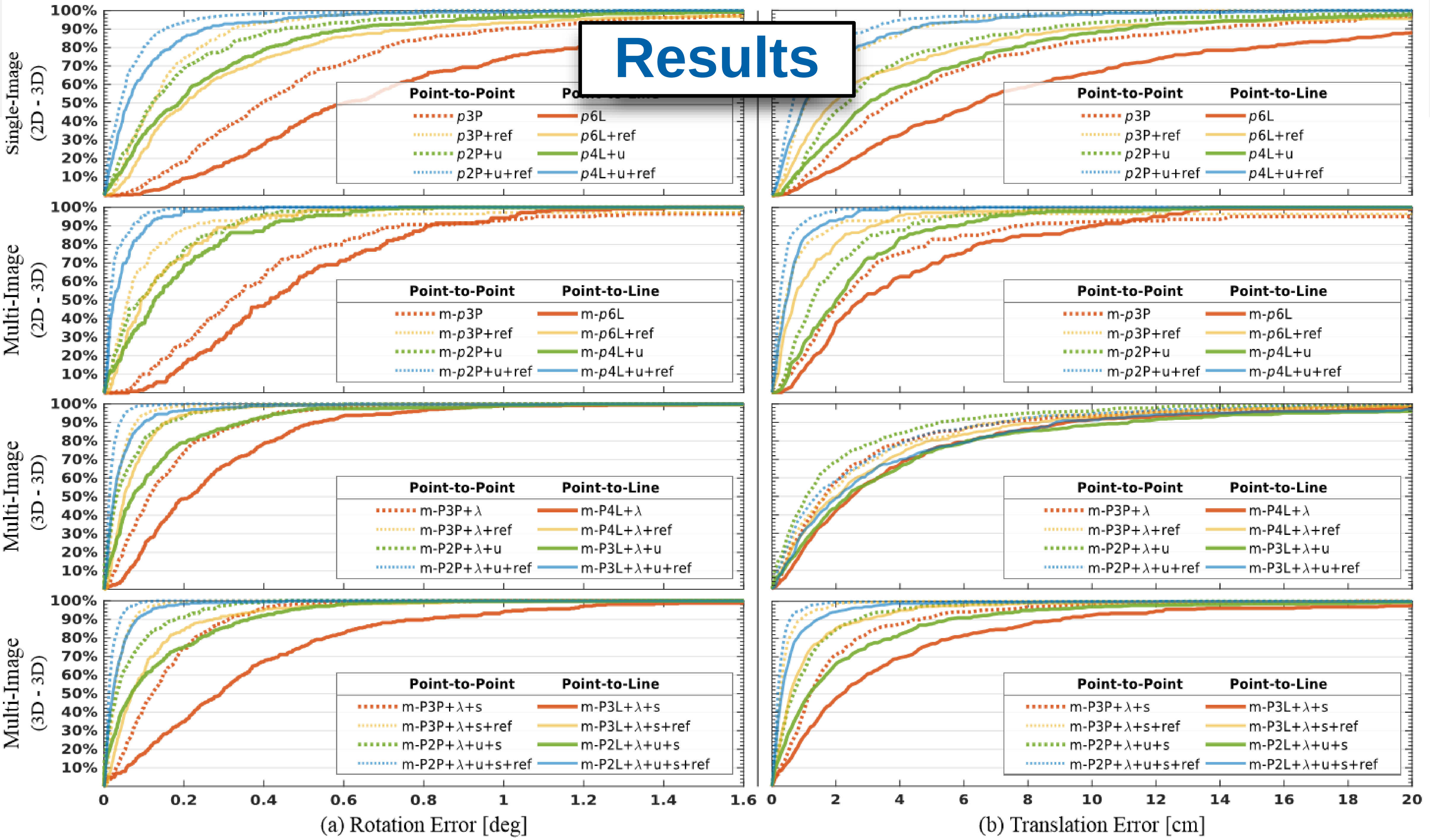
2. Proposed **several variants** with:
   - Query 3D point cloud (from multiple images),
   - known vertical direction,
   - known scale.

   We leverage existing minimal solvers [2-7].

   [2] Nister et al. 2007 [3] Lee et al. 2014 [4] Stewenius et al. 2005
   [5] Sweeney et al. 2015a [6] Sweeney et al. 2015b [7] Sweeney et al. 2014

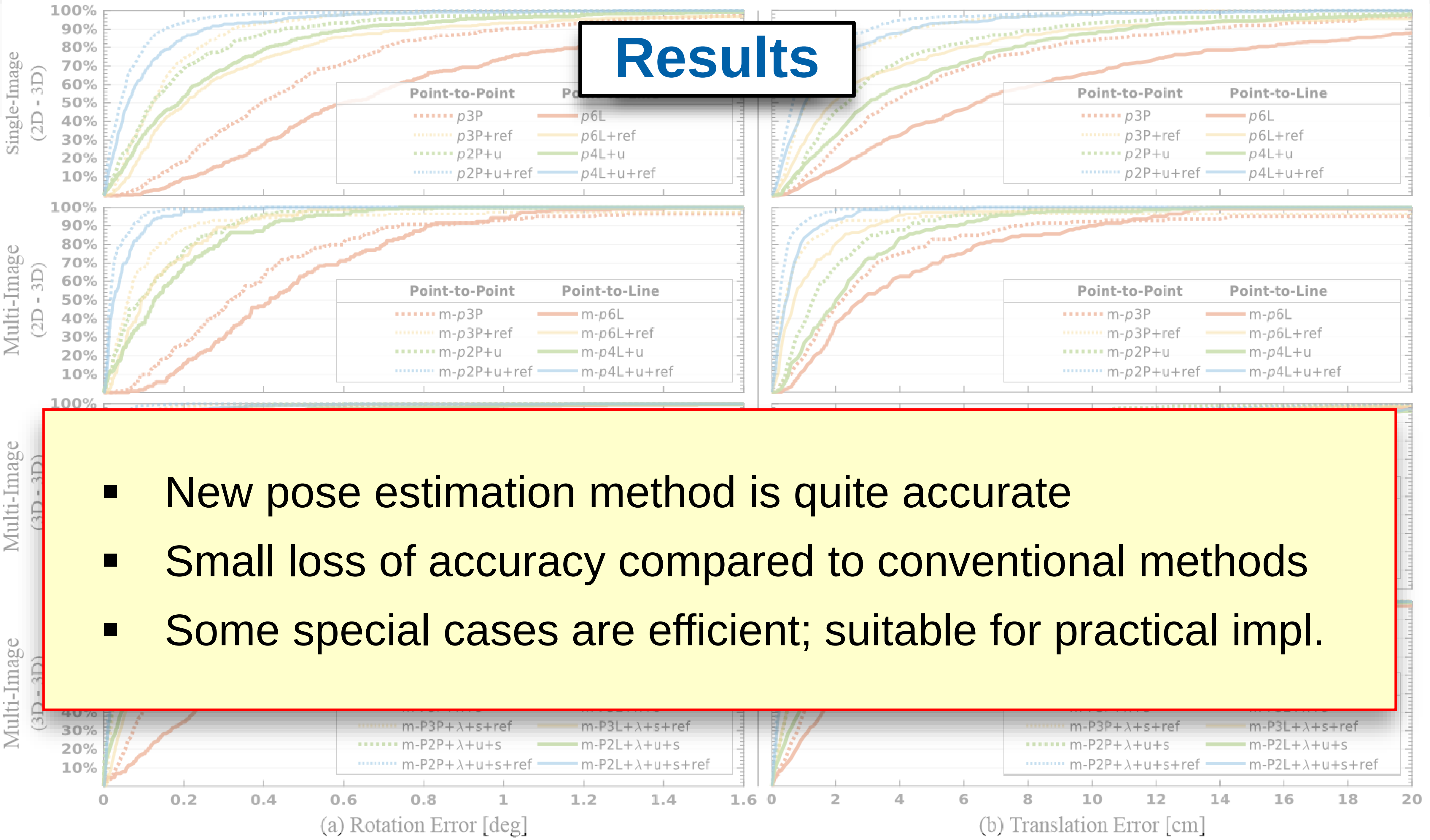
3. Most variants are **computationally efficient** and can be used with RANSAC.

Results



Single-Image (2D - 3D)
Multi-Image (2D - 3D)
Multi-Image (3D - 3D)
Multi-Image (3D - 3D)
100%
90%
80%
70%
60%
50%
40%
30%
20%
10%
Point-to-Point
Point-to-Line
p3P
p3P+ref
p2P+u
p2P+u+ref
p6L
p6L+ref
p4L+u
p4L+u+ref
m-p3P
m-p3P+ref
m-p2P+u
m-p2P+u+ref
m-p6L
m-p6L+ref
m-p4L+u
m-p4L+u+ref
m-P3P+λ
m-P3P+λ+ref
m-P2P+λ+u
m-P2P+λ+u+ref
m-P4L+λ
m-P4L+λ+ref
m-P3L+λ+u
m-P3L+λ+u+ref
m-P3P+λ+s
m-P3P+λ+s+ref
m-P2P+λ+u+s
m-P2P+λ+u+s+ref
m-P3L+λ+s
m-P3L+λ+s+ref
m-P2L+λ+u+s
m-P2L+λ+u+s+ref
0
0.2
0.4
0.6
0.8
1
1.2
1.4
1.6
(a) Rotation Error [deg]
0
2
4
6
8
10
12
14
16
18
20
(b) Translation Error [cm]

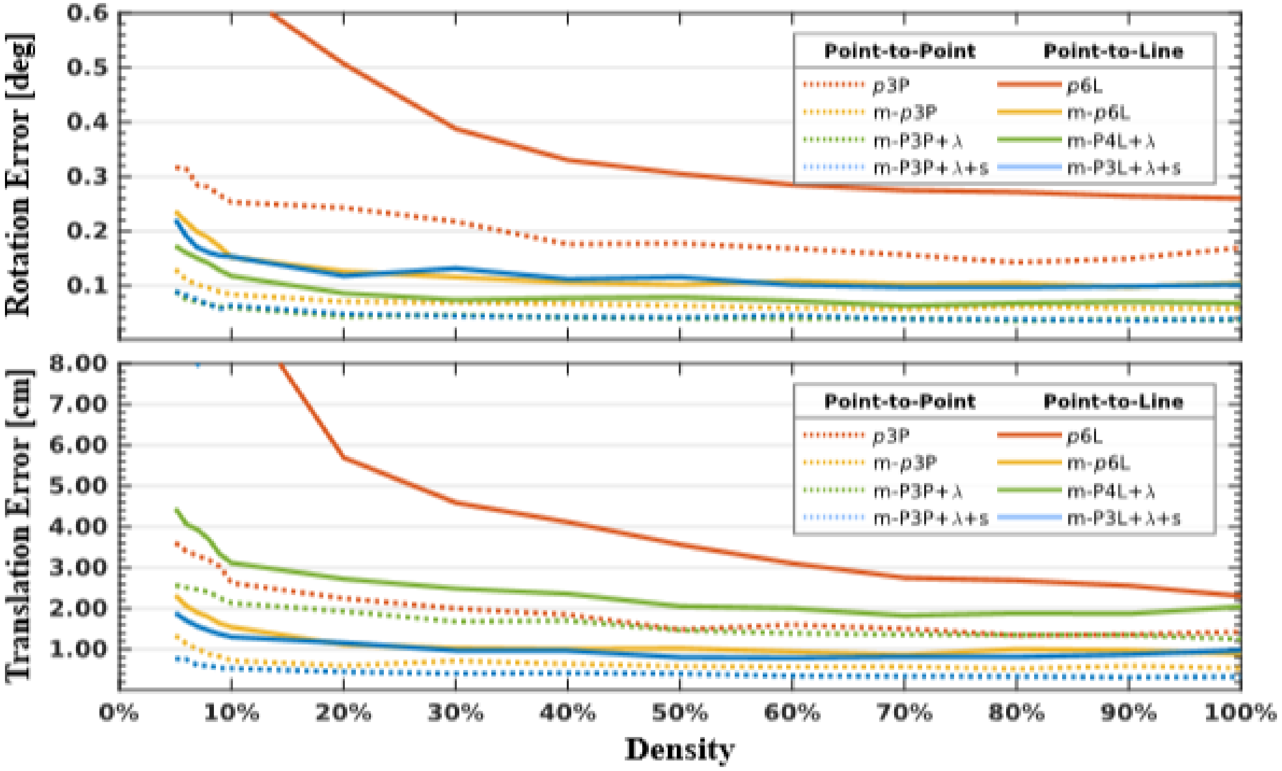
Results



- New pose estimation method is quite accurate
- Small loss of accuracy compared to conventional methods
- Some special cases are efficient; suitable for practical impl.

Additional Considerations

- Line Cloud Transformation must be permanent
- What is revealed during localization?
- Can the original 3D points be estimated from the 3D lines?
  - Sometimes. Densely sampled 3D points indicates where surfaces are likely to exist!
  - Solution: subsample the 3D points, pose estimation still works

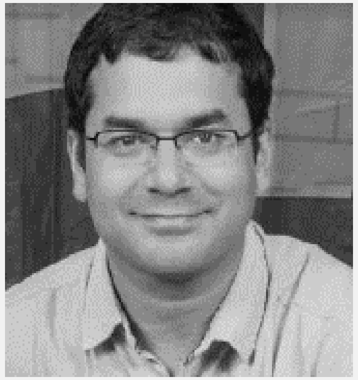


Privacy-Preserving Image Queries for Camera Localization

ICCV 2019

Pablo Speciale[1]

Johannes L. Schönberger[1]



Sudipta N. Sinha[2]

Marc Pollefeys[1,3]

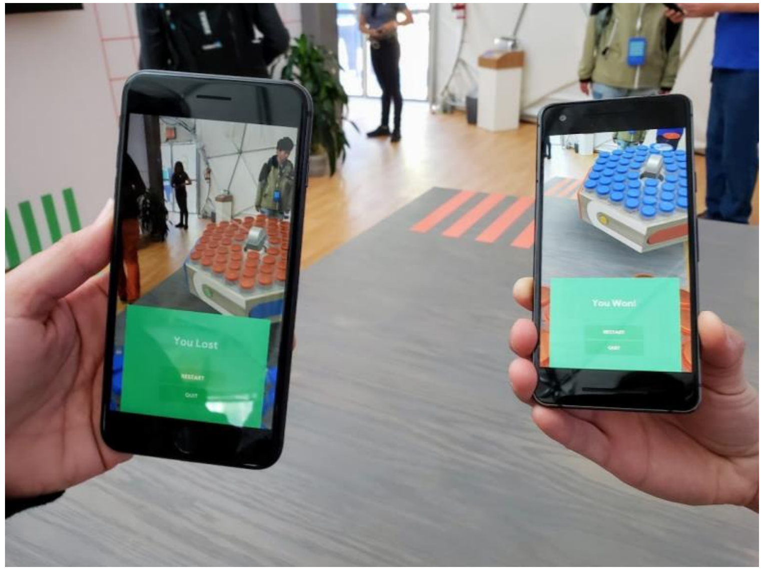
[1] Microsoft Mixed Reality & AI Group, Zurich

[2] Microsoft Research Redmond

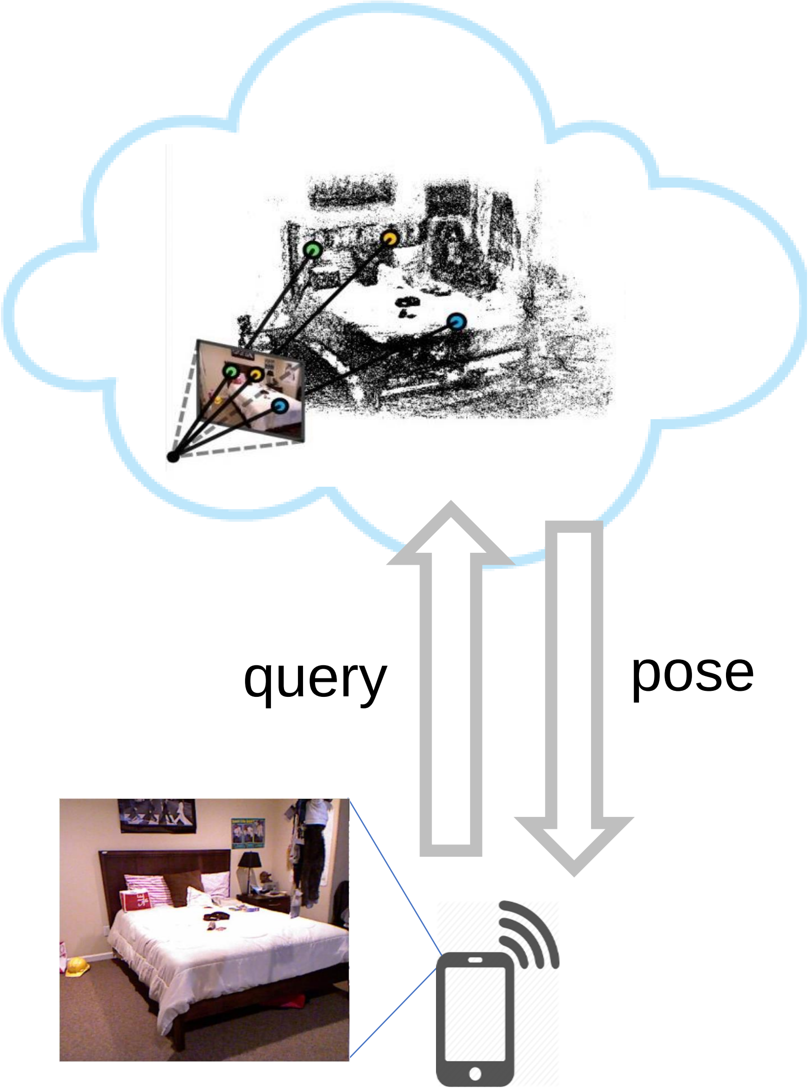
[3] ETH Zurich

Localization in the Cloud

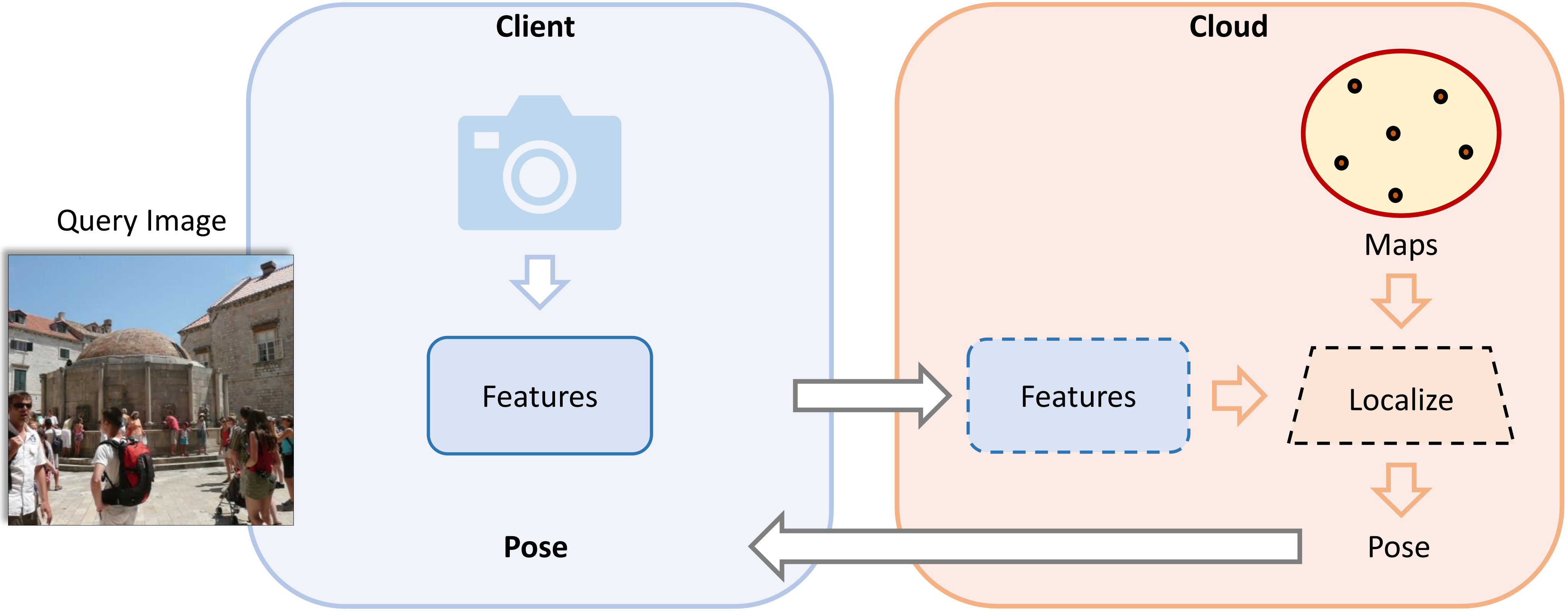
Microsoft ASA
(Azure Spatial Anchors)



Google AR Core
(Cloud Anchors)

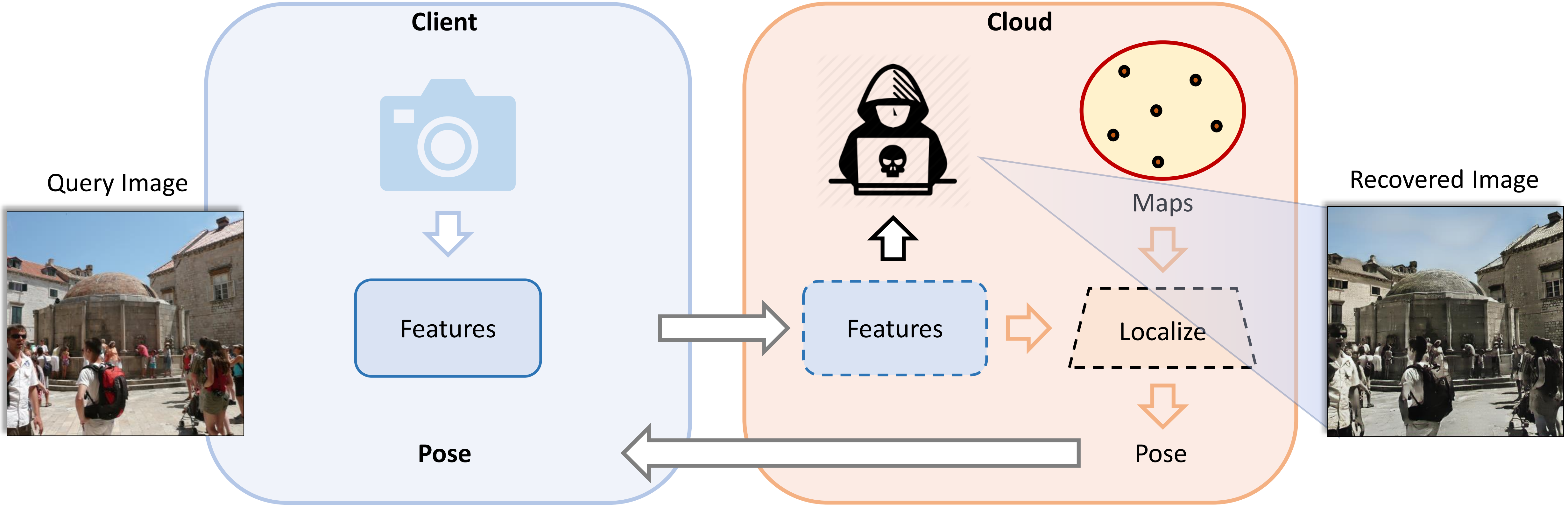


Privacy Risk in Cloud-based Localization



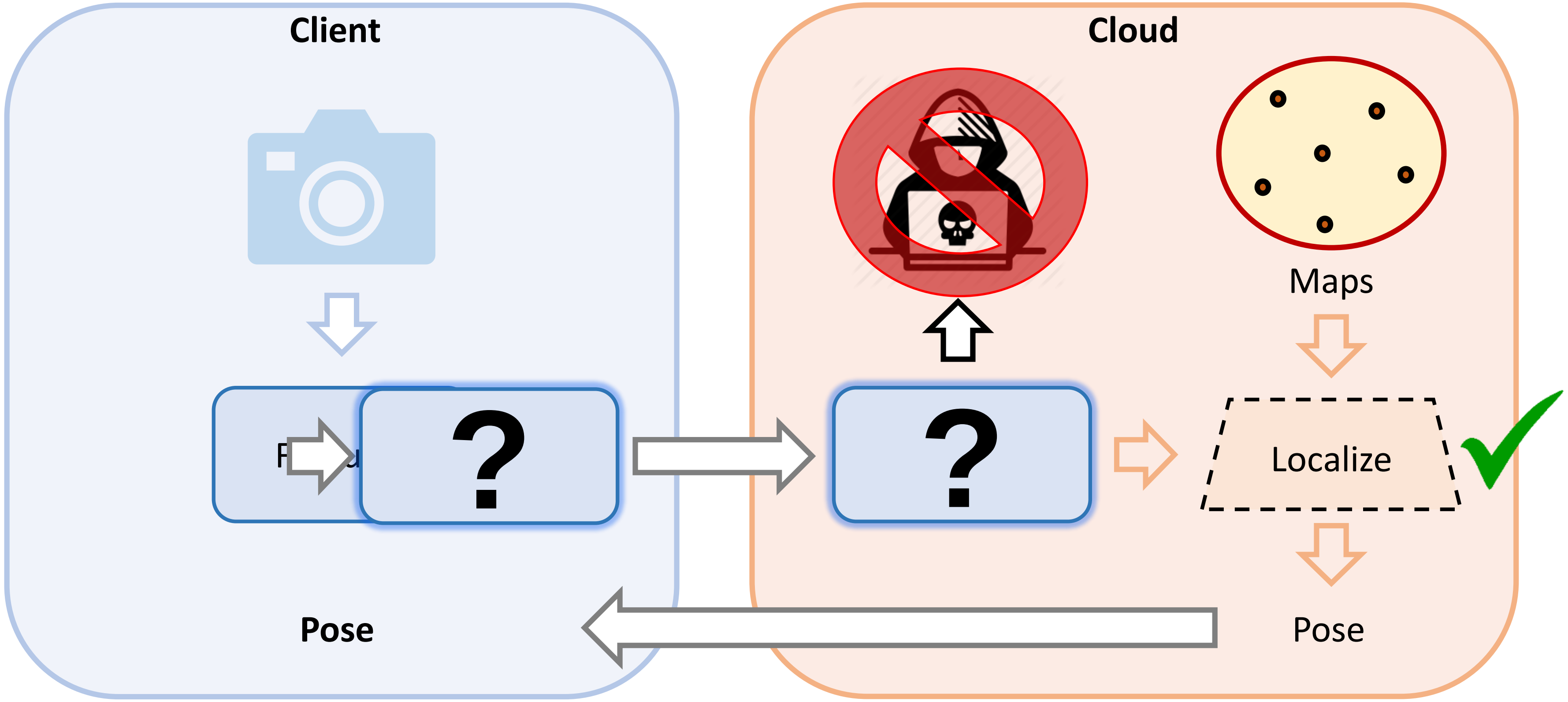
- Client sends image features to cloud
- Localization runs on cloud server
- Pose is sent back to Client

Privacy Risk in Cloud-based Localization



Adversary on cloud can invert features (recover the image)

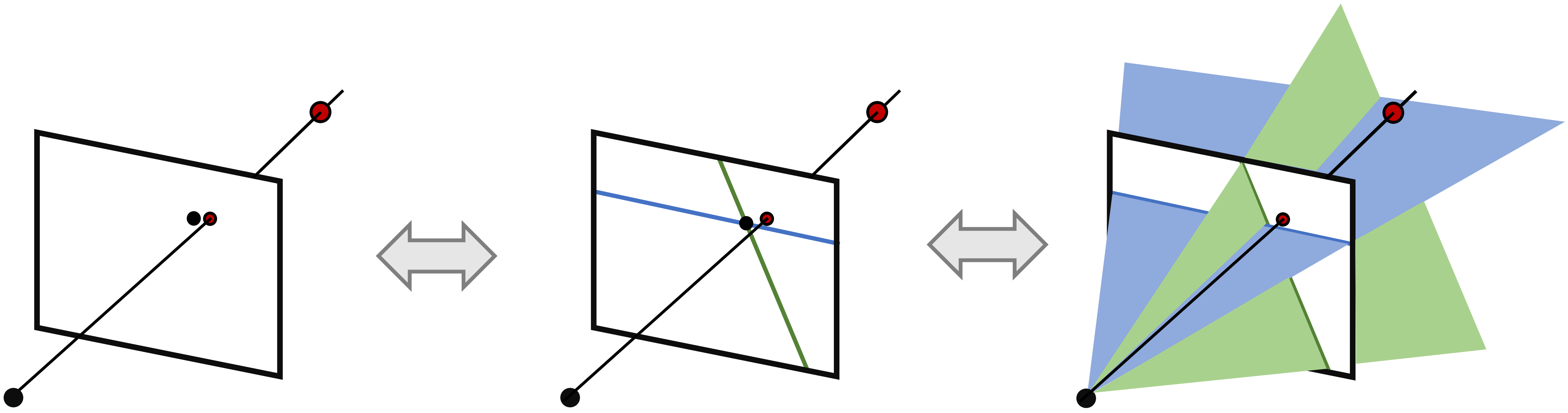
Privacy Risk in Cloud-based Localization



Our Goal:

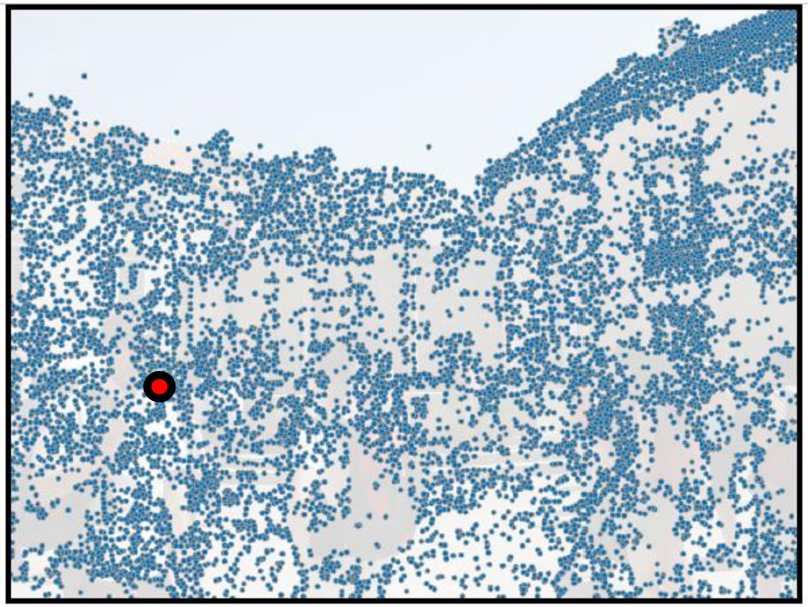
- Hide query features
- Prevent feature inversion on server
- Allow camera pose estimation

Key Insight

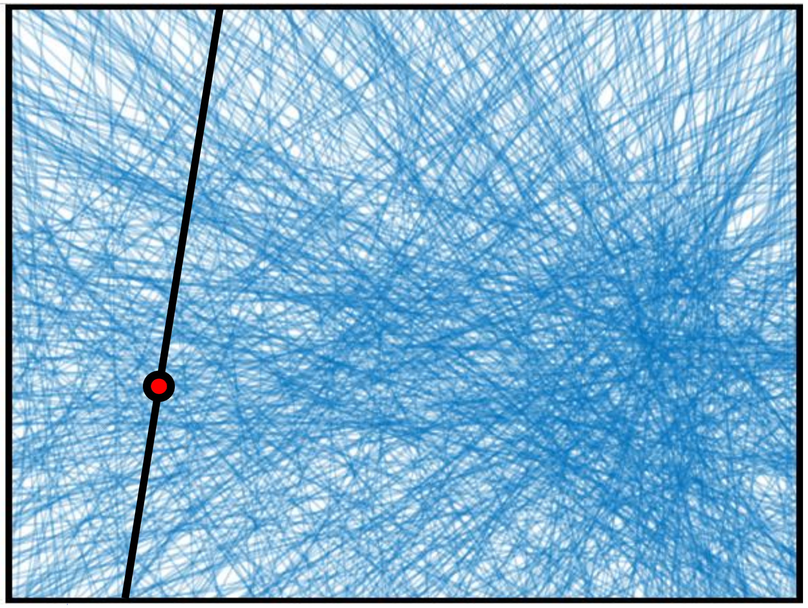


Proposed Idea

Query Image



2D Feature Points

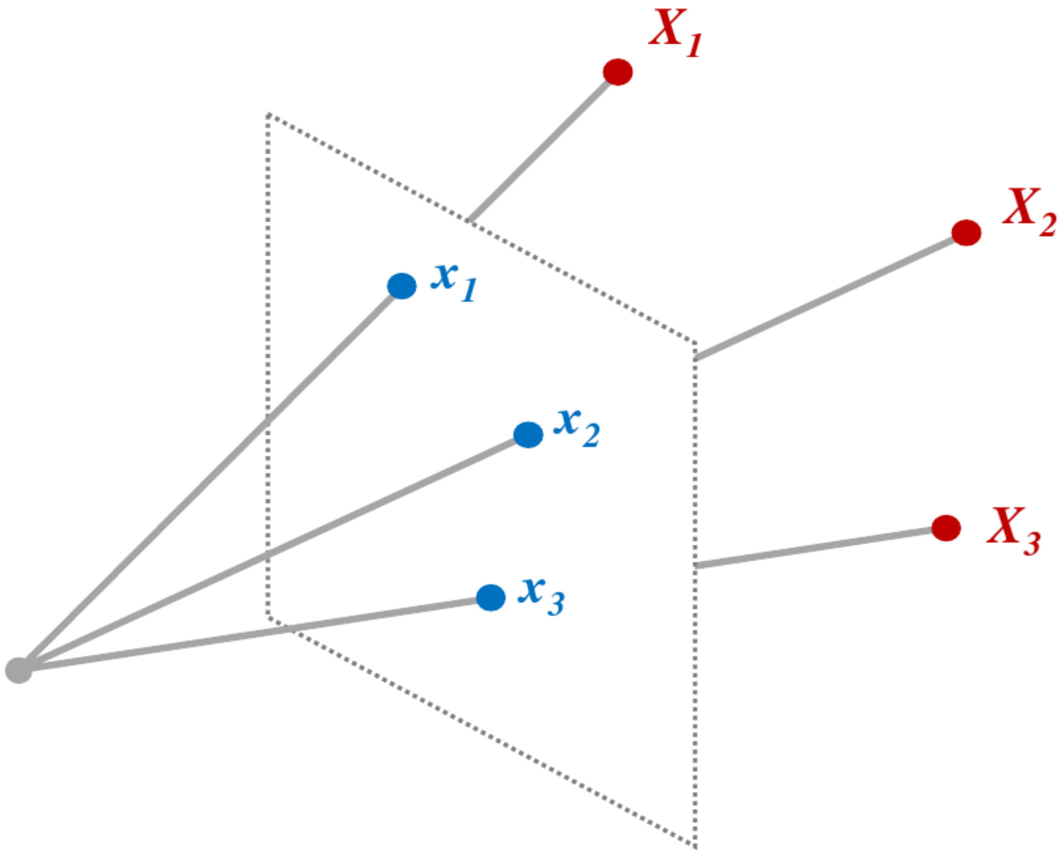


2D Feature Lines

- Select a randomly oriented 2D line through each 2D feature point
- Discard the 2D feature points
- Upload 2D features lines + descriptors to the cloud

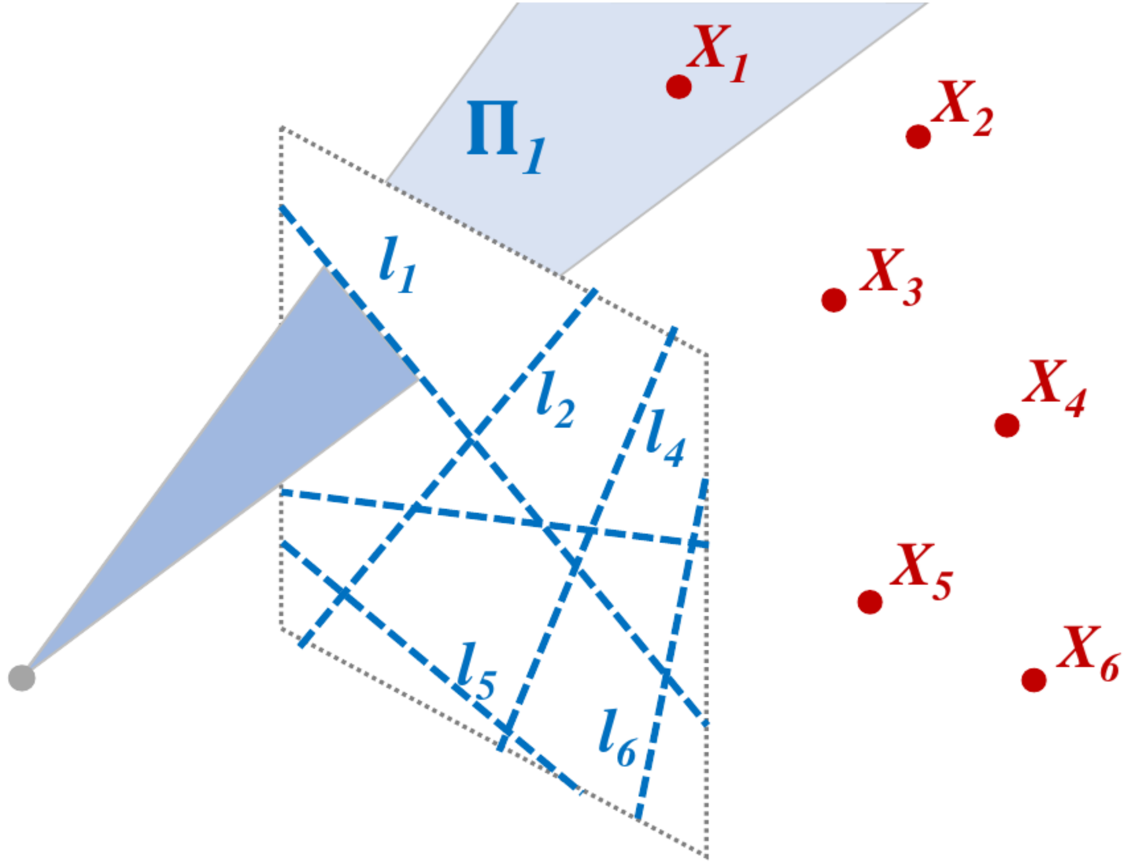
Camera Pose Estimation

Traditional Method (p3P)



Three 2D *image point* – 3D point correspondences

Proposed Method (l6P)



Six 2D *image line* – 3D point correspondences

Camera Pose Estimation

1. Our minimal problem can be cast as **Point-to-Plane problem** [1].
   [1] Ramalingam et al. 2013

2. Proposed **several variants** with:
   - known structure (multiple images),
   - known vertical direction,
   - known scale.

   We leverage existing minimal solvers [2-7].

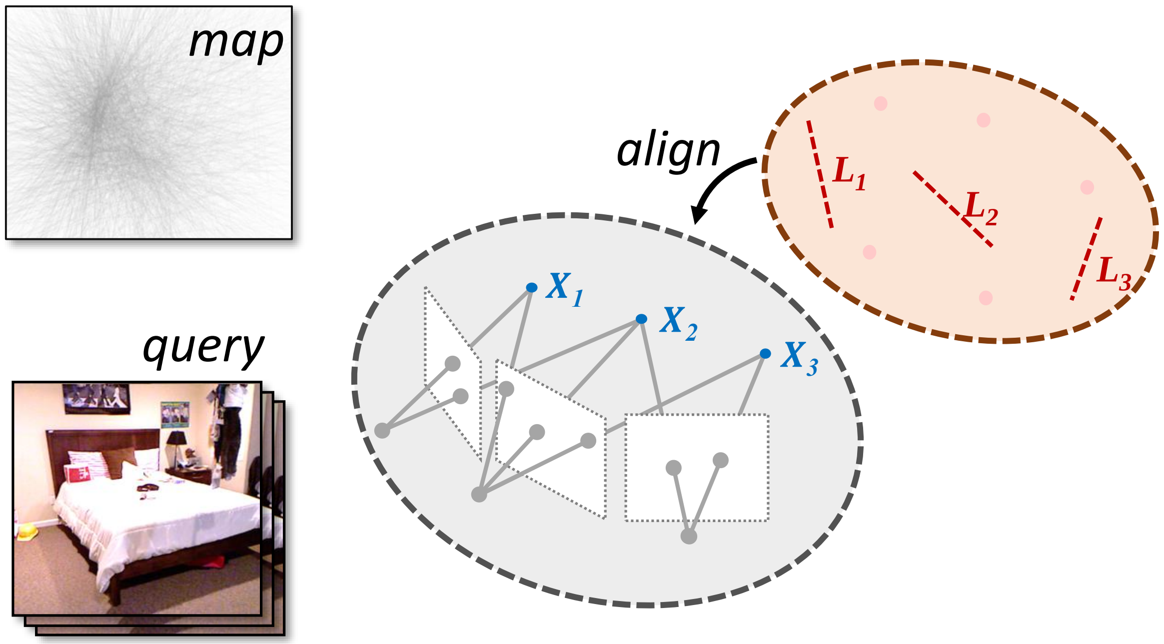
   [2] Camposeco et al. 2018 [3] Lee et al. 2016 [4] Stewenius et al. 2005

   [5] Sweeney et al. 2015a [6] Sweeney et al. 2015b [7] Sweeney et al. 2014

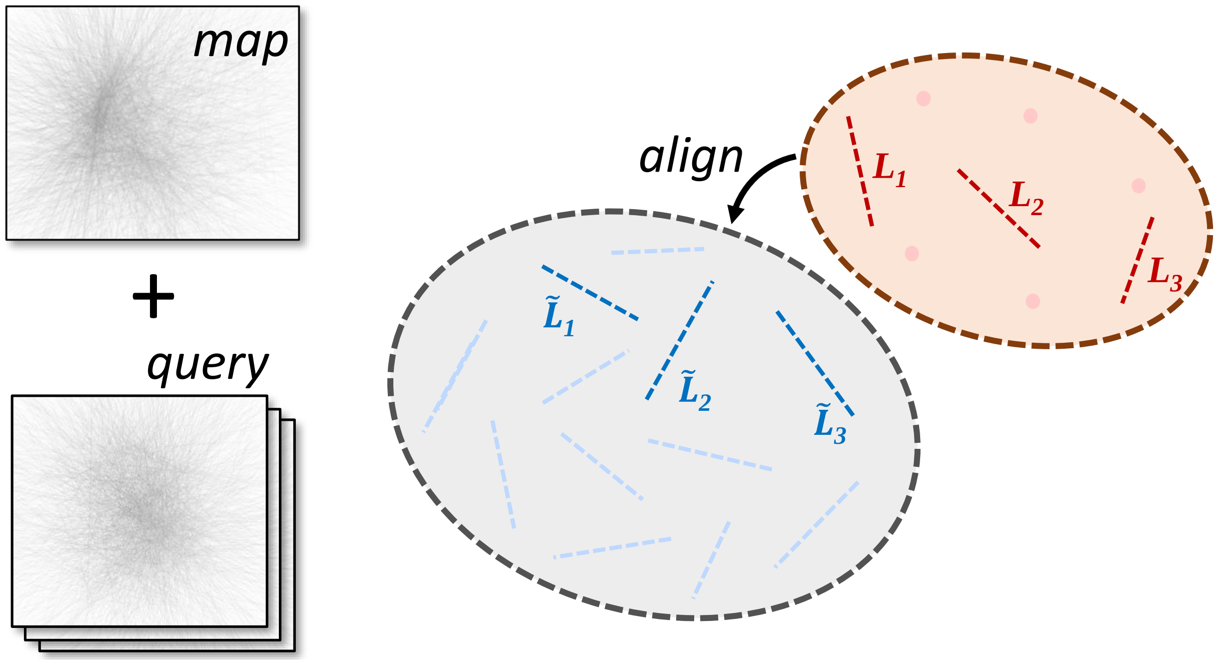
3. Most variants are **computationally efficient** and can be used with RANSAC.

Confidential Query + Confidential Map

Confidential Map [Speciale et al. CVPR 2019]

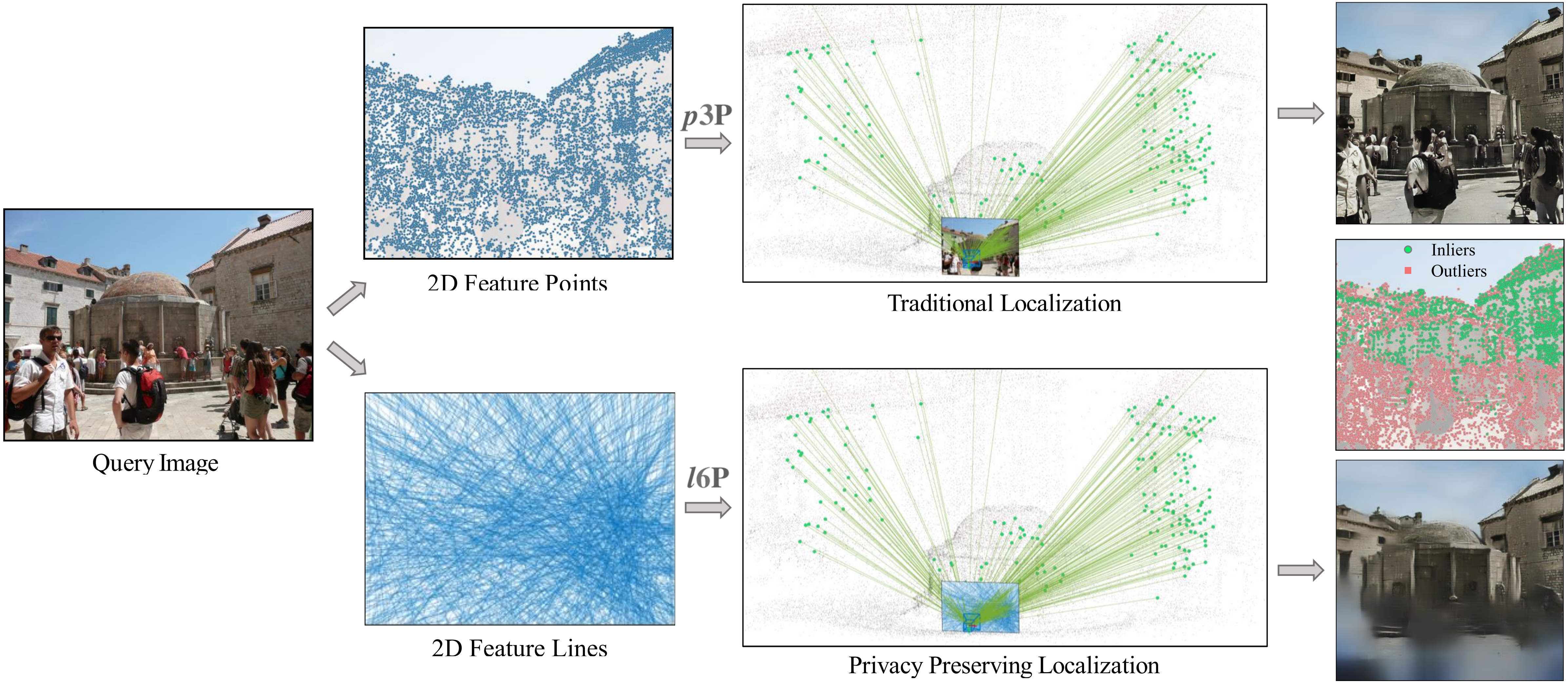


Confidential Query + Confidential Map



- Underlying problem: align 3D lines in query to 3D lines in map
- 6-pt generalized relative pose problem [Stewenius et al. 2005]
- 4-pt generalized relative pose + vertical [Sweeney et al. 2015]

What gets revealed after localization?



Query Image

2D Feature Points

2D Feature Lines

Traditional Localization

Privacy Preserving Localization

What gets revealed after localization?

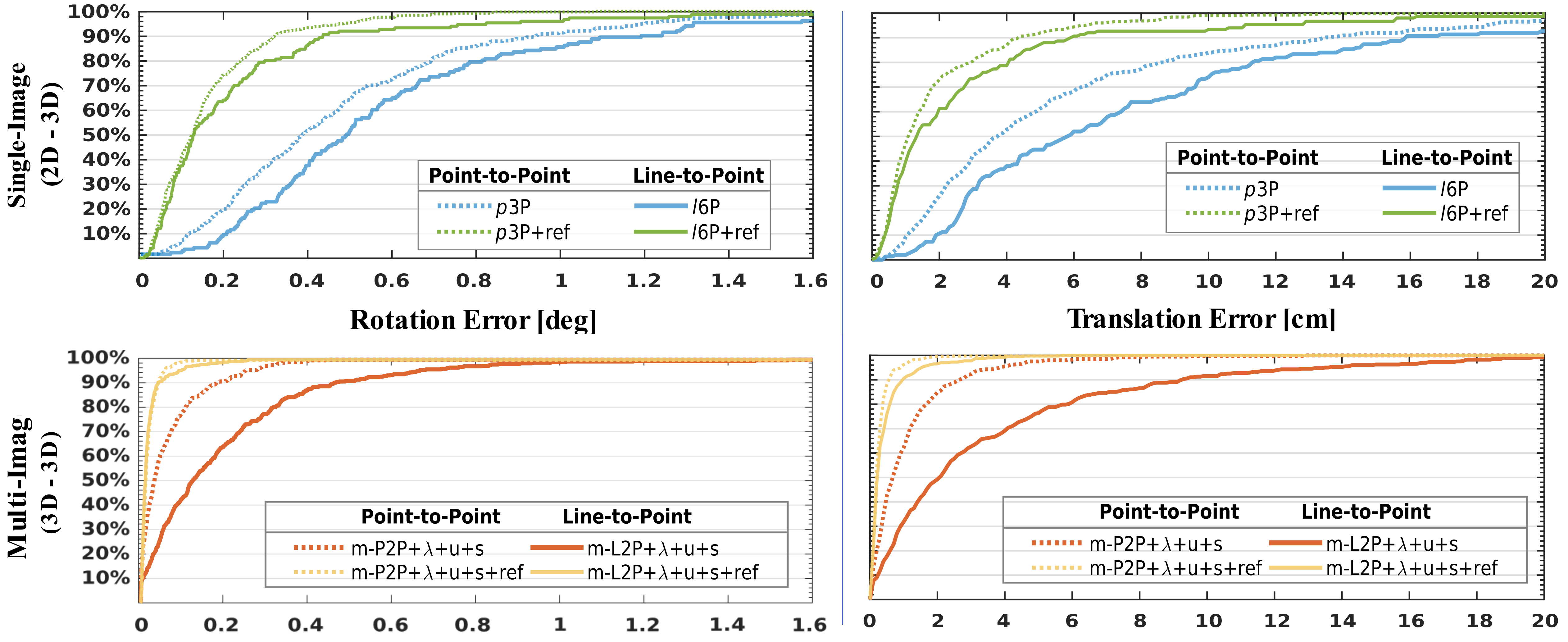
Query
(original image)

Image Inversion
(*all* features)

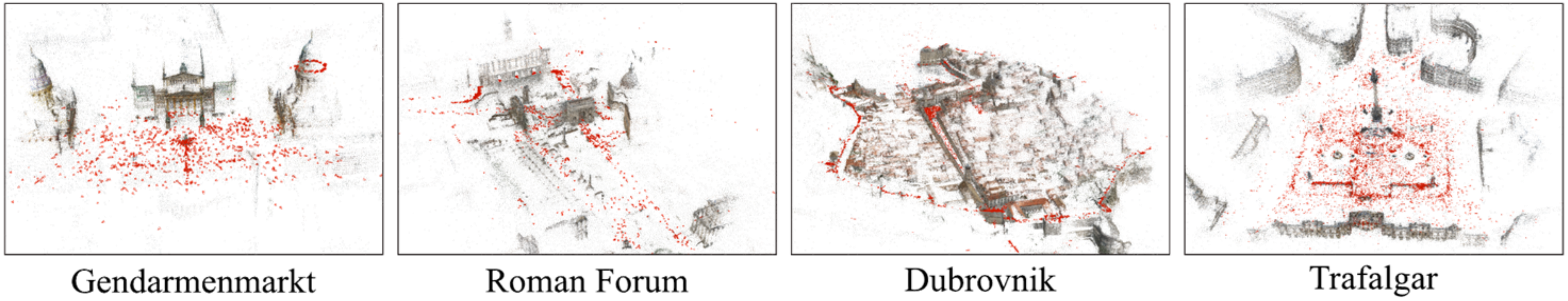
Image Inversion
(*only revealed* features)

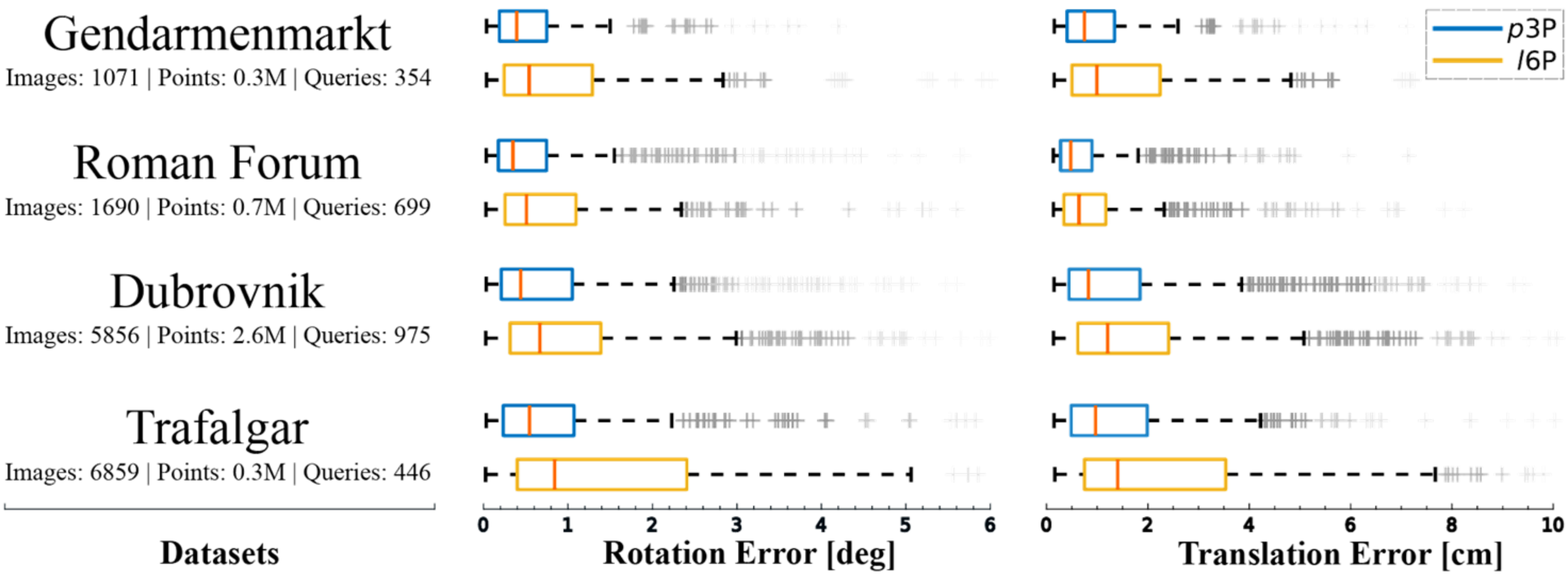
Results: Localization Accuracy

Accuracy/recall curves. Cumulative rotation and translation error histograms.



Results: Internet Photo Collection





Conclusions

- Highlighted new type of privacy issues in AR/Robotics applications; many other open problems …

- Proposed privacy-preserving camera localization techniques
  - where the **map is concealed,**
  - where the **query image is concealed**,
  - where both **map and query remain concealed**.

- Our techniques nicely map into **known minimal solvers**; many of which are accurate and **computationally efficient**.