# Rise of the Indoor Crowd: Reconstruction of Building Interior View via Mobile Crowdsourcing

Si Chen, Muyuan Li, Kui Ren, Xinwen Fu, Chunming Qiao

Dept of CSE, SUNY Buffalo

UbiSeC Lab

Ubiquitous Security & PrivaCy

Research Laboratory

University at Buffalo

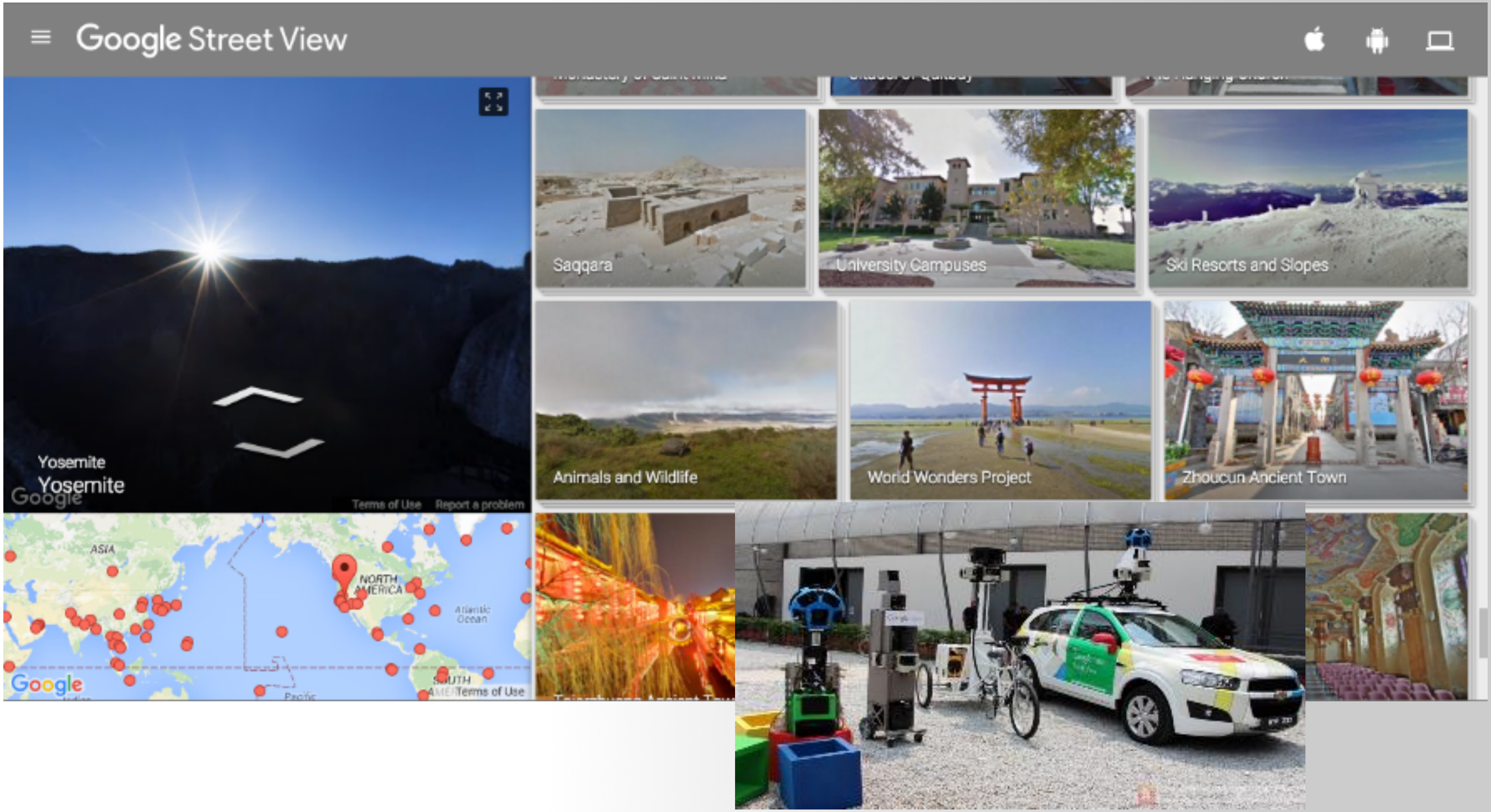
The State University of New York

# Outline

- Introduction
- System Model
- Design Details
- System Evaluation
- Conclusion

# Motivation






Establishing Large Scale Information Infrastructure in the Era of IoT!

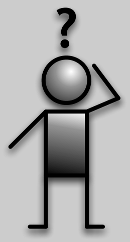
# Motivation



- Techniques and data collection approach developed for outdoor environments do not suit for indoor scenarios.


No GPS/localization signal

Complexity and Quantity

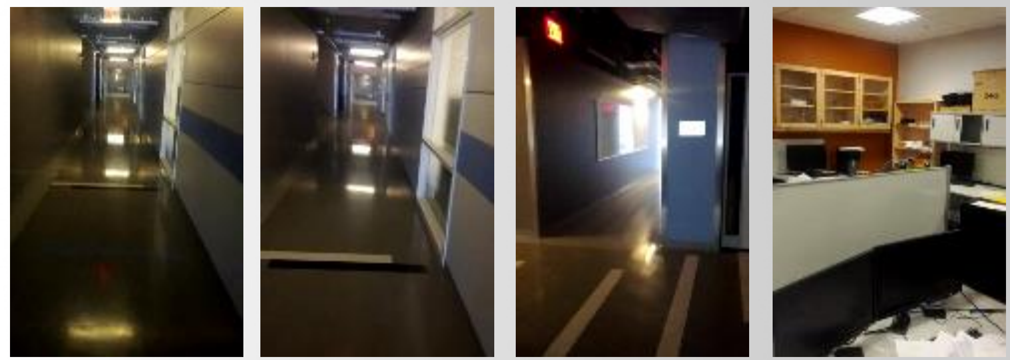



Existing outdoor street-view reconstruction techniques either cannot be directly applied to indoor environments or are prohibitively costly!

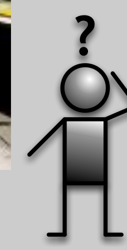
# Motivation

- Techniques and data collection approach developed for outdoor environments do not suit for indoor scenarios.


No GPS/localization signal



Complexity and Quantity




We need a practical approach to establish large-scale indoor interior view for buildings!

# Related Work



- Indoor visualization has been studied by robotics and computer vision communities:
  - Simultaneous localization and mapping (SLAM)


http://www.navvis.lmt.ei.tum.de/about/

  - Structure from Motion (SfM) and Multiview-stereo

http://www.cs.cornell.edu/~snavely/bundler/

# Our Approach: Mobile Crowdsourcing

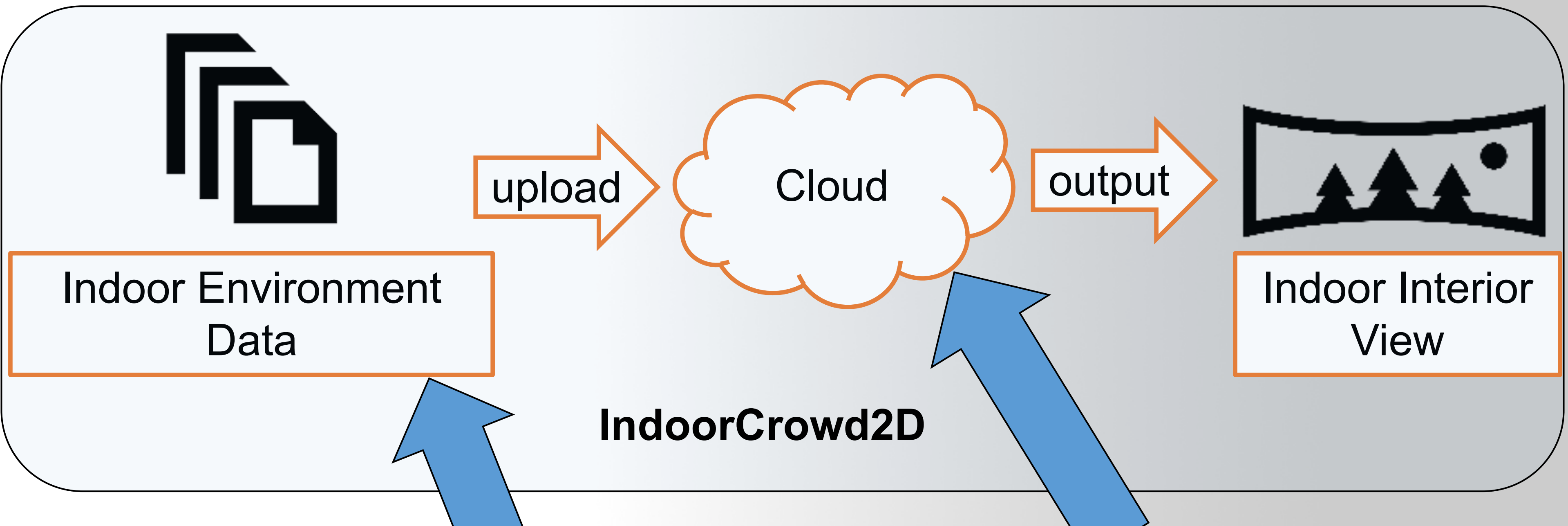
UbiSeC Lab
Ubiquitous Security & PrivaCy
Research Laboratory

- Crowdsourcing can provide the access to large quantity of mobile devices as well as people in an extremely cost-effective manner.

- High-resolution cameras and various sensors are built in with modern smart-devices.


# Our Approach: IndoorCrowd2D

A smart device empowered system utilizes the power of the crowd to reconstruct building interior-views.




**Mobile Crowdsourcing:**

1. **Off-the-shelf devices**
2. **Large quantity**
3. **Low cost**

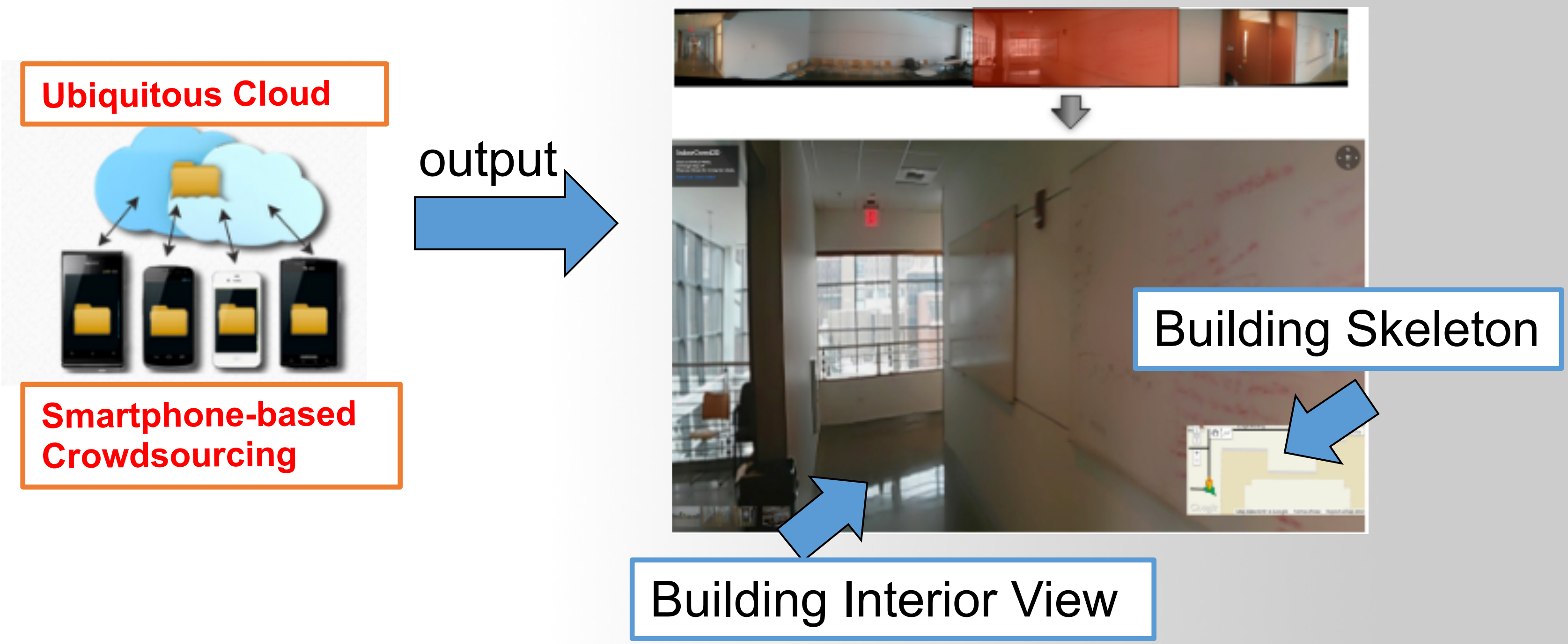
**Cloud Reconstruction:**

1. **Allow gradual building-up**
2. **Robust to noisy data**
3. **Scalable**

# Our Approach: IndoorCrowd2D



- System Output:
  - Panoramic images for visualizing building interior-views.
  - Navigable hallway skeletons for each building floors.



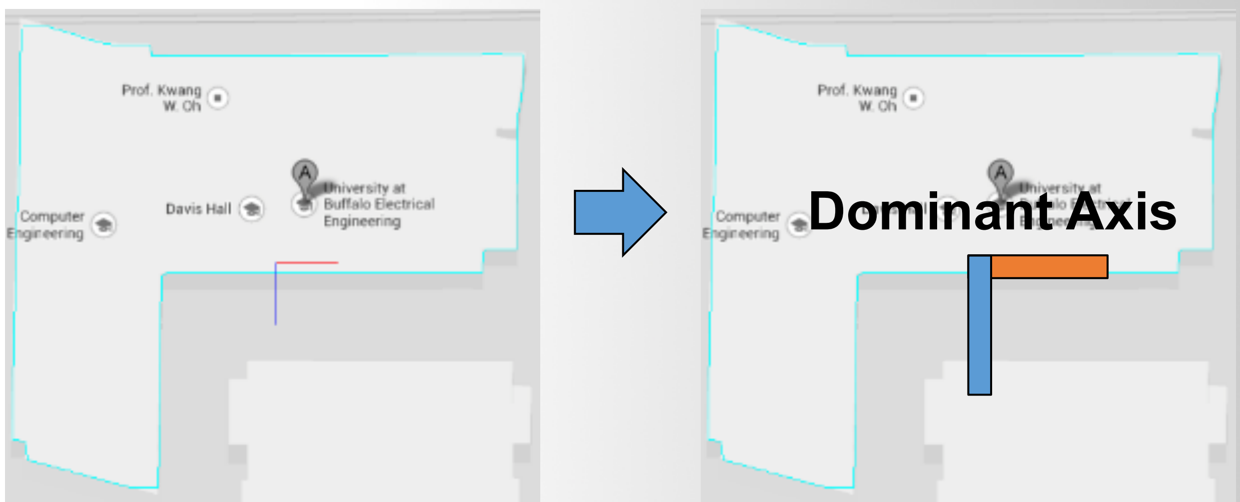
# Outline

- Introduction
- **System Model**
- Design Details
- System Evaluation
- Conclusion

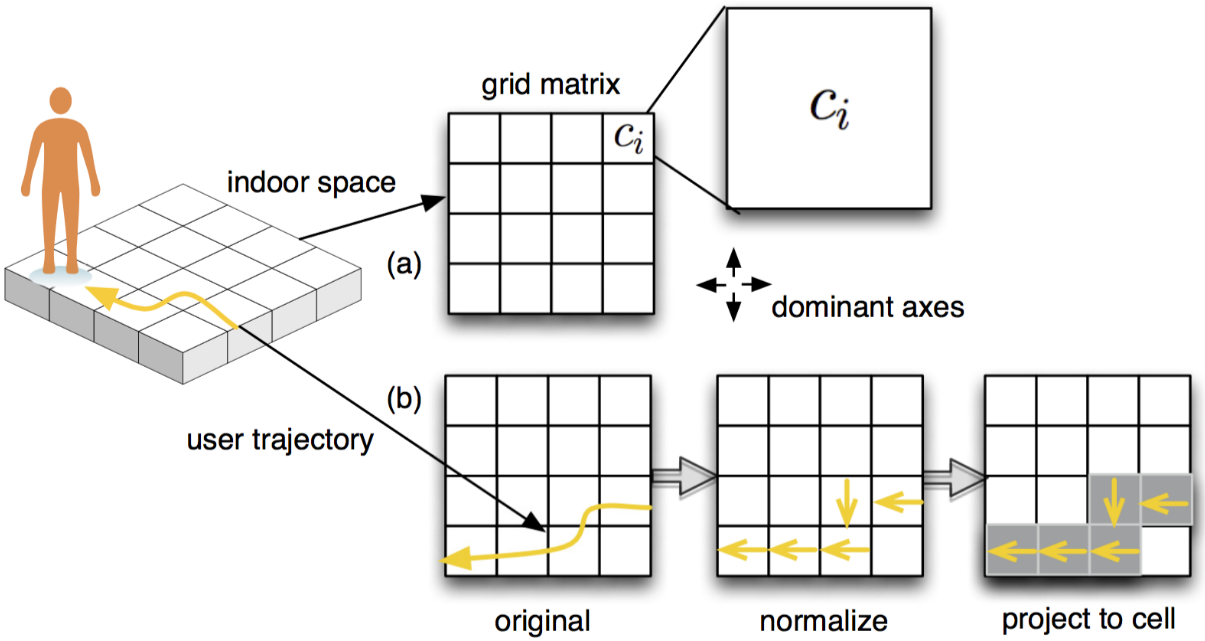
# System Model: Indoor Spatial Model

**Reduced Manhattan World (RMW) assumption:**

- Assume 2 perpendicular dominant axis for each building.
  - Each line segment (mostly corridors) inside the building is aligned to one of the two axes.
- The dominant axis is used to initialize the indoor spatial model.

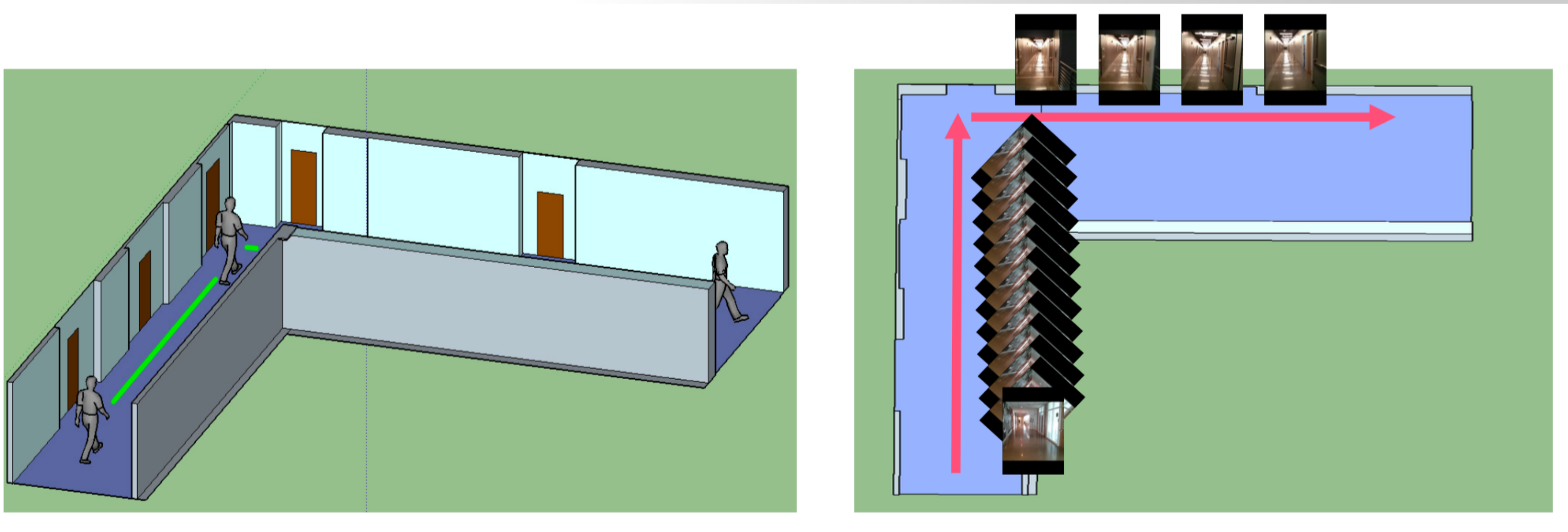


# System Model: Indoor Spatial Model




- Represent the indoor environment by a homogeneous grid matrix based on RMW.
- Store the user trajectory information by projecting it to the grid matrix.
- No indoor localization infrastructure is assumed during user photo shooting

# System Model: User Trajectory Model



- We define a customized image vector data structure to include
  - both user movement information obtained from the sensory data and
  - the corresponding image data

# System Model: User Trajectory Model

Image Vector Bundle (IVB):

- Each IVB corresponds to a user's movement during one camera shooting session
- Include relative spatial location, heading direction, image data, and timestamp.

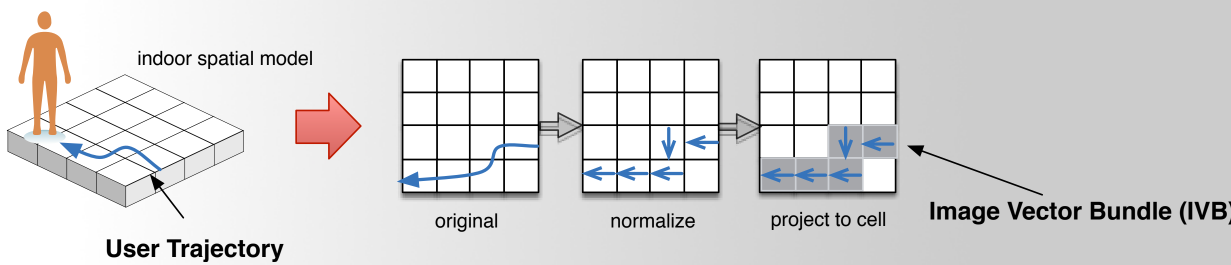
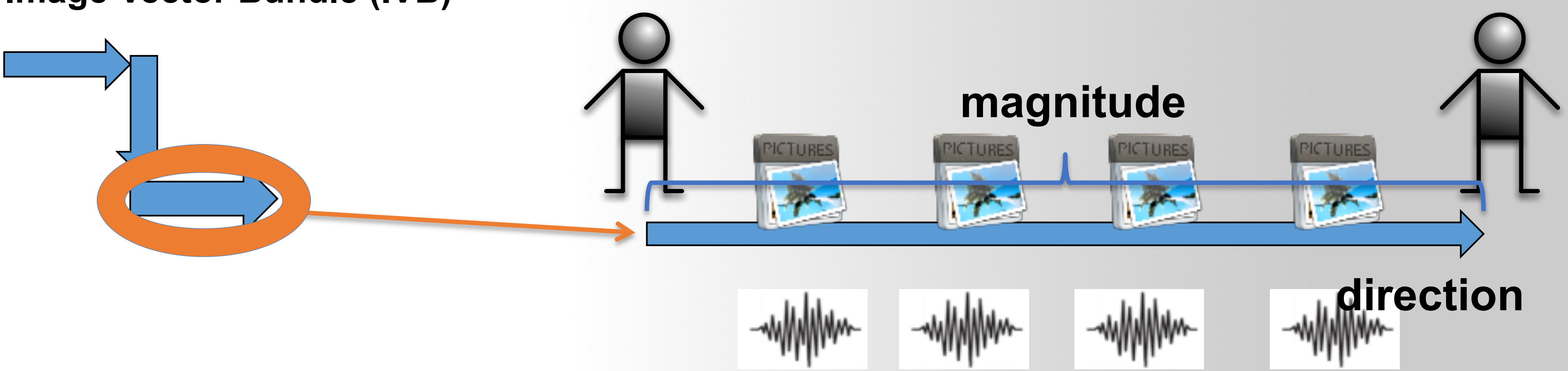
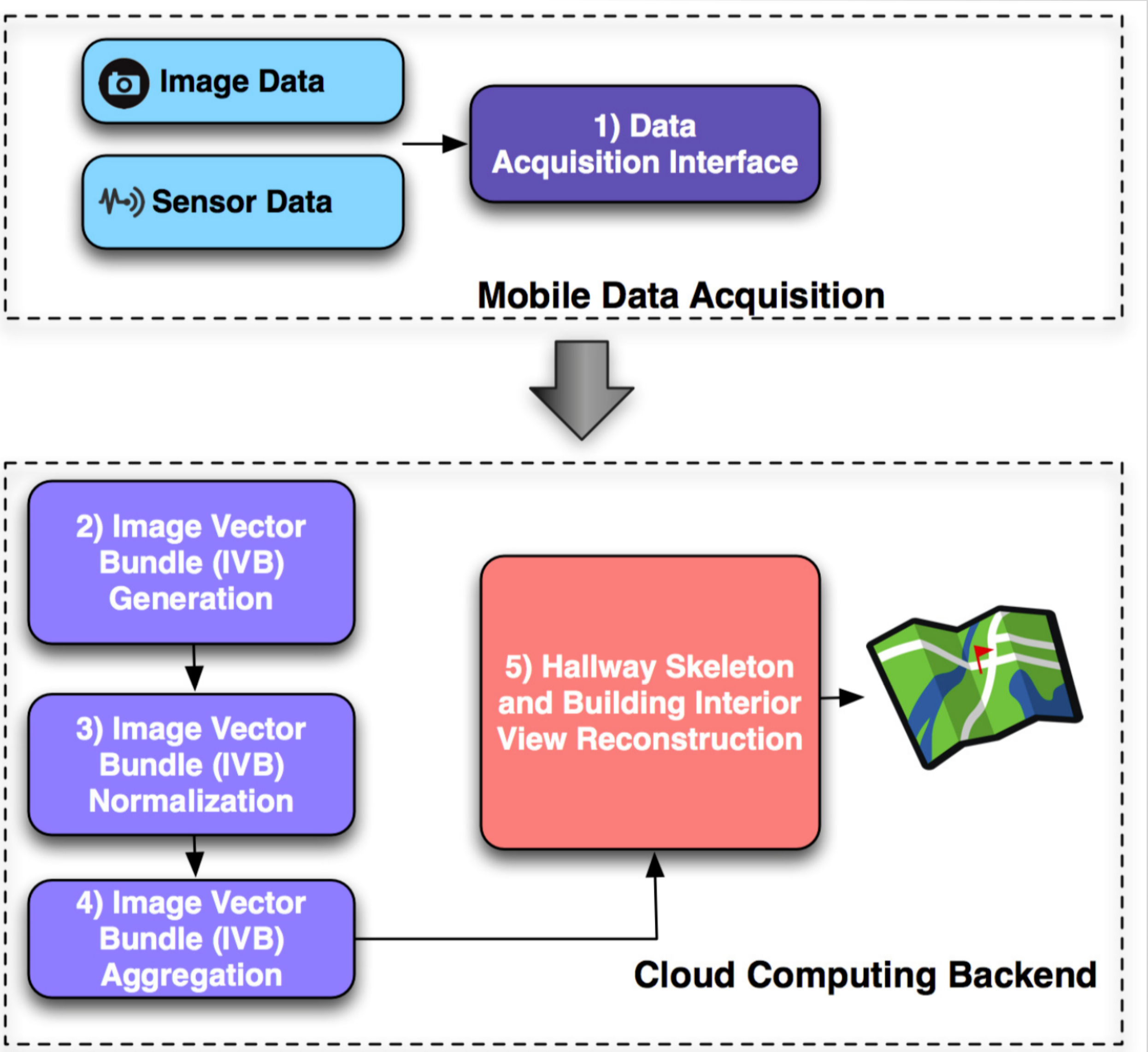



**Image Vector Bundle (IVB)**



# IndoorCrowd2D Architecture






Mobile App

Cloud Backend

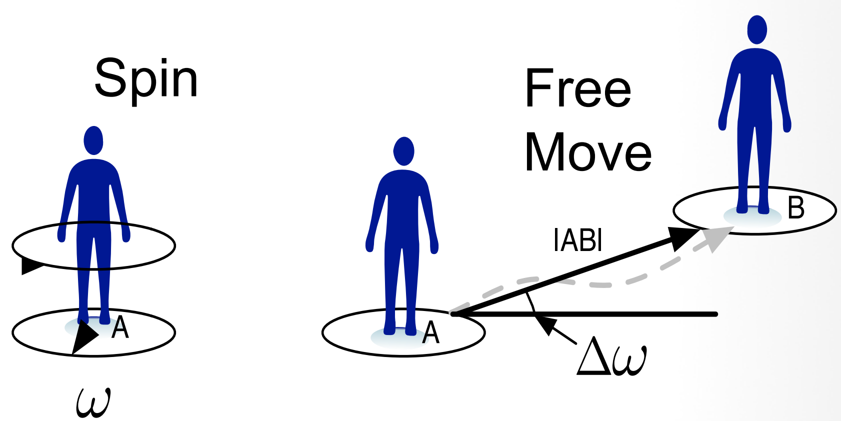
# Outline

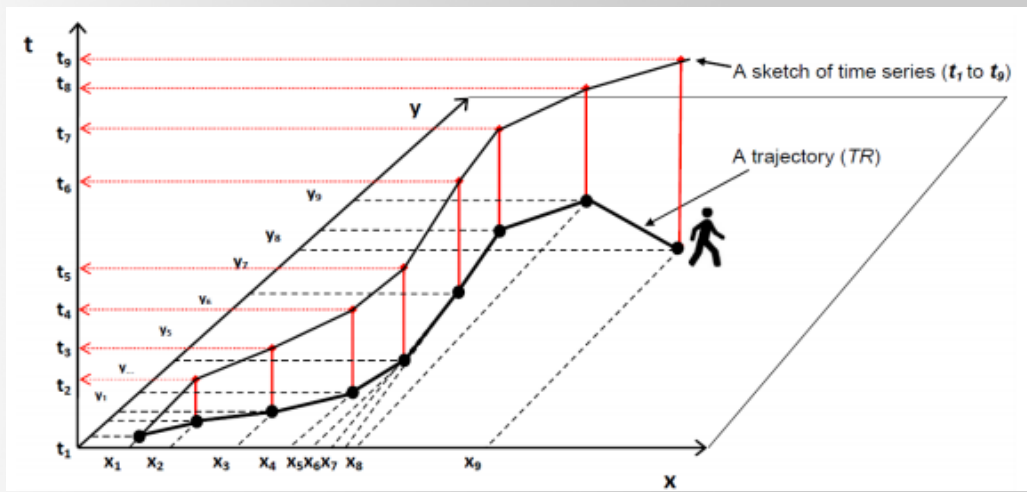
- Introduction
- System Model
- Design Details
  - Collecting Data from Untrained Users
  - Aggregating multiple user trajectories
  - Generating indoor interior-views
- Evaluation
- Conclusion

# Collecting Data from Untrained Users



- Users with their mobile phones choose to shoot the environment at their will.
- Our phone APP takes pictures and records the sensory data to capture camera positions, view directions, etc.



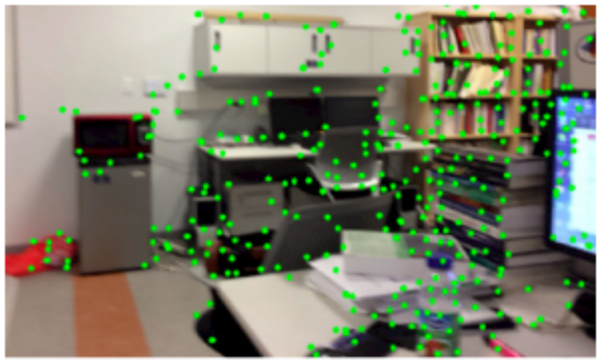



# Collecting Data from Untrained Users

UbiSeC Lab
Ubiquitous Security & PrivaCy Research Laboratory

- The crowdsourced data is inherently incomplete, opportunistic, and noisy.
- Our App guides user data collection through real-time data quality feedback for improved quality.

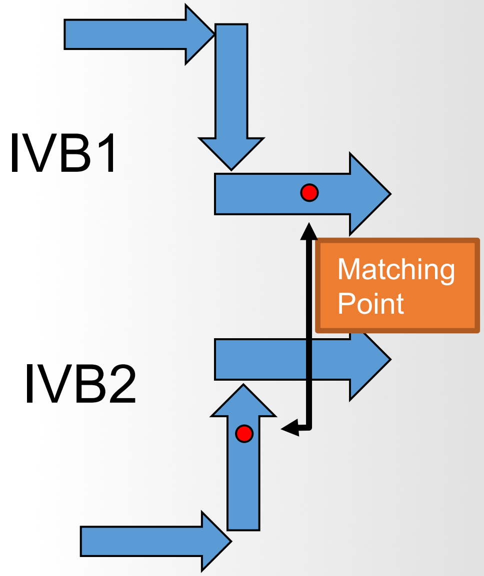
Three metrics are measured in real-time:
1. Linear acceleration,
2. Angular acceleration
3. The # of SURF features in each picture.


Example SURF feature points of a picture

# Aggregating multiple user trajectories

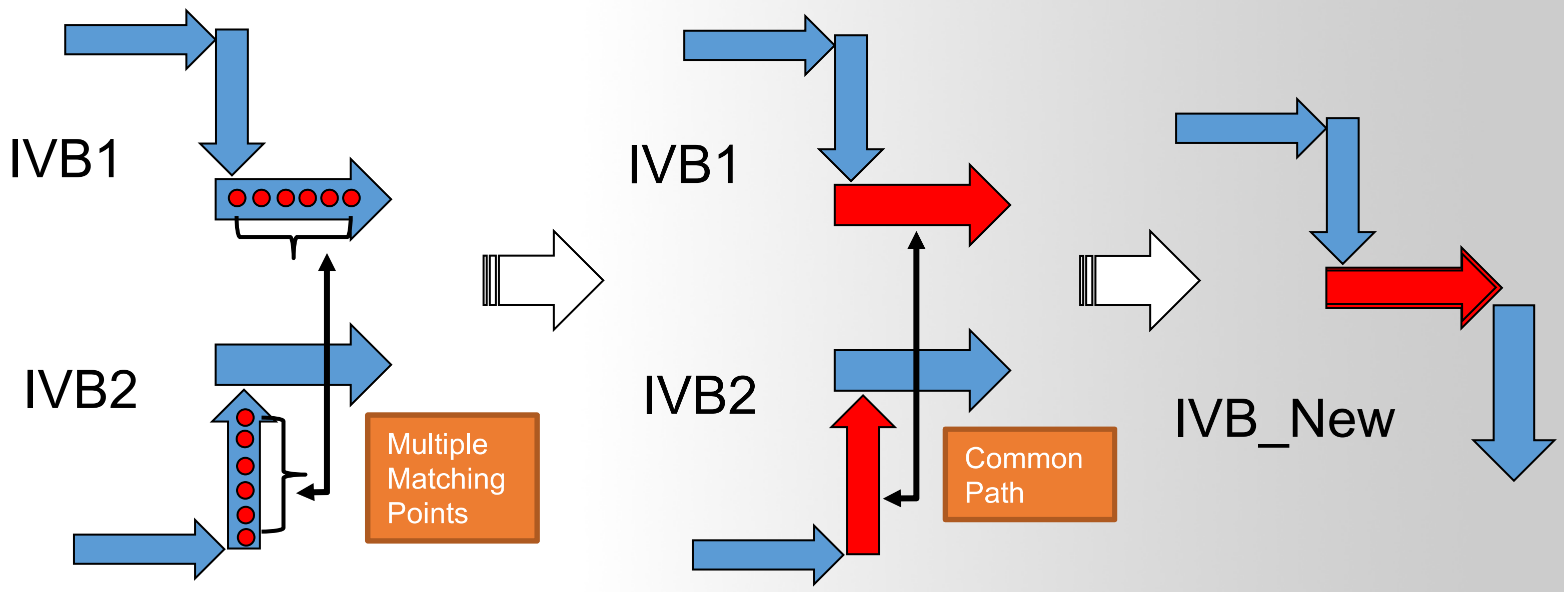
UbiSeC Lab
Ubiquitous Security & PrivaCy Research Laboratory

- At cloud side, we use a hierarchical approach to aggregate multiple user trajectories.
- We first try to find the matching point of two IVBs based on the following manner:
  - Construct a codebook of “visual features” using SURF for each image;
  - Quantize visual feature descriptors by the k-nearest neighbor (kNN) algorithm;
  - Use Euclidean distance as a similarity metric.



# Aggregating multiple user trajectories

UbiSeC Lab
Ubiquitous Security & PrivaCy Research Laboratory



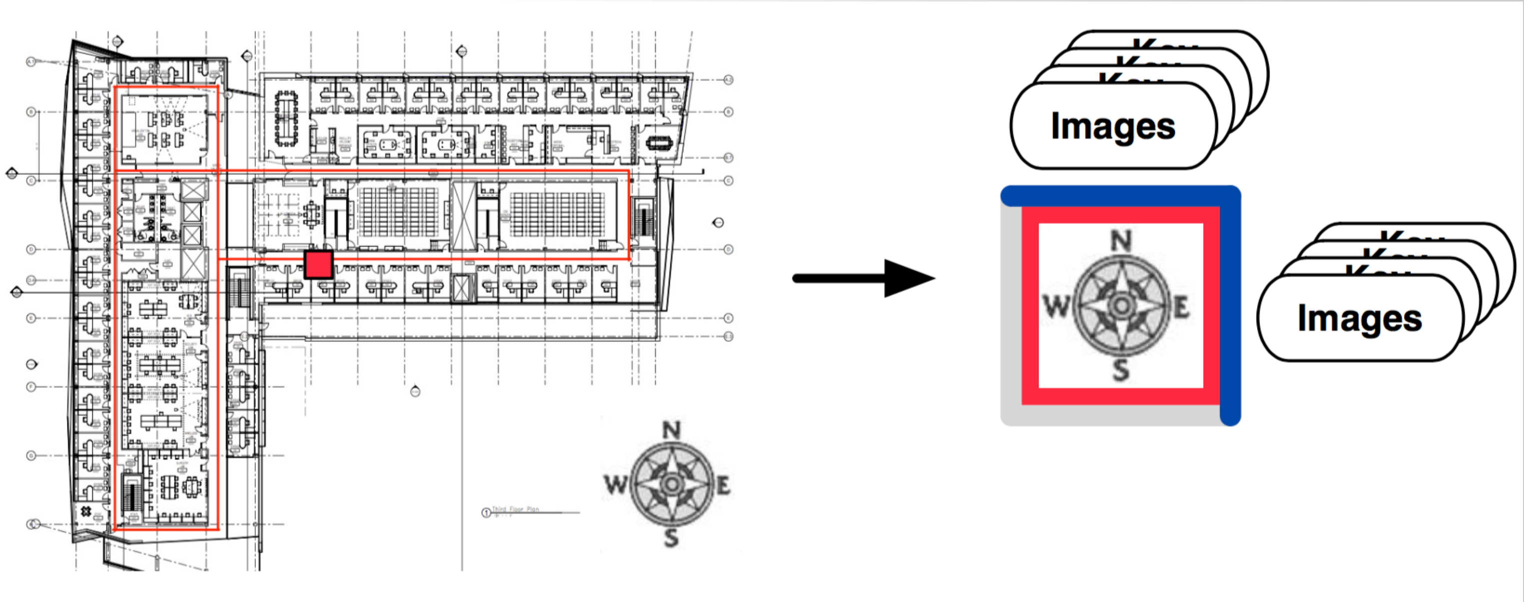

- We then use multiple matching points to calculate the common path between two IVBs based on the longest common subsequence (LCS) metric.
- We merged two IVBs into one larger IVB if two IVBs shares a common path.

# Aggregating multiple user trajectories

- We're able to aggregate multiple user trajectories by running this algorithm multiple times.
- The aggregated IVB can be used for representing the hallway skeleton of a building floor.


Aggregating 425 IVBs into one large IVB

# Generating indoor interior-views



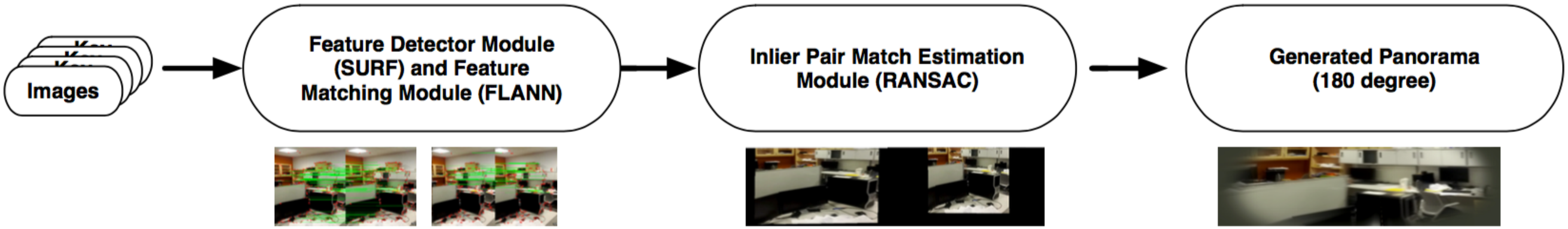

selected the crowdsourced images inside a particular cell

- To generate indoor interior view, we leverage the aggregated image vectore bundle.

- We first selected the images inside a particular grid cell from the aggregated IVB.

# Generating indoor interior-views






δ-building interior view generation algorithm

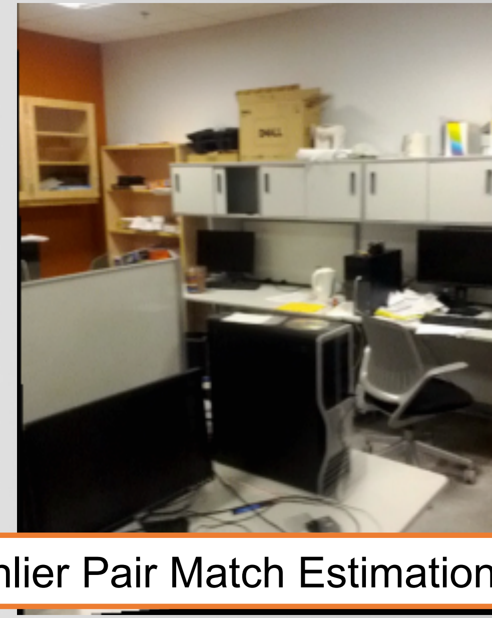
- We then further process the selected images by leveraging a δ-building interior view generation algorithm.

- δ-building interior view generation algorithm is a combination of several state-of-the-art panorama reconstruction algorithms.
  - Input: aggregated images from a grid cell
  - Output: an interactive panorama

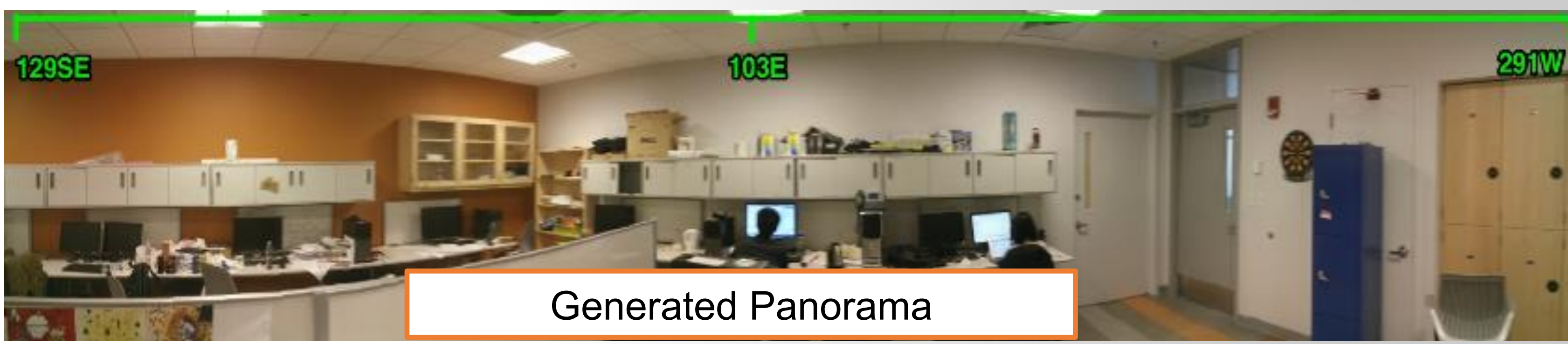
# Generating indoor interior-views

An example of the δ-building interior view generation process

Feature Detecting and Matching



Inlier Pair Match Estimation Module



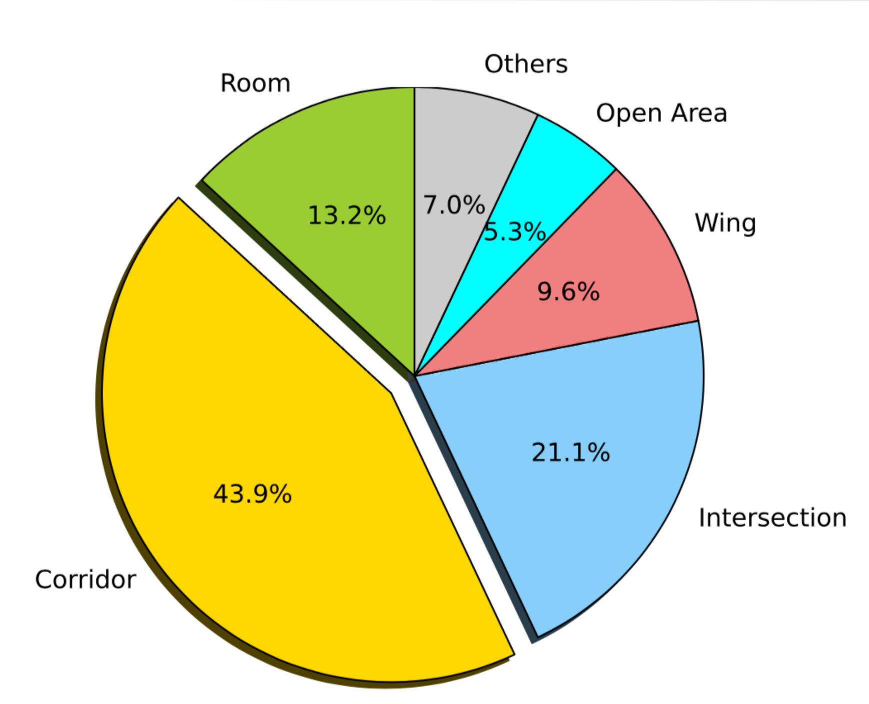

Generated Panorama

# Outline

- Introduction
- System Model
- Design Details
- **System Evaluation**
- Conclusion

# System Evaluation

## Evaluation Dataset




We use the following dataset to evaluate our IndoorCrowd2D prototype:

- Two campus building floors
- 55,453 images from 1,151 datasets uploaded by 25 untrained users.

## Evaluation Metrics

We use the precision, recall, and F-measure as evaluation metrics to evaluate the performance of IndoorCrowd2D

$$precision = \frac{|\mathcal{S}_{gen} \bigcap \mathcal{S}_{true}|}{|\mathcal{S}_{gen}|}$$

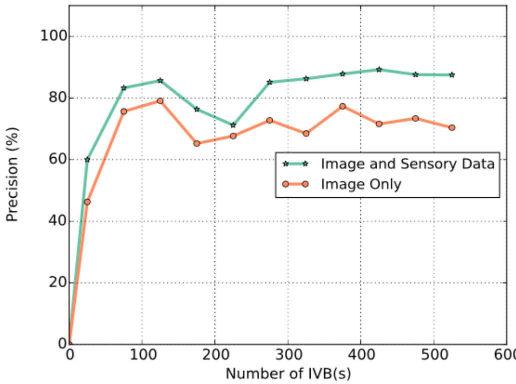
$$recall = \frac{|\mathcal{S}_{gen} \bigcap \mathcal{S}_{true}|}{|\mathcal{S}_{true}|}$$

$$F - measure = 2 * \frac{precision * recall}{precision + recall}$$

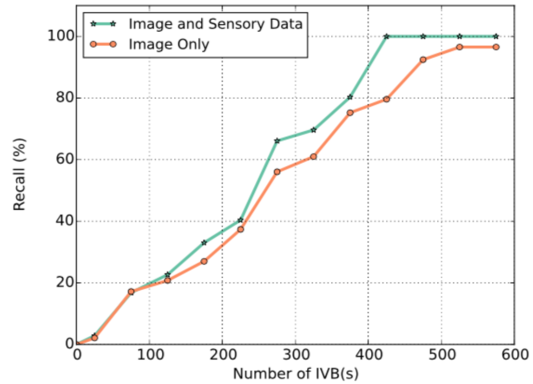
$\mathcal{S}_{true}$: Ground Truth Skeleton

$\mathcal{S}_{gen}$: Generated Skeleton

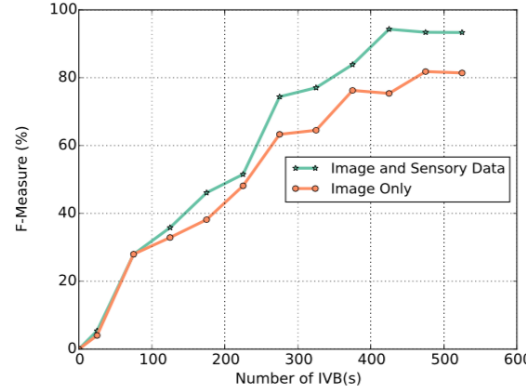
# System Evaluation



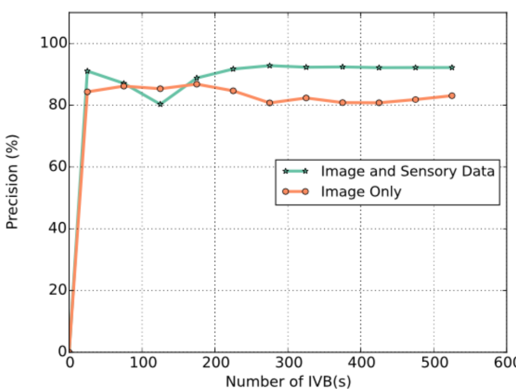

(a) Hallway skeleton precision for TB dataset



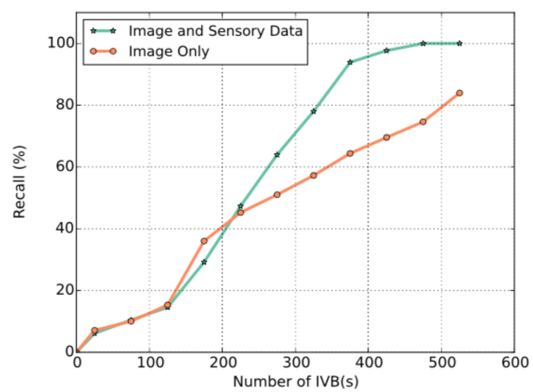

(b) Hallway skeleton recall for TB dataset



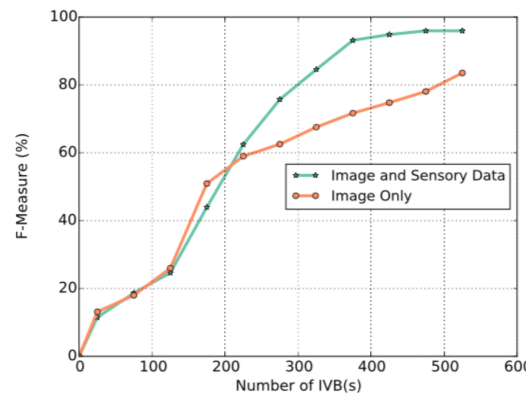

(c) Hallway skeleton F-measure for TB dataset




(d) Hallway skeleton precision for GYM dataset

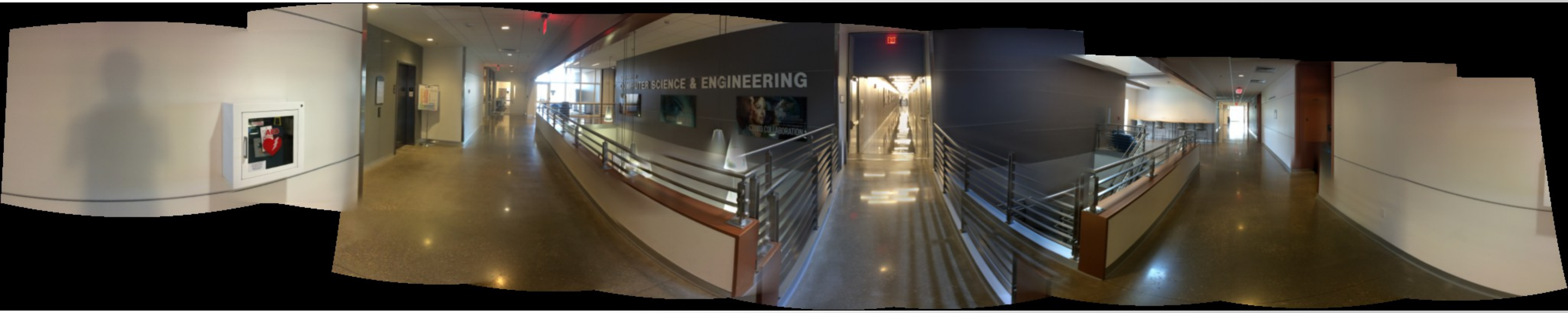



(e) Hallway skeleton recall for GYM dataset




(f) Hallway skeleton F-measure for GYM dataset

# System Evaluation




(a) Output Panorama


(b) Ground Truth Panorama

# Outline

- Introduction
- System Model
- Design Details
- System Evaluation
- **Conclusion**

# Conclusion and Future Works



- IndoorCrowd2D – a novel crowdsourcing system empowered by off-the-shelf smartphones for building interior view reconstructions.
    - IndoorCrowd2D is readily deployable in real-world scenarios.
    - IndoorCrowd2D is expected to provide indoor panorama and geo-data for each individual floor of any building around the world.
- IndoorCrowd2D serves an important stepping stone towards the ultimate goal of economically-viable massive indoor 3D model reconstruction.

# Q & A

# Longest common subsequence (LCS) metric

UbiSeC Lab
Ubiquitous Security & PrivaCy Research Laboratory

$$L\left(\mathbf{Z}_i^A, \mathbf{Z}_j^B\right) = \begin{cases} 0, if\ i = 0\ or\ j = 0; \\ 1 + L(\mathbf{Z}_{i-1}^A, \mathbf{Z}_{j-1}^B), \\ \qquad if\ d(\vec{z}_i^A, \vec{z}_j^B) \le \epsilon\ and\ |i - j| < \delta; \\ max(L(\mathbf{Z}_i^A, \mathbf{Z}_{j-1}^B), L(\mathbf{Z}_{i-1}^A, \mathbf{Z}_j^B)), \\ \qquad otherwise; \end{cases}$$

Where $Z^A$ and $Z^B$ are the two user trajectories with length of i and j, respectively. Parameter $\delta$ represents the maximum length difference between two user trajectories and $\varepsilon$ is the distance threshold.