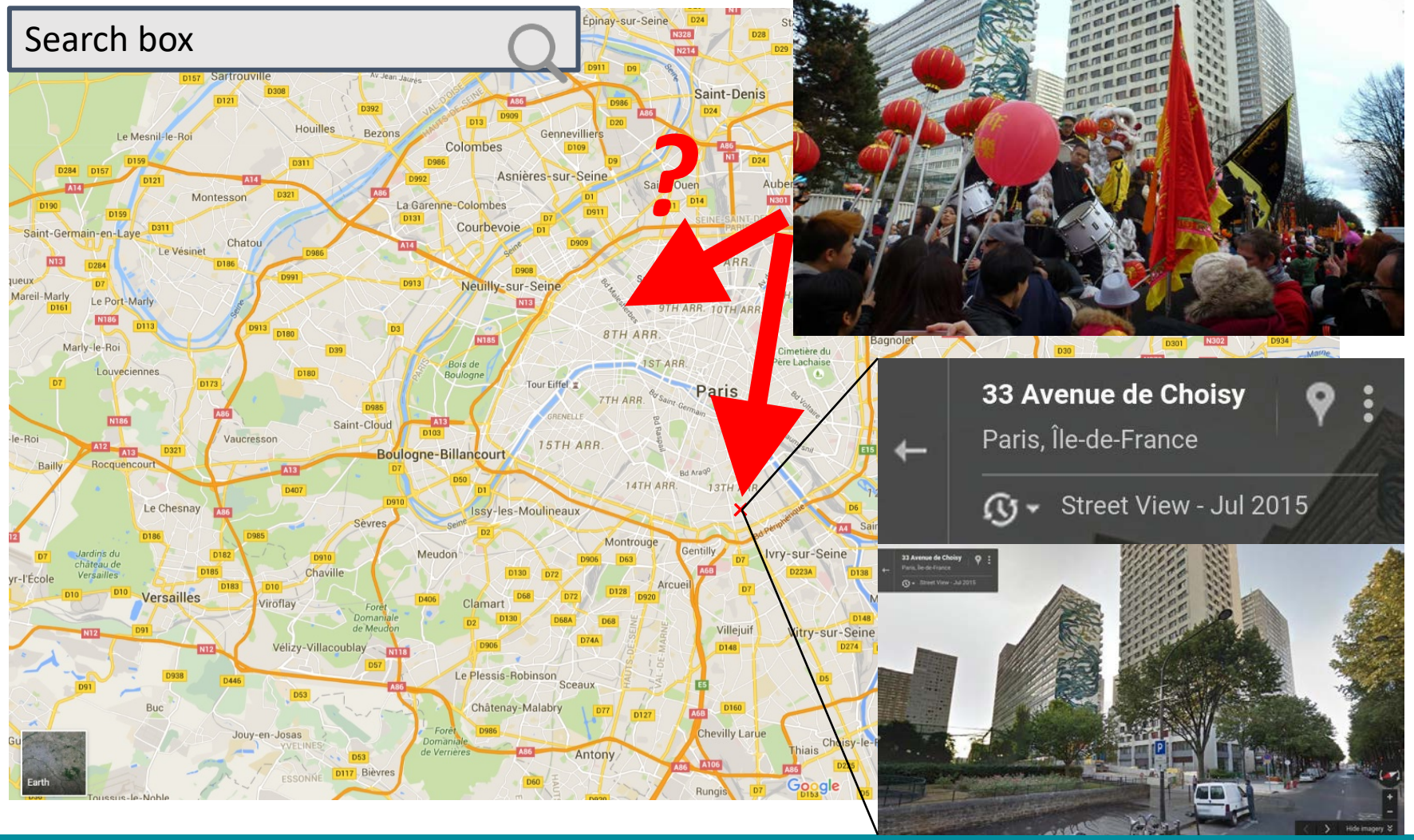
# CVPR 2021 tutorial on Cross-view and Cross-modal Visual Geo-Localization

Han-Pang Chiu, Rakesh (Teddy) Kumar

Center for Vision Technologies
SRI International, Princeton, NJ, USA
Email: han-pang.chiu@sri.com, rakesh.kumar@sri.com

June 20th, 2021

# Image based Geo-Localization



**Missing hiker found based on photo he texted from Los Angeles area mountains**

https://www.nbcnews.com/news/us-news/missing-hiker-found-based-photo-he-texted-los-angeles-area-n1264199?cid=sm_npd_nn_tw_ma

**April 15th, 2021**

# Image based Geo-Location Applications: Navigation and Mapping

## Visual Geo-localization

### Autonomous Vehicles and Robotics



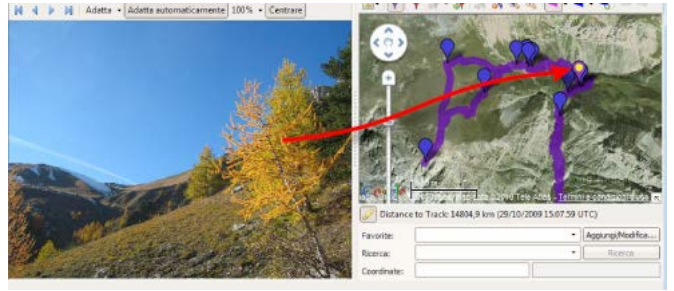

Robots


Appliances

### Augmented Reality and Person Localization

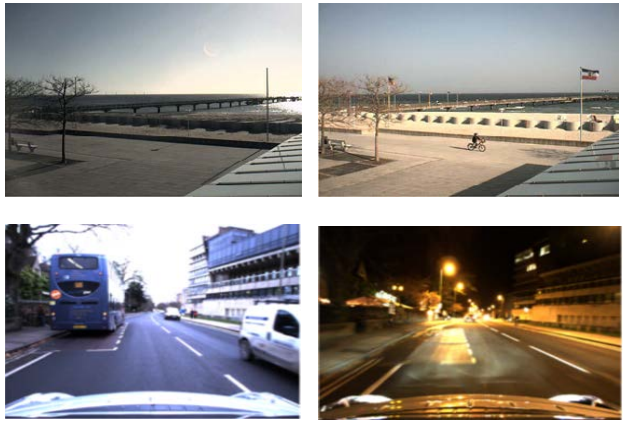

### 3D Modeling Outdoors & Indoors

### Geo-tag images



# Cross-Time, Cross-View, and Cross-Modal

Cross-Time

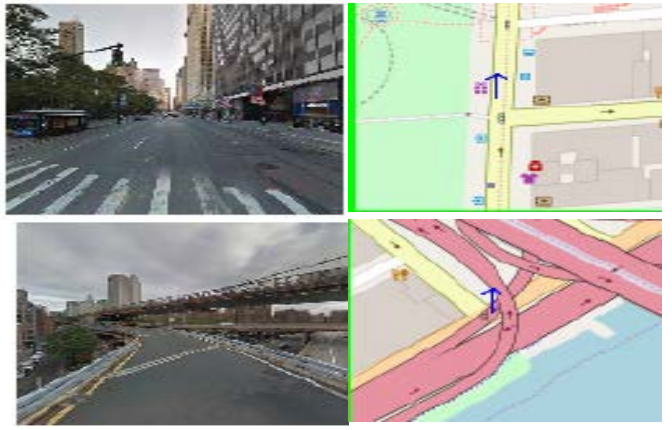


Sample Pairs (Ground RGB)

Cross-View

Sample Pairs (Ground-Aerial RGB)

Cross-Modal



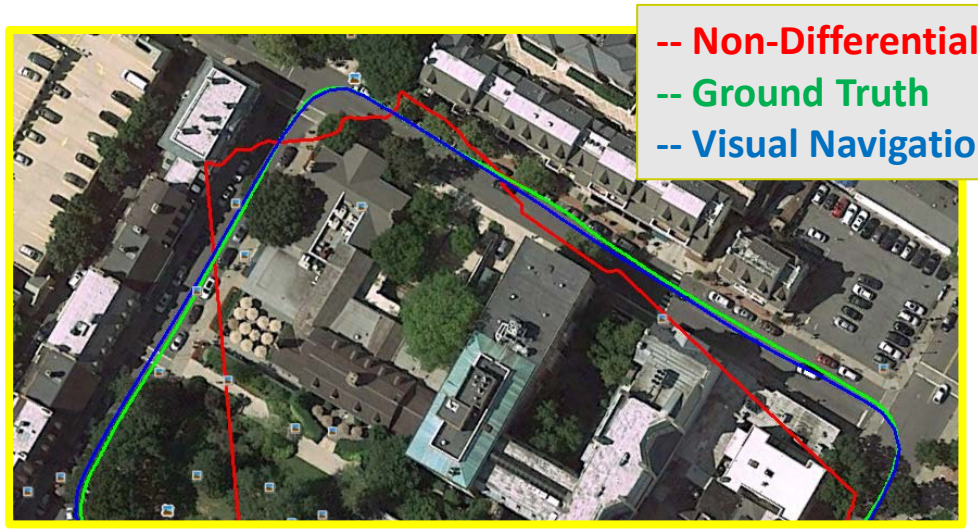
Sample Pairs (Ground RGB-OpenStreetMap)

Low — Availability of Geo-Referenced Database — High

Difficulty in Image-Based Visual Localization based on Reference Data

# Application: Navigation for Self-Driving Car

- **Challenge**: Automotive industry uses costly and bulky 3D LIDAR with GPS on each vehicle to geo-localize the vehicle within a 3D geo-referenced database, which is also constructed using LIDAR sensors.
- **Solution:** Using cameras instead of LIDARs for this geo-localization process can significantly reduce the sensor cost on each ground vehicle.
- **3D Geo-Referenced Database**:
  - Ground Camera Collection: Cross-Time
  - Aerial Camera Collection: Cross-View
  - LiDAR Collection: Cross-Modal





V. Murali et al., “Utilizing Semantic Visual Landmarks for Precise Vehicle Navigation”, IEEE ITSC 2017.
H. Chiu et al., “Sub-Meter Vehicle Navigation Using Efficient Pre-Mapped Visual Landmarks”, IEEE ITSC 2016.


6

# Application: Large-Scale Augmented Reality

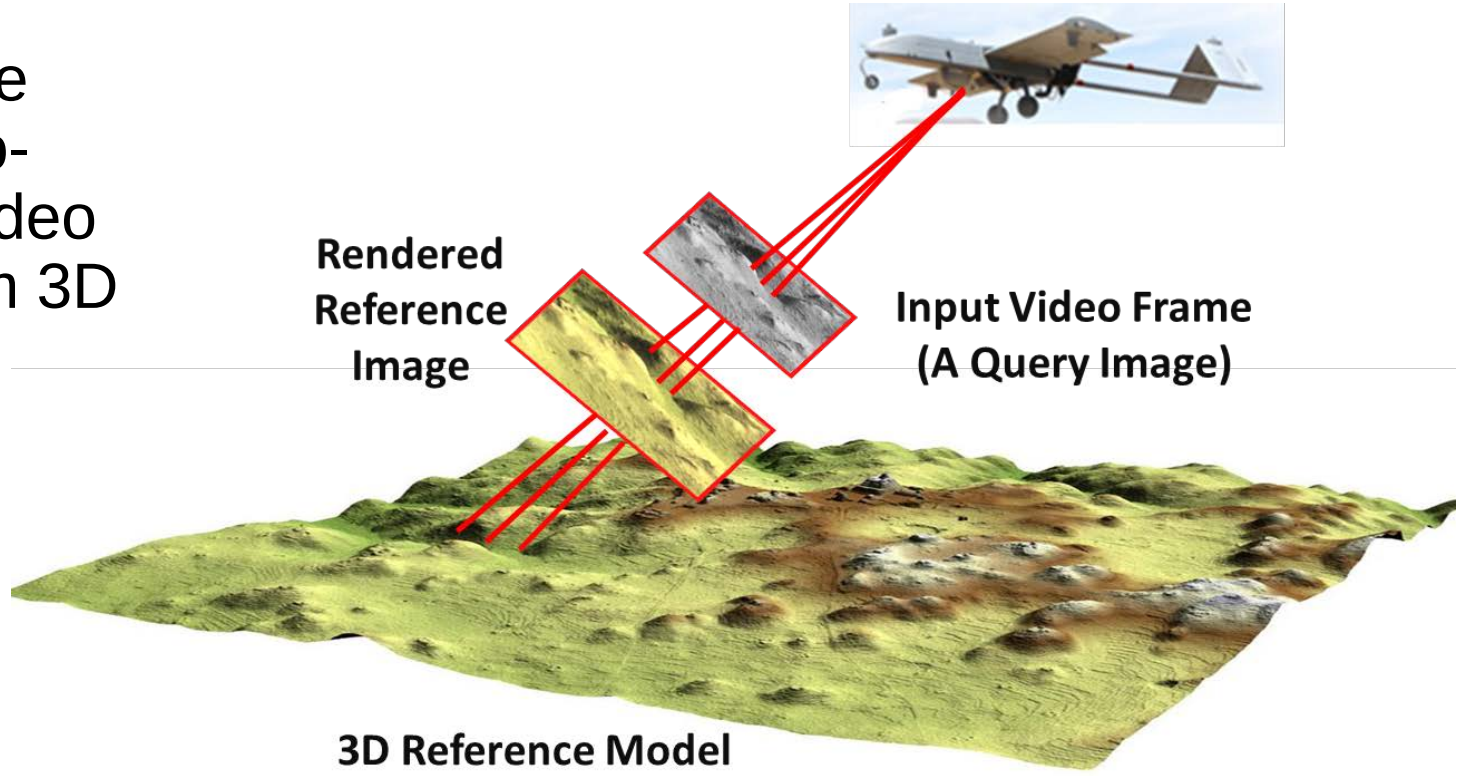
- **Challenge**: Estimating highly-accurate 3D pose (position and orientation) of the user's view is necessary for AR, which offers simulated insertions mixed with a live video feed of user's real view.
  - GPS or traditional visual navigation (SLAM) algorithms cannot provide required localization accuracy requirement to large-scale AR.
- **Solution:** Visual geo-localization to a pre-built 3D geo-referenced database offers an appealing solution to this application.


**H. Chiu et al., "Augmented Reality Driving Using Semantic Geo-Registration", IEEE VR 2018.**
**T. Oskiper et al., "Stable Vision-Aided Navigation for Large-Scale Augmented Reality", IEEE VR 2011.**


7

# Application: GPS-Denied Aerial Navigation

- **Challenge**: Provide accurate aerial navigation solutions under long-term GPS outage – navigation solution drifts quickly without GPS
- **Solution:** Incorporate absolute information through visual geo-registration from each input video frame to rendered image (from 3D referenced database) into navigation system.
- It provides accurate (3D RMS error < 10 meters) and consistent solutions on large-scale GPS-denied scenarios.




H. Chiu, A. Das, P. Miller, S. Samarasekera, R. Kumar, “Precise Vision-Aided Aerial Navigation”, IEEE IROS 2014.

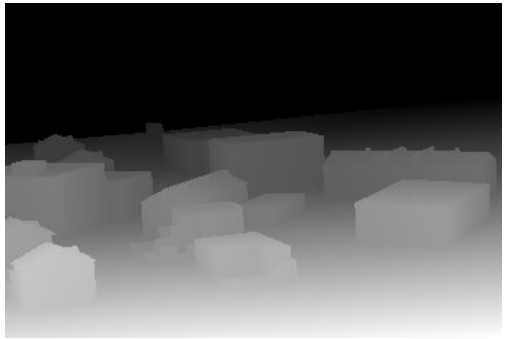

8

# Geo-registration of video to site model …

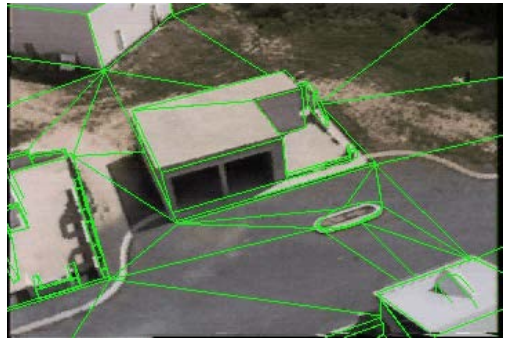
**Original Video**

(video)

**Site model**



**Geo-registration of video to site model**



(video)

(video)

**Re-projection of video after merging with model.**

# Topics:

- Cross view/ Cross time: Rakesh (Teddy) Kumar
  - Coarse Search using neural networks
  - Fine Geo-Localization and end-to-end solutions

- Same View and Cross View Geo-Localization: Mubarak Shah
  - Image geo-localization using constraint dominant sets for same view
  - GAN based approaches for alignment based on synthesis for cross view

- Large Scale Cross View Image Geo-localization: Chen Chen
  - Cross-view image geo-localization beyond one-to-one retrieval

- Cross Modal Visual Localization: Han Pang Chiu
  - Using 3D/ LIDAR data for visual localization
  - Using Maps for visual localization