FIT5046, Semester 1, Assignment 2: Research Paper Analysis Presentation

One-Shot Wayfinding Methods for Blind People via OCR and Arrow Analysis with 360-Degree Smartphone Camera

Authors:Yutaro Yamanaka, Seita Kayukawa, Hironobu Takagi, Yuichi Nagaoka, Yoshimune Hiratsuka & Satoshi Kurihara

Ziqi Pei ID: 33429472

Jiabao Li ID: 28885511

Introduction

Research area: vision technologies

Problem solved: Assisting blind people individuals in finding their way without needing to modify their surroundings or create maps beforehand.

Airport accessibility and navigation aids (Guerreiro et al., 2019)

Context-aware indoor auxiliary navigation (Li et al., 2016)

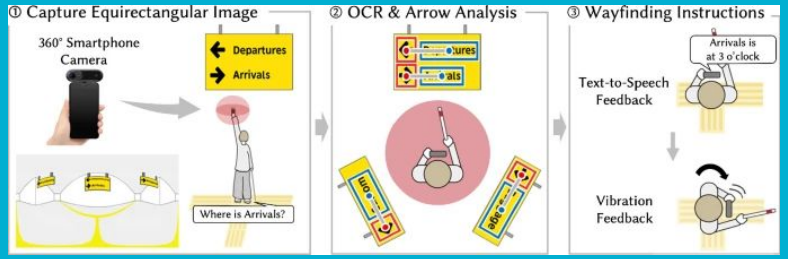
Novel:

- Take a single picture of the surroundings without needing to aim accurately.

-Identify text and arrows on signs to guide the way.

-Figure out the direction to the destination by following text arrows.

-No requirement to alter the environment or create a map in advance.



Solution

One-shot navigation method:

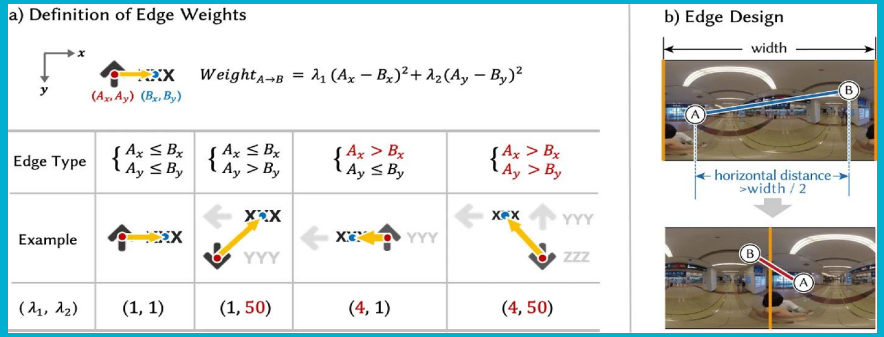
A camera takes a full-view image at once.

OCR recognizes text in logos.

A CNN spots arrows.

MST examines connections between text and arrows.

Estimates where you need to go based on what's detected.



θ = arctan(y/x) + α

θ (theta) represents the angle to be solved

arctan is the arctangent function, used to solve the corresponding angle based on the tangent value

y/x is the tangent value of a coordinate point (x,y), that is, the y coordinate divided by the x coordinate

α (alpha) is an additional angle value, which we usually call an offset or correction value

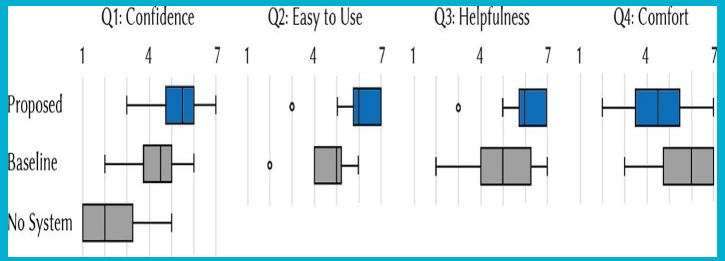
Jiabao Li

Evaluation

User Study:

1).Compared proposed 360° camera vs. baseline smartphone system

Process:
visually impaired people [2, 18, 29, 31], analyze the performance of the algorithm in different public scenarios, summarize success and failure cases, and discuss improvement plans.

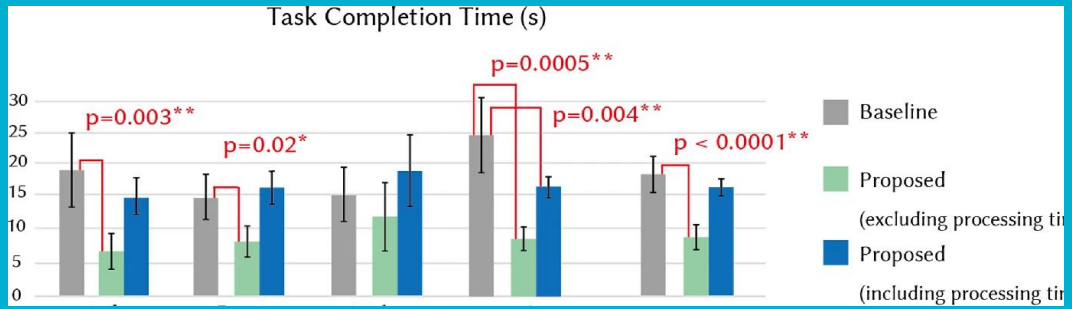


Results

Task Accuracy: Our results demonstrate that the proposed system achieved a task accuracy of 93.8%, surpassing the baseline accuracy of 75% with a p-value of 0.057.

Task Completion Time: The proposed system exhibited a significantly shorter task completion time compared to the baseline system ($p < 0.0001$). Even when accounting for processing time, the proposed system remained notably faster, especially when the target flag was positioned behind ($p = 0.004$).

Video Observation: Video recordings were used to analyze participants' behavior, particularly when they selected the incorrect tactile paving branch.



Ziqi Pei

Conclusion (and comparison)

	Key Contributions	Limitations and Challenges	Feature Research Direction
360-degree camera captures panoramic images in one shot	Innovative one-shot method to capture omnidirectional signage information	Requires users to stop at tactile paving intersections to capture images, which was not reproduced in the lab study	Conduct real-world user study in public buildings to evaluate practical usability
Optical character recognition (OCR) detects logo text	Long processing time for computer vision algorithms	Improve accuracy of OCR and arrow detection, especially for small and distant text/arrows	Guide users closer to signs before image capture to improve detection accuracy
Convolutional Neural Network (CNN) Detection of Arrows	Narrow spacing between arrows and wide spacing between text reduces arrow analysis accuracy	Reduced accuracy for distant or cluttered signs	Develop more advanced direction estimation techniques
Minimum spanning tree (MST) analysis of text and arrow relationships	Associating detected text with directional arrows	Cannot handle cases where signs do not directly face the 360-camera	Exploit signage boundaries and document layout analysis techniques to improve arrow-text association
Estimate the self-viewing direction of the destination	Estimating egocentric navigation direction without pre-mapping or environmental modifications	Lab study did not test error recovery when users choose wrong directions in complex buildings	Explore using detected cues like tactile paving directions to improve direction estimation
Voice and vibration feedback guides users	Provide multiple interface options (sonification, spatialized audio, vibration patterns, shape-changing devices) to suit different user preferences and skills	Clock-position instructions not intuitive for all users	Design seamless and practical user interface

References

(Assigned Paper)
Yamanaka, Y., Kayukawa, S., Takagi, H., Nagaoka, Y., Hiratsuka, Y., & Kurihara, S. (2021, November). One-shot wayfinding method for blind people via ocr and arrow analysis with a 360-degree smartphone camera. In International Conference on Mobile and Ubiquitous Systems: Computing, Networking, and Services (pp. 150-168). Cham: Springer International Publishing.

(Ziqi Pei)
Guerreiro, J., Ahmetovic, D., Sato, D., Kitani, K., & Asakawa, C. (2019, May). Airport accessibility and navigation assistance for people with visual impairments. In Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems (pp. 1-14).

(JiaBao Li)
Li, B., Munoz, J. P., Rong, X., Xiao, J., Tian, Y., & Arditi, A. (2016). ISANA: wearable context-aware indoor assistive navigation with obstacle avoidance for the blind. In Computer Vision–ECCV 2016 Workshops: Amsterdam, The Netherlands, October 8-10 and 15-16, 2016, Proceedings, Part II 14 (pp. 448-462). Springer International Publishing.