A Demo of DIAM: Drone-based Indoor Accessibility Mapping

Xia Su, Ruiqi Chen, Weiye Zhang, Jingwei Ma, Jon E. Froehlich
University of Washington
Seattle, Washington, USA

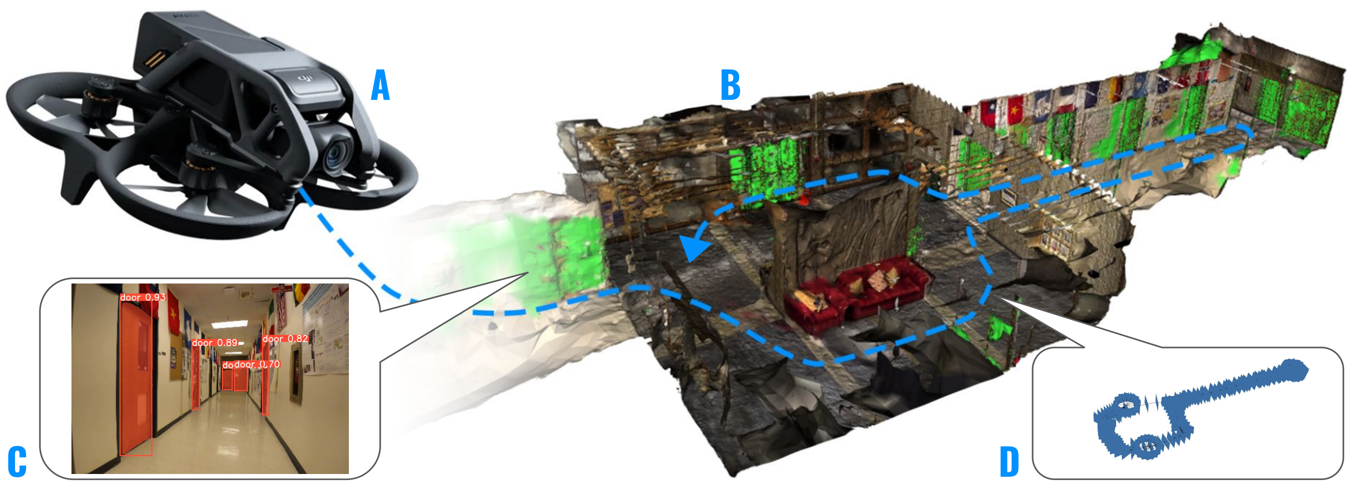

Figure 1: We introduce *Drone-based Indoor Accessibility Mapping* (DIAM), a drone-based indoor scanning system that semi-automatically maps indoor spaces and identifies key accessibility barriers/features such as stairs, elevators, and doors. (A) The DJI Avata drone used to scan indoor spaces; (B) An image-based 3D reconstruction with automatically recognized accessibility classes highlighted in green (*e.g.,* doors). (C) For computer vision, DIAM uses a few-shot, fine-tuned instance segmentation model to detect accessibility targets. (D) DIAM also estimates camera position and fly trajectory to help locate detected facilities.

ABSTRACT

Indoor mapping data is crucial for navigation and accessibility, yet such data are widely lacking due to the manual labor and expense of data collection, especially for larger indoor spaces. In this demo paper, we introduce *Drone-based Indoor Accessibility Mapping* (DIAM), a drone-based indoor scanning system that efficiently produces 3D reconstructions of indoor spaces with automatically recognized and located accessibility features/barriers such as stairs, elevators, and doors automatically. With DIAM, our goal is to scan indoor spaces quickly and generate a precise, detailed, and visual 3D indoor accessibility map. We describe DIAM's system design, present its technical capabilities, and discuss future use cases.

KEYWORDS

Indoor Accessibility Mapping, Indoor Scanning, Drone, 3D reconstruction

UIST '24, Oct 13–16, 2024, Pittsburgh, PA

ACM ISBN 978-x-xxxx-xxxx-x/YY/MM…$15.00
https://doi.org/10.1145/nnnnnnn.nnnnnnn

ACM Reference Format:
Xia Su, Ruiqi Chen, Weiye Zhang, Jingwei Ma, Jon E. Froehlich. 2024. A Demo of DIAM: Drone-based Indoor Accessibility Mapping. In *Proceedings of The ACM Symposium on User Interface Software and Technology (UIST '24).* ACM, New York, NY, USA, 3 pages. https://doi.org/10.1145/nnnnnnn.nnnnnnn

1 INTRODUCTION

Indoor accessibility mapping data, which helps disabled people navigate [4] and assess [10] remote spaces, are widely lacking [2], especially in terms of data quality and accessibility information [1, 9]. Traditional indoor mapping methods are either high-cost and time-consuming [7] or lack accessibility information [1].

In our research, we are exploring the use of consumer-level drones and state-of-the-art AI for efficient and low-cost indoor accessibility mapping. Drawing on prior work in drone-based mapping [6, 11], we present *Drone-based Indoor Accessibility Mapping* (DIAM), a drone-based indoor scanning system that efficiently produces 3D reconstructions of indoor spaces with automatically recognized and located accessibility features/barriers such as stairs, elevators, and doors automatically (1).

While our long-term vision is create a fully automated system—from autonomous flythroughs to automatic detections of accessibility barriers—our current system is a multi-step pipeline with some manual intervention. First, to map a space, the user performs a manual drone fly through with video recording. Second, DIAM

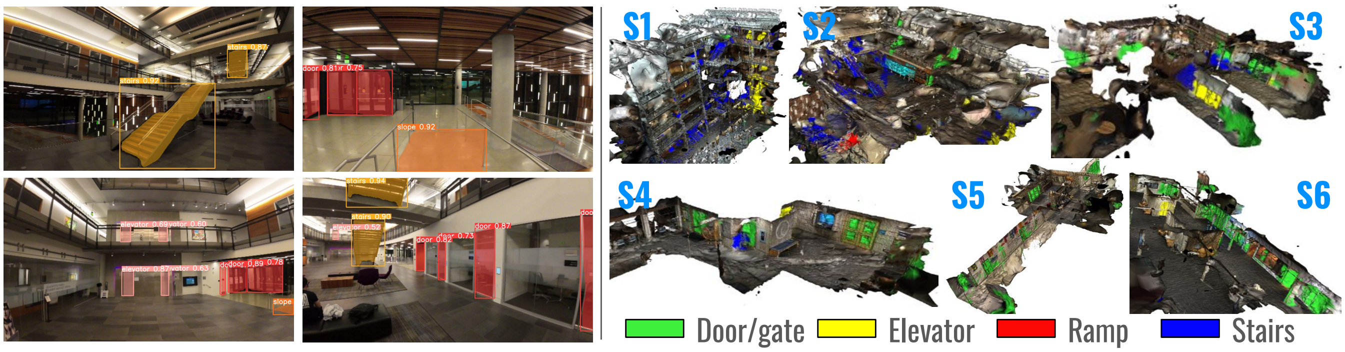

Figure 2: Left: DIAM's detection of indoor facilities. Right: DIAM reconstruction results across six atrium spaces.

Table 1: DIAM performance across six atrium spaces in automatically classifying four accessibility-related features: *door*, *elevator*, *ramp*, and *stairs*. Ground truth (GT) was collected manually. TP=*true positive*; FP=*false positive*.

Space	Door					Elevator					Ramp					Stairs				
	GT	TP	FP	Prec.	Rec.	GT	TP	FP	Prec.	Rec.	GT	TP	FP	Prec.	Rec.	GT	TP	FP	Prec.	Rec.
S1	8	7	1	0.88	0.88	12	12	0	1	1	0	-	-	-	-	6	6	0	1	1
S2	20	10	1	0.91	0.5	4	4	4	0.5	1	1	1	0	1	1	5	5	1	0.83	1
S3	10	8	0	1	0.8	4	4	0	1	1	0	-	-	-	-	1	1	0	1	1
S4	16	11	0	1	0.69	1	0	0	-	0	0	-	-	-	-	0	-	-	-	-
S5	6	6	0	1	1	1	1	0	1	1	0	-	-	-	-	1	1	1	0.5	1
S6	9	9	0	1	1	2	2	0	1	1	0	-	-	-	-	1	1	0	1	1

creates a 3D reconstruction of the indoor space using photogrammetry computed with the video data. Third, we apply the YOLOV8 [5] instance segmentation model fine tuned by few-shot learning to detect relevant accessibility barriers and features. Finally, we use raycasting to locate the segmentation results back into the constructed 3D model.

To evaluate DIAM, we mapped six indoor spaces with the DIAM pipeline, visually assess the 3D reconstructions, and calculate accessibility detection performance. As a UIST Demo, we plan to capture the conference center with DIAM and allow attendees to interactively inspect the indoor accessibility mapping results.

2 THE DIAM SYSTEM

DIAM is a multi-stage pipeline that includes drone-based RGB data collection, 3D reconstruction, and computer vision for accessibility-feature detection (*e.g., stairs, doors, ramps*).

Drone-based data collection. We use a *DJI Avata* drone (cost $759) to rapidly capture indoor space. When flying, we use the wide angle camera setting and capture videos at 4k 30FPS. In our experience, a 300 square meter 6-story atrium space takes 4 minutes.

3D reconstruction. DIAM transforms the drone-based video into image frames at 2 FPS and uses a state-of-the-art photogrammatric processing tool, called *Agisoft Metashape* [8], to perform camera pose estimation and 3D reconstruction. Figure 1B shows the reconstructed 3D model and Figure 1D shows the estimated camera pose and fly trajectory. The reconstruction was finished with *Sequential* reference preselection for photo alignment and *Depth Maps* as source data for model building.

Accessibility Feature Segmentation. To best detect accessibility-related issues in diverse indoor spaces, DIAM uses few-shot learning to fine-tune a YOLOV8[5] instance segmentation model. For each scene, we used Roboflow [3] to manually annotate 50 random video frames for four classes of key indoor accessibility features and barriers (drawn from prior work[10]): *door/gate*, *elevator*, *ramp*, and *stairs*. We then fine-tune a pre-trained YOLOV8 model for 30 epochs and use the best performing weights to predict on all frames used in 3D reconstruction, see Figure 2 left.

Localizing Accessibility Features with Ray-casting. Finally, DIAM converts the segmentation into 3D spatial information using ray tracing. Our algorithm utilizes the estimated camera parameters of *Agisoft Metashape* to map the segmentation masks to corresponding 3D positions in the reconstructed model. The casting results are highlighted with color to show the 3D position of the detected facilities (See Figure 1B and Figure 2 right).

Preliminary Evaluation We tested DIAM's performance in the atriums of six public buildings with varying sizes and layouts. For each space, we manually flew the DJI Avata drone to capture the space layout and key indoor facilities, then use the DIAM system to conduct 3D reconstruction and facility detection. On average, the fly footage is 171 seconds long (STD = 36.5s). The resulting 3D model are visualized with detected facilities highlighted in color (see Figure 2). See Table 1 for precision ($\frac{CorrectDetections}{AllDetections}$) and recall ($\frac{CorrectDetections}{AllFacilities}$) results.

3 FUTURE WORK AND CONCLUSION

In this demo paper, we introduced the DIAM system and conducted a preliminary technical evaluation in six spaces. The DIAM system showcases the technical feasibility of drone-based accessibility mapping but still needs further development to improve automation . In the future, we plan to design techniques for autonomous drone scans and to develop a full end-to-end toolchain that includes drone flying, CV processing, and model analysis. We also envision techniques to track changes in the environment over time (*e.g.,* via nightly scans) and ways for venues to easily share their DIAM-created maps online for visitors to plan and asses their visits.

REFERENCES

[1] Angela Constantinescu, Karin Müller, Claudia Loitsch, Sebastian Zappe, and Rainer Stiefelhagen. 2022. Traveling to unknown buildings: accessibility features for indoor maps. In *International Conference on Computers Helping People with Special Needs*. Springer, 221–228.

[2] Jon E. Froehlich, Anke M. Brock, Anat Caspi, João Guerreiro, Kotaro Hara, Reuben Kirkham, Johannes Schöning, and Benjamin Tannert. 2019. Grand challenges in accessible maps. *Interactions* 26, 2 (Feb. 2019), 78–81. https://doi.org/10.1145/3301657

[3] Roboflow Inc. 2023. Roboflow: End-to-End Computer Vision Tools. https://roboflow.com Accessed: 2024-07-16.

[4] Watthanasak Jeamwatthanachai, Mike Wald, and Gary Wills. 2016. Map data representation for indoor navigation a design framework towards a construction of indoor map. In *2016 International Conference on Information Society (i-Society)*. IEEE, 91–96.

[5] Glenn Jocher, Ayush Chaurasia, Alex Stoken, Jirka Borovec, Yonghye Kwon, Kalen Michael, Jiacong Fang, Zeng Yifu, Colin Wong, Diego Montes, et al. 2022. ultralytics/yolov5: v7. 0-yolov5 sota realtime instance segmentation. *Zenodo* (2022).

[6] Samer Karam, Francesco Nex, Bhanu Teja Chidura, and Norman Kerle. 2022. Microdrone-based indoor mapping with graph slam. *Drones* 6, 11 (2022), 352.

[7] Roi Otero, Susana Lagüela, Iván Garrido, and Pedro Arias. 2020. Mobile indoor mapping technologies: A review. *Automation in Construction* 120 (2020), 103399.

[8] Jin-Si R Over, Andrew C Ritchie, Christine J Kranenburg, Jenna A Brown, Daniel D Buscombe, Tom Noble, Christopher R Sherwood, Jonathan A Warrick, and Phillipe A Wernette. 2021. *Processing coastal imagery with Agisoft Metashape Professional Edition, version 1.6—Structure from motion workflow documentation*. Technical Report. US Geological Survey.

[9] Julian Striegl, Claudia Loitsch, Jan Schmalfuss-Schwarz, and Gerhard Weber. 2020. Analysis of Indoor Maps Accounting the Needs of People with Impairments. In *Computers Helping People with Special Needs: 17th International Conference, ICCHP 2020, Lecco, Italy, September 9–11, 2020, Proceedings, Part II* (Lecco, Italy). Springer-Verlag, Berlin, Heidelberg, 305–314. https://doi.org/10.1007/978-3-030-58805-2_36

[10] Xia Su, Han Zhang, Kaiming Cheng, Jaewook Lee, Qiaochu Liu, Wyatt Olson, and Jon E Froehlich. 2024. RASSAR: Room Accessibility and Safety Scanning in Augmented Reality. In *Proceedings of the CHI Conference on Human Factors in Computing Systems*. 1–17.

[11] Guyue Zhou, Ang Liu, Kang Yang, Tao Wang, and Zexiang Li. 2014. An embedded solution to visual mapping for consumer drones. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops*. 656–661.