# LastStep: Vision-based Bus Stop Localization and Mapping for Improving Accessibility for Blind Riders


Zheng Li
zhengli1@umbc.edu

Mahbubur Rahman
mahbub1@umbc.edu

Ryan Robucci
robucci@umbc.edu

Nilanjan Banerjee
nilanb@umbc.edu

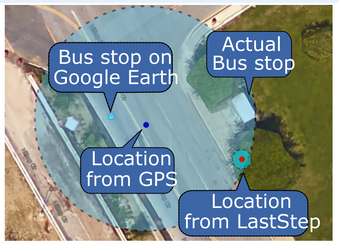
Department of Computer Science and Electrical Engineering, University of Maryland Baltimore County

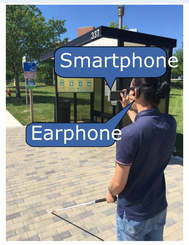

Eclipse Research Cluster
eclipse.umbc.edu

## Problem: Blind Riders Need Accurate Localization Tools

We propose a computer vision-based system, ***LastStep***, that can localize bus stops with a step-level accuracy.

- Minimizing manual search effort
- Utilizing exsiting infrastrutures
  - It complements RF-based systems where deployment is infeasible.
  - In conjunction with RF anchors, it potentially increases system reliability and/or accuracy.




Existing systems:

- GPS-based: large variance (5m), extra manual search required
- RF-based: environmental modification required

## System Architecture

*Back-end*

**Contextual Databas**

- Key points of signs
- Physical size of signs
- Fine-grained maps

**Web Interface**

- Display GE and GSV
- Label objects
- Generate bus stop fine-grained maps

*Front-end*

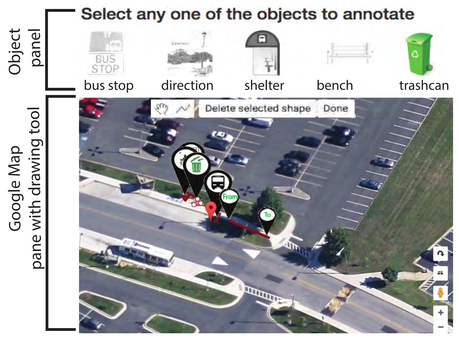
**Smartphone**

- Capture images
- Detect signs
- Estimate 3D projection
- Estimate gravity vector
- Localize on maps
- Provide verbal feedback

- The entire front-end is in one COTS smartphone.
- Minimal setup: parameters of signs can be shared between the same type (in a city/county).
- Scalable: using a combination of Google Earth and Google Street View data to generate bus stop maps.

## Crowdsourcing: Annotate the Topology Map

Human workers annotate objects in a bus stop with Google Earth data.



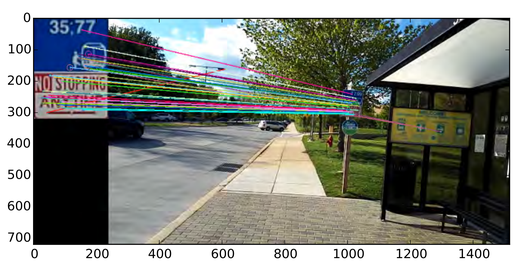

- They can "zoom in" to Google Street View to identify the objects.


Comparing the results obtained from Amazon Mechanical Turk to experts' annotations.

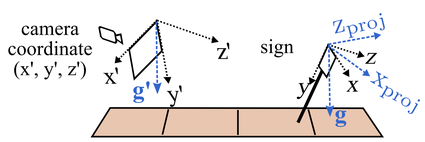
| Bus Stop | Location Difference | | Direction Difference | |
|---|---|---|---|---|
| | median (m) | average (m) | median (') | average (') |
| 1 | 0.42 | 0.47 | 40.16' | 41.17' |
| 2 | 0.79 | 3.06 | 2.35' | 1.04' |
| 3 | 0.27 | 0.79 | 17.51' | 3.18' |
| 4 | 0.51 | 0.68 | 5.16' | 8.92' |
| 5 | 2.67 | 3.23 | 9.00' | 10.32' |
| 6 | 2.07 | 3.17 | 64.58' | 19.32' |

## Computer Vision: Detect, Match, and Localize

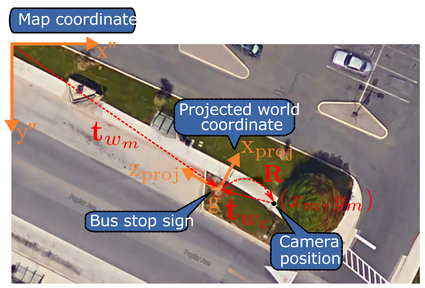
1. Detect a bus stop sign with SIFT feature



2. Estimate translation (distance), rotation matrix, and gravity vector




3. Position a user onto the map




Estimated position in the map:

$$p_m = \frac{k_w}{k_m}\left(R_\theta A P\left(-R^T\right) t_{w_c} + t_{w_m}\right)$$

where

$$p_{c_i} = K[R|t_{w_c}]p_{w_i}$$

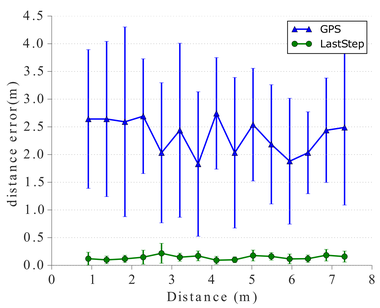
based on bus stop sign matching;

$P$ is the orthogonal projection matrix from 3D world coordinate to the 2D ground plane:

$$g \cdot v = 0$$

which comes from gravity vector estimation.


Two examples of positioning user's locations onto the annotated map




Our system estimates the distance with a smaller error and variance.

## Test on Emulated Human Participants

- Evaluating on three participants, emulated (blind-folded)
- Participants perform searching in 6 sites, with or without our system. Search time is recorded in unit of minute.

| Participant | | Sites | | | | | |
|---|---|---|---|---|---|---|---|
| | | I | II | III | IV | V | VI |
| A | w | 2.3 | 4.1 | 1.2 | 2.1 | 1.7 | 3.5 |
| | w/o | X | X | 8.7 | 10.5 | X | X |
| B | w | 1.6 | 3.5 | 1.3 | 1.9 | 2.3 | 3.1 |
| | w/o | X | X | 12.3 | X | 11.6 | X |
| C | w | X | 2.5 | 3.8 | 2.7 | 4.5 | 2.3 |
| | w/o | X | X | X | X | X | 5.4 |

- "X" means that participant admitted failure after 10 min.
- It shows a higher success rate (17/18) and higher speed (4 times faster) with our system.

## Conclusion

- We present LastStep, an accurate and infrastructure-modification-free localization system for blind riders.
- It performs localization with an step-level accuracy. (less than 0.5 meters error)
- With LastStep, participants can perform localization with a higher success rate and 4 times faster than manual searching.